S. V. Kozyrev
Steklov Mathematical Institute

Quinary lattice model of secondary structures of proteins

Talk given at the Conference on Mathematical Modelling in Biology, INM RAS, Moscow, Russia, October 12, 2012

We construct a new model of lattice polymers and describe for this model all possible secondary structures for sufficiently short lattice polymers.

This model belongs to a new class of lattice models with cooperative interaction.

S.V.Kozyrev, I.V.Volovich, Quinary lattice model of secondary structures of proteins, arXiv:1206.4424

Lattice polymer (of the length N). Conformation — a sequence of vertices in the cubic lattice $\mathbb{Z}^3$, i.e. the map

$$\Gamma : \{1, \ldots, N\} \to \mathbb{Z}^3,$$

without self intersections where neighbor natural numbers map to neighbor (i.e. distance one) vertices of the lattice.

Standard model of energy of lattice polymers — interaction of nearest neighbors

$$E_2(\Gamma) = - \sum_{1 \le i < j \le N} \delta(d(\Gamma(i), \Gamma(j))),$$

where $d(\cdot, \cdot)$ is the distance in $\mathbb{Z}^3$, $\delta(1) = 1$, $\delta(i) = 0$, $i > 1$.

does not describe secondary structures (special preferred conformations)

A.Yu.Grosberg, A.R.Khokhlov, Giant Molecules: Here, There, and Everywhere, 2nd ed., World Scientific Publishing Company, 2010.

A.Yu.Grosberg, Disordered polymers, Phys. Usp. (1997) V.40. P.125–158.

E.I.Shakhnovich, A.M.Gutin, Engineering of stable and fast-folding sequences of model proteins, Proc. Natl. Acad. Sci. USA, 1993, Vol. 90, pp. 7195-7199.

A.V.Finkelstein, O.B.Ptitsyn, Protein Physics. (2002). London–Amsterdam: Academic Press.

S. Istrail, F.Lam, Combinatorial algorithms for protein folding in lattice models: a survey of mathematical results, Communication in information and systems, 2009, Vol. 9, No. 4, P. 303–346.

A.N.Nekrasov, Analysis of the information structure of protein sequences: a new method for analyzing the domain organization of proteins. J. Biomol. Struct. Dyn. v.21(5), pp.615–624, (2004).

The quinary lattice model (a new form of cooperative interaction)

Conformation of a segment of a lattice polymer of length 5:

$$\Gamma_i = (\Gamma(i-2), \Gamma(i-1), \Gamma(i), \Gamma(i+1), \Gamma(i+2)), \quad i = 3, \ldots, N-2.$$

The energy is a sum of contributions from conformations of 5–tuples

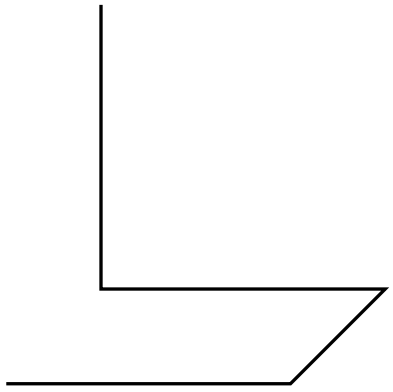
$$E_5(\Gamma) = -\sum_{i=3}^{N-2} \Phi(\Gamma_i).$$

The function Φ of conformations of 5-tuples is invariant with respect to lattice rotations and translations.

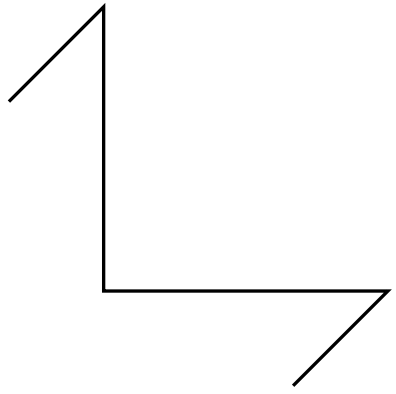
Φ is equal to zero for all conformations
except conformations denoted 1 and 2 for which

