Biopolimers helix-coil structure

JELENKA SAVKOVIC STEVANOVIC
Department of Chemical Engineering, Faculty of
Technology and Metallurgy, The University of Belgrade, 11000 Belgrade, Karnegijeva 4
SERBIA

Abstract: In this paper purine and pyrimidine bases and their parameters were studied. Molecular modelling of the protein transition state was provided. Probability of the helix-coil transition state was derived. Transition parameters were determined by fitting the temperature transition curves to the macroscopic experimental data. The obtained results in the frame of investigation of this paper show transition parameters values and transition dynamics for the poly-L-iso-leucyne and poly-valyne. The probabilistic factor, entropy and enthalpy changing were determined.

Keyword **:** Bioengineering, helix, coil, transition, entropy, enthalpy.

Received: June 19, 2021. Revised: March 26, 2022. Accepted: April 29, 2022. Published: July 2, 2022.

1. Introduction

The complex structural organization of an organism is specified by an inherited script conveying an enormous amount of coded information. A particular sequence of nucleotides says the same thing to one organism as it does to another, differences between organism reflect genetic programs of different nucleotides sequences [1]-[5].

Biomedical important it neither nucleotides nor their parent purine and pyrimidine bases in the diet are incorporated into human tissue nucleic acids or into purine or pyrimidine coenzymes. Even when a diet rich in nucleoproteins is ingested, human subjects form the constituents of tissue nucleic acids from amphibolic intermediates. This de novo synthesis permits purine and pyrimidine analogs with potential as anticancer drugs to be incorporated into DNA. The rates of synthesis of purine and pyrimidine oxy- and deoxyribonucleotides are subject to precise regulation [2],[3]. Mechanisms have evolved to ensure production of these compounds in quantities and at times appropriate to meet varying physiologic demand [4]-[6].

The structure and function of the purines and pyrimidines and their nucleosides and nucleotides were studied in numerous literature. Synthetic analogs of naturally occurring nucleotides find application in cancer chemotherapy as enzyme inhibitors and can replace the naturally occurring nucleotides in nucleic acids. Therapeutic attempts to inhibit the growth of cancer cells or certain viruses have often employed administration of analogs of bases, nucleosides, or nucleotides that inhibit the synthesis of either DNA or RNA. Allopurinol, a purine analog, is widely used in the treatment of gout.

Recombinant DNA technology, often referred to as genetic engineering, has revolutionized biology and is having an ever-increasing impact on clinical medicine. Much has been learned about human genetic disease from pedigree analysis and study of affected proteins, but in many cases where the specific genetic defect is unknown, these approaches can not be used. The approach circumvents these limitations by going directly to the DNA molecule for information.

Understanding recombinant DNA approach is important for several reasons. The information explosing occurring in this area is truly staggering. To understand and keep up with this field, one must have an appreciation of the fundamental concepts involved. There is now rational approach to understanding the molecular bases of a number of diseasese e.g. familial hypercholesterolemia, sickle cell disease, the thalassemias, cystic fibrosis, Huntington’s chorea. Using recominant DNA technology, human proteins can be produced in abundance for therapy (e.g. insulin, growth hormone, plasminogen activator). Proteins for

vaccines, e.g. hepatitis B, and for diagnostic tests, e. g. AIDS test, can be obtained. Recombinant DNA technology is used to diagnose existing diseases and predict the risk of developing a given disease. Gene therapy for sickle cell disease, the thalassemias, adenosine deaminase deficiency, and other diseases may be devised. Gene therapy has already been accomplished in mice, in whom hereditary hypogonadism has been corrected by transgenic injection of the gonadotropin releasing hormone gene into the fertilized ovum.

The three major pyrimidine bases present in the nucleotides of both prokaryotes and eukaryotes are cytosine -C, thymine-T, and uricil-U. The purine bases adenine-A and guanine-G are the two major purines found in living organisms. Two other purine bases, hypoxanthine and xanthine, occur as intermediates in the metabolism of adenine and guanine. In humans, a completely oxidizes purine base uric acid, is formed as the end product of purine catabolism [2]-[4]. In previous papers [7],[8] ability to use genetic information system have illustrated.

