A Model for Protein Folding and Bio-catalysis

M. Pitkänen,

October 1, 2020
Email: matpitka6@gmail.com.
http://tgdtheory.com/public_html/.
Recent postal address: Rinnekatu 2-4 A 8, 03620, Karkkila, Finland.

Contents

Abstract

The model for the evolution of genetic code leads to the idea that the folding of proteins obeys a folding code inherited from the genetic code. The flux connections between molecules containing dark matter in macroscopic quantum phase and characterized by two integers are the basic new physics element of the model.

After some trials one ends up with a general conceptualization of the situation with the identification o magnetic flux tubes as correlates of attention at molecular level so that a direct connection with TGD inspired theory of consciousness emerges at quantitative level. This allows a far reaching generalization of the DNA as topological quantum computer paradigm and makes it much more detailed. By their asymmetric character hydrogen bonds are excellent candidates for contracted magnetic flux tubes serving as correlates of attention at molecular level.

One can consider two models. For the first model the flux tubes between amino-acids are assumed to determine the protein folding.

1. The constant part of free amino-acid containing $O-H$, $O=$, and NH_2 would correspond to the codon XYZ in the sense that the flux tubes would carry the "color" representing the four nucleotides in terms of quark pairs. Color inheritance by flux tube reconnection makes this possible. For the amino-adics inside protein $O=$ and $N-H$ would correspond to YZ. Also flux tubes connecting the acceptor atoms of hydrogen bonds are required by the model of DNA as topological quantum computer. The long flux tubes between $O=$ atoms and their length reduction in a phase transition reducing Planck constant could be essential in protein-ligand interaction.

2. The model predicts a code for protein folding: depending on whether also $=O-O=$ flux tubes are allowed or not, $Y=Z$ or $Y=Z_c$ condition is satisfied by the amino-acids having $N-H--O=$ hydrogen bond. For $=O-O=$ bonds $Y-Y_c$ pairing holds true. If one identifies hydrogen bond with flux tube ($Y(n)=Z(n+k)$) the model works badly for both options. If one assumes only that the presence of a flux tube connecting amino-acids in either direction ($Y(n)=Z(n+k)$ or $Z(n)=Y(n+k)$) is a prerequisite for the formation of hydrogen bond, the model works. $Y=Z_c$ option predicts the average length of alpha bonds correctly. $Y=Z$ rule is however favored by the study of alpha helices for four enzymes: the possible average length of alpha helix is considerably longer than the average length of alpha helix if gene is the unique gene allowing to satisfy $Y=Z$ rule. The explicit study of alpha helices for four enzymes demonstrates that the failure to satisfy the condition for the existence of hydrogen bond fails rarely and at most for two amino-acids (for 2 amino-acids in single case only). For beta sheets there ar no failures for $Y=Z$ option.

3. The information apparently lost in the many-to-one character of the codon-amino-acid correspondence would code for the folding of the protein and similar amino-acid sequences could give rise to different foldings. Also catalyst action would reduce to effective base pairing and one can speak about catalyst code. The DNA sequences associated with alpha helices and beta sheets are completely predictable unless one assumes a quantum counterpart of wobble base pairing meaning that $N-H$ flux tubes are before hydrogen bonding in quantum superpositions of braid colors associated with the third nucleotides Z of codons XYZ coding for amino-acid. Only the latter option works. The outcome is very simple quantitative model for folding and catalyst action based on minimization of energy and predicting as its solutions alpha helices and beta strands.

Second model represents a diametrical opposite of the first model in the sense in that it assumes flux tube connections only between amino-acids and water molecules. These flux tubes mediate an attractive (repulsive) interaction in the case of hydrophily (hydrophoby) due to the behavior of magnetic (presumably) interaction energy as a function of Planck constant (or integers characterizing the level of dark matter) assignable to the flux tube. For hydrophoby (hydrophily) the interaction energy is minimized for long (short) flux tubes. The interaction between amino-acids is induced by this interaction in a manner analogous to how the interaction between electrons and ions induces secondary interaction between the members of a Cooper pair. The model explains the basic qualitative aspects of protein folding and the quantitative model of folding based on amino-acid-amino-acid flux tubes allows a generalization which is however discussed at numerical level.

The third proposal asks whether protein folding could be induced by the flux tube connections of protein with water's MB rather than between proteins as in the first two models. This model is certainly an idealization since S-S valence bonds are known to play an important

part in the folding. These flux tube connections could be accompanied by hydrogen bonds - even longer than usual if h_{eff} as spectrum for water as has been proposed. This involves more detailed ideas about the origin of hydrophobia and hydrophilia at the level of magnetic body (MB). Hydrophilic amino acids would tend to form flux tube connections with the MB of water unlike hydrophobic amino acids. The formation of flux tube connection would serve as a correlate for attention at molecular level.

Decade after writing this chapter the vision about the role of DNA in TGD Universe evolved with inspiration coming from the model of water memory and homeopathy and the realization that homeopathy might represent a core elemement in the functioning of immune system involving new physics in an essential manner. The key idea is that dark variants of amino-acid sequences would have coded for the 2-braiding of the magnetic flux tube patterns defining invader molecule as a dynamical process: dark proteins would mimic physically the braiding of invader molecule's magnetic body. Dark DNA sequences would have coded this braiding symbolically and their translation to dark amino-acids would transform symbolic representation to a concrete physical one. The emergence of ordinary DNA and amino-acids would have realized the same at biochemical level and amino-acid sequences representing the invader would serve as antigene attaching to the invader molecule. Not only the pattern produced in protein folding but also the temporal pattern of protein folding would be coded by DNA.

1 Introduction

The model for the evolution of the genetic code leads [K8] to the idea that the folding of proteins obeys a code inherited from the genetic code. One can imagine several variants of this code. One of them is that amino-acid behaves like the conjugate Y_c of the middle nucleotide of the codon XYZ coding for it. Conjugation for amino-acids would correspond to the hydrophilic-hydrophobic dichotomy. Also catalyst action could reduce to effective base pairing in this picture chemically and at the level of quarks associated with the flux tube to matter antimatter conjugation. The guess that amino-acid and its conjugate form pairs turned out to be wrong however and after various twists and turns I ended up with the hypothesis that the amino-acid in protein behaves like Y_cZ_c where Z corresponds to third nucleotide for some codon coding for the amino-acid.

There exists a wonderful book "Proteins: Structures and Molecular Properties" by Thomas E. Creighton published 1993 [I14] and I am grateful for Timo Immonen for possibility to use the book. In the following I freely refer to the general facts discussed in this book rather than referring separately to every detail.

1.1 Flux Tubes As Correlates Of Directed Attention At Molecular Level

After some trials one ends up with a general conceptualization of the situation with the identification of ("wormhole") magnetic flux tubes as correlates for attention at molecular level so that a direct connection with TGD inspired theory of consciousness emerges at quantitative level. Whether wormhole flux tubes or ordinary flux tubes are needed is not a completely settled question yet and the attribute "wormhole" will not be used in the sequel. This suggests a generalization of the DNA as topological quantum computer paradigm making it much more detailed.

There are too many uncertainties involved to allow anything except playing with the options that one is able to imagine. There are two kinds of flux tubes. Those between amino-acids and those between amino-acids and water molecules. The contractions of flux tubes in $\hbar$ changing phase transitions are expected to be important for protein folding and could also give rise to the interaction responsible for hydrophily and hydrophoby and be therefore highly relevant for protein folding. A basic question about which I became aware only about one year after working out the first draft of this chapter concerns the relative importance of these two kinds of flux tubes. The first model assumed that only amino-acid-amino-acid flux tubes are relevant and assumed strong selection rules inspired by DNA as TQC model. The second model which emerged year later represents an extreme in which only the flux tube connections between amino-acids and water molecules assumed to be responsible for hydrophily and hydrophoby induce the interactions between amino-acids as secondary interactions. This model works surprisingly well at qualitative level.

1.2 The Model Of Folding Code Based On Flux Tube Connections Between Amino-Acids

The first model assumes that only the flux tubes between amino-acids are relevant for protein folding.

1.2.1 What kind of atoms can be connected by flux tubes?

1. Hydrogen bonds play a key role in bio-catalysis but are not understood completely satisfactorily in the standard chemistry. Hence the basic question is whether hydrogen bonds can be regarded as or are accompanied by short (wormhole) magnetic flux tubes: note that the subject-object asymmetry of directed attention would correspond to donor-acceptor asymmetry of they hydrogen bond. If this is the case, the identification of the magnetic flux tube connection as a prerequisite for a hydrogen bond or as hydrogen bond becomes natural. At least the atoms able to form hydrogen bonds could form flux tube contacts so that the model would be very predictive and would conform with the known important role of hydrogen bonds in bio-catalysis.

2. The fact that hydrogen bonds connect base pairs suggests a generalization of the notion of base pairing stating that under some conditions amino-acids coded by XYZ and UY_cV can behave like base pairs. These amino-acic pairs correspond to pairs of amino-acid residues which are hydrophilic *resp.* hydrophobic and hydrophobic residue do not form hydrogen bonds in general. These flux tubes would thus be more general and in general long. The model for DNA as topological quantum computer requires this kind of flux tubes and they would in general connect atoms or molecules which act as acceptors in hydrogen bonding: $O =$ atom in amino-acid and aromatic ring are basic examples.

3. If one assumes that both $N-H$ and $O =$ associated with the constant part of the amino-acid can act as flux tube terminals and represent Z and Y nucleotides of the codon XYZ coding for the amino-acid, one obtains $Y = Z$ pairing of $O = -O =$ flux tubes are allowed and $Y = Z_c$ pairing if only hydrogen bond like pairings are allowed.

1.2.2 Color inheritance by a reconnection of flux tubes

1. There should exist some mechanism allowing amino-acids to inherit the base pairing property from the tRNAs associated with them so that one can identify amino-acid with the middle nucleotide of the codon coding it. If tRNA middle nucleotide is connected to $O =$ of the amino-acid, this becomes possible since the reconnection of flux tubes preserves the "color" of the flux tubes coded by (A,T,G,C) that is by the quark or anti-quark coding for the nucleotide. The temporary formation of a hydrogen bond between $N - H$ and $O =$ of two amino-acids as in the case of alpha helix would allow $N - H$ to inherit the conjugate of the color associated with $O =$. Alternative interpretation is that this hydrogen bond is possible only if the predetermined color of $N-H$ is consistent with the inherited one. The inheritance of flux tube color would be a completely general mechanism and even the donor atoms in the residues of amino-acids could inherit the color of $O =$ in this manner.

2. A possible interpretation for the fixing of the flux tube color is in terms of quantum measurement selecting one color from quantum superposition in the reconnection process. This would mean that the unitary process can bring superposition back and reconnection process can change the inherited color. The hydrogen bonds between water molecules could correspond to quantum superpositions of different colors. This superposition property might relate to the wobble base pairing phenomenon for the third nucleotide in tRNA.

1.2.3 Folding code

The identification of $N - H$ as a representation for the conjugate of the third nucleotide Z means that amino-acids would remember which codon coded them. If only hydrogen bond like flux tubes are allowed, flux tubes can connect only amino-acids satisfying $Y = Z_c$. If $= O - O =$ flux tubes are allowed $Y = Z$ rule favored by the model of DNA as topological quantum computer

follows. The isospin symmetry of the third nucleotide implies that both rules are quite flexible. If one identifies hydrogen bond with flux tube ($Y(n) = Z(n+k)$) the model works badly for both options. If one assumes only that the presence of a flux tube connecting amino-acids in either direction ($Y(n) = Z(n+k)$ or $Z(n) = Y(n+k)$) is a prerequisite for the formation of hydrogen bond, the model works. $Y = Z$ rule is favored by the study of five enzymes: the possible average length of alpha helix is considerably longer than the average length of alpha helix if gene is the unique gene allowing to satisfy $Y = Z$ rule. The explicit study of alpha helices and beta sheets for these enzymes demonstrates that the failure to satisfy the condition for the existence of hydrogen bond fails rarely and at most for two amino-acids (for 2 amino-acids in single case only).

$Y = Z$ rule could mean a solution of the basic problem of proteonics: Do genes determine the folding of proteins and how this would take place? The interpretation would be that the information loss suggested by the many-to-one character of the genetic code is only apparent. The apparently lost information which corresponds to the $A - G$ and $T - C$ symmetries of the third nucleotide codes for the hydrogen bonding and hence for the folding of the protein. The model in its most stringent form is easy to kill since in the case of alpha helices and beta sheets the hydrogen bonding fixes completely the DNA sequence coding for the protein. A weaker variant of the model based on quantum variant of wobble base pairing: in this case there are no conditions on DNA sequence. It turns out that only this variant works. Hence hydrogen bonded amino-acid behave as if they were coded by the unique codon consistent with $Y = Z$ rule.

1.2.4 Quantitative model

The quantitative model relies on the assumption that the contribution of a flux tube connecting two amino-acids to the potential energy depends only on the distance between the molecules in question. The extremals of the total interaction energy are same for any choice of the potential and only the absolute minimum of the interaction energy depends on the choice of the potential. The simplest potential corresponds to harmonic oscillator potential and would explain formation of alpha helices and beta sheets and with the fact that hydrophilic and hydrophobic residues tend to have a large distance and only few flux tube contacts. For large Planck constant also long flux tubes could correspond to attractive harmonic oscillator potential. Also the contribution of other interactions between neighboring amino-acids are expected to be present but are neglected in the simplest model. The model predicts alpha helices and beta sheets, and more generally, periodic structures, as solutions to energy minimization equations.

1.3 A Model For Protein Folding Based On Flux Tubes Between Amino-Acids And Water Molecules

This proposal represents a diametrical opposite of the first model in the sense in that it assumes flux tube connections only between amino-acids and water molecules. These flux tubes mediate an attractive (repulsive) interaction in the case of hydrophily (hydrophoby) due to the behavior of magnetic (presumably) interaction energy as a function of Planck constant (or integers characterizing the level of dark matter) assignable to the flux tube. For hydrophoby (hydrophily) the interaction energy is minimized for long (short) flux tubes. The interaction between amino-acids is induced by this interaction in a manner analogous to how the interaction between electrons and ions induces secondary interaction between the members of a Cooper pair. The model explains the basic qualitative aspects of protein folding and the quantitative model of folding based on amino-acid-amino-acid flux tubes allows a generalization which is however discussed at numerical level.

1.4 Protein folding, hydrophoby and hydrophily, and molecular attention

The third proposal asks whether protein folding could be induced by the flux tube connections of protein with water's MB rather than between proteins as in the first two models. This model is certainly an idealization since S-S valence bonds are known to play an important part in the folding. These flux tube connections could be accompanied by hydrogen bonds - even longer than usual if h_{eff} as spectrum for water as has been proposed. This involves more detailed ideas about

the origin of hydrophobia and hydrophilia at the level of magnetic body (MB) discussed more quantitatively in [**?**]. Hydrophilic amino acids would tend to form flux tube connections with the MB of water unlike hydrophobic amino acids. The formation of flux tube connection would serve as a correlate for attention at molecular level.

1.5 Postlude

The above summarized efforts are just the first attempts to apply TGD views in order to understand protein folding, and must be taken just as excercises without deeper vision about the meanings of protein folding and folding code assuming it exists.

Decade after writing this chapter the vision about the role of DNA in TGD Universe evolved with inspiration coming from the model of water memory and homeopathy and the realization that homeopathy might represent a core elemement in the functioning of immune system involving new physics in an essential manner. The key idea is that dark variants of amino-acid sequences would have coded for the 2-braiding of the magnetic flux tube patterns defining invader molecule as a dynamical process: dark proteins would mimic physically the braiding of invader molecule's magnetic body. Dark DNA sequences would have coded this braiding symbolically and their translation to dark amino-acids would transform symbolic representation to a concrete physical one. The emergence of ordinary DNA and amino-acids would have realized the same at biochemical level and amino-acid sequences representing the invader would serve as antigene attaching to the invader molecule. Not only the pattern produced in protein folding but also the temporal pattern of protein folding would be coded by DNA.

