Univerza *v Ljubljani*

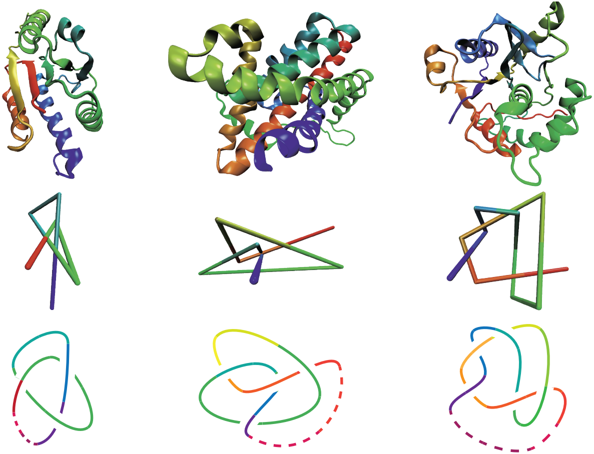
Bonded knots

Boštjan Gabrovšek

SGT & GWU, 12 November 2020

Motivation

We model a protein backbone with a closed embedded interval.



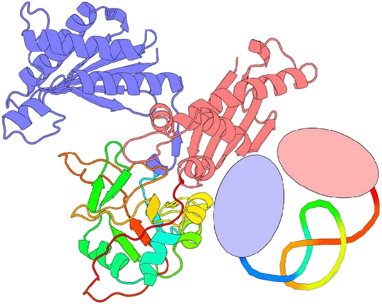
3_1 4_1 5_2

Problems & questions

1. How do we model a knotted protein / what is knottedness in a protein?
2. How do we distinguish/classify such structures?
3. Why are proteins knotted (evolutionary advantages)?
4. How do protein form knots?
5. Understand the (un)knotting process in microbiological processes.

Prelomnice

- 1994: existence of knotted proteins proposed (Mansfield)
- 1994: first knotted protein found (Liang, Mislow)
- 2000: first deep knot found, 3_1 in 4_1 (Taylor)
- 2014: knotted protein database knotprot.cent.uw.edu.pl



Protein Tp0642, deepest knot found up to date (Lim, Jackson, 2015)

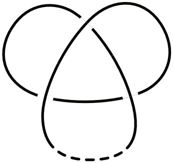
Advantages

Hypothesised (biological) advantages of knotted proteins:

- kinetic stability
- increases thermal
- prevention to being pulled into the proteasome
- knotted enzymes are often found in the proximity of proteins soon to be degraded and they face the danger of being degraded themselfs

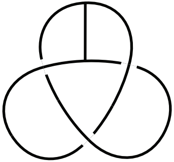
Topological models

In connection with knot theory, knotted protein have been so far modelled as:

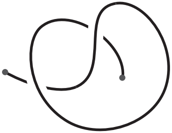


knots

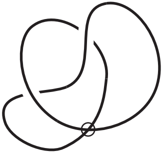
slipknots / lassos



Θ-curves



knotoids

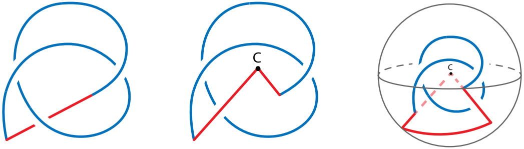


virtual knots

Closing a knot

Existence of an unambiguous closure method is still an open question, but common methods are of

- direct nature (Virnau, Mirny,...)



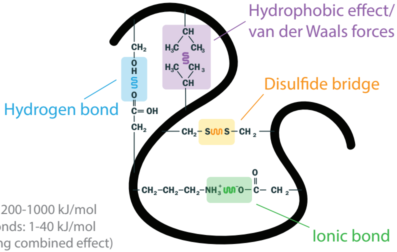
- probabilistic nature (Sulkowka, Millet,...)

PDB

3BJX: 6_1(64%), 0_1(27%), 4_1(6%), 3_1(2.5%)

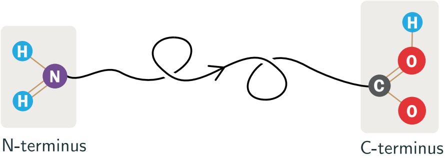
Bonds & Orientation

The three-dimensional protein structure also consists of bonds tying parts of the peptide backbone. These bonds have both a structural and functional role and can be of several types.



