DNA Topoisomerase Topology and Binding Atomic Entanglement for Fixed Integer Coefficients Polynomial Representation

Jeff Cromwell, PhD

Mathematical Learning Space Research Portfolio Spring Semester 2021 : Preprint 117

1 Abstract

Topoisomerase is found in the mitochondria of cells and generates ATP and has a role in programmed cell death and aging. Topoisomerases (or DNA topoisomerases) are isomerase enzymes that act on the topology of DNA such as over and under winding of DNA. The winding problem of DNA is because of the intertwined double-helical structure where with DNA replication and transcription, DNA is in a state of overwinding ahead of a replication fork with torsion and can prevent the ability of DNA or RNA polymerases to move down the DNA strand.

Separation of the phosphate backbone of either one or both the DNA strands permits the DNA to be (a) untangled or (b) unwound, and, at the end resealed again. Because the overall chemical composition and connectivity of the DNA is changeless, the DNA substrate and product are chemical isomers difference is in the global topology. Bacterial topoisomerases and human topoisomerases proceed via similar mechanisms for managing DNA supercoils. The double-helical configuration of DNA strands makes separation difficult for the helicase enzymes if other enzymes transcribe the sequences for protein encoding, or if chromosomes are to be replicated. In circular DNA, i.e. double-helical DNA is bent around and joined in a circle, the two strands are topologically linked, or knotted. Identical loops of DNA, i.e. different numbers of twists, are topoisomers, and cannot be interconverted without the breaking of DNA strands. Topoisomerases catalyze and guide the unknotting or unlinking of DNA by creating transient breaks in the DNA using a conserved tyrosine as the catalytic residue.

In the brief paper, the following sequences of Proteins 16k, Taxonomy 11k, Proteomes 3k and Structures 24, Pathways 7 and Genome3D 388. A binding mode for the anticancer drug camptothecin has been proposed on the basis of chemical and biochemical information combined with the three-dimensional structures of topoisomerase I-DNA complexes. Trefoil knots present in DNA binding and other architectures as well as DNA topoisomerases was examined with molecular properties for similarities. Human topoisomerase I is inhibited by camptothecin (CPT), a plant alkaloid with antitumour activity and was examined. K-31 secondary protein structures for 21 crystals were described and knots with 1CMX Knot, 4RLV Knot, 2K0A Knot, and 2K0A Knot presented.

*The Mathematical Learning Space Research Portfolio
Email address: http://mathlearningspace.weebly.com/ (Jeff Cromwell, PhD)

2 Introduction

DNA topoisomerases regulate the number of topological links between two DNA strands (i.e. change the number of superhelical turns) by catalysing (a) transient single- or (b) double-strand breaks, (c) crossing the strands through one another, then (d) resealing the breaks. These enzymes have several functions: (a) to remove DNA supercoils during transcription, replication (b) strand breakage during recombination for chromosome condensation and (c) in mitosis, disentangle intertwined DNA. Topoisomerases can relieve negative supercoils, i.e. positive and negative supercoils, or induce positive and negative supercoiling in DNA. The enzymes can promote catenation, i.e. two single circular DNA strands are linked together after replication and decatenation separation of two linked, closed, circular chromosomes, and can untangle the entanglement of linear chromosomes. [1] [401][2][3]

The double-helical configuration of DNA strands makes separation difficult for the helicase enzymes if other enzymes transcribe the sequences for protein encoding, or if chromosomes are to be replicated. In circular DNA, i.e. double-helical DNA is bent around and joined in a circle, the two strands are topologically linked, or knotted. Otherwise identical loops of DNA, i.e. different numbers of twists, are topoisomers, and cannot be interconverted without the breaking of DNA strands. Topoisomerases catalyze and guide the unknotting or unlinking of DNA by creating transient breaks in the DNA using a conserved tyrosine as the catalytic residue. [1] [401] [2][3]

DNA topoisomerases are divided into two classes.