In this paper a complex structural bioinformations have modelling.

2. Nucleobases transfer

Synthetic analogs of nucleobases, nucleosides, and nucleotides are widely used in the medical sciences and clinical medicine. In the past, most of these uses have depended upon the role of nucleotides as components of nucleic acids for cellular growth and division. For a cell to divide, its nucleic acids must be replicated. This requires that the precursors of nucleic acids - the normal purine and pyrimidine deoxy- be readily available.

One of the most important components of the oncologict's pharmacopeia is the group of synthetic analogs of purine and pyrimidine nucleotides and nucleosides.

The pharmacologic approach has been to use an analog in which either the heterocyclic ring structure of the sugar moiety has been altered in such a way as to induce toxic effects when the analog becomes incorporated into various cellular constituents. Many of these effects results from inhibition by the drug of specific enzyme activities necessary for nucleic acids synthesis or from the incorporation of metabolites of the drug into the nucleic acids where they alter the base pairing essential to accurate transfer of information. Examples of these would be the 5-fluoro or 5-iododerivatives of uracil or deoxyuridine, which serve as thymine or thymidine analogs, respectively.

Both 6-thioguanine and 6-mercaptopurine, in which naturally occurring hydroxyl groups are replaced with thiol groups at the 6 position, are widely used clinically.

The analogs in which the purine or pyrimidine ring contains extra nitrogen atoms, such as 5- or 6- azauridine or azacytidine and 8-azaguanine, also have been tested clinically.

The purine analog 4- hydroxypyrazolo pyrimidine (allopurinol) is widely marketed as an inhibitor of de novo purine biosynthesis and of xanthine oxidase. It is used for the treatment of hyperuricemia and gout. Nucleosides containing arabinose rather than ribose as the sugar moieties, e.g. cytarabine (arabynosil cytosine-Ara-C), are used in chemotherapy of cancer and viral infections.

Azathiopirine, which is catabolized to 6-mercptopurine, is useful in organ transplantion as a supressor of events involved in immunologic rejection.

A series of nucleoside analogs with antivirial activities has been studied for several years, one 5-iododeoxyuridine, is effective in the local treatment of herpetic keratitis, an infection of the cornea by herpes virus.

Numerous analogs of purine and pyrimidine ribonucleotides have been synthesized so as to generate nonhydrolyzable di- or triphosphates for use in vitro. These analogs allow the investigator to determine wheither given biochemical effects of nucleoside di- or triphosphates require hydrolysis or wheither their effects are mediated by occupying specific nucleotide binding sites on enzymes or regulatory proteins.

The reduction of the ribonucleoside diphosphates to deoxyribonucleoside diphosphates is subject to complex regulation.

3. DNA organization

DNA is a complex biopolymer that is organized as a doublehelix . The fundamental organizational element is the sequence of purine (adenine or guanine and pyrimidine cytosine or thymine bases. These bases are attached to the C-1' position of the sugar deoxyribose, and the bases are linked together through joining of the sugar moieties at their 3'and 5'positions via a phosphordiester bond. The alternating deoxyribose phosphate groups form the backbone of the double helix [2],[3]. These 3'-5' linkages also define the orientation of a given strand

of the DNA molecule, and since the 2-strands run in opposite directions, they are said to be antiparralel.