$$\Phi(1) = \Phi(2) = 1.$$

Conformation 1



Conformation 2



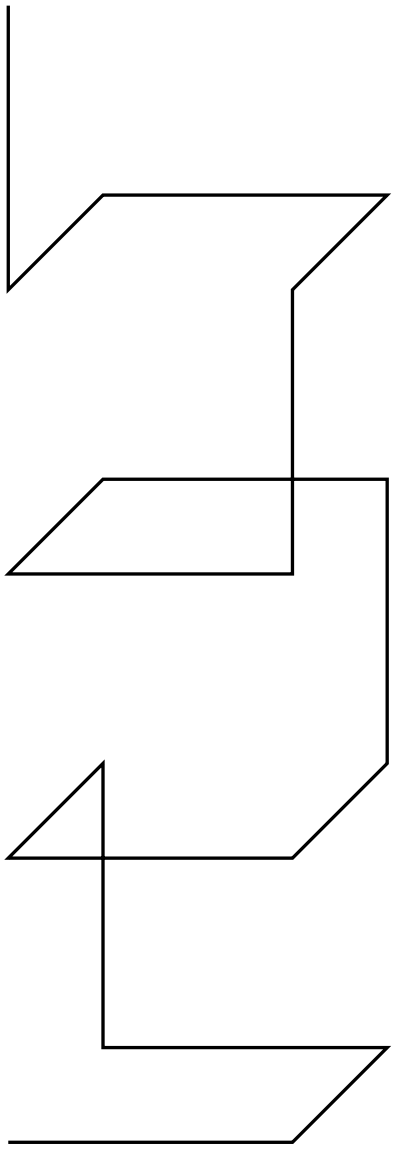
Conformations of a lattice polymer which
are minima of the energy E_5,
i.e. a conformation of any 5-tuple of neighbor monomers
either of the type 1 or of the type 2 are called *minimal*.

We consider these conformations as models
of secondary structures in proteins.

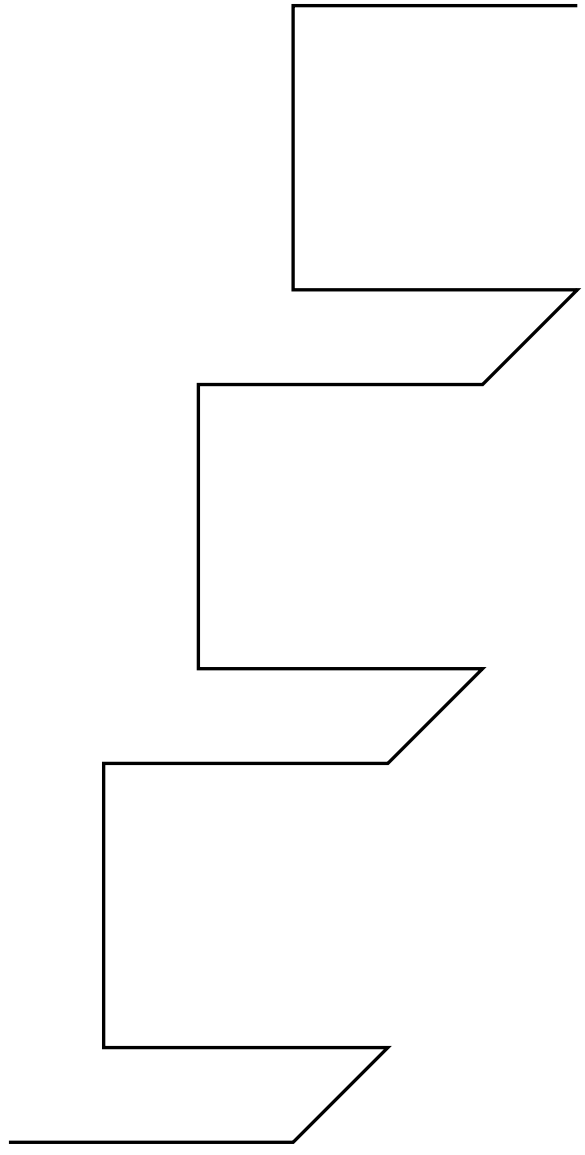
Examples of minimal conformations (see the next slide) —
lattice models of α-helix and β-strand