1. type I enzymes (5.99.1.2; topoisomerases I, III and V) break single-strand DNA
2. type II enzymes (5.99.1.3; topoisomerases II, IV and VI) break double-strand DNA

Type I topoisomerases are ATP-independent enzymes (except for reverse gyrase)subdivided by structure and reaction mechanisms: (a) type IA (Topo IA; bacterial and archaeal topoisomerase I, topoisomerase III and reverse gyrase) and (b) type IB (Topo IB; eukaryotic topoisomerase I and topoisomerase V). These enzymes provide relaxing positively and/or negatively supercoiled DNA, except for reverse gyrase into positive supercoils into DNA. This function is vital for the processes of (a) replication, (b) transcription, and (c) recombination. Unlike Topo IA enzymes, Topo IB enzymes do not require a single-stranded region of DNA or metal ions for their function. The type IB family of DNA topoisomerases includes (1) eukaryotic nuclear topoisomerase I, (2) topoisomerases of poxviruses, and (3) bacterial versions of Topo IB. [1] and are in the superfamily of DNA breaking-rejoining enzymes that have the same fold in their C-terminal catalytic domain and the overall reaction mechanism with tyrosine recombinases. The C-terminal catalytic domain in topoisomerases is linked to a divergent N-terminal domain and provides no sequence or structure similarity to the N-terminal domains of tyrosine recombinases. This family of DNA topoisomerase I enzymes has both type IA enzymes from bacteria and type IB enzymes from eukaryotes and viruses. [1] [401] [2][3]

Type II family passes a region of duplex (two strands) from the same molecule or a different molecule by way of a double stranded gap. Type II cleaves both strands of DNA from the double-stranded break. [1] [401] [2][3]

2.1. Camptothecin (CPT)

Human topoisomerase I is inhibited by camptothecin (CPT), a plant alkaloid with antitumour activity. The crystal structures of human topoisomerase I comprising the core and carboxyl-terminal domains in covalent and noncovalent complexes with 22-base pair DNA duplexes reveal an enzyme that "clamps" around essentially B-form DNA. The core domain and the (a) first eight residues of the carboxyl-terminal domain of the enzyme, (b) active-site nucleophile tyrosine-723, (c) share significant structural similarity with the bacteriophage family of DNA integrases. A binding mode for the anticancer drug camptothecin has the basis of chemical and biochemical information combined with the three-dimensional structures of topoisomerase I-DNA complexes. [1] [401] [470]

Topoisomerase is found in the mitochondria of cells. The mitochondria generate ATP as well as playing a role in programmed cell death and aging. The mitochondrial DNA of animal cells is a circular, double-stranded DNA that requires the activity of topoisomerase to be replicated. The classes of topoisomerase found in the mitochondria are I, IIβ, IIIα. [1]

3 Mathematical Background: DNA Topology

DNA topology is the tertiary conformations of DNA, such as (a) supercoiling, (b) knotting, and (c) catenation. Topology of DNA can be disrupted by most metabolic processes: (a) RNA polymerase can cause positive supercoils by over-winding the DNA in front of the enzyme, and can also cause (b) negative supercoils by under-winding the DNA behind the enzyme. DNA polymerase has the same effect in DNA replication with positive and negative supercoiling balance out the entire global topology of the DNA where the overall topology remains the same. However, as the DNA replication or transcription fork moves (1) forward and (2) positive supercoiling increases, the (3) DNA strands wrap tighter and tighter around each other which (4) makes more difficult for the polymerase to move forward. It is important for the local topology of DNA ahead of and behind the polymerase to be relieved so that replication and cell division can proceed. This is what DNA topoisomerases are used for. [1] There are three main types of topology: (a) Supercoiling (b) Knotting and (c) Catenation. [2][3]