It would be fascinating if the vision about the role of flux tube connections would generalize to interactions of all molecules in living matter. The mere selection rules would mean hidden simplicity behind extremely complex looking interactions in living matter. The model for protein folding and catalytic action described in the original version of this chapter was the first attempt in this direction. At the end of the chapter an improvement of the model inspired by recent considerations is suggested.

Several persons have helped me in writing this chapter. I want to express my gratitude to Ulla Mattfolk for informing about the idea of protein folding code and to Dale Trenary for interesting discussions, for suggesting proteins which could allow to test the model, and for providing concrete help in loading data help from protein data bank. Also I want to thank Timo Immonen for discussions and for loaning the excellent book "Proteins: Structures and Molecular Properties" of Creighton. I am also grateful for Pekka Rapinoja for writing the program transforming protein data file to a form readable by MATLAB.

The appendix of the book gives a summary about basic concepts of TGD with illustrations. Pdf representation of same files serving as a kind of glossary can be found at `http://tgdtheory.fi/tgdglossary.pdf` [**?**].

2 A Model For Flux Tubes

Biochemistry represents extremely complex and refined choreography. It is hard to believe that this reduces to a mere unconscious and actually apparent fight for chemical survival. In TGD Universe consciousness would be involved even at the molecular level and magnetic body would be the choreographer whose dance would induce the molecular activities. This picture combined with the idea of standard plugs and terminals at which flux tubes end, leads to a to a picture allowing to formulate a model for protein folding.

2.1 Flux Tubes As A Correlates For Directed Attention

Molecular survival is the standard candidate for the fundamental variational principle motivating the molecular intentional actions. There is entire hierarchy of selves and the survival at the higher level of hierarchy would force co-operation and altruistic behavior at the lower levels. One might hope that this hypothesis reduces to Negentropy Maximization Principle [K2], which states that the information contents of conscious experience is maximized. If this picture is accepted, the evolution of molecular system becomes analogous to the evolution of a society.

Directed attention is the basic aspect of consciousness and the natural guess would be that directed attention corresponds to the formation of magnetic flux tubes between subject and target. The directedness property requires some manner to order the subject and target.

1. The ordering by the values of Planck constant is what first comes in mind. The larger space-time sheet characterized by a larger value of Planck constant and thus at a higher level of evolutionary hierarchy would direct its attention to the smaller one.

2. Also the ordering by the value of p-adic prime characterizing the size scale of the space-time sheet could be considered but in this case directedness could be questioned.

3. Attention can be directed also to thoughts. Could this mean that attention is directed from real space-time sheets to p-adic space-time sheets for various values of primes but not vice versa? Or could the direction be just the opposite at least in the intentional action transforming p-adic space-time sheet to real space-time sheet? Perhaps directions are opposite for cognition.

The generation of (wormhole) magnetic flux tubes could be the correlate for the directed attention, not only at molecular level, but quite generally. Metaphorically, the strands of braid would be the light rays from the eyes of the perceiver to the target and their braiding would code the motions of the target to a topological quantum computation like activity and form a memory representation at least. The additional aspect of directed attention would be the coloring of the braid strands, kind of coloring for the virtual light rays emerging from the eyes of the molecular observer. In the case of DNA this can induce a coloring of braid strands emerging from amino-acids and other molecules so that it would indeed become possible to assign to free amino-acid the conjugate of the codon XYZ coding for it.

Attention can be also redirected. For this process there is a very nice topological description as a reconnection of flux tubes. What happens is that flux tubes $A \to B$ and $C \to D$ fuse for a moment and become flux tubes $A \to D$ and $C \to B$. This process is possible only if the strands have the same color so that the values of the quark charges associated with A and B are the same.

1. Reconnection process can modify TQC programs. For instance, in the case of the flux tubes coming from nucleotides X and X_c and ending to the lipid layer this process means that X and X_c and corresponding lipids become connected and genome builds memory representation about this process via similar link.

2. Reconnection process makes also possible what might be called color inheritance allowing amino-acids to inherit the conjugate colors of the nucleotides of the codon coding it.

3. DNA would have memory representation about molecular processes via these changing braiding topologies, and one could say that these molecular processes reflect the bodily motions of the magnetic body. Entire molecular dynamics of the organism could represent an enormous TQC induced by the motor activities of the magnetic body. At the level of sensory experience similar idea has been discussed earlier [K5]: out of body experiences (OBEs) and illusions such as train illusion could be understood in terms of motor action of magnetic body inducing virtual sensory percepts.

Attention can be also switched on and off. Here the structure of the lipid ends containing two nearby situated $=O$: s suggest the mechanism: the short flux tube connecting $=O$: s disappears by reconnection mechanism with a pair of hydrogen bonded water molecules leading to a shortcut of the connecting flux tubes to $=O--H_2O$ hydrogen bonds. The minimization of Coulomb interaction energy at each end implies that re-appearance of the flux tubes creates a short flux tube with the original strand color.

2.2 Does Directed Attention Generate Memory Representations And TQC Like Processes

Directed attention induces braiding if the target is moving and changing its shape. This gives rise to a memory representation of the behavior of the object of attention and also to a TQC like process. A considerable generalization of TQC paradigm suggests itself.

Tqc could be induced by the braiding between DNA and lipids, DNA and proteins via folding processes, DNA RNA braiding and braiding between DNA and its conjugate, DNA and protein braiding. The outcome of TQC would be represented as the temporal patterns of biochemical concentrations and rates and there would be hierarchy of p-adic time scales and those associated with the dark matter hierarchy.

For instance, the protein content of lipid membranes is about 50 per cent and varies between 25-75 per cent so that protein folding and lipid flow could define TQC programs as self-organization patterns. The folding of protein is dynamical process: alpha helices are created and disappear in time scale of 10^{-7} seconds and the side chains of protein can rotate.

The details of the TQC like process depend on what one assumes. The minimal scenario is deduced from the transcription and translation processes and from the condition that magnetic body keeps control or at least keeps book about what happens using genome as a tool. The picture would be essentially what one might obtain by applying a rough model for web in terms of nodes and links. The reader is encouraged to use paper and pencil to make the following description more illustrative.

1. Assume that mRNA and DNA remain connected by flux tubes after transcription and that only reconnection process can cut this connection so that mRNA inherits the conjugate colors of DNA. Assume same for mRNA and tRNA. Assume that amino-acid associated with tRNA has similar flux tube connections with the nucleotides of tRNA. Under these assumptions amino-acid inherits the conjugate colors of DNA nucleotides via the connection line DNA-mRNA-tRNA-amino-acid faith-fully if all links are correspond to quark pairs rather than their superpositions. Wobble pairing for Z nucleotide could actually correspond to this kind of superposition.

2. One can consider several options for the amino-acid-acid DNA correspondence but trial-and-error work showed that a realistic folding code is obtained only if X, Y, and Z correspond to $O-H$, $O=$, and NH_2 in the constant part of free amino-acid. During translation the formation of the peptide bond between amino-acids dehydration leads to a loss of $O-H$ and one H from NH_2. The flux tube from tRNA to $O-H$ becomes a flux tube to water molecule inheriting the color of X so that $O=-NH_2$ of the amino-acid inside protein represents the conjugate of YZ.

3. Hydrogen bonding between $O=$ and NH of n: th and $n+k$: th amino-acids inside alpha helices and n: th and $n+1$: th amino-acids inside beta strands reduces effectively to base pairing characterized by $Y=Z$ rule. Assuming that flux tube is only a prerequisite for the formation of hydrogen bond, $Y(n)=Z(n+k)$ or $Z(n)=Y(n+k)$ allows the existence of hydrogen bond. The identification of hydrogen bond with flux tube gives a more stringent condition $Y(n)=Z(n+k)$. The first option is favored. Either condition is extremely restrictive condition on the gene coding the amino-acid unless one assumes quantum counterpart of wobble base pairing for mRNA or tRNA-amino-acid pairing in the case of Z nucleotide (as one indeed must do). Note that the $O=$ atom of the amino-acid is in a special role in that it can have hydrogen bond flux tubes to donors and flux tube connections with $O=$: s of other amino-acids, the residues of amino-acids containing acceptors (say $O=$ or aromatic ring), and with the aromatic rings of say ATP.

4. The recombination process for two conjugate DNA-mRNA-tRNA-amino-acid links can transform the flux tubes in such manner that one obtains link between the $=O$: s of amino-acids A_1 and A_2 characterized by Y and Y_c. Besides hydrogen bonding this mechanism could be central in the enzyme substrate interaction. The process would pair tRNAs corresponding to Y and Y_c together to give DNA-mRNA-tRNA-tRNA-mRNA-DNA link providing a memory representation about amino-acid pairing A_1-A_2. One could say that magnetic body creates with the mediation of the genome dynamical TQC programs to which much of the bio-molecular activity reduces. Not all however, since two amino-acid pairs A_1-A_2 and A_3-A_4 can recombine to A_1-A_4 and A_3-A_2 without DNA knowing anything about it. Magnetic body would however know.

5. The constant part of non-hydrogen bonded amino-acid inside protein would behave like Y_cZ_c if amino-acid is coded by XYZ. The $COOH$ end of protein would behave like $X_cY_cZ_c$.

Also flux tubes connecting the residue groups become possible and protein does not behave like single nucleotide anymore. By color inheritance everything resulting in the reconnection process between $O =$ and NH_2 and residues reduces in a well-defined sense to the genetic code.

2.3 Realization Of Flux Tubes

The basic questions about flux are following. Where do they begin, where do they end, and do they have intermediate plugs which allow temporary cutting of the flux tube.

2.3.1 Where do flux tubes begin from?

The view about magnetic body as a controller of biological body using genome as a control tool suggests that DNA is to a high degree responsible for directed attention and other molecules as targets so that flux tubes emanate from DNA nucleotides. The reason would be that the aromatic cycles of DNA correspond to larger value of Planck constant. Some chemical or geometric property of DNA nucleotides or of DNA nucleotides of DNA strand could raise them to the role of subject. Aromatic cycle property correlates with the symmetries associated with large value of Planck constant and is the best candidate for this property.

If this picture is accepted then also some amino-acid residues might act as subjects/objects depending on the option. Phe, His, Trp, Tyr contain aromatic cycle. The derivatives of Trp and Tyr act as neurotransmitters and His is extremely effective nucleophilic catalyst. This would make possible more specific catalytic mechanisms through the pairing of Phe, His, Trp, and Tyr with residues having flux tube terminals.

This raises the question about the physical interaction determining the color of the strand emerging from the aromatic cycle. The interaction energy of quark at the end of flux tube with the classical electromagnetic fields of nuclei and electrons of the ring should determine this. The wormhole contact containing quark/antiquark at the throat at space-time sheet containing nuclei and electrons could also de-localize inside the ring. One of the earliest hypothesis of TGD inspired model for living matter was that wormhole Bose-Einstein condensates could be crucial for understanding of the behavior of biomolecules [K6]. Wormhole throats with quark and antiquark at their throats appear also in the model of high T_c superconductivity [K1]. As far as couplings are considered, these wormhole contacts are in many respects analogous to the so called axions predicted by some theories of elementary particle physics. The wormhole contact like property is by no means exceptional: all gauge bosons correspond to wormhole contacts in TGD Universe.

The only manner for the electronic space-time sheet to feed its electromagnetic gauge flux to larger space-time sheets using exactly two wormhole contacts is to use wormhole contacts with $\overline{u}$ and d at their "upper" throat (T, G). For proton one would have $\overline{d}$ and u at their "upper" throat (A, C). The presence of electron or proton at nucleotide space-time sheet near the end of flux tube might allow to understand the correlation. The transfer of electrons and protons between space-time sheets with different p-adic length scale is basic element of TGD based model of metabolism so that there might be some relation.

2.3.2 Acceptors as plugs and donors as terminals of flux tubes?

Standardization constraint suggests that flux tubes are attached to standard plugs and terminals. The explicit study of various biological molecules and the role of water in biology gives some hints.

1. An attractive idea is that $= O$ serves as a plug to which flux arrives and from which it can also continue. For the minimal option suggested by hydrogen bonding $O =$ could be connected to two donors and $O =$ could not be connected to $O =$. The assumption that the flux tube can connect also two $O =$: s represents a hypothesis going outside the framework of standard physics. A stronger assumption is that all acceptors can act as plugs. For instance, the aromatic rings of DNA nucleotides could act as acceptors and be connected to a sequence of $O =$ plugs eventually terminating to a hydrogen bond.

2. Donors such as $O - H$ would in turn correspond to a terminal at which flux tube can end. One might be very naive and say that conscious bio-molecules have learned the fundamental

role of oxygen and water in the metabolism and become very attentive to the presence of $=O$ and $O-H$. $=O$ appears in $COOH$ part of each amino-acid so that this part defines the standard plug. $=O$ appears also in the residues of Asp, Glu, Asn, Gln. $O-H$ groups appear inside the residues of Asp, Glu and Ser, Thr.

3. Hydrogen bonds $X-H--Y$ have the basic defining property associated with directed attention, namely the asymmetry between donor X and acceptor Y. Hence there is a great temptation consider the possibility that hydrogen bonds correspond to short flux tubes, that flux tubes could be seen as generalized hydrogen bonds. Quite generally, Y could be seen as the object of directed attention of X characterized by larger value of Planck constant. The assumption that two $O=$: s, or even two acceptors of a hydrogen bond, can be connected by a flux tube means more than a generalization of hydrogen bond the connection with a donor would correspond only to the final step in the sequence of flux tubes and plugs giving rise to a directed attention.

4. This hypothesis makes the model rather predictive. For instance, $N-H$, NH_2, $O-H$ and much less often $C-H$ and $S-H$ are the basic donors in the case of proteins whereas $O=$, $-O-$, $-N=S-S$, $-S^-$ and aromatic rings are the basic acceptors. Reconnection process should be involved with the dynamics of ordinary hydrogen bonding. Reconnection process implies inheritance of the flux tube color and means a realization of the symbol based dynamics. It turns out that this hypothesis leads to a model explaining basic qualitative facts about protein folding.

2.4 Flux Tubes And DNA

The model of DNA as topological quantum computer gives useful guide lines in the attempt to form a vision about flux tubes. It was assumed that braid strands defined by "wormhole magnetic" flux tubes join nucleotides to lipids and can continue through the nuclear or cell membrane but are split during TQC. The hydrophilic ends of lipids attach to water molecules and self-organization patterns for the water flow in gel phase induce a 2-D flow in the lipid layer which is liquid crystal defining TQC programs at the classical level as braidings. The flow indeed induces braiding if one assumes that during topological computation the connection through the cell membrane is split and reconnected after the halting of TQC.

The challenge is to understand microscopically how the flux tube joins DNA nucleotide to the phospholipid [I9]. Certainly the points at which the flux tubes attach should be completely standard plugs and the formation of polypeptide bonds is an excellent guide line here. Recall that phospholipid, the TQC dancer, has two hydrophobic legs and head. Each leg has at the hydrophilic end O=C-O-C part joining it to glyceride connected to monophosphate group in turn connected to a hydrophilic residue R. The most often appearing residues are serine, inositol, ethanolamine, and choline. Only three of these appear in large quantities and there is asymmetry between cell exterior and interior.

Let us denote by $=O_1$ and $=O_2$ the two oxygens (maybe analogs of right and left hemispheres!) in question. The proposal is that DNA nucleotide and $=O_1$ are connected by a flux tube: the asymmetry between right and left lipid legs should determine which of the legs is "left leg" and which $O=$ is the "left brain hemisphere". $=O_2$, the "holistic right brain hemisphere", connects in turn to the flux tube coming from the other symmetrically situated $=O_2$ at the outer surface of the second lipid layer. Besides this $=O_1$ and $=O_2$ are connected by a flux tube serving as switch on both sides of the membrane.

During TQC the short $O=-O=$ flux tube would experience reconnection with a flux tube acting as hydrogen bond between water molecules so that the connection is split and $O=$: s form hydrogen bonds. The reversal of this reconnection creates the connection again and halts the computation. The lipid residue R couples with the flow of the liquid in gel phase. Since $=O$ is in question the quark or antiquark at the end can correspond to the DNA nucleotide in question. The necessary complete correlation between quark and antiquark charges at the ends of flux tubes associated with $=O_1$ and $=O_2$ can be understood as being due to the minimization of Coulomb interaction energy.