The quinary lattice model possesses
nontrivial secondary structures

α-helix



β-strand



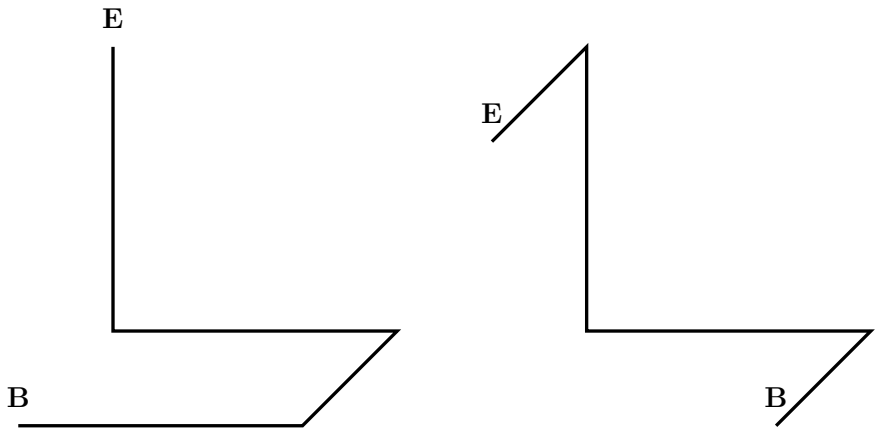
α–helix is a right lattice helix with the symmetry with respect to the rotation counterclockwise by $3\pi/2$ and translation by one upwards (according to the thumb rule).

β–strand is symmetric with respect to translations.

Let us describe minimal conformations (secondary structures) of lattice proteins

We consider for conformations 1 and 2 their *directions*.
The beginning B and the end E of the conformations.

$\overrightarrow{1}$, $\overrightarrow{2}$ and $\overleftarrow{1}$, $\overleftarrow{2}$ — conformations 1 and 2
which are considered from B to E and from E to B correspondingly (i.e. considered as oriented graphs).


E
B
E
B

Minimal conformation of a lattice polymer generates a sequence of symbols $\overrightarrow{1}$, $\overrightarrow{2}$, $\overleftarrow{1}$, $\overleftarrow{2}$ of conformations of 5-tuples of neighbor monomers (when we read the sequence of monomers in the polymer from the beginning to the end).

Which sequences $\Gamma_3\Gamma_4\ldots\Gamma_{N-2}$ of the symbols $\overrightarrow{1}$, $\overrightarrow{2}$, $\overleftarrow{1}$, $\overleftarrow{2}$ can be generated by minimal conformations of a lattice polymer?

It is not always possible to combine a couple of symbols into one conformation (due to geometric restrictions).
Two consecutive 5–tuples of monomers intersect:

$$\Gamma_i = (\Gamma(i-2), \Gamma(i-1), \Gamma(i), \Gamma(i+1), \Gamma(i+2)), \ldots,$$

$$\Gamma_{i+1} = \qquad \ldots, (\Gamma(i-1), \Gamma(i), \Gamma(i+1), \Gamma(i+2), \Gamma(i+3)).$$

This implies the selection rules for the conformations Γ_i, Γ_{i+1} of the consecutive 5–tuples.

Lemma 1) Possible pairs of neighbor conformations of 5–tuples in the sequence $\Gamma_3\Gamma_4\ldots\Gamma_{N-2}$ related to some minimal conformation of a lattice polymer are described by the following table

	$\overrightarrow{1}$	$\overleftarrow{1}$	$\overrightarrow{2}$	$\overleftarrow{2}$
$\overrightarrow{1}$	$-$	$+$	$+$	$-$
$\overleftarrow{1}$	$+$	$-$	$-$	$-$
$\overrightarrow{2}$	$-$	$-$	$-$	$+$
$\overleftarrow{2}$	$-$	$+$	$+$	$-$

(i.e. for any pair of symbols denoted by $+$ in the table above there exists a minimal conformation of a lattice polymer of length 6).