Since the processes of replication or transcription is essential, DNA must be compact with these three types of topologies. Transcription or replication DNA requires freedom from these three states. In replication, the newly replicated duplex of DNA and the original duplex of DNA become intertwined and needs complete separation for genomic integrity of cell division. As a transcription bubble proceeds, (a) DNA ahead of the transcription fork becomes overwound, or positively supercoiled, while (b) DNA behind the transcription bubble becomes underwound, or negatively supercoiled. As replication occurs, (a) DNA ahead of the replication bubble becomes positively supercoiled, while (b) DNA behind the replication fork becomes entangled forming precatenanes. Topologically linked circular molecules(i.e. catenanes) have a positive supercoiled form during the process of replication of circular plasmids. The (1) unlinking of catenanes is performed by type IIA topoisomerases is more efficient unlinking positive supercoiled DNA. The conformational properties of negative vs. positive supercoiled catenanes affects their (2) features in respect to their corresponding (3) enzymatic reaction catalyzed by topoisomerases with (4) positive supercoiled DNA provides a (5) sharp DNA bend in the first bound DNA segment permits topoisomerase to (6) bind successfully for an enzymatic reaction to the following segment in a specific inside-to-outside matter. Negative supercoiled DNA does not provide such a bend and the access of the enzyme to the first segment is nearly impossible and inhibits unlinking. At the end of replication, when daughter chromosomes must be fully disentangled before mitosis occurs the topological problem is addressed by Topoisomerase IIA. [1] [2][3]

4 Results

Table 1 has the molecular properties for a sample of 30 of 100 from the 16K Proteins of DNA topoisomerases. [1000]

	Stability Index	Binding Potential	ALiphatic	f.1	CpH5	CpH7	CpH9
48	1.99	0.13	4.63	-0.03	23.14	12.11	6.80
27	2.45	0.14	4.85	-0.02	33.96	26.44	23.59
18	0.94	0.09	2.52	-0.03	17.81	11.90	8.04
46	2.37	0.14	4.76	-0.03	27.01	17.63	12.75
59	1.68	0.12	4.99	-0.02	28.40	17.32	10.95
37	3.79	0.21	5.84	-0.08	36.11	17.16	4.81
67	9.57	0.60	9.48	-0.23	59.51	16.79	-2.47
28	2.57	0.15	3.58	-0.04	21.92	9.59	3.88
100	1.61	0.12	2.87	-0.05	18.74	7.02	-0.98
9	2.12	0.13	4.43	-0.03	20.40	9.12	3.56
52	5.51	0.36	11.30	-0.13	64.26	41.26	23.02
50	3.45	0.25	8.96	-0.08	49.84	30.17	17.82
71	15.18	0.41	18.70	-0.08	40.27	16.56	0.99
97	1.66	0.13	2.78	-0.05	18.74	7.02	-0.97
55	5.51	0.36	11.30	-0.13	64.26	41.26	23.02
82	1.79	0.13	2.73	-0.05	19.81	8.03	0.03
72	6.04	0.38	10.46	-0.13	71.33	48.72	36.51
95	1.68	0.12	2.93	-0.05	19.81	8.78	1.00
8	7.49	0.53	10.11	-0.21	77.82	48.99	35.16
44	2.38	0.15	4.08	-0.04	32.41	24.41	20.49
76	1.11	0.11	2.89	-0.04	15.59	8.71	5.36
34	2.18	0.15	4.54	-0.05	33.81	21.26	12.61
10	1.46	0.10	3.51	-0.03	13.93	5.27	-2.11
62	1.63	0.13	4.44	-0.03	29.56	16.77	9.03
61	1.36	0.12	4.23	-0.03	34.96	23.49	15.28
66	3.58	0.27	7.53	-0.09	56.87	35.59	21.12
7	5.14	0.39	8.15	-0.16	58.92	32.96	19.68
58	1.93	0.12	4.80	-0.03	24.58	14.05	7.11
73	6.03	0.37	10.66	-0.13	71.33	48.72	36.51
86	1.75	0.12	2.78	-0.05	20.63	8.99	0.26

Table 1: For the Aliphatic index with the four aminos of Alanine, Valine, Isoleucine, and Leucine for the side chain as a positive measure for the increase of thermostability of globular proteins. The Binding Potential of the sequence is the potential of a peptide to bind to membranes or other proteins as receptors with a high binding potential if greater than 2.48. The stability of a protein is based on a value less than 40 with a value above 40 unstable. F.1 is a measure of hydrophobicity as stabilization force solvent dependent in protein folding. The charge at three different pH values of 5, 7 and 9 is the s the net charge of a protein sequence based on the Henderson-Hasselbalch equation defined at pH using one of the 9 different pKa scales. Here the scale chosen is Lehninger. [1001]

Table 2 has the atomic spatial structures for DNA TOPOISOMERASE.[500]