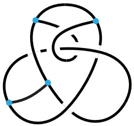
covalent bonds: 200-1000 kJ/mol
non-covalent bonds: 1-40 kJ/mol
(can have a strong combined effect)

The protein backbone also has a natural orientation

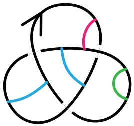


Spatial graphs

We can model *a protein with bonds* as:



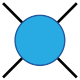
3-valent spatial graph



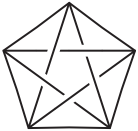
bonded knot

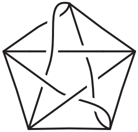
We distinguish between *non-rigid* graphs and *rigid* graphs.

non-rigid vertex



rigid vertex

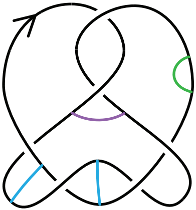




Non-rigid bonded knots (G., 2019)

A *(non-rigid) colored bonded knot* is the triple $(K, \mathcal{B}, c)$, where:

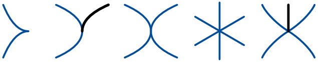
- $K \hookrightarrow \mathbb{R}^3$ is an oriented *knot*,
- $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ is the set of *bonds* properly embedded into $\mathbb{R}^3 - K$,
- $c : \mathcal{B} \to \mathbb{N}$ is the *coloring function*.



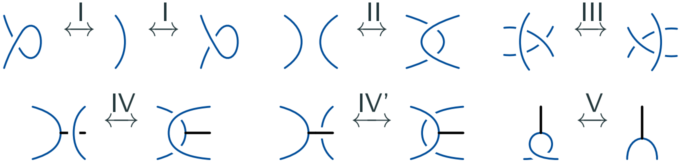
Two knots are *equivalent* if they are ambient isotopic.

Reidemeister moves

A *diagram* of a bonded knots K is a *regular projection* of K to a plane.

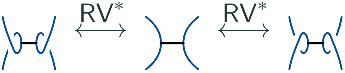
Forbidden positions: 

Reidemeister moves:



Theorem: Non-rigid vertex (ambient) isotopy is generated by moves I–V.

In order to study rigid vertex isotopy, we replace move V by either:



Rigid bonded knots

Let $\mathcal{D}$ be the set of all colored bonded knot diagrams.

Rigid (colored) bonded knots are equivalence classes

$$\bar{\mathcal{L}} = \mathcal{D}/\sim,$$

where $D_1 \sim D_2$ iff thy are connected through planar isotopy and a finite sequence of moves I–IV and RV (or RV*).

Rigid bonded knots are easier to study, but *non-rigid knots* better reflect spatial isotopy (and are better models of bonded proteins).

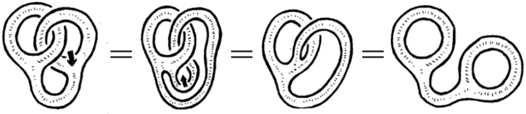
Invariants of spatial graphs

Rigid graphs:

- *Yamada polynomial*
- Kauffman's *T invariant*

Non-rigid graphs:

- *topological invariants* of $\mathbb{R}^3 \setminus G$ (are weak)

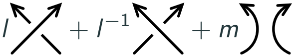


- *tangle-replacement* invariants

The HOMFLYPT skein module of bonded knots

Let

- $\mathcal{L}$ be the set of all non-rigid bonded knots,
- R be a commitative ring with units l in m (also let $l^2 + 1$ and $l^2 \pm ml + 1$ be invertible in R),
- $R[\mathcal{L}]$ be the free R-modul generated by $\mathcal{L}$,
- $S(R, l, m)$ be the submodule generated by expressions



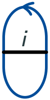
The *HOMFLYPT skein module* is the quotient module

$$\mathcal{H}(R, l, m) = R[\mathcal{L}]/S(R, l, m)$$

By taking $\bar{\mathcal{L}}$ to be the set of rigid bonded knots, we similarly define the rigid HOMFLYPT skein module $\bar{\mathcal{H}}(R, l, m)$