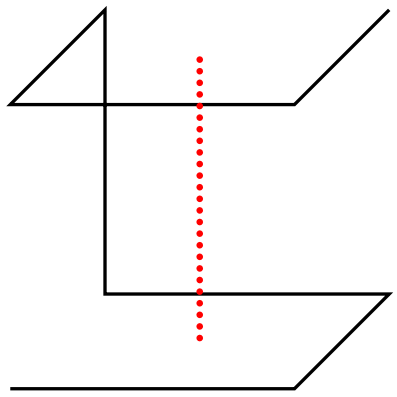
2) Any conformation described by a triple of symbols from $\{\overrightarrow{1},\ \overrightarrow{2},\ \overleftarrow{1},\ \overleftarrow{2}\}$ permitted by the above table corresponds to some minimal conformation of a lattice polymer of length 7 except the triples $\overrightarrow{2}\,\overleftarrow{2}\,\overrightarrow{2}$, $\overleftarrow{2}\,\overrightarrow{2}\,\overleftarrow{2}$..

Examples

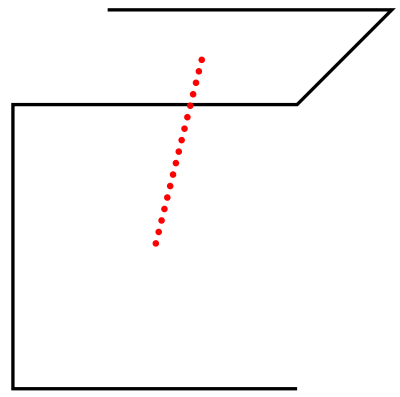
Lattice α-helix and β-strand — iterations of α and β conformations.

Central lines of α and β conformations
(red dotted line)
connect the centers of the corresponding faces of the cube containing the corresponding conformation.

Conformation $\alpha = \overrightarrow{1}\,\overrightarrow{2}\,\overleftarrow{2}\,\overleftarrow{1}$



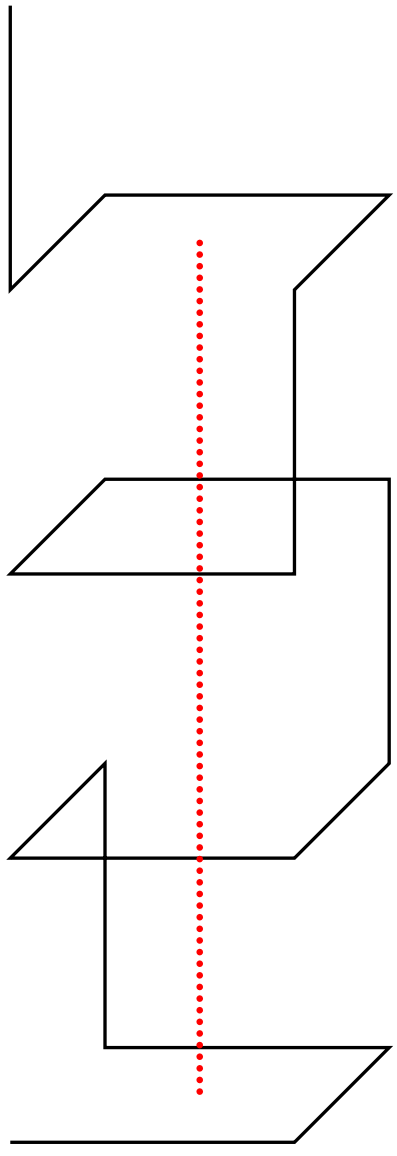
Conformation $\beta = \overrightarrow{1}\,\overleftarrow{1}$