If one is ready to accept magnetic flux tubes between all acceptors then the aromatic rings of nucleotides known to be acceptors could be connected by a flux tube to the $O=$ atom of the lipid

or to some intermediate O = atom. The phosphate groups associated with nucleotides of DNA strand contain also $= O$, which could act as a plug to which the flux tube from the nucleotide is attached. The detailed charge structure of the aromatic ring(s) should determine the quark-nucleotide correspondence. The connection line to the lipid could involve several intermediate O = plugs and the first plug in the series would be the O = atom of the monophosphate of the nucleotide.

There is a strong temptation to assume that subset of XYP molecules, $X = A, G, T, C$, $Y = M, D, T$ act as standard plugs with X and phosphates connected by flux tubes to a string. This would make possible to engineer braid strands from standard pieces connected by standard plugs. DNA nucleotide XMP would have flux tube connection to the aromatic ring of X and the $O =$ of last P would be connected to next plug of the communication line. If so, a close connection with metabolism and topological quantum computation would emerge. Phosphorylation would be an absolutely essential for both metabolism and buildup of connection lines acting as braid strands. $O = -O =$ flux tubes could also act as switches inducing a shortcut of the flux tube connection by reconnecting with a hydrogen bond connecting two water molecules. This is an essential step in the model for how DNA acts as topological quantum computer.

This picture would fit with the fact that XYP molecules, in particular AMP, ADP, and ATP, appear in bio-molecules involved with varying functions such as signalling, control, and metabolism. $= O$ might act as a universal plug to which flux tubes from electronegative atoms of information molecules can attach their flux tubes. This would also provide a concrete realization of the idea that information molecules (neurotransmitters, hormones) are analogous to links in Internet [K3]: they would not represent the information but establish a communication channel. The magnetic flux tube associated with the information molecule would connect it to another cell and by the join to $= O$ plug having flux tube to another cell, say to its nucleus, would create a communication or control channel.

2.5 Introns And DNA-Protein Attachment

An example is the situation in which protein acts as an enzyme attaching on DNA. Suppose that this process effectively reduces to a base pairing between amino-acid and DNA nucleotide. Protein can attach to any portion of DNA. The simplest interaction is the attachment to the gene coding for the amino-acid itself but much more general enzymatic interactions are possible. It must be however noticed that DNA sequence coding for given amino-acid sequences is considerably longer than amino-acid sequence: the sequence coding for 10 amino-acids is about 10 nm long whereas the corresponding straight amino-acid strand is about 4.7 nm long. It is known that DNA can change its conformation from strand during enzyme-DNA action [I14], and the contraction of DNA strand might make possible to have enzyme-DNA interaction involving fusion along several subsequent amino-acids. This kind of mechanism might work also in the case that attachment region corresponds to several exons. There is however no need to assume that subsequent amino-acids are form a contact with DNA.

One can of course ask whether genes containing introns tend to code for proteins which are used for topological quantum computations. Introns, perhaps the repeating sequences with no obvious function, would have at least this useful function but very probably much more useful ones too (they are now known to be transcribed to RNA and TGD suggest that language corresponds to intronic gene expression). The emergence of introns might be somewhat like the emergence of information society.

The folding of proteins tends to be conserved in the evolution whereas primary structure can change quite a lot apart from some amino-acids critical for enzymatic action. This confirms with the effective base pairing interaction between amino-acids and DNA to be discussed later and would mean that DNA-amino-acid TQC programs are rather robust against mutations.

3 Model For The Folding Code Based On Interactions Mediated By Flux Tubes Between Amino-acids

The model for the protein folding to be discussed in this section relies on the hypothesis that dark flux tube connections between amino-acids and their contractions in $\hbar$ changing phase transitions

determine the dynamics of the folding. A model in which flux tubes between amino-acids and water molecules alone induce the interactions between amino-acids will be discussed in separate section. A realistic model might involve both kind of flux tubes.

3.1 4-D Spin Glass Energy Landscape And Code Of Catalytic Action

There is a proposal that protein folding corresponds to a motion in a fractal spin glass energy landscape in presence of external perturbations due to the presence of water and leading to the bottom of some deep valley [I12, I12]. In TGD framework 3-D spin glass landscape is replaced by 4-D one [K4]. The vacuum degeneracy of Kähler action implies 4-D spin glass energy landscape in the sense that quantum jump sequences lead to space-time sheets representing asymptotic self organization patterns depending only weakly on the initial conditions (with respect to subjective time measured as quantum jumps). Proteins would be like skilled musicians possessing a repertoire of motor activities represented by deep valleys in 4-D spin glass landscape.

This picture generalizes to the functioning of living matter in various scales and the quantum dynamics of brain is a natural application giving also connection with p-adicity since ultra-metric topology is naturally associated with the space of valley bottoms. In the case of catalytic reactions a quantum jump changing Planck constant for some magnetic flux tubes connecting some living biomolecules (DNA, RNA, amino-acids, water(?), ..) and changing the lengths of these flux tubes could be the basic mechanism leading from a given valley to a new one and the reduction of the genetic code to single nucleotide or di-nucleotide code would code this quantum jumps.

To me this proposal for the folding code - or rather, the code of entire biocatalysis - looks so beautiful that it deserves to be killed this should be easy for a professional biochemist. If the hypothesis survives, it would provide a royal road to the understanding of the catalytic biochemistry.

3.2 Flux Tubes And Amino-Acids

3.2.1 Matter antimatter asymmetry at the level of interactions of amino-acids

The first thing that I learned was that in the case of amino-acid belonging to protein interior second nucleotide Y in the codon XYZ coding for amino-acid is what matters. Only $Y = A, G$ amino-acid residue can form hydrogen bonds and is hydrophilic and thus interacts strongly with water and DNA and RNA. In T, C case the formation of hydrogen bonds is impossible or rare (ser, thr). In their interactions with water these amino-acids are passive, or rather-avoid water- and tend to interact with each other. This division is fundamental for the understanding of the interactions of amino-acids. The division of amino-acids to hydrophobic *resp.* non-hydrophobic ones corresponds to the assignment of quarks to A and G and antiquarks to T and C so that strong matter antimatter asymmetry is in question. Similar asymmetry appears in cosmology: in TGD Universe antimatter resides inside cosmic strings in the interior of big voids containing matter as galaxies at their boundaries so that one can understand why antimatter is not visible.

3.2.2 Flux tubes can connect with all electronegative atoms

The model for di-nucleotide precursor code [K8] involves precursors for which 3 precursors contain only oxygen ions or double bonded oxygens. The only possible conclusion is that oxygen can connect to any DNA letter (quark or antiquark) and that first letter-precursor correlation is a selection of the most probable alternative. Also in water oxygen atoms should form flux tube contacts with each other and amino-acids and DNA. Also nitrogen atoms might form similar flux tube connections and this was assumed in the model. Same would apply to sulphur appearing in met and tyr and to electronegative atoms in general.

3.2.3 What can one learn from the formation of alpha helices and beta sheets?

Assume that hydrogen bonds correspond to flux tubes. The formation of peptide bonds by the elimination of H_2O-molecules and generation of hydrogen bonds between $N - H$ and $O =$ is an essential step in the formation of alpha helices and beta sheets. Second observation is that free amino-acids decompose naturally into three parts corresponding to $O = COH$, R, and NH_2. One

can also count $O=$ as a separate unit so that there would be four units in this case. This suggests that amino-acid could correspond to the entire DNA codon XYZ coding for it. In this case there would be 2 flux tubes per amino-acid and one can consider the following options.

1. Y could correspond to either R or $O=$. If hydrogen bonds correspond to flux tubes, $R-Y$ correspondence is not realistic. The reason is that R should be either donor or accept and hydrophobic amino-acids do not possess neither property. Hence only $O=$ can corresponds to Y.

2. $O-H$ could correspond to Z, $O=$ to Y, and NH_2 to X. For this option the amino-acid in protein would correspond to XY. If one identifies hydrogen bonds as special case of flux tubes, the hydrogen bonds of alpha helix would obey $X-Y_c$ rule which seems too restrictive.

3. $O-H$ could correspond to X, R or $O=$ to Y, and NH_2 to Z. For this option the amino-acid in protein would correspond to YZ. In this case the hydrogen bond of alpha helix would obey $Y=Z_c$ rule which by the isospin symmetry of the last nucleotide of the codon might be flexible enough.

3.2.4 Interactions of proteins with ions and electrons

Proteins interact also with electrons and ions. Typical process are the addition or removal of proton, electron, ion such Ca^{++}, or molecule such as O_2. These interactions are not well understood. For instance, the interactions involve the transfer of electrons between ligand protein and protein inducing oxidation (electron is given), reduction (electron is received) or redox reaction (both reduction and oxidation take place). In metabolism redox process is central. These reactions are reversible and it is difficult to understand how electrons are able make their long journey from the interior of the ligand so fast and avoiding dissipative effects. The formation of cyclotron Bose-Einstein condensates of bosonic ions and electronic Cooper pair condensates at the magnetic flux tubes connecting ligand and protein might provide the solution of the mystery. Note that the new nuclear physics predicted by TGD predicts nuclei which can have anomalous em charge associated with the color fluxtubes connecting nucleons to nuclear string so that fermionic ions Na^+, Cl^-, K^+ could have exotic bosonic counterparts.

3.3 Trying To Identify The Folding Code

The basic question is what kind of generalized pairings are realistic for amino-acids. The identification of hydrogen bonds as flux tubes leads to rather unique identification of the pairing and excludes the naively expected $Y-Y_c$ pairing.

3.3.1 A trial for the folding code

Protein folding code is something which is expected to exist but is not understood [I13]. This inspired a work which led to several trials for the folding code. Also a natural generalization to a code for catalysis emerged. In the most plausible candidate for the code flux tubes are identified as correlates of directed attention at molecular level. By their asymmetry hydrogen bonds are identified as a special case of flux tubes. Free amino-acid behaves like $X_cY_cZ_c$ and the amino-acid inside protein like Y_cZ_c. There are *two flux tubes* per amino-acid corresponding to $N-H$ and $O=$ representing Z_c and Y_c.

This leaves two options for pairing.

1. If $O=$ can act as a terminal for hydrogen bond and long flux tube then $N-H$ and Y can connect simultaneously to $O=$ and one has $Y=Z$ pairing.

2. If $O=$ can act as a terminal for only single flux tube representing Y then reconnection process for $N-H$ and $O=$ flux tubes creates the hydrogen bond and $Y=Z_c$ pairing for amino-acids results

Both pairings are highly flexible so that obvious inconsistencies with the data about alpha helices and beta sheets are avoided. A highly non-trivial and testable prediction of both pairings

is that the two identical proteins coded by different DNA sequences can have different foldings since the allowed pairings are not identical. Thus amino-acids would remember at the level of the braidings which DNA sequence coded them. This prediction can be avoided only Z flux tube corresponds to a quantum superposition of the nucleotides coding for the amino-acid in question so that one has quantum superposition over quark pairs associated with the third nucleotide.

The two-point mutations possibly carried out intentionally by the magnetic body controlling the genome conserving amino-acid pairings by hydrogen bonds and thus perhaps also folding and the catalytic properties should transform $Y = Z_c$ ($Y = Z$) pair to an allowed pair of this kind so that quite wide repertoire of allowed 2-point mutations is available for this option.

3.3.2 $Y = Z_c$ or $Y = Z$ pairing might work

The isospin symmetry of the third nucleotide implies that $Y = Z_c$ pairing is quite flexible. Roughly, the rule would allow flux tube connections only between amino-acids for which Y and Z correspond to quark and antiquark. The amino-acid pairs can be classified to three types. The amino-acid pairs for which both amino-acids can act as acceptors and donors, the pairs for which amino-acids can act only as an acceptor or donor, and the pairs for which no flux tubes are possible.

There are two options to be considered.

Option 1: Flux tube in either direction between amino-acids is prerequisite for the existence of the hydrogen bond. In this case the condition is $Y(n) = Z(n + k)$ or $Z(n) = Y(n + k)$.

Option 2: Hydrogen bond is identified as a flux tube. The condition is $Y(n) = Z(n + k)$ and thus stronger than for the first option.

Table 1 of Appendix summarizes the allowed and non-allowed pairings for $Y = Z_c$ and $Y = Z$ pairings. To understand the tables some notation conventions must be introduced.

1. Let X_{ij} denote the amino-acids in i: th and j: the column of the code table. $i, j = 1, 2$ corresponds to hydrophobic amino-acid residues and i, j=3, 4 to hydrophilic amino-acid residues.

2. For $Y = Z_c$ option the sets t, T, U, V, W, X are defined as $t = \{phe\}$, $T = X_{12} - t$, $U = \{tyr, his, asn, asp, cys, arg, ser, gly\}$, $V = \{trp, gln, lys, glu, gly\}$, $W = \{gln, lys, glu, trp, arg, gly\}$, and $X = \{tyr, his, asn, asp, cys\}$.

3. For $Y = Z$ option the sets t, T, U, V, W, X are defined as t=$\{$met$\}$, $T = X_{12} - t$, $U = \{trp, gln, lys, glu, arg, gly\}$, $V = \{tyr, his, asn, asp, cys\}$, $W = \{tyr, his, asn, asp, cys, arg, ser, gly\}$, and $X = \{gln, lys, glu, trp\}$. ser has been excluded from V since it appears also in the second column of the code table.

Some clarifying comments about the table are in order.

1. Pro is an exception since Z nucleotide cannot be represented in this case and Pro can act as donor. This has not been taken into account in the tables.

2. The codons coding for the paired amino-acid give additional strong limitations on the pairing unless Z corresponds to quantum superposition of quark pairs associated with the third nucleotide for the codons coding for the amino-acid.

3. Depending on option either phe-phe or Met-Met hydrogen bonding is forbidden so that for hydrophobic amino-acids almost all pairings are possible. This might allow to select between the two options or kill both. The special role of met suggest that $Y = Z$ pairing might be the right option. Also the model for DNA as TQC assumes that O = associated with lipids can act as a plug to which two flux tubes terminate. On the other hand, phe is also exceptional in the sense that it is the only amino-acid in X_{12} which has aromatic ring and can act as an acceptor.

4. The amino-acids which can act simultaneously as donors and acceptors are of special interest as far interactions between catalyst sites of protein and ligand are considered. Second flux tube could be involved with the structure of the catalyst site and second flux tube with the bonding of between catalyst sites. This kind of amino-acids correspond to $T \times T$, $U \times$

U, $X_{12} \times W$. For both options hydrophobic amino-acid can be connected with any other hydrophobic amino-acid. In the case that the two amino-acids are connected by two flux tubes one has stronger conditions giving $(Y_1, Z_1) = (Z_2, Y_2)_c$ or $(Y_1, Z_1) = (Y_2, Z_2)$.

5. $T \times t$, $U \times V$, and $T \times X$ correspond to pairings for which amino-acids can act as donor or acceptor only. The triplets abc in which (a, b) belongs to one of these sets should not appear in alpha helices. For instance, for $Y = Z$ pairing hydrogen bonded $xmety$ triplets with x, y in X_{12} should not be possible.

6. The hydrogen bonds of alpha helices and beta sheets provide a test for the model. For instance, the appearance of gly in the hydrophobic portions of alpha helices is consistent with both $Y = Z_c$ and $Y = Z$ pairing. The alpha helix appearing as an example in [I14] is consistent with both options.

2. Flux tube is identified as hydrogen bond

Table 2 of Appendix summarizes the allowed and non-allowed pairings for $Y(n) = Z_c(n + k)$ and $Y(n) = Z(n + k)$ pairings in this case. The notational conventions are following.

1. Let X_{ij} denote the amino-acids in i: th and j: th column of the code table. $i, j = 1, 2$ corresponds to hydrophobic amino-acid residues and i, j=3, 4 to hydrophilic amino-acid residues. Only the sets X_{12} and X_{23} are of interest.

2. For $Y = Z_c$ option the sets t_1, t_2, V, W are defined as $t_1 = \{phe, pro\}$, $t_2 = \{met, pro\}$, $V = \{trp, gln, lys, glu\}$, and $W = \{tyr, his, asn, asp, cys\}$.

