Knot theory in understanding Proteins

Shantha Bhushan

Discussion meeting on Knot Theory -IISER Mohali Dec 2013

Protein molecule

- Protein is a polymer formed by a chain of amino acids.
- Proteins perform various functions- enzyme, transport, scaffolding give examples and pictures
- There are about 82,300 structures in the PDB (Protein Data Base)

Structure of protein

Protein backbone

Backbone is the chain of carbon atoms (C-alpha) starting at Carboxyl end (C- terminus) and ending at the amide end (N terminus). This is a piecewise linear open chain.

Terminology

- Folding , misfolding, unfolding
- Native state, denatured state
- Conformation

Main Areas of Research in Proteins

- How does folding take place-pathways
- Predict structure
- Classification of protein structures into folds
- Relate function to structure

Topological Barriers

- Final structure must determine the folding pathway
- Threading of protein chain through a region that is entirely surrounded by other parts
- Topology of backbone
- For small proteins fast folding rate correlates with low contact order (which is parameter describing the average separation along the protein chain between those parts that are in contact in the native folded configuration)

Why is topology of backbone interesting

- Insight into protein folding phenomena
- Relate non trivial topology with structure and function
- Creation of new molecules with novel properties

Possible role of the backbone

Suggested that knots (and slipknots) could play a role in

- against degradation by the proteasome
- provide structural stability in transporter proteins
- enhance thermal and mechanical stability or
- alter enzymatic activity

Role of knots

- Soler and Faisca (2013) showed that knots do give kinetic stability as the unfolding rate of knotted protein is much slower.
- Experiments show that denatured protein has knotted backbone
- Simulation shows that probable topology of backbone in denatured state is knotted

- Effects of Knots on Protein Folding Properties. Miguel A. Soler, Patrícia F. N. Faísca . Published: September 04, 2013.DOI: 10.1371/journal.pone.0074756

Detecting knots

Taylor method keep end points fixed.

Startomg with this PL curve, simplify it – Take three consecutive points
If this triangle does not intersect the curve, eliminate the middle vertex .

After no more smoothening can take place, join the two ends to obtain a PL knot

Millett Dobay Stasiak (MDS) method

Some knotted backbone

6_1 Knot

Miscellaneous:					
Ubiquitin Hydrolase UCH-L3	H.sapiens	1xd3	229	52	13-212 (12) 4
	S.cerevisiae (synth.)	1cmx	214	31	14-228 (6) 4,1
Ubiquitin Hydrolase UCH-L1	H.sapiens	2etl	219	52	10-216 (13)
S-adenosylmethionine synthetase	E.coli	1fug	383	31	33-260 (32)
	rattus norv.	1qm4	368	31	46-281 (29) 1
	H. sapiens	2p02	380	31	59-302 (21)
Class II ketol-acid reductoisomerase	Spinacia oleracea	1yve	513	41	321-533 (62)
	E.coli	1yrl	487	41	222-437 (52) 2
Transcarbamylase	B.fragilis	1js1	324	31	169-267 (57)
	X.campestris	1yh0	328	31	173-277 (57) 2
Methyltransferase	H.sapiens	2ha8	159	31	103-148 (30)
	P. gingivalis	2i6d	231	31	177-223 (9)

Protein	Species	PDB code	Length	Knot	Knotted core
YbeA-like	E.coli	1ns5	153	31	69-121 (32)
	T.maritima	1o6d	147	31	68-117 (30)
	S.aureus	1vh0	157	31	73-126 (31)
	B.subtilis	1to0	148	31	73-125 (32)
tRNA(m1G37)-methyltransferase TrmD	H.influenza	1uaj	241	31	85-130 (92)
	E.coli	1p9p	235	31	90-130 (89)
	S.cerevisiae	2v3k	219	31	175-225 (27)
SpoU-like RNA 2'-O ribose mtf.	T.thermophilus	1v2x	191	31	96-140 (51)
	H.influenza	1j85	156	31	77-114 (42)
	T.thermophilus	1ipa	258	31	190-234 (29)
	E.coli	1gz0	242	31	173-215 (28)
	A. aeolicus	1zjr	197	31	100-144 (58)
	S. viridochromog.	1x7p	267	31	209-251 (31)
YggJ C-terminal domain-like	H.influenza	1nxz	246	31	166-217 (30)
	B.subtilis	1vhk	235	31	168-226 (27) 1
	T.thermophilus	1v6z	227	31	104-203 (25) 3
Hypothetical protein MTH1 (MT0001)	A.M. thermoautotr.	1k3r	262	31	48-234 (28)