Basic pairing is one of the most fundamental concepts of DNA structure and function. Adenine and thymine always pair, by hydrogen bonding, as do guanine and cytosine. These base pairs are said to be complementary, and the guanine content of a fragment of double stranded DNA will always equal its cytosine content, likewise the thymine and adenine contents are equal. Interaction can be reduced by heating the DNA to denaturate it. Base pairing and hydrophobic base stacking interactions hold the 2-DNA strands together. The laws of base pairing predict that 2- complementary DNA strands will reanneal exactly in register upon renaturation, as happen when the temperature of the solution is slowly reduced to normal. Indeed, the degree of base-pair matching or mismatching can be estimated from the temperature required for denaturation – renaturation. Segments of DNA with high degrees of base – pair matching require more energy input heat to accomplished denaturation, or, to put it another way, a closely matched segment will withstand more heat before the strand separate.

This reaction is used to determine whether there are significant differences between 2 DNA sequences, and it underlines the concept of hybridization, which is fundamental to the processes which be described hear.

There are about $3\ 10^9$ base pairs (bp) in each human haploid genome. If an average gene length is $3\ 10^3$ bp (three kilobases, kb), the genome could consist of 10^6 genes, assuming that there is no overlap and that transcription proceeds in only one direction.

The double – helical DNA is packaged into a more compact structure by a number of proteins, most notably the basic proteins called *histones*. The DNA presents within nucleus of a cell, if simply extended, would be about a meter long. The chromosomal proteins compact this long length of DNA so that it can be packaged into a nucleus with a volume of a few cubic microns.

Gene organization is very interesting. In general, prokaryotic genes consist of a small regulatory region (100-500 bp) and a large protein-coding segment (500-10000 bp). Several genes are often controlled by a single regulatory unit. Most mammalian genes are more complicated, in that the coding region are interrupted by non-coding regions that are eliminated when the primary RNA transcript is processed into mature messenger RNA (mRNA). The coding regions (those regions that appear in the mature RNA species) are called *exons*, and noncoding regions, which interpose or intervene between the exons, are called *introns*. Introns always removed from precursors RNA before transport into the cytoplasm occurs. The process by which introns are removed from precursor RNA and by which exons are ligated together is called RNA splicing. Incorrect processing of the primary transcript into the mature mRNA can results in disease in humans. Regulatory regions for specific eukaryotic genes are usually located in the DNA that flanks the transcription initiation site at its 5'end (5' flanking–sequence DNA). Occasionally, such sequences are found within the gene itself or in the region that flanks the 3' end of the gene. In cells of mammalian, each gene has its own regulatory region

4. Protein structure

The reduction of the ribonucleoside diphosphates to deoxyribonucleoside diphosphates is subject to complex regulation.

Long strands some amino acids such as glycine, valyne, leucyne, and iso-leucyne have freedom driving rotation. Glugacon consists of 29 amino acids residuals and as well as insulin which consists 51 amino acids counting in peptide hormone. All atoms of the main polypeptide strand between two neighbors α- carbon atoms are on the same surface. R –group represent branch strands i.e. that is are amino acids residuals. If these groups are equal that is homopolymers, and if these groups different then they copolymers, proteins. Proteins are biopolymers whose polymer strands consist from α-L-amino acids residuals connected interrelation by peptide branches. In this paper fundamental concept of energy and entropy of the helix-coil transition was studied.

Under random coils have a longer degree of conformational freedom than the ordered double-stranded structure, the ordered structure is disrupted with an increase in temperature. The double-to-single strand transition represents thermal denaturation or melting. Under ordinary conditions, native DNA in an aqueous solution takes a double-stranded structure, known as β-form. The double stranded is maintained by two main forces hydrogen bonds between complementary pairs on opposite strands, and

stacking interactions between neighboring base pairs. A model protein molecule has shown in Fig. 1.

Let it denoted with i number of amino acids in helix position, and with j number of group in which these residuals distributed, than probability conformation is $\sigma^j s^i$. Probability distribution function P in all combinations i and j is defined:

$$P = \sum_{j,i} \sigma^j s^i \qquad (1)$$

The mean value of amino acids in helix position will be equal:

$$< m >= \frac{\sum_{i,j} i\sigma^j s^i}{P} = \frac{d \ln P}{d \ln s} \qquad (2)$$

where σ -coil and s -helix parameters.
and helix contain θ is equal:

$$\theta = \frac{1}{n} \frac{d \ln P}{d \ln s} \qquad (3)$$

With large values n with maximum length of the chain value λ_0.