	Accession	Name
1	1a31	HUMAN RECONSTITUTED DNA TOPOISOMERASE I IN COVALENT COMPLEX WITH A 22 BASE PAIR DNA DUPLEX
2	1a35	HUMAN TOPOISOMERASE I/DNA COMPLEX
3	1a36	TOPOISOMERASE I/DNA COMPLEX
4	1a41	TYPE 1-TOPOISOMERASE CATALYTIC FRAGMENT FROM VACCINIA VIRUS
5	1ej9	CRYSTAL STRUCTURE OF HUMAN TOPOISOMERASE I DNA COMPLEX
6	1k4s	HUMAN DNA TOPOISOMERASE I IN COVALENT COMPLEX WITH A 22 BASE PAIR DNA DUPLEX
7	1k4t	HUMAN DNA TOPOISOMERASE I (70 KDA) IN COMPLEX WITH THE POISON TOPOTECAN AND COVALENT COMPLEX WITH A 22 BASE PAIR DNA DUPLEX
8	1lpq	Human DNA Topoisomerase I (70 Kda) In Non-Covalent Complex With A 22 Base Pair DNA Duplex Containing an 8-oxoG Lesion
9	1nh3	Human Topoisomerase I Ara-C Complex
10	1ois	YEAST DNA TOPOISOMERASE I, N-TERMINAL FRAGMENT
11	1r49	Human topoisomerase I (Topo70) double mutant K532R/Y723F
12	1rr8	Structural Mechanisms of Camptothecin Resistance by Mutations in Human Topoisomerase I
13	1rrj	Structural Mechanisms of Camptothecin Resistance by Mutations in Human Topoisomerase I
14	1sc7	Human DNA Topoisomerase I (70 Kda) In Complex With The Indenoisoquinoline MJ-II-38 and Covalent Complex With A 22 Base Pair DNA Duplex
15	1seu	Human DNA Topoisomerase I (70 Kda) In Complex With The Indolocarbazole SA315F and Covalent Complex With A 22 Base Pair DNA Duplex
16	1t8i	Human DNA Topoisomerase I (70 Kda) In Complex With The Poison Camptothecin and Covalent Complex With A 22 Base Pair DNA Duplex
17	1tl8	Human DNA topoisomerase I (70 kDa) in complex with the indenoisoquinoline AI-III-52 and covalent complex with a 22 base pair DNA duplex
18	1vcc	AMINO TERMINAL 9KDA DOMAIN OF VACCINIA VIRUS DNA TOPOISOMERASE I RESIDUES 1-77, EXPERIMENTAL ELECTRON DENSITY FOR RESIDUES 1-77
19	2b9s	Crystal Structure of heterodimeric L. donovani topoisomerase I-vanadate-DNA complex
20	2f4q	Crystal Structure of Deinococcus radiodurans topoisomerase IB
21	2h7f	Structure of variola topoisomerase covalently bound to DNA
22	2h7g	Structure of variola topoisomerase non-covalently bound to DNA
23	3igc	Smallpox virus topoisomerase-DNA transition state
24	3m4a	Crystal structure of a bacterial topoisomerase IB in complex with DNA reveals a secondary DNA binding site

Table 3 has the Crystal Stability Index Binding Potential ALiphatic f.1 CpH5 CpH7 and CpH9. [1001]