3. For $Y = Z$ option the sets t_1, t_2, V, W are defined as $t_1 = \{met, pro\}$, $t_2 = \{phe, pro\}$, $V = \{tyr, his, asn, asp, cys\}$, and $W = \{trp, gln, lys, glu\}$. ser has been excluded from V since it appears also in the second column of the code table.

Some clarifying comments about **Table 1** are in order.

1. The two options are related by the duality $t_1 \leftrightarrow t_2$, $V \leftrightarrow W$. Pro appears in the list because it contains no $N - H$ group and cannot therefore act as donor. The fact that Pro often appears as first amino-acid in alpha helix conforms with this.

2. The codons coding for the paired amino-acid give additional strong limitations on the pairing unless Z corresponds to quantum superposition of quark pairs associated with the third nucleotide for the codons coding for the amino-acid. This could be interpreted as counterpart of wobble base pairing.

3. Met (contains S), pro, and phe (only amino-acid with aromatic ring in X_{12}) are exceptional for both options. $X_{12} \times t_1$ and $X_{34} \times t_2 = O - -(H - N)$ hydrogen bonding is forbidden. This poses strong conditions at the boundaries of hydrophilic and hydrophobic regions.

One might hope that either of these models could give a solution to the basic problem of proteonics whether genes code for the protein folding and how: the apparently lost information in the mapping of codons to amino-acids codes for the folding determined hydrogen bonds and more general flux tubes. The hydrogen bonds of alpha helices and beta sheets provide a test for the model. In absence of quantum counterpart of wobble base pairing for Z both models allows to deduce from the mere amino-acid sequence and hydrogen bonding the DNA sequence coding for the protein in the case of alpha helices and presumably also beta sheets. This is of course a testable prediction. For non-hydrogen bonded portions of protein this might not be possible and an interesting question is whether they tend to consist of amino-acids in sets t, V and $t \cup X$ so that hydrogen bonds are not allowed. In any case this would mean a solution to the basic problem of proteomics whether genes code for the protein folding and how: the apparently lost information in the mapping of codons to amino-acids codes for the folding determined hydrogen bonds and more general flux tubes.

3.3.3 Tests for $Y = Z$ and $Y = Z_c$ pairings

The test consists of deducing the number N of pairs which did not satisfy the condition $(a(ii), a(ii+4))$ not equal to (t, t), or does not belong to $(V \times V)$ or to $t \times V)$. From this the average length L of portions satisfying alpha helix conditions $k = 4$ can be deduced as $L = N/N_{tot}$, where N_{tot} is the number of amino-acids in the sequence.

The test was carried out for one structural unit of asparagine synthetase [I2], xylose isomerase [I11], hydrolase [I8], glutathione s-transferase [I7], and restriction endonuclease BamHI [I3].

1. Option 1: Flux tube from in direction is prerequisite for the formation of hydrogen bond

From **Table 6** of Appendix one finds that the test for values of k different from $k = 4$ for helix gave also surprisingly large values of $L(k)$ for $Y = Z$ option. The average length of alpha helix is 10 amino-acids so that both options could work. $Y = Z_c$ option gives results rather near to this value.

One can apply test also to individual alpha helices. For asparagin synthethase alpha helices correspond to the intervals [7, 28], [76, 84], [130, 155], [170, 177], [76, 84], [170, 177], [182, 194], [256, 268], [277, 284], [297, 305], [309, 314], and [320, 326] in the standard numbering of amino-acids. The test was done for $k = 3, 4, 5, 6$ assuming that the upper end of tested interval is 6 units higher. $N = (0, 0, 0, 0)$ results for both options for all intervals except for the interval $[7, 28]$ for $Y = Z_c$ for which one obtains $N = (4, 2, 3, 3)$. Hence $Y = Z$ option is favored.

In the case of remaining enzymes only long enough alpha helices were tested and **Table 4** of Appendix gives the results

The conclusions are following.

1. From **Table 4** it seems clear that $Y = Z_c$ option does not work satisfactorily whereas $Y = Z$ option has rather few failures.

2. In the case of xylose isomerase and $(Y = Z)$ option with $k = 4$ there are four helices for which failure occurs for single amino-acid. The prediction is that the corresponding hydrogen bonds are actually absent.

3. The worst failure occurs for glutathione s-transferase and involves two amino-acids which are at positions n and $n + 4$. The hydrogen bonds are predicted to not exist between met-glu and glu-asp in met-glu-asp.

Beta sheets consist of beta strands which can be regarded as $(n, n + 1)$ helices so that stability conditions correspond to $k = 1$. As the **Table 5** of Appendix shows, there are no failures for Y=Z option whereas $Y = Z_c$ option has several failures and very bad failure for glutathione s-transferase (3 failures for 4 units long strand).

One might think that loops could contain amino-acids for which the hydrogen bonds to neighbors are not possible. The test for BamHI showed that this is not the case. Only single loop failed for $Y = Z$ option for $k = 1, 2..., 4$ and this occurred for $k = 1$.

The remaining test is for whether the $Y = Z$ pairing indeed can fix the DNA sequence completely. BamHI begins as *met glu val glu lys glu phe ile.....* For beta sheet beginning from second amino-acid requires that the Y=Z rules holds true for subsequent codons in DNA sequence *aag ctt cct taa ttc cgg aag...* [I4]. By comparing the Z of a given codon in beta sheet to the Y of the next codon inside beta sheet one finds that the $Y(n) = Z(n + 1)$ or $Z(n) = Y(n + 1)$ fails. Similar conclusion follows from an analogous check for the first alpha helix. Situation is saved if the quantum counterpart of wobble base pairing is at work so that the flux tube from tRNA to $N - H$ would in superposition of colors (quark pairs) corresponding to superposition nucleotides Z in codons XYZ for given X and Y coding for the amino-acid in question. Hydrogen bonded amino-acid sequence would behave as if it were coded by the unique DNA sequence. Note that for a given amino-acid X is unique except for leu and arg and Y is unique except for ser. The $N - H$: s and $O =$: s for which hydrogen bonds are lacking could form hydrogen bonds with water molecules and $O =$: s could have long flux tubes with other $O =$: s in the protein.

2. Option 2: flux tube is identified as hydrogen bond

Tables 7 and **8** of Appendix summarize the results of the test for $Y = Z$ and $Y = Z_c$ option when flux tube is identified as hydrogen bond. For the first option the average length of hydrogen

bonded interval would be around 5 amino-acids for $k = 4$ helix for $Y = Z$ and somewhat shorter for $Y = Z_c$. BamHI is exceptional since in this case the length is 16.8 (10.4) amino-acids. for $Y = Z$ ($Y = Z_c$). There is no clear difference between the two alternatives in the case of alpha helices and neither alternative looks promising in this case.

3.3.4 Are $=O-O=$ flux tubes present?

$Y = Z$ option for the folding code assumes that flux tubes can connect acceptor atoms by flux tubes. The pairing would be $Y - Y_c$ pairing considered in the original model as the only possible pairing. In amino-acids only $O =:$ s not acting as acceptors for ordinary hydrogen bonds could have flux tube connections of this kind with each other or other molecules.

1. In the case of amino-acids $Y - Y_c$ pairing would be between amino-acid in X_{12} and amino-acid in X_{34} part of the code table. These connections would be typically associated with the portions of the protein between alpha helices and beta sheets. The $k : th$ amino-acid ($k = 3, 4$ or 5) following Pro would be an exception to this rule and this kind of flux tubes could be involved with the long scale stabilization of proteins.

2. The $O =$ atom would effectively behave like Y_c. Depending on whether it corresponds to quark or anti-quark, the corresponding amino-acid would be typically hydrophilic or hydrophobic- or rather - able to form hydrogen bonds or not. Since hydrophilic and hydrophobic residues tend to avoid each other the flux tubes in question should be rather long. The phase transitions increasing Planck constant might make this possible. This would bring in a strong long range correlation between the dynamics of the amino-acid residues belonging to the first and third (second and fourth) column of the code table.

3. $=O-O=$ flux tubes could be also between different proteins. In the case of protein-ligand complex the Planck constant changing phase transition reducing the length of this kind of flux tube could bring proteins together after which a recombination process the hydrogen bond connecting two water molecules would transform the bond to hydrogen bonds of $O =:$ s with water molecules.

4. The phase transition increasing $\hbar$ would allow hydrophobic amino-acids to increase their distance from water molecules in a controlled manner. This could be essential for folding and make possible the formation of pockets connected by flux tubes of large $\hbar$ to water. In quantum models for evolution of consciousness these pockets are believed to play a prominent role. Molecular sex in this sense would mean a correlation tending to keep partners at large distance except when $\hbar$ reducing phase transition occurs.

3.3.5 Evolution and amino-acid pairings

The evolution at the molecular level corresponds to the emergence of increasingly complex molecules using as basic building blocks amino-acid chains and non-translated residues attached to them in the post-translational processing of the amino-acid chains. Also increasingly complex reaction paths emerge. Molecular survival and the competition for the metabolic resources at molecular level could be seen as the basic driving force of this evolution.

Typically, in the original situation the enzymes would have received the substrate molecules from the environment but sooner or later this would have become difficult. The solution would have been a synthesis of the substrate from simpler ingredients by starting from some precursor.

If molecules (with magnetic bodies included) are conscious entities able to direct attention, one can imagine that magnetic body controlling them with the mediation of genome and able to actively modify it, could help through modifications of the genome to create to the catalyst a binding site able to bind the precursor. Immune system is doing this very intensively. If the enzyme binding the precursor already exists, a combination of genes coding for the enzyme and the enzyme having the metabolites as ligands could allow to achieve this. All this would reduce to the motor activities of magnetic body, in particular reconnection of flux tubes, a kind of dance of Shiva. Genome would not be anymore a sequence of DNA developing through random mutations under selection pressures.

In this framework amino-acids would have appeared before their precursors and possessed some function in RNA world, say the catalysis of join of RNA_2 di-nucleotides to the increasing chain as proposed in [K8]. Competition might have led to a situation in which RNA_2 learned to catalyze selectively the generation of amino-acids from much simpler precursors (three of the proposed precursors contain only C, $=O$, and O^-) giving rise to positive feedback implying an exponential amplification of RNA and amino-acid populations. The reduced genetic code would have been present at two levels. The reader can decide whether this is a shortcoming of the model or a fundamental biochemical duality.

Can one make any clear cut predictions about preferred mutations?

1. In TGD framework mutations are not expected to be always random point mutations but could be even a result of a purposeful action of the magnetic body. Chemical similarity is expected to be conserved in good mutations. This is known to be the case. For $Y = Z$ or $Y = Z_c$ pairing the simplest mutations should leave both Y and Z invariant so that only the first nucleotide X can suffer a mutation.

2. Also bi-local mutations of the second and third nucleotides of codons forming $Y = Z$ ($Y = Z_c$) pair and conserving this property might occur and could be crucial for the coherence of the organisms. As found, the formation of flux tube between amino-acids A_1 and A_2 induces a flux tube between nucleotides Y and Z at the corresponding genes. This flux tube could force the possibly intentional mutations to occur as simultaneous mutations of the two genes so that $Y = Z$ ($Y = Z_c$) condition remains true for amino-acids connected by flux tube.

3. A new element is that isospin rotation of Z nucleotide ($A \leftarrow G$, $T \leftarrow C$) which does not affect amino-acid, affects its folding so that same protein might have different folding patterns and different catalytic properties corresponding to different codons coding for it. This would mean a breaking of the central dogma a the level of magnetic body. Some examples are in order. The mutations Ala/Ser, Ser/Thr, Ile/Val/Leu, Asp/Glu do not change Y. Lys/Arg (A/G)), Tyr/Phe (A/U), gly/Ala (G/C), ... are also prevalent and one might hope that they correspond to binary mutations in some important cases.

4. Folding is known to be more conserved than amino-acid sequence [I14]. Since folding is a collective property of gene, local chemistry might not be enough and the proposed non-local conservation laws might be needed. Two-point mutations would also correlate the mutations of the binding sites of protein and ligand. For the model assuming two flux tubes per amino-acid, the prediction would be conserved $Y = Z$ ($Y = Z_c$) pairs in genes coding for protein and ligand and these pairs might allow to deduce the paired points. This is consistent with the fact that hydrophobic (-philic) regions tend to be paired in the protein-ligand complex. The paired nucleotides need not belong to the same strand since genes are evenly distributed between strand and its conjugate and characterized by A, G surplus.

5. If the flux tubes can connect also side chains, the situation becomes more complex. There is a temptation to think that these flux tubes would connect only the nearby amino-acids of the same peptide and do not therefore affect the large scale dynamics of folding. This would be the case if the value of Planck constant associated with these flux tubes is smaller than for the flux tubes connecting amino-acids as basic units. If flux tubes can begin from the aromatic side chains, the replacement of an aromatic side chain with an aromatic side chain is favored (also chemical similarity explains this). The most basic facts about folding do not provide obvious support for the idea about flux tubes between residues.

   (a) Hydrophobic residues tend to cluster in dense packing in protein interior (antimatter at quark level) and Val (T), Leu (T), Ile (T), Phe (T), Ala (C), and gly (G) make 63 percent of the interior of protein: the special role of gly (matter rather than antimatter at quark level) is due to the reduction of the side chain to hydrogen atom.

   (b) Asp (A), Glu (A), Lys (A) and Arg (G) with ionized residues are mostly at the surface of protein and make 23 per cent of protein surface and 4 per cent of interior. As noticed earlier, matter and antimatter at quark level tend to be far from each other. This is consistent with $Y = Z$ pairing between nearby amino-acids and absence of flux tubes between matter and antimatter if there are two flux tubes per amino-acid.

(c) Polar groups tend to be paired by hydrogen bonds and oppositely charged groups tend to be near each other. Acidic Cys residues tend to be in positions where they can form $S-S$ bonds. This can be explained as being induced by $Y-Y$ pairing in the proposed scenario. Aromatic residues tend to have favorable electrostatic interactions with each other and with S, O and amino groups.

4 A Simple Quantitative Model For Protein Folding And Catalyst Action Assuming Flux Tubes Between Amino-Acids

Levinthal paradox states that if protein folding is a process in which protein checks for all possible conformations, folding would take astrophysical time. Small single domain proteins with lengths below 100 residues however fold in single step in millisecond time scale and longest folding times are measured in days. This suggests that protein folding is in some sense guided dynamical process and flux tubes would be the natural guides.

It is possible to construct a simple quantitative model for protein folding and catalyst action assuming a long range interaction mediated by flux tubes between amino-acids obeying base pairing rule in some sense. A further assumption is that hydrogen bonds correspond to flux tubes. There are two options to consider.

1. If there is only single flux tube per amino-acid the rule implies that conjugate amino-acids are connected by a flux tube: this is conflict with the empirical facts.

2. If there are two flux tubes per amino-acid base pairing predicts that amino-adic pairing obeys $Y = Z$ or $Y = Z_c$ rule depending on whether $O =:$ s can act as intermediate plugs for flux tubes or not.

The model is consistent with quantum criticality, and the general vision about 4-D spin glass landscape. The extremals are not completely deterministic just as vacuum extremals of Kähler action and only absolute minimization of energy selects minima. The cautious interpretation is that absolute minimization of energy stabilizes various unstable patterns generated spontaneously by ordinary chemical interactions such as alpha helices and beta sheets. The interpretation of hydrogen bond in terms of flux tube suggests more bold interpretation.

The principle is flexible enough to carry out this purpose but also poses strong constraints on how these patterns integrate to higher level structures. The disappearance of a subset of flux tubes does not spoil the extremal property although it increases its non-determinism and makes folding less predictable and in the case of binding sites it reduces the selectivity of catalyst action. The interpretation would be in terms of molecular ageing. The density of flux tubes can be seen as an analog for the resolution of quantum measurement which is in a fundamental role in quantum TGD, as well as a direct correlate for cognitive and sensory resolutions. The model extends to a model of catalyst dynamics if one the relative motion of reactant molecules is slow in the time scale of folding dynamics so that adiabaticity assumption makes sense. In the following I often use the basic data which can be found from [I14] without explicit reference.