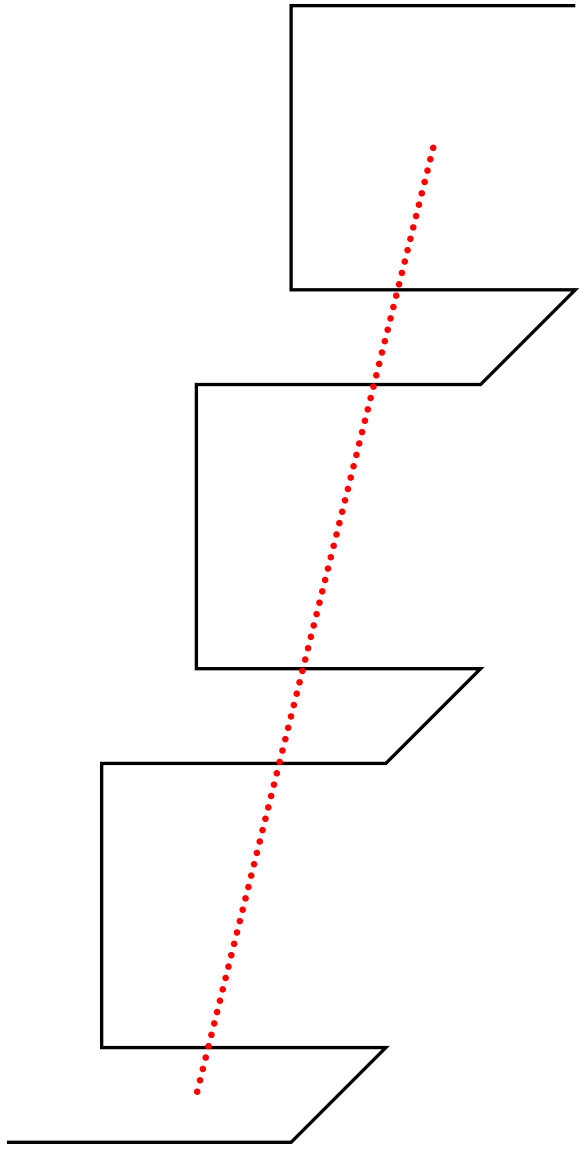
Iterations of α or iterations of β
— straight central lines,

mixture of α and β
— broken central line, angle 135°.