	Crystal	Stability Index	Binding Potential	ALiphatic	f.1	CpH5	CpH7	CpH9
1	1a31	41.82	2.30	71.97	-0.80	30.52	16.98	6.76
2	1a35	40.72	2.25	73.46	-0.76	32.49	18.19	7.07
3	1a36	40.62	2.37	73.16	-0.81	38.82	23.20	11.82
4	1a41	37.49	1.62	93.80	-0.22	19.15	13.13	8.77
5	1ej9	41.78	2.25	71.88	-0.79	29.94	15.19	3.99
6	1k4s	43.26	2.33	70.86	-0.84	31.16	16.19	4.88
7	1k4t	41.70	2.41	70.62	-0.87	40.27	24.20	12.63
8	1lpq	41.32	2.39	71.45	-0.85	38.27	22.20	10.72
9	1nh3	42.68	2.33	69.98	-0.84	28.97	14.98	4.67
10	1ois	47.91	2.07	71.77	-0.74	11.93	4.38	-0.02
11	1r49	42.59	2.43	71.02	-0.86	37.16	21.20	9.77
12	1rr8	44.01	2.32	70.89	-0.83	30.16	15.19	3.90
13	1rrj	41.70	2.40	70.62	-0.86	40.27	24.20	12.63
14	1sc7	42.50	2.42	70.37	-0.88	38.50	22.20	10.63
15	1seu	41.70	2.40	70.62	-0.86	40.27	24.20	12.63
16	1t8i	41.70	2.41	70.62	-0.87	40.27	24.20	12.63
17	1tl8	41.70	2.41	70.62	-0.87	40.27	24.20	12.63
18	1vcc	40.41	1.82	79.61	-0.44	2.83	-0.30	-2.15
19	2b9s	38.60	2.20	73.45	-0.61	23.24	6.97	-2.96
20	2f4q	35.29	2.19	82.75	-0.46	20.34	12.38	7.73
21	2h7f	36.33	1.74	88.03	-0.33	28.72	19.91	15.73
22	2h7g	36.50	1.73	88.59	-0.32	29.65	20.91	16.73
23	3igc	36.41	1.75	88.31	-0.33	28.72	19.91	15.73
24	3m4a	34.80	2.21	84.11	-0.45	23.05	14.62	9.73

Table 2: For the Aliphatic index with the four aminos of Alanine, Valine, Isoleucine, and Leucine for the side chain as a positive measure for the increase of thermostability of globular proteins. The Binding Potential of the sequence is the potential of a peptide to bind to membranes or other proteins as receptors with a high binding potential if greater than 2.48. The stability of a protein is based on a value less than 40 with a value above 40 unstable. F.1 is a measure of hydrophobicity as stabilization force solvent dependent in protein folding. The charge at three different pH values of 5, 7 and 9 is the s the net charge of a protein sequence based on the Henderson-Hasselbalch equation defined at pH using one of the 9 different pKa scales. Here the scale chosen is Lehninger. [1003]

5 Knot Analysis

Knots are the basic objects entangled in closed or open chains. Several types of knots have been found in proteins: (a) trefoil, (b) figure-8, (c) 52, and (d) Stevedore's knot. An unknotted loop is called the trivial knot, or the unknot, and is denoted 01. The first number denotes the minimal number of crossings a given knot in a projection (i.e. minimal number of crossings in a projection of a trefoil onto a plane is 3). The 31, 52 and 61 knots are chiral, i.e. they differ from their mirror images, and their complete characterization requires the determination of their chirality, plus (+) or a minus (-) sign next to the symbol of a knot. The 41 knot is an example of an achiral knot, i.e. it is identical to its mirror image and cannot be assigned a chirality. Knots are defined uniquely on both open and closed chains such as proteins, one must choose how to connect the two loose ends, so that a closed chain is formed. The knot type detection is accomplished with a polynomial knot invariant known as Alexander polynomial obtained by a planar diagram of a knot with the projection of a knot on some two-dimensional plane. Alexander polynomial is different for all prime knots with eight or fewer crossings and sufficient to detect knots to appear in proteins with the most complicated knot found in proteins with six crossings. [1][100] [101] [102] [103]

Figure 1 has the knot polynomials. [1007]

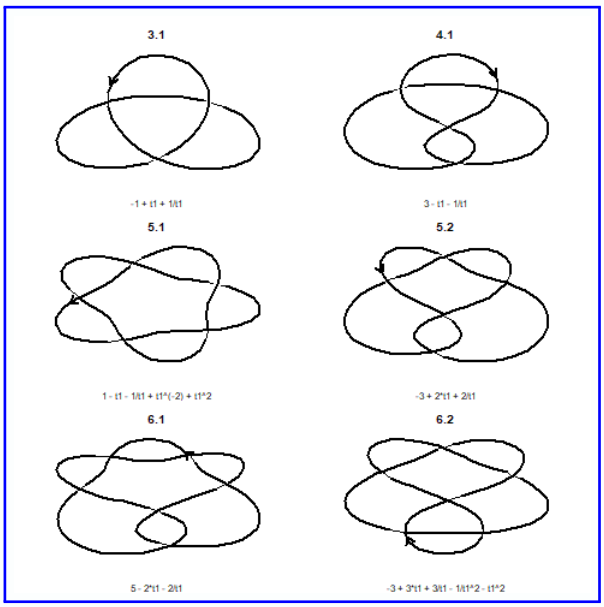

Figure 1: Knot Polynomials

for DNA binding, Table 4 has the crystal structure definitions of 5wtuA, 2rh3A , and 4lrvA. [1000]