4.1 The Model

Let us assign potential energy to the flux tube connecting i: th and $k(i)$: th amino-acid and depending only on the distance $r_{i,k(i)}$. What comes in mind first is the potential energy of harmonic oscillator:

$$V(r) = \frac{kr^2}{2} . \tag{4.1}$$

$k > 0$ corresponds to harmonic oscillator. Also $k < 0$ is possible in which case the distance between amino-acid and its conjugate tends to be maximized in equilibrium: this option turns out to be the more plausible one and conforms also with the notion of quantum criticality. Besides this there

is the constraint that the distances between amino-acid and its follower are constant: $r_{i+1,i} = R$. Using Lagrange multipliers this gives rise to the action

$$L \quad = \quad -E = -\frac{k}{2}\sum_i r^2_{i,k(i)} + \sum_i \lambda_i r^2_{i+1,i} \quad . \tag{4.2}$$

Energy is the negative of this action for static solutions. One could consider also adding kinetic term to this action to describe the dynamics of folding. This action is hoped to give only a qualitative view about folding and the ordinary chemical interactions should fix the details of the folding and select between different folding patterns. Several amino-acid chains could be present and have mutual long range interactions.

If $N-H$ and $O=$ both can be connected by flux tubes, each amino-acid gives two terms to the energy corresponding to the flux tube beginning from $N-H$ and flux tube ending at $O=$.

The extremals of this action satisfy

$$\frac{\partial L}{\partial r^k_i} \quad = \quad 0 \quad , \quad i = 1, ..., N \quad . \tag{4.3}$$

1. If there is only *single flux tube per amino-acid*, this gives the conditions

$$\begin{aligned} \lambda_{i+1}\overline{r}_{i+1,i} - \lambda_{i-1}\overline{r}_{i,i-1} &= -k\overline{r}_{i,k(i)} \quad , \\ r_{j+1,j} &= R \quad . \end{aligned} \tag{4.4}$$

   The geometric content of these conditions is that the vectors $\overline{r}_{i,k(i)}$, $\overline{r}_{i+1,i}$, and $\overline{r}_{i,i-1}$ are in the same plane.

2. If there are *two flux tubes per amino-acid* ($= O_i - -(N-H)_{k_1(i)}$ and $(N-H)_i - -(O=)_{k_2(i)}$)

$$\begin{aligned} \lambda_{i+1}\overline{r}_{i+1,i} - \lambda_{i-1}\overline{r}_{i,i-1} &= -k[\overline{r}_{i,k_1(i)} + \overline{r}_{i,k_2(i)}] \quad , \\ r_{j+1,j} &= R \quad . \end{aligned} \tag{4.5}$$

   In this the resultant of the vectors $\overline{r}_{i,k_1(i)} + \overline{r}_{i,k_2(i)}$ would be in the plane determined by $\overline{r}_{i+1,i}$ and $\overline{r}_{i,i-1}$. Note that due to the lack of $N-H$ in Pro it can happen that there is only single flux tube per amino-acid.

Long range interactions of amino-acids with their conjugates would dictate the local folding of the amino-acid chain but extremum property alone does not say much about the lengths of the flux tubes.

Suppose that $\overline{r}_i$, $\overline{r}_{i,k(i)}$, $\overline{r}_{i,i-1}$, λ_{i-1} are known. Can one solve λ_{i+1} and $\overline{r}_{i+1,i}$? Since the vectors are in the same plane, the linear dependence does not fix the direction of $\overline{r}_{i+1}$ in this plane but only the value of λ_i in this plane once $\overline{r}_{i+1}$ is fixed or vice versa. Therefore the direction in the plane remains un-determined and equations of motion are not fully deterministic as far as extremals are considered. Absolute minimization however eliminates this non-determinism by maximizing the distances $r_{i,i(k)}$ for $k > 0$ option. The expressions for λ_i result from elementary linear algebra by introducing dual basis of non-orthogonal basis defined by $\overline{r}_{i,k(i)}$ and $\overline{r}_{i,i-1}$.

1. In the case that there is *single flux tube per amino-acid*, one has

$$\begin{aligned} &\lambda_{i+1} = -k\overline{e}_{i+1}\cdot\overline{r}_{i,k(i)} \quad , \quad \lambda_{i-1} = -k\overline{e}_{i-1}\cdot\overline{r}_{i,k(i)} \quad , \\ &\overline{e}_{i+1}\cdot\overline{r}_{i+1,i)} = 1 \quad , \qquad \overline{e}_{i+1}\cdot\overline{r}_{i,i-1)} = 0 \quad , \\ &\overline{e}_{i-1}\cdot\overline{r}_{i+1,i)} = 0 \quad , \qquad \overline{e}_{i-1}\cdot\overline{r}_{i,i-1)} = 1 \quad . \end{aligned} \tag{4.6}$$

The non-determinism does not make it easy to find absolute minimum since non-determinism corresponds to circle $(S^1)^{2N}$ for amino-acid sequence with N flux tube pairings. These conditions do not make sense when $\overline{r}_{i+1,i)}$ and $\overline{r}_{i,i-1)}$ are parallel: in this case the force must be parallel to $\overline{r}_{i+1,i)}$.

2. For *two flux tubes per amino-acid* one has a slightly more complex expression for these conditions:

$$\begin{aligned} &\lambda_{i+1} = -k\overline{e}_{i+1}\cdot[\overline{r}_{i,k_1(i)} + \overline{r}_{i,k_2(i)}] \ , \quad \lambda_{i-1} = -k\overline{e}_{i-1}\cdot[\overline{r}_{i,k_1(i)} + \overline{r}_{i,k_2(i)}] \ , \\ &\overline{e}_{i+1}\cdot\overline{r}_{i+1,i)} = 1 \ , \qquad\qquad \overline{e}_{i+1}\cdot\overline{r}_{i,i-1)} = 0 \ , \\ &\overline{e}_{i-1}\cdot\overline{r}_{i+1,i)} = 0 \ , \qquad\qquad \overline{e}_{i-1}\cdot\overline{r}_{i,i-1)} = 1 \ . \end{aligned} \tag{4.7}$$

The strong resemblance with the dynamics defined by Kähler action predicting spin glass degeneracy associated with vacuum extremals of Kähler action and removed by small deformations to non-vacuum extremals raises the hope that the model indeed catches something essential about the notions of 4-D spin glass degeneracy and quantum criticality.

4.2 Basic Mathematical Consequences

Consider first the basic consequences of the variational equations.

1. Absolute minimization of energy is very powerful selection principle and expected to choose highly symmetric configurations such as α helices, β sheets, and more complex structures. If combined with adiabaticity assumption it could also allow to understand the dynamics of binding between two proteins and protein and DNA/RNA.

2. The extremals of $k > 0$ action are mirror images of $k < 0$ action so that the energy minimum for $k > 0$ is energy maximum for $k < 0$. If energy minimization is applied also the choice of $Y - Y$ flux tubes, the connected amino-acids should be as near as possible which favors alpha helices and beta sheets. In light of this $k > 0$ option looks the realistic one. It could however be that for large distances the sign of the potential energy changes. For $k > 0$ option long flux tubes are not favored by energy minimization. The simplest cure would be large value of Planck constant changing the scale of the potential. If the potential energy changes sign at large distances the situation changes also and $r_{i,k(i)}$ would be as large as possible subject to the condition from fixed chain length.

3. If the amino-acid is not paired, it does not experience the long range force and one has

$$\lambda_{i+1}\overline{r}_{i+1,i} - \lambda_{i-1}\overline{r}_{i,i-1} \quad = \quad 0 \ . \tag{4.8}$$

Situation becomes non-deterministic and the portions of the amino-acid chain for which the amino-acids do not have a pair behave like random coils. This is encouraging since this kind of portions are present in folded amino-acids. The absence of $N - H$ from Pro allows to understand the very special role of Pro as being associated with turns of alpha helices and beta sheets.

4. The disappearance of some flux tubes does not destroy a given solution of the conditions but makes it increasingly non-deterministic. The interpretation as a degradation or ageing at molecular level conforms with the interpretation of braiding as a basic characteristic of life. An attractive interpretation of the density of flux tubes is as correlate for resolution for cognition and sensory perception and motor action as counterpart of measurement resolution which is fundamental notion of quantum TGD.

4.3 Model For The Helical Structures

α helix [I14, I1], which is only one member of a rich family of helical structures possible for amino-acid chains, serves as the first test for the model. As a matter fact, the specific properties of α helix are not relevant for the model discussed.

1. α helix has nearly vertical $NH---O=$ hydrogen bond between i: th and $i-4$: th amino-acid. Also $(i, i-3)$ and $(i, i-5)$ bondings are possible. There are 3.6 residues per turn so that the basic structural unit has 5 turns and consists of 18 amino-acids. One residue corresponds to a vertical translation of 1.5 Angstrom. The chain contains single amino-acid per length of about 3.8 Angstrom and the angular separation of subsequent amino-acids is 100 degrees in the planar projection.

2. Isolated α helices are not stable but can be stabilized by secondary coiling: their lifetime is of order $10^{-5} - 10^{-7}$ seconds. If the flux tubes are associated with hydrogen bonds, the instability would be naturally due to a reconnection process involving water molecules.

Consider now the model.

1. Assume that hydrogen bond is accompanied by a special case of a flux tube resulting in the reduction of the value of Planck constant. Short flux tubes (hydrogen bonds) would connect $i-k$: th, i: th and $i+k$: th amino-acids, $k = 3, 4$ or 5. The forces between $i-k$: th and i: th and i: th and $i+k$: th amino-acid compensate each other exactly for an ideal helix so that the conditions are satisfied identically. This kind of mechanism work also for more general helices. $Y = Z$ ($Y = Z_c$) pairing poses special conditions on the helical structures themselves and also on the genes coding for these structures.

2. gly helices are consistent with both $Y = Z$ and $Y = Z_c$ pairings. The spontaneous generation of unstable helices in sequences consisting of mere gly could be understood as the instability of gly-gly flux tubes against reconnection with hydrogen bonds connecting surrounding water molecules. Also the sequences consisting of mere Pro can give rise to unstable helices. Pro does not possess $N-H$ and the residue cannot act as a donor in hydrogen bond. This suggests that the residue of Pro can have flux tubes connecting it to $O=$ but not identifiable as ordinary hydrogen bond.

There are also more complex structures formed form helices [I14]. For coiled coils of two or more alpha helices consisting of repeating heptad unit of 7 amino-acids first and fifth amino-acids tend to be conjugates so that horizontal flux tubes connecting first and fifth amino-acids of neighboring could be responsible for the stability and make also possible the hydrophobic bonding between first and fourth residues. Collagen [I6] is a triplet helix and appears as a basic constituent of bones, tendons, skin, ligaments, blood vessels, and supporting membraneous tissues. The units of collagen triple helix consists of very long repetitive sequences of type $(gly-XY)_n$, with a preponderance of Pro for X (also Lys residues are possible). gly-Pro-Y and gly-X-Hyp appear often: here X and Y are arbitrary amino-acids (Hyp denotes hydroxyprolin with $O=$ replaced with OH: this transforms Pro from acceptor to donor). Heating of collagen triple helix unfolds it and converts it to gelatin, in which polypeptide chains are dissociated, unraveled and disordered. Cooling regenerates these conformations for short stretches.

Consider as a example collagen triplet helix [I6] having $gly-Pro-Y$ as a repeating unit. Assume $Y = Z$ or $Y = Z_c$ pairing. $Y-Y$ hydrogen bonds are possible if Y belongs to the group T or U. Only phe ($Y = Z_c$) or met ($Y = Z$) is excluded from T. $Y = Z_c$ corresponds to $U = \{tyr, his, asn, asp, cys, arg, ser, gly\}$ and $Y = Z$ to $U = \{trp, gln, lys, glu, arg, gly\}$. This prediction might kill the model. glys can be connected for both options.

1. The first model goes like follows. alpha helix structure is guaranteed by hydrogen bonds between the Y: s inside each collagen unit ($k = 3$). The amino-acids gly_i, $i = 1, 2, 3$, are connected by almost horizontal flux tubes cyclically as $gly_1 - gly_2$, $gly_2 - gly_3$, $gly_3 - gly_1$. This cyclic bonding would induce the coiling of alpha helices. The free $O=$: s of Pros could act as acceptors in the hydrogen bonds with the surrounding water molecules (for instance). For gly-X-Hyp one would have similar structure but Hyp would act as donor in the hydrogen

bonds with water molecules. The objection is that if long hydrogen bonds are possible they would have been observed.

2. Second model is based on the philosophy that coiling is a long range effect and must be due to $= O - O =$ flux tubes. gly ($Y = G$) and Pro ($Y = C$) can be connected for both options but only by single flux tube by the special properties of Pro: this bonding would give $n, n+4$ hydrogen bond of alpha helix. The simultaneous presence of $n, n+3$ $Y - Y$ bonds and $n, n+4$ pro-gly bonds might be made possible by coiling. Hence the free $O =$ in gly could be connected with a similar $O =$ in the neighboring strand. $gly_1 - gly_2$, $gly_2 - gly_3$, $gly_3 - gly_1$ cannot form a closed cycle but the repeating helical pattern $gly_1 - gly_2$, $gly_3 - gly_1$, $gly_2 - gly_3$ is possible and could produce the coiling.

4.4 Model For β Sheets

beta strands are typically 4-5 amino-acids long structures. Hydrogen bonds are of type $(n, n+1)$ and β strands have 2 amino-acids per turn so that $\overline{r}_{i-1,i}$ and $\overline{r}_{i,i+1}$ span a vertical plane and the equations of the model are trivially satisfied. beta strands as such are not stable. beta sheets [I5] consisting of β strands which can be either parallel or antiparallel and are glued together by the interactions between residues. beta sheets are also slightly twisted which relates to the chirality of amino-acids. In the antiparallel case strand returns back and forms at the ends of sheet a loop so that so called β hairpin is formed. In parallel case the strand returns as alpha helix to the lower end of the sheet. At the time of writing of [I14] the mechanism of formation of β sheets was not understood.

If horizontal flux tubes between neighboring strands assignable to hydrogen bonds or =O-O= flux tubes between the residues are responsible for the stabilization of the beta sheet structure, then given residue must have two hydrogen bonds with same length to the amino-acids at right and left so that the contributions from right and left side to the force compensate each other and the force is automatically vertical as implied by the twisting angle of π per amino-acid in beta sheet. For self connecting flux tubes inside loops the force would be in the plane of loop and if the force is repulsive loop like structure is expected.

The slight twisting of beta sheet represents a challenge for the model. TGD predicts large parity breaking and thus the twisting and preferred helicity at the level of principle but it is not clear whether the simplest model can explain the twisting.

4.5 Secondary Protein Structures

Protein structures are divided into four classes on basis of their secondary structures [I14, I10]. All these structures are consistent with the general model.

1. (α) containing only α helices, which must stabilize each other by horizontal flux tubes.

2. (β) containing only β sheets both usually antiparallel, which appear always in pairs packing against each other. Horizontal flux tubes connecting the β sheets must act as stabilizers.

3. ($\alpha+\beta$) proteins can contain only single β sheet, usually antiparallel, with α helices clustering together at one or both ends of the β sheet. Antiparallel β sheet stabilizes itself.

4. (α/β) in which sheets and helices interact and often alternate along the polypeptide chain. Single parallel β sheet and so called β barrel, kind of sandwich like structure, are basic examples here. The most spectacular barrel consists of 4+4 parallel β strands with α helices outside the barrel.

Concerning the organization of alpha helices and beta sheets to higher level structures the simplest guess is that the large Planck constant flux tubes connecting random coil portions of the amino-acid sequence with each other or with free $O =$ accompanying Pros. The mere assumption that a given portion of coil has only long flux tubes to distant parts of the protein could explain random coil character. The failure of $Y = Z$ condition implies this too. The notion of long hydrogen bond is somewhat questionable and long flux tubes connecting $= O$: s look more favorable. Also free $O =$: s inside alpha helices and beta strands could be connected in this manner.