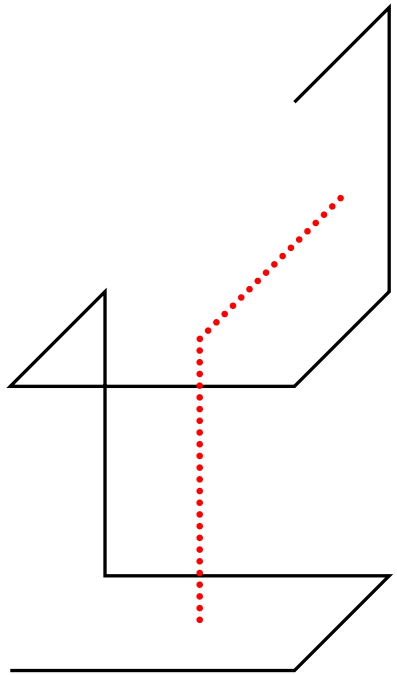
α-helix



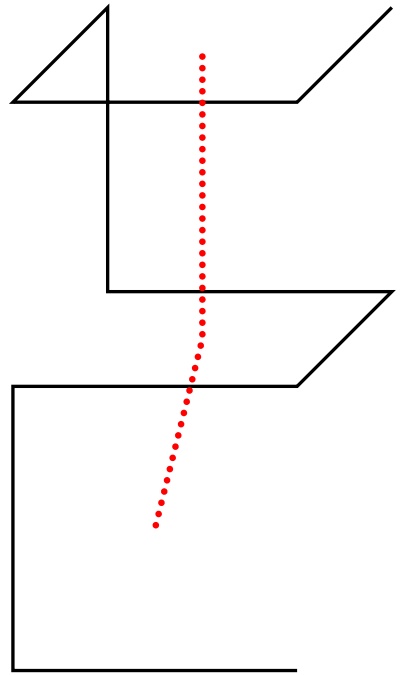
β-strand



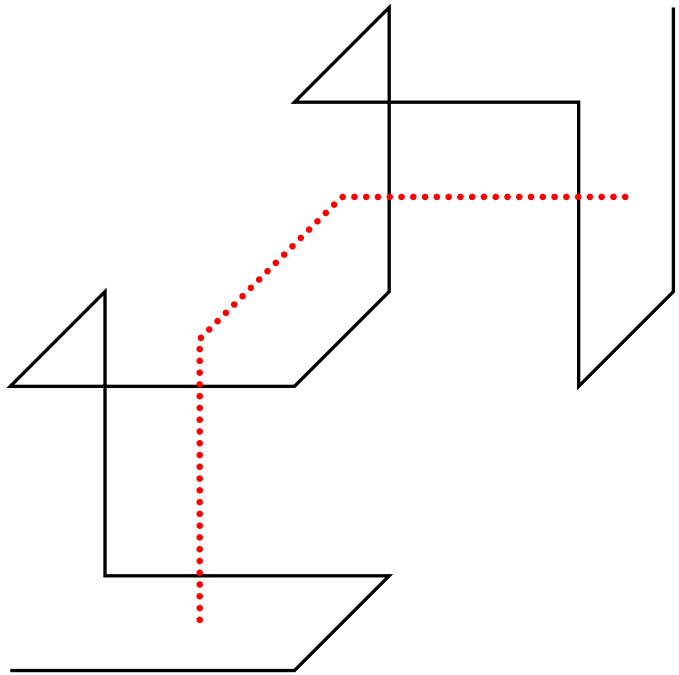
Conformation $\alpha\beta$



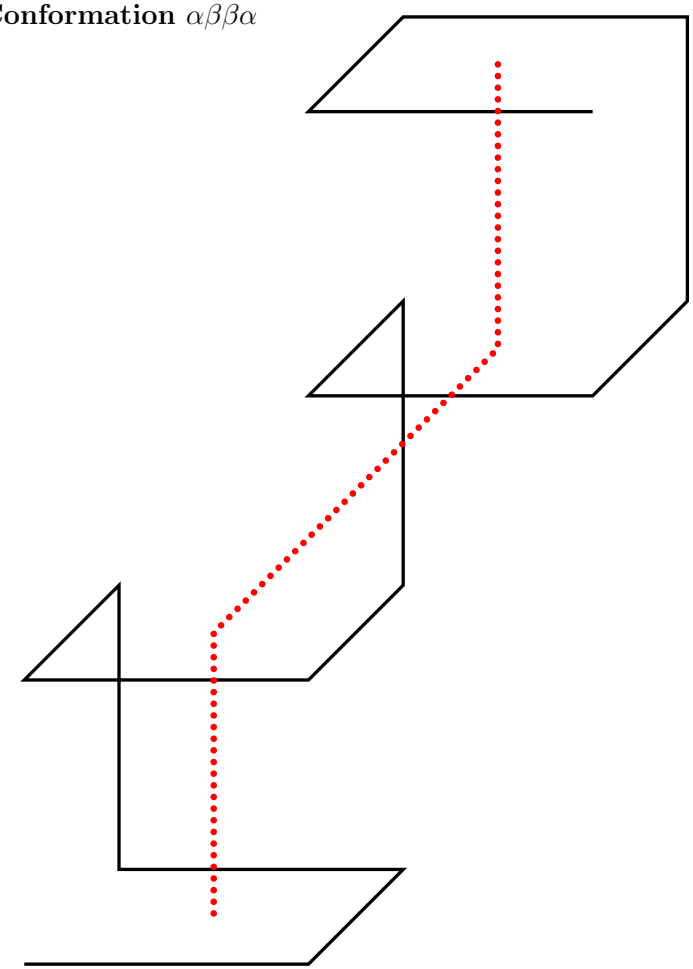
Conformation $\beta\alpha$



Conformation $\alpha\beta\alpha$



Conformation $\alpha\beta\beta\alpha$



Theorem

1) Any minimal conformation of a lattice polymer with the length $N > 6$ has the following form:
The corresponding sequence $\Gamma_3\Gamma_4 \ldots \Gamma_{N-2}$ of conformations of 5–tuples can be obtained from some sequence of α and β structures, $\alpha = \overrightarrow{1}\,\overrightarrow{2}\,\overleftarrow{2}\,\overleftarrow{1}$, $\beta = \overrightarrow{1}\,\overleftarrow{1}$ by elimination of a finite number of symbols $\overrightarrow{1}$, $\overrightarrow{2}$, $\overleftarrow{1}$, $\overleftarrow{2}$ in the beginning and the end of the sequence.