K -31	5wtuA	Crystal structure of dnde g21/24k mutant involved in dna phosphorothioation
K -31	2rh3A	Crystal structure of plasmid ptic58 virc2
K -31	4lrvA	Crystal structure of dnde from escherichia coli b7a involved in dna phosphorothioation modification

for the 4LRV A chain the polynomial is [1007]

$$2 * l^2 + l^2 * m^2 - l^4 \quad (1)$$

$$-1 + x + 1/x \quad (2)$$

$$1/t - 1/t^4 + t^(-3) \quad (3)$$

where (a) Homfly, (b) Alexander and (c) Jones polynomial. Figure 2 has 5WTU Knot, 2RH3 Knot, 4LRV Knot and 4LRV Knot.[100] [101] [102] [103][1007]

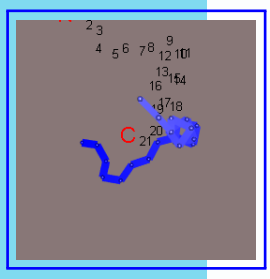

Figure 2: 5WTU Knot

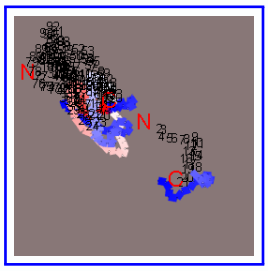

Figure 3: 2RH3 Knot

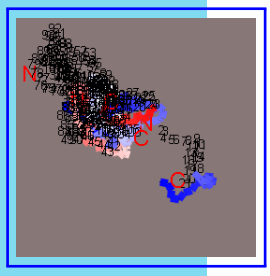

Figure 4: 4LRV Knot

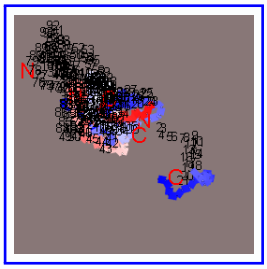

Figure 5: 4LRV Knot

For the trefoil knot sequences of DNA binding of Table 5, the same molecular properties are given by [1007]

	Stability Index	Binding Potential	ALiphatic	f.1	CpH5	CpH7	CpH9
1	42.51	1.43	104.72	0.00	24.27	16.21	12.98
2	50.81	1.30	88.10	0.00	7.91	4.67	2.22
3	41.67	1.55	99.82	0.00	84.02	51.46	40.56

Table 3: For the Aliphatic index with the four aminos of Alanine, Valine, Isoleucine, and Leucine for the side chain as a positive measure for the increase of thermostability of globular proteins. The Binding Potential of the sequence is the potential of a peptide to bind to membranes or other proteins as receptors with a high binding potential if greater than 2.48. The stability of a protein is based on a value less than 40 with a value above 40 unstable. F.1 is a measure of hydrophobicity as stabilization force solvent dependent in protein folding. The charge at three different pH values of 5, 7 and 9 is the s the net charge of a protein sequence based on the Henderson-Hasselbalch equation defined at pH using one of the 9 different pKa scales. Here the scale chosen is Lehninger. [1003]

Table 6 has the K-31 secondary protein structures for crystals.[100] [101] [102] [103]