The HOMFLYPT skein module of <u>rigid</u> bonded knots

We define the following *elementary bonded knots* with color i:

$\Theta_i =$ 

, $\bar{\Theta}_i = i$ 
, $H_i = i$ 
$\bar{H}_i = i$ 
,

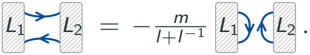
Theorem (G., 2019)

The HOMFLYPT skein module of *rigid* bonded knots $\mathcal{H}$ is freely generated by

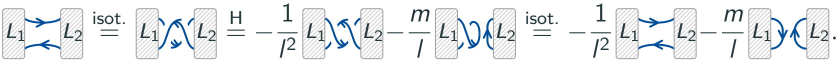
$$\mathcal{B} = \left\{ \prod_{i=1}^{k} \Theta_i^{m_i} \bar{\Theta}_i^{\bar{m}_i'} H_i^{n_i} \bar{H}_i^{n_i'} \mid k \in \mathbb{N};\ \vec{m}, \vec{m}', \vec{n}, \vec{n}' \in \mathbb{N}_0^k \setminus \vec{0} \right\} \cup \{U\}.$$

Idea of proof

Lemma:



Indeed,

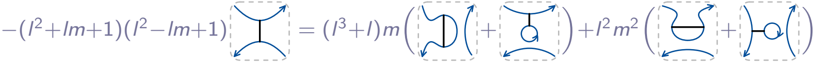


Idea of proof (generating set)

First, we show that $\mathcal{B}$ is the generating set taking these steps:

1. isolate the bond,
2. show that this bond can be "cut out" and expressed as a linear combination of knots and Θ's and H's,
3. repeat the process until no bonds left.

Using the HOMFLYPT relation, we can compute:



Using the lemma, we get:

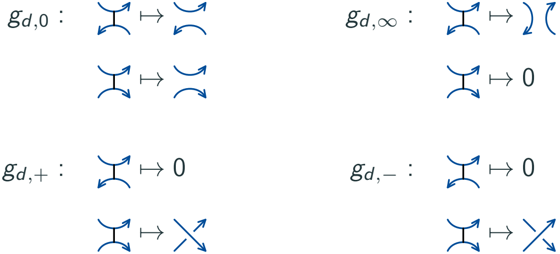
Similarly, we can get:

Idea of proof (freeness)

Second, we show that $\mathcal{H}$ is free.

We show $\quad \sum_{A\in\mathcal{B}} r(A)\,A = 0 \quad \Rightarrow \quad r(A) = 0, \forall A.$

We define R-linear maps $\hat{\mathcal{L}}_d \to \hat{\mathcal{L}}_{d-1}$, which locally replace the last d-th bond of each generator with a non-bond:



The four maps can be extended R-linearly to maps $R\hat{\mathcal{L}}_d \to R\hat{\mathcal{L}}_{d-1}$, which induce the maps on the module:

$$g^*_{d,0}, g^*_{d,\infty}, g^*_{d,+}, g^*_{d,-} : \hat{\mathcal{H}}_d(R, l, m) \to \hat{\mathcal{H}}_{d-1}(R, l, m).$$

Idea of proof (freeness)

We apply the maps on $\sum_{A\in\mathcal{B}} r(A)\,A = 0$. E.g., applying $g_{d,0}$, we get

$$0 = \sum_{B\in\hat{\mathcal{B}}_{d-1}} B\left(r(B\Theta_{c_d})\,g^*_{1,0}(\Theta_{c_d}) + r(B\bar{\Theta}_{c_d})\,g^*_{1,0}(\bar{\Theta}_{c_d}) + r(BH_{c_d})\,g^*_{1,0}(H_{c_d}) + r(B\bar{H}_{c_d})\,g^*_{1,0}(\bar{H}_{c_d})\right)$$

$$= \sum_{B\in\hat{\mathcal{B}}_{d-1}} B\left(r(B\Theta_{c_d}) + r(B\bar{\Theta}_{c_d}) + \frac{-(l+l^{-1})}{m} r(BH_{c_d}) + \frac{-(l+l^{-1})}{m} r(B\bar{H}_{c_d})\right).$$