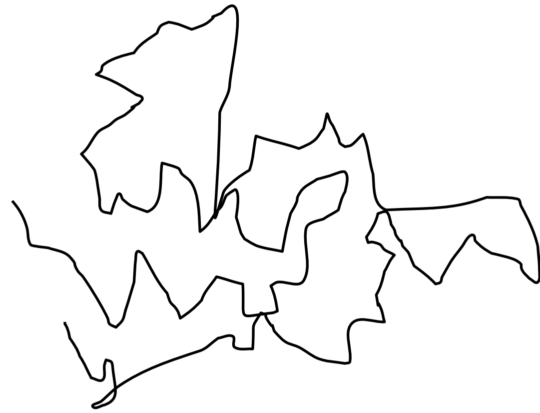

Fig. 1 Simple protein molecule model

$$\theta \approx \frac{d \ln \lambda_0}{d \ln s} \qquad (4)$$

where

$$\lambda_0 = \frac{-(1+s) + \sqrt{(1-s)^2 + 4\sigma s}}{2}, \qquad (5)$$

and taking into account that σ much less then 1.

$$\theta \approx \frac{1}{2} + \frac{s-1}{2} \sqrt{(1-s) + 4\sigma s}$$

Further is derived:

$$\frac{d\theta}{d \ln s} \approx \frac{d^2 \ln \lambda_0}{d(\ln s)^2}$$

And taking into account that $\sigma << 1$ and $s \approx 1$ it is obtained:

$$\frac{d\theta}{d \ln s} \approx \frac{1}{4\sqrt{\sigma}} \qquad (6)$$

It is obviously that transition velocity proportional $\sigma^{-1/2}$.

5. Parameters determination

The free energy can be related to the probability of the chain being in conformation, and the fraction of molecules that has accomplished the transition evaluated as:

$$1-\theta = \frac{1}{1+e^{-\Delta G/RT}} \qquad (7)$$

If enthalpy change which following helix growth is denoted ΔH_s, and entropy change ΔS_s, than enthalpy changes and entropy changes in the transition region can be expressed:

$$\ln s = -\frac{\Delta H_s}{RT} + \frac{\Delta S_s}{R} \qquad (8)$$

When helix-coil transition as result temperature changing can write

$$\frac{d\theta}{dT}\Big|_{T=T} = \frac{1}{4RT^2{}_c}\frac{\Delta H_s}{\sqrt{\sigma}} \qquad (9)$$

where is T_c crossing temperature. Changing temperature near crossing point and measuring helix content, can determine $d\theta/dT$ and calculates value $\Delta H_S/\sqrt{\sigma}$. By calorimetric methods can determine value ΔH_S independent of θ and these data can be used for calculating $\sqrt{\sigma}$. When length of strand is not enough θ can determine according equation (3).

Amino acids residual in helix state strands of several groups. Let ν describes number of groups in the strand from n acids residual can determine by equation:

$$\nu = \frac{1}{n}\frac{\partial \ln P}{\partial \sigma} \qquad (10)$$

If quantity λ_0 from equation (5) set in equation (10) at $s=1$ (crossing point), including that $\sigma << 1$, dependence ν from σ is:

$$\nu \approx \frac{\sqrt{\sigma}}{2}$$

The helix growth (s) and coil (σ) parameters can be determined by fitting temperature progress curve s(T) with experimental data (equation 8).

$$\ln s = a - \frac{b}{T} \qquad (9)$$

where parameters a and b were determined fitting with experimental data [9],[10]. Computational procedure for parameters a and b determination based on regression method was used [11],[12].

The obtained results for enthalpy and entropy changes ΔH_s and ΔS_s are shown in Table 1 and Table 2.