K -31	6nd4N	Conformational switches control early maturation of the eukaryotic small ribosomal subunit
K -31	6qx9BP	Structure of a human fully-assembled precatalytic spliceosome (pre-b complex).
K -31	2efvA	Crystal structure of a hypothetical protein(mj0366) from methanocaldococcus jannaschii
K -31	5z576	Cryo-em structure of the human activated spliceosome (late bact) at 6.5 angstrom
K -31	5z566	Cryo-em structure of a human activated spliceosome (mature bact) at 5.1 angstrom.
K -31	5nrlS	Structure of a pre-catalytic spliceosome
K -31	5o9zy	Cryo-em structure of a pre-catalytic human spliceosome primed for activation (b complex)
K -31	5zwm5	Cryo-em structure of the yeast pre-b complex at an average resolution of 3.4 4.6 angstrom (tri-snrnp and u2 snrnp part)
K -31	5zwo5	Cryo-em structure of the yeast b complex at average resolution of 3.9 angstrom
K -31	2ofzA	Ultrahigh resolution crystal structure of rna binding domain of sars nucleopcapsid (n protein) at 1.1 angstrom resolution in monoclinic form.
K -31	5zd5D	Sf3b spliceosomal complex bound to e7107
K -31	5wtuA	Crystal structure of dnde g21/24k mutant involved in dna phosphorothioation
K -31	5ooms	Structure of a native assembly intermediate of the human mitochondrial ribosome with unfolded interfacial rrna
K -31	5lqwY	Yeast activated spliceosome
K -31	5ifeD	Crystal structure of the human sf3b core complex
K -31	5gm6J	Cryo-em structure of the activated spliceosome (bact complex) at 3.5 angstrom resolution
K -31	5sybA	Crystal structure of human phf5a
K -31	1cmxA	Structural basis for the specificity of ubiquitin c-terminal hydrolases
K -31	4lrvA	Crystal structure of dnde from escherichia coli b7a involved in dna phosphorothioation modification
K -31	2k0aA	1h, 15n and 13c chemical shift assignments for rds3 protein
K -31	3mlgA	2ouf-2x, a designed knotted protein
K -31	3mt5A	Crystal structure of the human bk gating apparatus
K -31	3o0pA	Pilus-related sortase c of group b streptococcus
K -31	2rh3A	Crystal structure of plasmid ptic58 virc2

Figure 3 has the following knots with 1CMX Knot, 4RLV Knot, 2K0A Knot, and 2K0A Knot. [1007][100] [101] [102] [103]

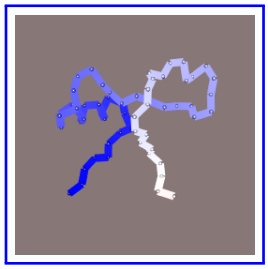

Figure 6: 1CMX Knot

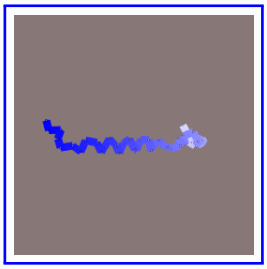

Figure 7: 4RLV Knot

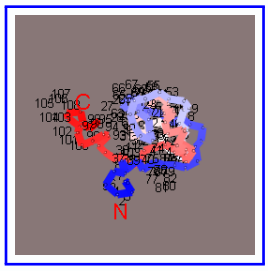

Figure 8: 2K0A Knot

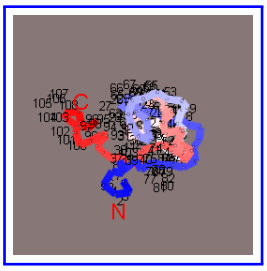

Figure 9: 2K0A Knot

6 Conclusion

The overall function of DNA topoisomerase is to manage the topological state of the DNA in the cell. In this paper, trefoil knots present in DNA binding and other architectures as well as DNA topoisomerases was examined with molecular properties for similarities. Human topoisomerase I has been shown to be inhibited by camptothecin (CPT), a plant alkaloid with antitumour activity and was examined. K-31 secondary protein structures for 21 crystals were described and knots with 1CMX Knot, 4RLV Knot, 2K0A Knot, and 2K0A Knot presented.

7 References

[1] Wikipedia contributors. "Topoisomerase." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 29 Mar. 2021. Web. 23 Jun. 2021.

[2] Wikipedia contributors. "Type I topoisomerase." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 8 Jun. 2021. Web. 23 Jun. 2021.

[3] Wikipedia contributors. "Type II topoisomerase." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 29 Mar. 2021. Web. 23 Jun. 2021.

[4] Wikipedia contributors. "Topoisomerase IV." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 29 Mar. 2021. Web. 23 Jun. 2021.