Plausible functions

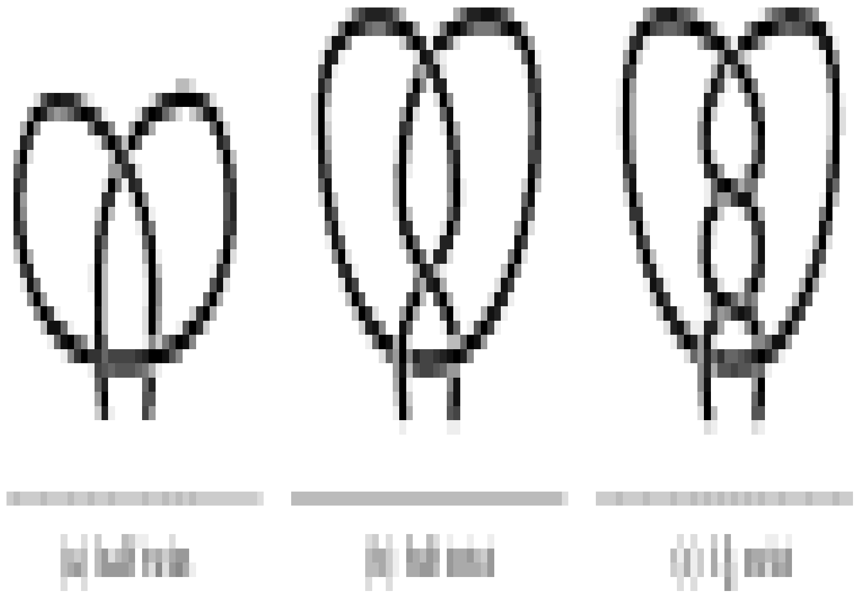
- Deep trefoil knot in chromophore binding region of Deinococcus radiourans
- Deep knot of methyltrasnferase in catalytic domain
- Deep 4_1 knot in acetohydroxy reductoisomerase
- 5_2 knot in UCHL 3 in humans and yeast possible function against degradation

Shallow Knots

Does it matter if the knot is deep?

Taylors hypothesis

All protein knots must be twist knots. The backbone forms a twist and the termini thread through the loop. (add pics)



Fingerprint of 5_2

Knot

Knotting fingerprint

Cysteine

- . Cysteine is the sole amino acid whose side chain can form covalent bonds, yielding disulfide bridges with other cysteine side chains: --CH_2-S-S-CH_2-- .

Disulphide bond (S-S)

Circular proteins

Fig. 1. Three-dimensional structures of naturally occurring proteins. Clockwise from the upper left the proteins are: bacteriocin AS-48 (1E68) from Enterococcus faecalis, microcin J25 (1HG6) from Escherichia coli, MCoTI-II (1HA9, 1IB9) from bitter melon seeds, RTD-1 (1HVZ) from the leukocytes of Rhesus macaques, kalata B1 (1KAL) from several plants of the Rubiaceae and Violaceae plant families, and SFTI-1 (1SFI, 1JBL) from the seeds of the common sunflower. Disulfide bonds are shown in yellow. PDB access codes are given in parentheses after the name of the protein.

Structure of B1 kalata

Circulin

Kalata family

Rigid vertex graphs

RV graphs

- RV graph associated with a protein molecule where atoms of the backbone are vertices and edges are the corresponding bonds.
- Rigid vertex correspond to vertices with disulphide bond (S-S)

Parity invariant

- Parity to distinguish the two foldings

Understanding the topology due to disulphide bridges

- For non circular backbone , Disuphide bond gives rise to psuedoknot which can be thought of RV graph .
- We can distinguish the folds by distinguishing the Rv graphs. Use RV graph to differentiate between protein folds with strong presence of disulphide bonds
- Use this to correlate known function of protein

Thank you

Linked Protein Chains

Keratin