4.6 Model For Protein-Protein Binding Sites

Binding sites obey geometric complementarity and are known to resemble protein interior being closely packed. This is also taken to mean that amino-acid chains run parallel to the surface although this statement is not made explicitly in [I14]: one could see binding sites as part of interior which is in a direct contact with exterior, somewhat like a sensory organ like eye. The interface between similar sized proteins is large and tends to be flat (not expected if proteins make sharp turns at the interface rather than running parallel to the surface). Various bonds eliminate electromagnetic interactions at the interface.

The basic mechanism of binding would be based on the reduction of Planck constant for the flux tubes connecting amino-acids. The high flexibility of $Y = Z$ and $Y - Z_c$ pairings -especially in the hydrophobic regions in the center of the binding site where it allows all but met-met and phe-phe flux tubes- makes it an excellent candidate for a folding code.

The question is whether complementary of bonded amino-acids should induce the geometric complementary of the binding sites in the proposed model.

1. The binding sites could be connected by only very few flux tubes or flux tubes could connect all amino-acids in a pairwise manner: the first extreme is highly flexible whereas second extreme would produce maximal selectivity. Complementary can thus be partial and its degree is predicted to correlate with the selectivity. The interpretation of disappearance of flux tubes as molecular ageing conforms with the gradual loss of selectivity implying reduced performance of immune system.

2. From the example of [I14] about the interface of identical proteins in the quaternary structure of dimer one learns that the geometrically and physically conjugate interfaces of identical monomers pair to form sandwich like structures via so called isologous and heterologous pairings such that valleys and hills fit. The interfaces are reported to resemble closely packed protein interiors and contain hydrophobic residues in the center and hydrophilic residues at periphery. In the case of identical monomers $Y - Z$ and $Y - Z_c$ pairing is possible for a very wide class of amino-acids. The prediction in the case of identical monomers would be that catalyst sites contain only very few amino-acids in the sets V and t defined previously.

3. Also the flux tubes between $= O$ atoms could be in key role in the protein-ligand interaction. The interfaces can be thought of as cutting protein along its interior: in center there are hydrophobic amino-acids and in periphery hydrophilic ones. The $= O - O =$ flux tubes would connect periphery of A (B) to the center of B (A). The reduction of Planck constant for would reduce the length of these flux tubes and bring protein and ligand close to each other so that hydrogen bond formation between residues could being. In this process the flux tube connecting $O =: s$ could by reconnection transform to two hydrogen bonds connecting $O =:$ s to water molecules. After the catalysis the reverse of this process would occur.

4. For single flux tube between $O =:$ s of amino-acid and ligand the force would be along the line $r_{i,k(i)}$ connecting them, In the improbable case that the amino-acids of protein and ligand are connected by *two* hydrogen bond like flux tubes the force is in the direction of $\overline{r}_{i,k_1(i)} + \overline{r}_{i,k_2(i)}$. The force is predicted to be in the plane spanned by $\overline{r}_{i+1,i}$ and $\overline{r}_{i,i-1}$ for protein and in the corresponding plane for ligand. This is true if the amino-acid sequence at the surface is slightly curved in the direction of the conjugate amino-acid or in opposite direction. This condition is guaranteed by the geometric complementarity.

5. The mechanism for the formation of ligand-protein pairs would be very simple: the binding sites of protein and ligand could be coded by same gene or its mutation respecting the Y so that the formation of copies of gene in DNA would be the simplest mechanism to guarantee the prerequisites for geometric conjugation. Geometric conjugation would result automatically if the flux tubes between interior and periphery of binding site determine its shape.

6. Slow enough relative motion of molecules induces an adiabatic variation of the shapes of the binding sites so that lock and key mechanism becomes dynamical. The simplest possibility is that binding site and its conjugate behave like two eyeballs gazing each other as proteins

move with respect to each other. This is possible if binding sites are separated from the rest of the protein by random pieces of chain. The analogy with eye might be actually deeper: I have proposed long time ago that directed attention in vision has as a space-time correlate flux tubes of topological light rays or both of these. Wormhole magnetic flux tubes might indeed connect perceiver and the object perceived and serve as correlates of attention in macroscopic length scales.

7. Also the hydrogen bonds between residues are important for the protein folding. The donor atoms of the residues can inherit the conjugate of the color of $O =$ and acceptor atoms can inherit the color of $N - H$ by temporary reconnection. Therefore also the hydrogen bonds between residues of hydrophilic residues containing both donor and acceptor atoms would be restricted by the colors of atoms and would reflect genetic code.

8. Geometric and physical conjugation (acids and basics combine in the interface) means that a virtual protein A+B is cut to pieces along the surface in the interior defining the interfaces. Could this chopping of bigger proteins to smaller ones able to bind allow a realization at the level of genome in the sense that glued portions of protein would originate from same gene or its reversed version and thus satisfy $Y = Z_c$ or $Y = Z$ rule approximately? Could also protein interior involve pairings analogous to catalyst and ligand pairings? This would partially explain why protein folding is more sensitive to the mutations in the interior of protein.

5 A Model For Protein Folding Based On Flux Tube Connections Between Water Molecules And Amino-acids

The overall feelings about the model just discussed are somewhat mixed.

1. The ideas about flux tube as a correlates for a directed attention and about the connection between hydrogen bond formation and flux tube contraction involving change of Planck constant are attractive. It would be nice if flux tubes between amino-acids could force the portions of amino-acid sequences to form representation about each other in their own geometry. What would be also nice that the notions of finite measurement resolution and cognitive resolution which are fundamental notions of quantum TGD would have direct correlates at the level of flux tube dynamics.

2. The model for protein folding involving only flux tube connections between amino-acids satisfying the proposed selection rules has however failures. This could be due to simple fact that the proposed selection rules are quite too restrictive. Also the flux tube connections between amino-acids and water are important and might even determine the folding patters to a high degree via the induced secondary interactions between amino-acids.

Second model for protein folding to be discussed represents an extreme in which the flux tube connections between amino-acids and water molecules determined the dynamics of the folding. It seems that this model leads to a realistic qualitative picture about folding. Also quantitative model can be constructed as a straightforward generalization of the model involving only the flux tube connections between selected amino-acids.

5.1 Could There Be New Physics Behind Hydrophily And Hydrophoby?

One could accept just as a fact that magnetic flux tubes to the magnetic body of water mediate an interaction which is attractive or repulsive between water molecules and amino-acids and attractive between DNA molecules and water. Accepting that this induces interaction between amino-acids one could proceed to model building without any mention about TGD.

One could also try to dig deeper and ask what might be the origin of this interaction.

1. **Option I**: Could one understand the interaction in terms of phase transitions changing the Planck constant of the magnetic flux tube. The interaction would be repulsive (attractive) would result if the interaction energy increases (decreases) when Planck constant is reduced.

Magnetic interaction energy is certainly the best candidate and could also imply the equivalence of the divisor code and dark baryon code.

2. **Option II**: Could hydrophily and hydrophoby be described in terms of em interactions of quarks representing nucleotides in the model of DNA as TQC. For instance, could amino-acids and water molecules be characterized by charges which are of opposite sign for water molecules and hydrophilic molecules and of same sign for water molecules and hydrophobic molecules.

For **Option I**, which represents completely new physics (using the standards of TGD!), the situation looks promising. The magnetic interaction energy assignable to the flux tube is a function of the integers (n_a, n_b) characterizing the corresponding page of the book like structure associated with generalized imbedding space - in particular of the Planck constant of the flux tube - and the minimization is performed by keeping the charges of the quarks possibly at its ends fixed. This new physics fits also nicely with the idea that magnetic body controls the living matter by utilizing phase transitions changing Planck constant.

What comes in mind in the case of **Option II** is that the ends of the flux tube carry opposite charges correlating with the codon coding for the amino-acid and giving rise to ordinary gauge interactions. Unfortunately this scenario does not seem to work.

1. It was already found that (denoting codons by XYZ) only $Y = A, G$ type amino-acid residue can form hydrogen bonds and is hydrophilic and thus interacts strongly with water and DNA and RNA. If water end of flux tube corresponds to anti-quarks the attractive interaction between quark and anti-quark at the ends of flux tube could relate to hydrophily. For hydrophobic amino-acids one would have interaction between identical quarks and already Fermi statistics would cause repulsion. In DNA as TQC model based on the coding of A, G and T, C in terms of quarks u, d and their anti-quarks hydrophily-hydrophoby dichotomy corresponds to matter-antimatter dichotomy for quark assigned to the ends of the flux tube. Quarks and anti-quark have opposite charges. Hence the flux tube ends of hydrophilic amino-acids could correspond to quarks and water and hydrophobic ends of flux tubes to anti-quarks. Therefore the DNA as TQC model would predict the needed behavior of the forces. In the case of Gly containing only hydrogen as residue the flux tube might be simply absent.

2. DNA codons A, T, C, G are bases and thus polar and hydrophilic. In the case of DNA charge conjugation for quarks corresponds to the puridine-pyrimidine complementarity corresponding to conjugation of nucleotides. The rule applying in the case of amino-acids would predict T, C to be hydrophobic nucleotides which does not make sense. Therefore it seems that hydrophily and hydrophoby cannot reduce to the interactions of dark quarks and that they only represent conjugation of nucleotides symbolically.

5.2 An Improve Model For Protein Folding

To begin with let us summarize some basic facts about protein folding.

1. Hydrophily and hydrophoby play a key role in protein folding and dictate to a high degree the resulting folding patterns. This suggests that one cannot neglect the role of water in the process.

2. Protein folding proceeds from short to long length scales starting with the formation of secondary structures such as alpha helices, beta sheets, and random coil portions and is followed by the formation of tertiary and higher structures.

3. The formation of hydrogen bonds is in a decisive role in the formation of secondary structures. The mechanism leading to their formation might be contraction of magnetic flux tube by a phase transition changing Planck constant.

4. The folding patterns do not depend strongly on the precise primary structure, that is precise amino-acid decomposition which suggests that instead of the detailed chemistry the forces between quarks and anti-quarks mediated by flux tubes is what matter so that hydrophily

and hydrophoby would become the basic characterizers of the interaction. The phase transitions changing Planck constant would indeed represent this kind of universal interactions independent of the chemistry.

5. In the first approximation amino-acids could be labeled by a variable telling whether it is hydrophobic, hydrophilic, or neither or these (Gly). This approximation would be broken by special amino-acids which appear in edges if beta sheets (Pro) and Cys which often appear as S-S boded pair in junctions. By bringing in forces depending on the angles between tangent vectors of successive amino-acids and on amino-adics themselves this tendency could be modeled.

5.3 A Model For Which The Magnetic Body Of Water Is Involved

The alternative approach to protein folding starts from the general vision about magnetic body containing dark matter as a controller of visible matter in living system. The protein and its magnetic body would be regarded as a living system in itself.

1. Magnetic body must have large number of flux tube contacts to the visible matter. An excellent candidate for the magnetic body is that assignable with water and having flux tube connections to DNA and both hydrophilic and hydrophobic amino-acids. The magnetic body could control and at least fasten the self-organization process leading to the folding pattern which - by applying standard argument - would otherwise take astronomical time otherwise. The two-step attractive connections between all hydrophilic amino-acids would be possible via the magnetic body of water. The non-hydrophilic amino-acids not in direct contact with water are known to be more like passive structural stuff responsible for a fixed structure but not so relevant for the functioning of the bio-molecule. Hydrophily and hydrophoby would reflect the dependence of interaction energy on the value of Planck constant associated with the flux tube mediating the interaction.

2. This picture implies a straightforward modification of the earlier model. The simplest model would minimize a potential function V expressible as a sum $V = V_1 + V_2 + V_3$ of three terms. V_1 would be sum of the values of a universal two-particle potential function $V_{phi,phi}(r)$ for arguments $r_{ij} = |r_i - r_j|$ varying over all hydrophilic amino-acid pairs and giving rise to an attractive force. V_2 would be a sum of a universal two-particle potential function $V_{pho,pho}(r)$ for arguments $r_{ij} = |r_i - r_j|$ varying over all hydrophobic amino-acid pairs. V_3 would be would be sum of the values of a universal potential function $V_{phi,pho}(r)$ for arguments $r_{ij} = |r_i - r_j|$ varying over all pairs of hydrophilic and hydrophobic amino-acids. This potential function would induce a repulsive force. Besides this a constraint force due to the fact that amino-acids form a sequence would be present.

3. The resultant of the forces along lines connecting amino-acids would be parallel to the amino-acid sequence in the mechanical equilibrium. Hydrogen bonds and other bonds are indeed formed between neighboring hydrophilic amino-acids and the contraction of the flux tubes connecting the amino-acids in question to the magnetic body of water could be the mechanism. The model seems to be consistent with the basic qualitative facts about folding. The quantitative testing of the model would require determination of the conformations minimizing the potential function subject to the constraint provided by amino-acid sequence. Here of course the freedom to choose the three functions provides a considerable flexibility and symmetry arguments might allow to pose conditions on the form of these functions.

4. One could also include to the potential function describing a direct interaction with water molecules depending on parameters like pH affecting the folding pattern. The resultant for a given amino-acid would be sum of forces directed from a hydrophilic amino-acids to neighboring water molecules. It is not clear whether the normal component of this force could be compensated by the induced forces between amino-acids in a typical equilibrium configuration and the formation of hydrogen bonds involving the contraction of the flux tube could be the manner to achieve this.

The alternative model is more complicated numerically than the model discussed and it would require a considerable amount of work to test it. In particular, the three universal potential functions involve free parameters even if one makes simplifying assumptions about their functional form (say simple behavior under scaling).

6 A Model For Protein Folding And Catalytic Action

It would be fascinating if the vision about the role of flux tube connections would generalize to interactions of all molecules in living matter. The mere selection rules would mean hidden simplicity behind extremely complex looking interactions in living matter. The model for protein folding and catalytic action discussed in [K7] is the first attempt in this direction. In the following this model is briefly summarized and the improvement of the model inspired by recent considerations is suggested.

6.1 Earlier Model For The Folding Code

The model for the evolution of the genetic code led [K8] to the idea that the folding of proteins obeys a code inherited from the genetic code. One can imagine several variants of this code. One of the is that amino-acid behaves like the conjugate Y_c of the middle nucleotide of the codon XYZ coding for it. Conjugation for amino-acids would correspond to the hydrophilic-hydrophobic dichotomy. Also catalyst action could reduce to effective base pairing in this picture chemically and at the level of quarks associated with the flux tube to matter antimatter conjugation. The guess that amino-acid and its conjugate form pairs turned out to be wrong however and after various twists and turns I ended up with the hypothesis that the amino-acid in protein behaves like Y_cZ_c where Z corresponds to third nucleotide for some codon coding for the amino-acid.

It however turned that the model as such is probably too restrictive and not fully consistent in the particular cases studied. In the following this model is discussed briefly and later an improved model for protein folding is proposed.

6.1.1 Flux tubes as correlates of directed attention at molecular level

After some trials one ends up with a general conceptualization of the situation with the identification of ("wormhole") magnetic flux tubes as correlates for attention at molecular level so that a direct connection with TGD inspired theory of consciousness emerges at quantitative level. Whether wormhole flux tubes or ordinary flux tubes are needed is not a completely settled question yet and the attribute "wormhole" will not be used in the sequel. This allows a far reaching generalization of the DNA as topological quantum computer paradigm and makes it much more detailed. The final outcome is very simple quantitative model for both protein folding and catalyst action based on minimization of energy, which seems to be consistent with basic experimental facts as well as general ideas.

6.1.2 What kind of atoms can be connected by flux tubes?

1. Hydrogen bonds play a key role in bio-catalysis but are not understood completely satisfactorily in the standard chemistry. Hence the basic question is whether hydrogen bonds can be regarded as or are accompanied by short (wormhole) magnetic flux tubes: note that the subject-object asymmetry of directed attention would correspond to donor-acceptor asymmetry of they hydrogen bond. If this is the case, the identification of the magnetic flux tube connection as a prerequisite for a hydrogen bond or as hydrogen bond becomes natural. At least the atoms able to form hydrogen bonds could form flux tube contacts so that the model would be very predictive and would conform with the known important role of hydrogen bonds in bio-catalysis.