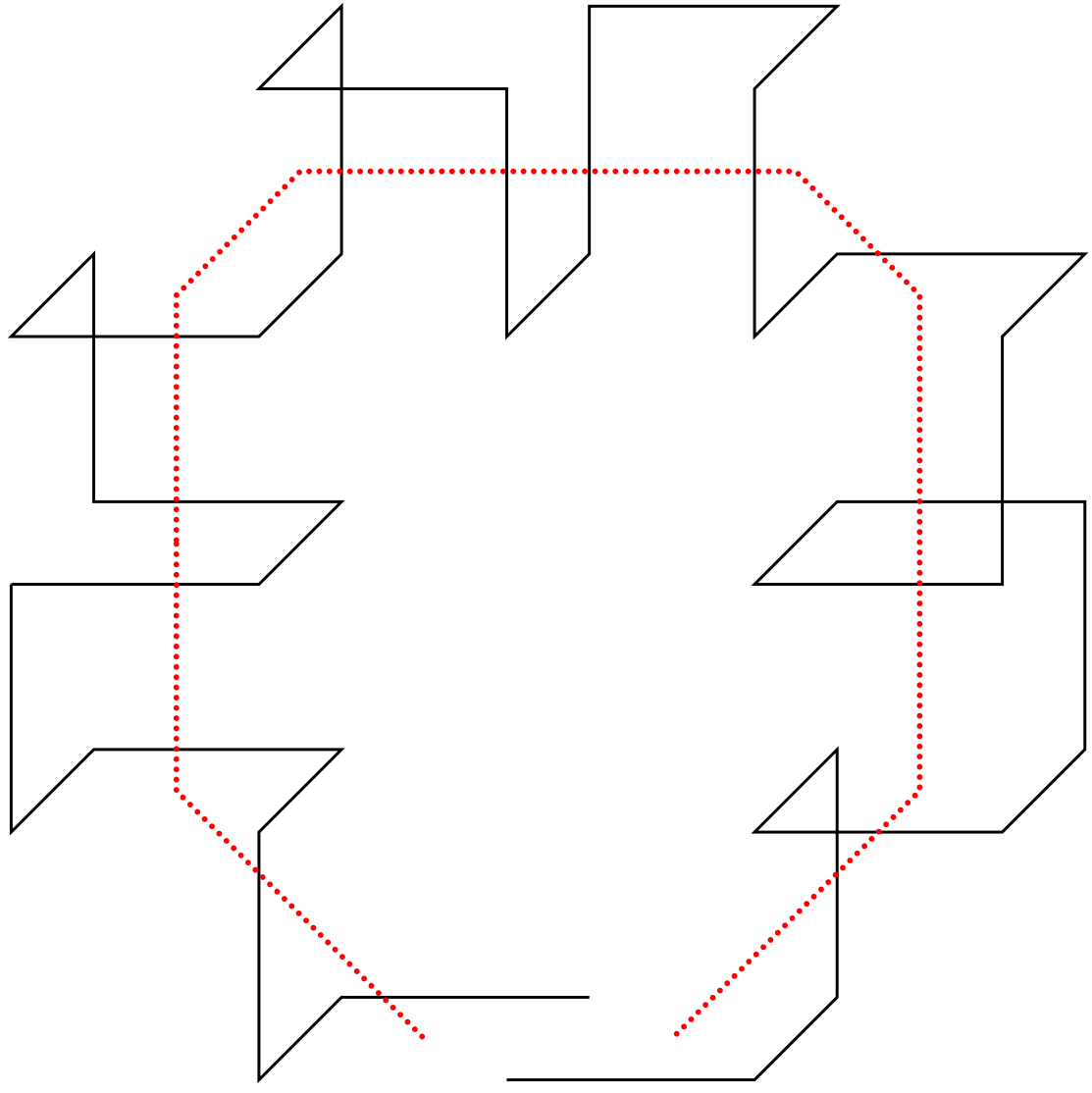
2) All conformations of a lattice polymer with the length $6 < N < 39$ obtained as above do not contain self intersections.
There exists a sequence of conformations of 5–tuples corresponding to the conformation of a lattice polymer of the length 39 with self–intersections.

An arbitrary minimal conformation of a lattice polymer is a combination of α and β structures,

all such combinations with the length of the polymer $6 < N < 39$ are possible.

Conformation $\overleftarrow{1}\beta\alpha\alpha\beta\alpha\alpha\beta\alpha\alpha\beta\overrightarrow{1}$ (38 vertices)

continuation possesses intersections



Conclusion

The quinary model of a lattice polymer is proposed.

New class of lattice models with cooperative interaction.

All minima of energy are combinations of
lattice α-helices and β-strands.

All conformations which are energy minima for
lattice polymers of length < 39 are described.

This new model allows to describe
the effect of existence of secondary structures of proteins.

Relation to real proteins?