Table 1 shows parameters for poly-L-valyn, and Table 2 for iso-leucyn.

Table 1. The obtained parameters for poly-L-valyn in aqueous solution in temperature region 273,16-343,16K

Parameter	Entropy, ΔS_s,KJ/molK	Enthalpy, ΔH_s,KJ/mol	Confidential ε
a=-1.5807710	13,145691	-	0.400 10^{-3}
b=0.0132824	-	0.1104564	0.400 10^{-3}

Table 2. The determined parameters for poly-L-isoleucyn in aqueous solution in temperature region 273,16- 343,16K

Parameter	Entropy, ΔS_s,KJ/mol K	Enthalpy, ΔH_s,KJ/mol	Confidential, ε
a=2,333083	19,401918	-	0.614 10^{-2}
b=-0.0040275	-	0.0334926	0.614 10^{-2}

6. Conclusion

In this paper bioinformation processing and gene transfer were examined. Modelling method of bioinformation in biosynthesis was studied.

The transition parameters for poly-L-isoleucyn and poly-L-valyn in aqueous solution were determined by numerical method. Probability conformation distribution function was derived. Helix growth parameter was defined. Enthalpy changes and entropy changes in the transition region 273,16-343,16K were determined.

Transition velocity and maximum length of the chain were developed. These parameters contribute precision regulatory mechanism.

Notation

a-parameter

b-parameter
G – free energy change, kJ/mol
H-enthalpy, kJ/mol
S- entropy, kJ/mol K
s – helix growth
A-adenine
C-cytosine
G-guanine
T-thymine
U-uracil

Greek Symbols

Δ -change
ε -error
λ -chain length, nm
ν -number of groups in strand
σ **-**coil parameter
θ -helicity

Abbreviation
DNA-deoxyribonucleic acid
RNA-ribonucleic acid

References

[1]J.Savkovic-Stevanovic, Nucleotides recomination states, *Trasaction Biology and Biomedicine,***7** (3) July, 2010,253-262,2010,ISSN 1109-9518.
[2]J.F.Henderson, A.R.P. Paterson, *Nucleotide metabolism: An introduction,* Academic Press,1973.
[3] B. Lewin, *Genes II*, Wiley & Sons, N.J.1985.
[4] A. M. Michelson, *The chemistry of nucleosides and nucleotides,* Academic Press,1963.
[5]Prusoff, D.C.Ward, Nucleoside analogs with antivirial activity, *Biochem. Pharmacol.* **25**,1233,1976.
[6] B.N.Ames, et al., Uric acid provides an antioxidant defense in humans against oxidant- and radical –caused aging and cancer: A hypothesis. *Proc. Natl. Acad. Sci* ,**78,**6858, 1981.
[7] J.Savkovic-Stevanovic, Human tissue, *Comput.Ecol.Eng*.,**5**(1) 43-49, 2009, ISSN 1452-0729.
[8] J.Savkovic-Stevanovic, Human tissue information processing, *WSEAS 2nd International Conference Biomedical Electronics and Biomedical Informatics, ID 618-272*, Moscow, Russia, August 20-22, pp.64-69, 2009.
[9] J.E. Alter, R.H. Andreatta, G. T.Taylor, H.A. Scheraga, *Macromolecules,* **6**,564, 1973.
[10] R. A. Fredrickson, M. C. Chang, S. P. Powers S., H. A.Scheraga, *Macromolecules* **14**, 625,1981.
[11] Lj. Vilendecic-Grkinic, J. Savkovic - Stevanovic, Autocorrelation model of the poly-L-glycine, *IES 2006- The 17th Seminar and Symposium on Information and Expert Systems in the Process Industries* ,pp.101-109, Belgrade, 25-26, October, 2006.
[12]Savkovic-Stevanovic J., *Chemical Process Modeling and Simulation*, 201-243, Faculty of Technology and Metallurgy, The University of Belgrade, Belgrade, 1995.