2. The fact that hydrogen bonds connect base pairs suggests a generalization of the notion of base pairing stating that under some conditions amino-acids coded by XYZ and UY_cV can behave like base pairs. These amino-acic pairs correspond to pairs of amino-acid residues which are hydrophilic *resp.* hydrophobic and hydrophobic residue do not form hydrogen

bonds in general. These flux tubes would thus be more general and in general long. The model for DNA as topological quantum computer requires this kind of flux tubes and they would in general connect atoms or molecules which act as acceptors in hydrogen bonding: $O =$ atom in amino-acid and aromatic ring are basic examples.

3. If one assumes that both $N-H$ and $O =$ associated with the constant part of the amino-acid can act as flux tube terminals and represent Z and Y nucleotides of the codon XYZ coding for the amino-acid, one obtains $Y = Z$ pairing of $O = -O =$ flux tubes are allowed and $Y = Z_c$ pairing if only hydrogen bond like pairings are allowed.

6.1.3 Color inheritance by a reconnection of flux tubes

1. There should exist some mechanism allowing amino-acids to inherit the base pairing property from the tRNAs associated with them so that one can identify amino-acid with the middle nucleotide of the codon coding it. If tRNA middle nucleotide is connected to $O =$ of the amino-acid, this becomes possible since the reconnection of flux tubes preserves the "color" of the flux tubes coded by (A, T, G, C) that is by the quark or anti-quark coding for the nucleotide. The temporary formation of a hydrogen bond between $N-H$ and $O =$ of two amino-acids as in the case of alpha helix would allow $N-H$ to inherit the conjugate of the color associated with $O =$. Alternative interpretation is that this hydrogen bond is possible only if the predetermined color of $N-H$ is consistent with the inherited one. The inheritance of flux tube color would be a completely general mechanism and even the donor atoms in the residues of amino-acids could inherit the color of $O =$ in this manner.

2. A possible interpretation for the fixing of the flux tube color is in terms of quantum measurement selecting one color from quantum superposition in the reconnection process. This would mean that the unitary process can bring superposition back and reconnection process can change the inherited color. The hydrogen bonds between water molecules could correspond to quantum superpositions of different colors. This superposition property might relate to the wobble base pairing phenomenon for the third nucleotide in tRNA.

6.1.4 Folding code

The identification of $N-H$ as a representation for the conjugate of the third nucleotide Z means that amino-acids would remember which codon coded them. If only hydrogen bond like flux tubes are allowed, flux tubes can connect only amino-acids satisfying $Y = Z_c$. If $= O - O =$ flux tubes are allowed $Y = Z$ rule favored by the model of DNA as topological quantum computer follows. The isospin symmetry of the third nucleotide implies that both rules are quite flexible. If one identifies hydrogen bond with flux tube ($Y(n) = Z(n+k)$) the model works badly for both options. If one assumes only that the presence of a flux tube connecting amino-acids in either direction ($Y(n) = Z(n+k)$ or $Z(n) = Y(n+k)$) is a prerequisite for the formation of hydrogen bond, the model works. $Y = Z$ rule is favored by the study of five enzymes: the possible average length of alpha helix is considerably longer than the average length of alpha helix if gene is the unique gene allowing to satisfy $Y = Z$ rule. The explicit study of alpha helices and beta sheets for these enzymes demonstrates that the failure to satisfy the condition for the existence of hydrogen bond fails rarely and at most for two amino-acids (for 2 amino-acids in single case only).

$Y = Z$ rule could mean a solution of the basic problem of proteonics: Do genes determine the folding of proteins and how this would take place? The interpretation would be that the information loss suggested by the many-to-one character of the genetic code is only apparent. The apparently lost information which corresponds to the $A-G$ and $T-C$ symmetries of the third nucleotide codes for the hydrogen bonding and hence for the folding of the protein. The model in its most stringent form is easy to kill since in the case of alpha helices and beta sheets the hydrogen bonding fixes completely the DNA sequence coding for the protein. A weaker variant of the model based on quantum variant of wobble base pairing: in this case there are no conditions on DNA sequence. It turns out that only this variant works. Hence hydrogen bonded amino-acid behave as if they were coded by the unique codon consistent with $Y = Z$ rule.

6.1.5 Quantitative model

The quantitative model relies on the assumption that the contribution of a flux tube connecting two amino-acids to the potential energy depends only on the distance between the molecules in question. The extremals of the total interaction energy are same for any choice of the potential and only the absolute minimum of the interaction energy depends on the choice of the potential. The simplest potential corresponds to harmonic oscillator potential and would explain formation of alpha helices and beta sheets and with the fact that hydrophilic and hydrophobic residues tend to have a large distance and only few flux tube contacts. For large Planck constant also long flux tubes could correspond to attractive harmonic oscillator potential. Also the contribution of other interactions between neighboring amino-acids are expected to be present but are neglected in the simplest model. The model predicts alpha helices and beta sheets, and more generally, periodic structures, as solutions to energy minimization equations.

The model fails to catch completely the basic rules of protein folding, and the predictions are not fully consistent with empirical facts in the cases studied. A model in which the hydrophilic and hydrophobic interactions are mediated by flux tubes between magnetic bodies of the molecule and water molecule and in this manner induce long range interactions between amino-acids - somewhat like the attractive interactions of electrons with ions induce attractive interaction between the members of a Cooper pair - looks more attractive. This model is however computationally much heavier and is not discussed in [K7]. In the sequel a formulation of this model is discussed.

6.2 Hydrophily And Hydrophoby Number Theoretically

Amino-acids can be classified to hydrophilic and hydrophobic ones whereas all DNA codons are hydrophilic. Hydrophily and hydrophoby are believed to relate to the standard chemistry alone and this might be the case. One can however just for fun ask whether hydrophily and hydrophoby could have a connection with divisor code, formation of flux tubes connecting the molecule to water molecules, and phase transitions changing the value of Planck constant and changing the length of flux tube. I have discussed this idea already in the model of protein folding [K7].

To simplify the model assume that only single dark page is associated with water molecule and labeled by (n_a^W, n_b^W). Of course, several levels characterized by different integers are also possible and this would bring in additional flexibility. Both hydrophoby and hydrophily would mean interaction mediated by the flux tubes to the magnetic body of water with the sign of the force differing for hydrophilic and hydrophobic amino-acids. There is no need to assume that quarks and anti-quarks generate the interaction. Gly for which the residue is just hydrogen atom does not allow classification as a hydrophilic or hydrophobic which would suggest that it does not have any flux tube connections with the magnetic body of the water. The interaction mediated by flux tubes between amino-acids and water molecules would be analogous to the interaction induced by the interaction between electrons and ions inducing attractive interaction between the members of Cooper pair. It would induce attractive interaction between hydrophilic amino-acids and repulsive interaction between hydrophilic and hydrophobic amino-acids favoring the formation of hydrophilic outer surfaces and hydrophobic inner surfaces.

One could understand hydrophily/hydrophoby dichotomy number theoretically for both options. The discussion of the first option makes clear that also second option is possible to realize.

1. Assume that n_a^W is divisible by all integers n_a^{DNA} associated with DNA codons and thus involves suitable powers of primes $p \leq 19$. It could contain also an integer factor which is product of primes larger than $p = 19$. This is necessary for achieving hydrophily of DNA codons.

2. Hydrophily of DNA codons also requires n_b^W must be proportional to the product of coprime integers n_b^W (primes for the simplest option) assignable to DNA codons. n_b^W could involve also a factor proportional to second integer expressible as product of primes $p > 19$. The simplest option is that this integer equals to 1.

3. For hydrophobic amino-acids integers n_b^A must be of form $mn_b^A = n_b^{DNA} m_b$ such that m_a does not divide n_b^W and n_b^W. This is enough to guarantee that magnetic flux tubes in either direction are impossible so that hydrophoby is guaranteed in the proposed sense. This

definition extends also to other molecules and can be expressed in terms of the integers (n_a, n_b) labeling the magnetic body of the molecule.

4. Second option is obtained by assigning the integer m_b only to Gly which is neither hydrophilic nor hydrophobic.

6.3 Could There Be New Physics Behind Hydrophily And Hydrophoby?

One could accept just as a fact that magnetic flux tubes to the magnetic body of water mediate an interaction which is attractive or repulsive between water molecules and amino-acids and attractive between DNA molecules and water. Accepting that this induces interaction between amino-acids one could proceed to model building without any mention about TGD.

One could also try to dig deeper and ask what might be the origin of this interaction.

1. **Option I**: Could one understand the interaction in terms of phase transitions changing the Planck constant of the magnetic flux tube. The interaction would be repulsive (attractive) would result if the interaction energy increases (decreases) when Planck constant is reduced. Magnetic interaction energy is certainly the best candidate and could also imply the equivalence of the divisor code and dark baryon code.

2. **Option II**: Could hydrophily and hydrophoby be described in terms of em interactions of quarks representing nucleotides in the model of DNA as TQC. For instance, could amino-acids and water molecules be characterized by charges which are of opposite sign for water molecules and hydrophilic molecules and of same sign for water molecules and hydrophobic molecules.

For **Option I**, which represents completely new physics (using the standards of TGD!), the situation looks promising. The magnetic interaction energy assignable to the flux tube is a function of the integers (n_a, n_b) -in particular of the Planck constant of the flux tube- and the minimization is performed by keeping the charges of the quarks possibly at its ends fixed. This new physics fits also nicely with the idea that magnetic body controls the living matter by utilizing phase transitions changing Planck constant.

What comes in mind in the case of **Option II** is that the ends of the flux tube carry opposite charges correlating with the codon coding for the amino-acid and giving rise to ordinary gauge interactions. Unfortunately this scenario does not seem to work.

1. In [K7] it was found that (denoting codons by XYZ) only $Y = A, G$ type amino-acid residue can form hydrogen bonds and is hydrophilic and thus interacts strongly with water and DNA and RNA. If water end of flux tube corresponds to anti-quarks the attractive interaction between quark and anti-quark at the ends of flux tube could relate to hydrophily. For hydrophobic amino-acids one would have interaction between identical quarks and already Fermi statistics would cause repulsion. In DNA as TQC model based on the coding of A, G and T, C in terms of quarks u, d and their anti-quarks hydrophily-hydrophoby dichotomy corresponds to matter-antimatter dichotomy for quark assigned to the ends of the flux tube. Quarks and anti-quark have opposite charges. Hence the flux tube ends of hydrophilic amino-acids could correspond to quarks and water and hydrophobic ends of flux tubes to anti-quarks. Therefore the DNA as TQC model would predict the needed behavior of the forces. In the case of Gly containing only hydrogen as residue the flux tube might be simply absent.

2. DNA codons A, T, C, G are bases and thus polar and hydrophilic. In the case of DNA charge conjugation for quarks corresponds to the puridine-pyrimidine complementarity corresponding to conjugation of nucleotides. The rule applying in the case of amino-acids would predict T, C to be hydrophobic nucleotides which does not make sense. Therefore it seems that hydrophily and hydrophoby cannot reduce to the interactions of dark quarks and that they only represent conjugation of nucleotides symbolically.

6.4 An Improved Model For Protein Folding

To begin with let us summarize some basic facts about protein folding.

1. Hydrophily and hydrophoby play a key role in protein folding and dictate to a high degree the resulting folding patterns. This suggests that one cannot neglect the role of water in the process.

2. Protein folding proceeds from short to long length scales starting with the formation of secondary structures such as alpha helices, beta sheets, and random coil portions and is followed by the formation of tertiary and higher structures.

3. The formation of hydrogen bonds is in a decisive role in the formation of secondary structures. The mechanism leading to their formation might be contraction of magnetic flux tube by a phase transition changing Planck constant.

4. The folding patterns do not depend strongly on the precise primary structure, that is precise amino-acid decomposition which suggests that instead of the detailed chemistry the forces between quarks and anti-quarks mediated by flux tubes is what matter so that hydrophily and hydrophoby would become the basic characterizers of the interaction. The phase transitions changing Planck constant would indeed represent this kind of universal interactions independent of the chemistry.

5. In the first approximation amino-acids could be labeled by a variable telling whether it is hydrophobic, hydrophilic, or neither or these (Gly). This approximation would be broken by special amino-acids which appear in edges if beta sheets (Pro) and Cys which often appear as S-S boded pair in junctions. By bringing in forces depending on the angles between tangent vectors of successive amino-acids and on amino-adics themselves this tendency could be modeled.

The earlier approach to protein folding inspired by DNA as TQC idea did not start from this picture but assumed that direct flux tube connections between amino-acids rather than the interactions induced by flux tube connections with the magnetic bodies of water molecules were responsible for the folding. The model did not lead to any spectacular results and the proposed rules were not fully consistent in the cases studied.

6.5 The Model For Which The Magnetic Body Of Water Is Involved

The improved approach to protein folding starts from the general vision about magnetic body containing dark matter as a controller of visible matter in living system. The protein and its magnetic body would be regarded as a living system in itself.

1. Magnetic body must have large number of flux tube contacts to the visible matter. An excellent candidate for the magnetic body is that assignable with water and having flux tube connections to DNA and both hydrophilic and hydrophobic amino-acids. The magnetic body could control and at least fasten the self-organization process leading to the folding pattern which - by applying standard argument - would otherwise take astronomical time otherwise. The two-step attractive connections between all hydrophilic amino-acids would be possible via the magnetic body of water. The non-hydrophilic amino-acids not in direct contact with water are known to be more like passive structural stuff responsible for a fixed structure but not so relevant for the functioning of the bio-molecule. Hydrophily and hydrophoby would reflect the dependence of interaction energy on the value of Planck constant associated with the flux tube mediating the interaction.

2. This picture implies a straightforward modification of the earlier model. The simplest model would minimize a potential function V expressible as a sum $V = V_1 + V_2 + V_3$ of three terms. V_1 would be sum of the values of a universal two-particle potential function $V_{phi,phi}(r)$ for arguments $r_{ij} = |r_i - r_j|$ varying over all hydrophilic amino-acid pairs and giving rise to an attractive force. V_2 would be a sum of a universal two-particle potential function $V_{pho,pho}(r)$ for arguments $r_{ij} = |r_i - r_j|$ varying over all hydrophobic amino-acid pairs. V_3 would be would

	A and D	*A or D*	*no flux tubes*
$X_{12} \times X_{12}$	$T \times T$	$T \times t$	$t \times t$
$X_{34} \times X_{34}$	$U \times U$	$U \times V$	$V \times V$
$X_{12} \times X_{34}$	$X_{12} \times W$	$T \times X$	$t \times X$

Table 1: General structure of pairings for $Y = Z_c$ and $Y = Z$ options. *A and D* means that both amino-acids can act as acceptors and donors. *A or D* that only acceptor or donor property is possible.

be sum of the values of a universal potential function $V_{phi,pho}(r)$ for arguments $r_{ij} = |r_i - r_j|$ varying over all pairs of hydrophilic and hydrophobic amino-acids. This potential function would induce a repulsive force. Besides this a constraint force due to the fact that amino-acids form a sequence would be present.

3. The resultant of the forces along lines connecting amino-acids would be parallel to the amino-acid sequence in the mechanical equilibrium. Hydrogen bonds and other bonds are indeed formed between neighboring hydrophilic amino-acids and the contraction of the flux tubes connecting the amino-acids in question to the magnetic body of water could be the mechanism. The model seems to be consistent with the basic qualitative facts about folding. The quantitative testing of the model would require determination of the conformations minimizing the potential function subject to the constraint provided by amino-acid sequence. Here of course the freedom to choose the three functions provides a considerable flexibility and symmetry arguments might allow to pose conditions on the form of these functions.

4. One could also include to the potential function describing a direct interaction with water molecules depending on parameters like pH affecting the folding pattern. The resultant for a given amino-acid would be sum of forces directed from a hydrophilic amino-acids to neighboring water molecules. It is not clear whether the normal component of this force could be compensated by the induced forces between amino-acids in a typical equilibrium configuration and the formation of hydrogen bonds involving the contraction of the flux tube could be the manner to achieve this.

6.5.1 Could one regard amino-acids and DNAs of given type as analog of species?

An interesting idea raised by the work with the model for protein folding is that the magnetic bodies amino-acids or DNA codon of a given type could behave like single phase on their respective page of the book so that the mutual interactions of their magnetic bodies could affect considerably the behavior of this phase to first order although amino-acids themselves are at different positions and one might expect only small correlations between their motions. Whether the dynamics of amino-acids of given type in protein folding are strongly correlated could be tested.