[5] Wikipedia contributors. "Topoisomerase inhibitor." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 25 Apr. 2021. Web. 23 Jun. 2021.

[10] Wikipedia contributors. "Knot theory." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 3 Feb. 2021. Web. 17 Jun. 2021.

[11] Wikipedia contributors. "Alexander polynomial." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 6 May. 2021. Web. 17 Jun. 2021.

[20] PUB00081702

[21] PUB00005230

[22] PUB00081703

[23] PUB00081704

[100] Dabrowski-Tumanski P, Rubach P, Goundaroulis D, Dorier J, Sułkowski P, Millett KC, Rawdon EJ, Stasiak A, Sulkowska JI, (2018) KnotProt 2.0: a database of proteins with knots and other entangled structures, Nucleic Acids Research, gky1140, https://doi.org/10.1093/nar/gky1140

[101] Jamroz M, Niemyska W, Rawdon EJ, Stasiak A, Millett KC, Sulkowska JI (2014) KnotProt: a database of proteins with knots and slipknots, Nucleic Acids Research, 43, D306-D314DOI: 10.1093/nar/gku1059

[102] Sulkowska JI, Rawdon EJ, Millett KC, Onuchic JN and Stasiak A (2012) Conservation of complex knotting and slipknotting patterns in proteins, Proc. Natl. Acad. Sci. U.S.A. 109, E1715–E1723, DOI: 10.1073/pnas.1205918109

[103] Dorier J, Goundaroulis D, Benedetti F, Stasiak A (2018) Knoto-ID: a tool to study the entanglement of open protein chains using the concept of knotoids, Bioinformatics, bty365, DOI: 10.1093/bioinformatics/bty365.

[400] Kanehisa, Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K.; KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 45, D353-D361 (2017).

[401] Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M.; KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 44, D457-D462 (2016).

[402] Kanehisa, M. and Goto, S.; KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27-30 (2000).

[420] GeneCards Version 3: the human gene integrator

[430] DrugBank 5.0: a major update to the DrugBank database for 2018.

[440] COSMIC: the Catalogue Of Somatic Mutations In Cancer.

[450] Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders

[460] The ClinicalTrials.gov Results Database — Update and Key Issues

[470] PubChem Substance and Compound databases

[480] The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible

[490] MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search.

[500] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne.(2000) The Protein Data Bank Nucleic Acids Research, 28: 235-242.

[601] Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-D613. doi:10.1093/nar/gky1131

[602] Ashburner et al. Gene ontology: tool for the unification of biology. Nat Genet. May 2000;25(1):25-9.

[603] The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. Jan 2019;47(D1):D330-D338

[701] InterPro Consortium Member databases.

[1000] R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[1001] Fang H. dcGOR: an R package for analysing ontologies and protein domain annotations. PLoS Computational Biology 10(10):e1003929, 2014. http://dx.doi.org/10.1371/journal.pcbi.1003929

[1002] Grant, B.J. et al. (2006) Bioinformatics 22, 2695–2696.

[1003] Osorio, D., Rondon-Villarreal, P. and Torres, R. Peptides: A package for data mining of antimicrobial peptides. The R Journal. 7(1), 4-14 (2015).

[1004] Martin Becker and Stefan Klößner (2017). PearsonDS: Pearson Distribution System. R package version 1.1. https://CRAN.R-project.org/package=PearsonDS

[1005] Juliane Schafer, Rainer Opgen-Rhein, Verena Zuber, Miika Ahdesmaki, A. Pedro Duarte Silva and Korbinian Strimmer. (2017). corpcor: Efficient Estimation of Covariance and (Partial) Correlation. R package version 1.6.9. https://CRAN.R-project.org/package=corpcor

[1006] Luna A et al. rcellminer: Exploring Molecular Profiles and Drug Response of the NCI-60 Cell Lines in R. Bioinformatics. 2015 Dec. http://www.ncbi.nlm.nih.gov/pubmed/26635141

[1007] Federico Comoglio and Maurizio Rinaldi (2016). Rknots: Topological Analysis of Knotted Proteins, Biopolymers and 3D Structures. R package version 1.3.2. https://CRAN.R-project.org/package=Rknots