In certain sense one could speak of single species formed by amino-acids of given type and folding as long range interaction could be seen as an outcome of self-organizing interaction between members of various species and between species themselves plus short range constraints due to the fact that amino-acids form a sequence. The question applies to DNA and RNA codons and also to larger units such as genes formed to which one could assign their own page of the book. Water would represent the page to which all DNAs can send flux tubes. Even the notion of biological species could involve common dark space-time sheet(s) where the magnetic bodies of the members of species are and interact making the members of species to behave like single coherent unit.

6.6 Appendix: Tables related to the model of protein folding

Table 6 represents the results of the test when flux tube from $Y(n)$ to $Z(n+k)$ or from $Z(n)$ to $Y(n+k)$ is prerequisite for hydrogen bond.

$A \times D$	*no flux tubes*
$X_{12} \times X_{12}$	$X_{12} \times t_1$
$X_{34} \times X_{34}$	$X_{34} \times V$
$X_{12} \times X_{34}$	$X_{12} \times W$
$X_{34} \times X_{12}$	$X_{34} \times t_2$

Table 2: General structure of pairings for $Y = Z_c$ and $Y = Z$ options. A and D refer to acceptor $(O =)$ and donor $N - H$ respectively. Only non-allowed hydrogen bonded pairs are listed.

protein	L(3)	L(4)	L(5)	L(6)
asparagin synthethase				
$Y = Z_c$	11.8	15.0	12.2	13.2
$Y = Z$	55.0	47.1	47.1	47.1
xylose isomerase				
$Y = Z_c$	10.2	9.7	12.4	11.3
$Y = Z$	24.8	24.8	16.5	26.4
hydrolase				
$Y = Z_c$	13.8	18.4	16.6	12.8
$Y = Z$	55.3	20.8	33.2	27.7
glutathione s-transferase				
$Y = Z_c$	12.4	12.4	13.1	15.0
$Y = Z$	35.0	35.0	26.3	30.0
BamHI				
$Y = Z_c$	9.7	8.5	10.7	10.7
$Y = Z$	30.4	23.7	30.4	35.5

Table 3: The average number $L(k)$ of amino-acids in the portion of amino-acid sequence satisfying the conditions making possible $(n, n+k)$ hydrogen bonding for $k = 3, 4, 5, 6$ for $Y = Z_c$ and $Y = Z$ option in the case that flux tube can connect $Y(n)$ to $Z(n + k)$ or $Z(n)$ to $Y(n + k)$.

REFERENCES

Biology

[I1] Alpha helix. Available at: http://en.wikipedia.org/wiki/Alpha_helix.

[I2] Asparagine Synthetase. Available at: http://www.pdb.org/pdb/files/11as.pdb.

[I3] Bamhi. Available at: http://www.rcsb.org/pdb/files/1bam.pdb.

[I4] Bamhi. Available at: http://rebase.neb.com/cgi-bin/seqsget?X55285.

[I5] Beta sheet. Available at: http://en.wikipedia.org/wiki/Beta_sheet.

[I6] Collagen. Available at: http://en.wikipedia.org/wiki/Collagen.

[I7] Glutathione S-transferase. Available at: http://www.pdb.org/pdb/files/14gs.pdb.

[I8] Hydrolase(O-glycosyl). Available at: http://www.pdb.org/pdb/files/107l.pdb.

[I9] Phospholipids. Available at: http://en.wikipedia.org/wiki/Phospholipids.

[I10] Secondary structures. Available at: http://en.wikipedia.org/wiki/Secondary_structure.

[I11] Xylose Isomerase. Available at: http://www.pdb.org/pdb/files/1a0c.pdb.

[I12] Buchanan M. Shape is all. *New Scientist*, 2157, 1998.

[I13] Rackovsky S. On the nature of the protein folding code. *Proc Natl Acad Sci USA*, 90(2), 1993. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=45720. See also *Protein folding* at http://en.wikipedia.org/wiki/Protein_folding.

[I14] Creighton TE. *Proteins: Structures and Molecular Properties*. Freeman, New York, 1993.

Books related to TGD

[K1] Pitkänen M. Bio-Systems as Super-Conductors: part I. In *Quantum Hardware of Living Matter*. Available at: http://tgdtheory.fi/pdfpool/superc1.pdf, 2006.

[K2] Pitkänen M. Negentropy Maximization Principle. In *TGD Inspired Theory of Consciousness*. Available at: http://tgdtheory.fi/pdfpool/nmpc.pdf, 2006.

[K3] Pitkänen M. Quantum Model for Nerve Pulse. In *TGD and EEG*. Available at: http://tgdtheory.fi/pdfpool/pulse.pdf, 2006.

[K4] Pitkänen M. Quantum Theory of Self-Organization. In *Bio-Systems as Self-Organizing Quantum Systems*. Available at: http://tgdtheory.fi/pdfpool/selforgac.pdf, 2006.

[K5] Pitkänen M. TGD Based Model for OBEs. In *TGD Inspired Theory of Consciousness*. Available at: http://tgdtheory.fi/pdfpool/OBE.pdf, 2006.

[K6] Pitkänen M. Wormhole Magnetic Fields. In *Quantum Hardware of Living Matter*. Available at: http://tgdtheory.fi/pdfpool/wormc.pdf, 2006.

[K7] Pitkänen M. A Model for Protein Folding and Bio-catalysis. In *Genes and Memes: Part I*. Available at: http://tgdtheory.fi/pdfpool/foldcat.pdf, 2019.

[K8] Pitkänen M. Evolution in Many-Sheeted Space-Time. In *Genes and Memes: Part I*. Available at: http://tgdtheory.fi/pdfpool/prebio.pdf, 2019.

alpha helix	$N(Y = Z)$	$N(Y = Z_c)$
xylose isomerase		
[74, 93]	(2, 1, 0, 1)(met-lys)	(1, 5, 1, 0)
[111, 129]	(0, 1, 0, 1)(asn-asp)	(2, 1, 0, 0)
[159, 179]	(0, 1, 0, 0)(asp-tyr)	(5, 4, 1, 3)
[201, 223]	(1, 1, 1, 2)(met-tyr)	(4, 3, 4, 3)
[245, 255]	(2, 0, 0, 0)	(0, 1, 1, 0)
[278, 287]	(0, 1, 0, 0)(his-tyr)	(0, 0, 0, 1)
[314, 327]	(0, 0, 2, 1)	(2, 0, 1, 1)
[349, 374]	(0, 0, 4, 0)	(3, 4, 1, 2)
[376, 386]	(1, 0, 0, 0)	(2, 1, 1, 1)
[393, 399]	(0, 0, 0, 0)	(0, 0, 0, 1)
[404, 414]	(0, 0, 0, 0)	(1, 1, 1, 1)
[424, 435]	(0, 0, 1, 0)	(2, 1, 2, 0)
hydrolase		
[39, 50]	(0, 0, 0, 1)	(2, 1, 2, 1)
[60, 79]	(0, 1, 0, 0)(asn-asp)	(2, 2, 3, 1)
[93, 113]	(1, 1, 0, 1)(val-gly)	(3, 0, 2, 1)
[115, 121]	(1, 0, 0, 0)	(0, 0, 0, 0)
[126, 134]	(0, 0, 1, 0)	(1, 0, 0, 0)
[143, 155]	(0, 0, 0, 0)	(0, 0, 0, 1)
glutathione s-transferase		
[12, 24]	(0, 0, 1, 0)	(0, 0, 0, 0)
[65, 76]	(0, 0, 1, 0)	(0, 1, 0, 0)
[83, 108]	(3, 2, 3, 1)(met-glu-asp)	(0, 0, 1, 0)
[111, 134]	(0, 0, 1, 2)	(3, 1, 4, 2)
[150, 166]	(1, 1, 2, 0)(asp-leu)	(1, 1, 1, 0)
[174, 184]	(0, 0, 0, 0)	(0, 0, 0, 1)
[187, 194]	(0, 0, 0, 0)	(0, 0, 0, 0)
BamHI		
[10, 18]	(0, 0, 0, 0)	(0, 0, 0, 0)
[20, 34]	(0, 0, 0, 0)	(0, 1, 1, 1)
[58, 72]	(1, 0, 1, 1)	(0, 0, 0, 0)
[79, 84]	(0, 0, 0, 0)	(0, 0, 0, 0)
[117, 132]	(0, 0, 0, 0)	(0, 0, 0, 0)
[146, 150]	(1, 1, 0, 0)	(0, 0, 0, 0)
[159, 169]	(0, 0, 0, 0)	(0, 0, 0, 0)
[200, 205]	(0, 0, 0, 0)	(0, 0, 0, 0)

Table 4: The test for alpha helices of four enzymes. The first column gives the range of amino-acids defining the alpha helix in question. The vectors in second and third column give the numbers of failures for $k = 3, 4, 5, 6$ for $(n, n + k)$ helix ($k = 4$ is the most interesting value). The amino-acid-pairs for which the hydrogen bond does not exist for $Y = Z$ option are given.

beta sheet	$N(Y=Z)$	$N(Y=Z_c)$
asparagin synthethase		
[113, 122]	0	2
[233, 240]	0	0
[245, 255]	0	0
[290, 297]	0	3
xylose isomerase		
[43, 47]	0	0
[96, 100]	0	2
[134, 140]	0	1
[262, 267]	0	0
[291, 295]	0	0
hydrolase		
[14, 19]	0	0
[25, 28]	0	0
glutathione s-transferase		
[3, 7]	0	0
[28, 32]	0	3
[54, 58]	0	0
[60, 63]	0	0
BamHI		
[2, 8]	0	0
[46, 48]	0	0
[70, 72]	0	0
[95, 100]	0	0
[105, 112]	0	0
[138, 144]	0	0
[174, 180]	0	0
[183, 185]	0	1

Table 5: The test for beta sheets of four enzymes. The first column gives the range of amino-acids defining the beta sheet in question. The vectors in second and third column give the numbers of failures for $k=1$ for $(n, n+1)$ helix.

protein	L(3)	L(4)	L(5)	L(6)
asparagin synthethase				
$Y=Z$	4.5	4.9	4.6	4.9
$Y=Z_c$	4.2	4.7	5.3	4.6
xylose isomerase				
$Y=Z$	3.7	4.3	3.5	3.9
$Y=Z_c$	4.0	3.0	4.1	3.8
hydrolase				
$Y=Z$	5.7	5.1	4.2	4.3
$Y=Z_c$	4.6	5.0	6.1	5.1
glutathione s-transferase				
$Y=Z$	5.1	5.8	4.6	4.0
$Y=Z_c$	4.5	3.6	4.0	4.7
BamHI				
$Y=Z$	11.3	16.8	15.6	12.9
$Y=Z_c$	16.8	10.4	12.1	12.9

Table 6: The average number $L(k)$ of amino-acids in the portion of amino-acid sequence satisfying the conditions making possible $(n, n+k)$ hydrogen bonding for $k=3,4,5,6$ for $Y=Z_c$ and $Y=Z$ option in the case that flux tube can connect $Y(n)$ to $Z(n+k)$.

alpha helix	$N(Y = Z)$	$N(Y = Z_c)$
asparagin synthethase		
[7, 28]	(4, 4, 4, 2)	(6, 4, 3, 5)
[76, 84]	(1, 0, 1, 1)	(2, 2, 1, 1
[130, 155]	(5, 5, 4, 5)	(2, 2, 2, 1)
[170, 177]	(0, 0, 1, 1)	(1, 1, 1, 1)
[182, 194]	(1, 1, 1, 1)	(2, 2, 1, 1)
[256, 268]	(1, 3, 0, 3)	(3, 0, 3, 0)
[277, 284]	(0, 0, 0, 0)	(0, 0, 0, 0)
[297, 305]	(0, 0, 0, 0)	(1, 0, 0, 0)
[309, 314]	(0, 1, 0, 0)	(2, 0, 1, 0)
[320, 326]	(1, 1, 0, 0)	(0, 0, 0, 0)
xylose isomerase		
[74, 93]	(6, 1, 6, 4)	(2, 5, 4, 3)
[111, 129]	(3, 3, 4, 2)	(4, 4, 4, 3)
[159, 179]	(3, 4, 6, 3)	(4, 3, 1, 3)
[201, 223]	(3, 6, 5, 4)	(5, 7, 4, 4)
[245, 255]	(1, 1, 1, 1)	(0, 1, 0, 0)
[278, 287]	(2, 1, 1, 0)(his-tyr)	(0, 0, 0, 1)
[314, 327]	(1, 2, 3, 3)	(2, 3, 2, 2)
[349, 374]	(4, 3, 8, 6)	(5, 8, 1, 3)
[376, 386]	(4, 0, 2, 3)	(1, 4, 3, 0)
[393, 399]	(0, 0, 0, 0)	(1, 0, 0, 0)
[404, 414]	(1, 1, 2, 1)	(3, 3, 1, 1)
[424, 435]	(1, 2, 0, 2)	(3, 2, 3, 1)
hydrolase		
[39, 50]	(2, 3, 0, 2)	(2, 1, 3, 1)
[60, 79]	(4, 2, 2, 4)(asn-asp)	(4, 4, 5, 2)
[93, 113]	(2, 1, 4, 4)(val-gly)	(3, 2, 1, 1)
[115, 121]	(0, 1, 0, 0)	(0, 1, 0, 0)
[126, 134]	(0, 1, 1, 1)	(1, 0, 0, 0)
[143, 155]	(0, 1, 1, 1)	(1, 0, 1, 1)
glutathione s-transferase		
[12, 24]	(1, 4, 2, 4)	(5, 2, 3, 1)
[65, 76]	(3, 1, 1, 0)	(2, 1, 1, 1)
[83, 108]	(5, 3, 5, 5)	(5, 3, 3, 2)
[111, 134]	(4, 5, 5, 4)	(4, 5, 3, 3)
[150, 166]	(2, 2, 3, 0)(asp-leu)	(4, 3, 1, 3)
[174, 184]	(0, 0, 0, 0)	(1, 1, 1, 1)
[187, 194]	(1, 1, 0, 1)	(0, 2, 0, 1)
BamHI		
[10, 18]	(0, 0, 0, 0)	(0, 0, 0, 0)
[20, 34]	(4, 1, 1, 2)	(2, 5, 4, 2)
[58, 72]	(3, 1, 2, 4)	(3, 5, 4, 1)
[79, 84]	(0, 0, 0, 0)	(0, 0, 0, 0)
[117, 132]	(0, 0, 0, 0)	(0, 0, 0, 0)
[146, 150]	(1, 1, 0, 0)	(0, 0, 0, 0)
[159, 169]	(0, 0, 0, 0)	(0, 0, 0, 0)
[200, 205]	(0, 0, 0, 0)	(0, 0, 0, 0)

Table 7: The test for alpha helices of four enzymes in the case of $Y(n) = Z(n + k)$ option. The first column gives the range of amino-acids defining the alpha helix in question. The vectors in second and third column give the numbers of failures for $k = 3, 4, 5, 6$ for $(n, n + k)$ helix ($k = 4$ is the most interesting value).

beta sheet	$N(Y = Z)$	$N(Y = Z_c)$
asparagin synthethase		
[113, 122]	1	5
[233, 240]	3	0
[245, 255]	1	0
[290, 297]	0	1
xylose isomerase		
[43, 47]	0	1
[96, 100]	2	4
[134, 140]	2	0
[262, 267]	2	1
[291, 295]	0	1
hydrolase		
[14, 19]	1	2
[25, 28]	0	0
glutathione s-transferase		
[3, 7]	0	1
[28, 32]	1	3
[54, 58]	4	1
[60, 63]	0	0
BamHI		
[2, 8]	0	0
[46, 48]	0	0
[70, 72]	2	0
[95, 100]	0	0
[105, 112]	4	0
[138, 144]	0	0
[174, 180]	0	0
[183, 185]	1	2

Table 8: The test for beta strands of four enzymes for $Y(n) = Z(n+1)$ option. The first column gives the range of amino-acids defining the beta sheet in question. The vectors in second and third column give the numbers of failures for $k = 1$ for $(n, n+1)$ helix.