For the other three maps we get:

$$0 = \sum_{B\in\hat{\mathcal{B}}_{d-1}} B\left(\frac{-(l+l^{-1})}{m} r(B\Theta_{c_d}) + r(BH_{c_d})\right).$$

$$0 = \sum_{B\in\hat{\mathcal{B}}_{d-1}} B\left(\frac{l^2-m^2l^2+1}{l^3m} r(B\bar{\Theta}_{c_d}) + r(B\bar{H}_{c_d})\right)$$

$$0 = \sum_{B\in\hat{\mathcal{B}}_{d-1}} B\left(\frac{-(l+l^{-1})}{m} r(B\bar{\Theta}_{c_d}) + r(\bar{H}_{c_d}B)\right).$$

The 4×4 system has an invertible determinant. By induction on the number of bonds, we conclude that the module is free.

Computing the invariant

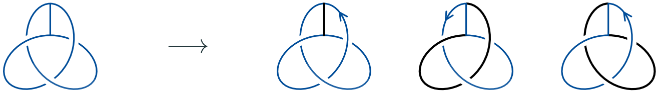
Let K be a bonded knot. The expression $[K]_{\bar{\mathcal{B}}}$, K written in the basis of the skein module, is an *invariant of colored bonded knots*.

We can *compute* the invariant $[K]_{\bar{\mathcal{B}}}$ by the following set of instructions:

1. isolate the bonds using move IV,
2. cut out the bonds (using the relations in the freeness proof),
3. compute the HOMFLYPT polynomial P of the remaining part of the classical knot.

Example 1

The Θ-curve $\Theta 3_1$ has three associated bonded knots.

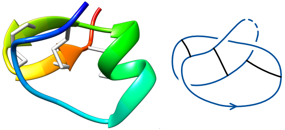


 $_{\bar{\mathcal{B}}} = (1^{-2}m^2 - l^{-2})$  $+\, l^{-3}m$ 

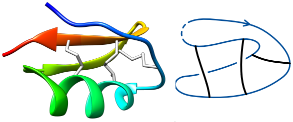
 $_{\bar{\mathcal{B}}} =$  $_{\bar{\mathcal{B}}} = (l^2m^2 - 2l^2 + m^2 - 1)$  $+\, (lm^3 - 2lm)$ 

Example 2

Toxins from venomous organisms form disulfide-rich peptides.



CN29 toxin (Mexican Nayarit Scorpion)



ADWX-1 toxin (Chinese scorpion)

$$\left[K_{\mathrm{CN29}}\right]_{\bar{\mathcal{B}}} = \frac{1}{(1+l^2)^2(l^2+ml+1)^2(l^2-ml+1)^2}\Big(l^6m^4(-1-3l^2-3l^4-l^6+l^2m^2+2l^4m^2)$$

$$+\, l^5m^3(1+3l^2+3l^4+l^6-m^2-6l^2m^2-6l^4m^2-l^6m^2+l^2m^4+3l^4m^4)$$

$$+\, l^7m^5(-1-l^2+l^2m^2) + l^6m^6(-1-2l^2+l^2m^2)$$

$$+\, l^6m^4(-1-2l^2-l^4-m^2-l^2m^2+l^4m^2+l^2m^4) + l^5m^5(-1-3l^2-2l^4+l^2m^2+l^4m^2)\Big)$$

$$\left[K_{\mathrm{ADWX\text{-}1}}\right]_{\bar{\mathcal{B}}} = \frac{1}{(1+l^2)^2(l^2+ml+1)^2(l^2-ml+1)^2}\Big(l^6m^4(-1-2l^2-l^4+l^4m^2)$$

$$+\, l^7m^5(-4-4l^2+2l^2m^2) + l^7m^5(-1+l^4) + l^7m^5(-2-2l^2+l^2m^2)$$

$$+\, l^4m^4(1+2l^2+l^4-2l^2m^2-3l^4m^2+l^4m^4) + l^6m^4(-2-4l^2-2l^4+2l^4m^2)\Big)$$

The HOMFLYPT skein module of non-rigid bonded knots

Theorem (G., 2019)

The HOMFLYPT skein module of *rigid* bonded knots $\mathcal{H}$ is freely generated by all finite products of $\Theta_i's$:

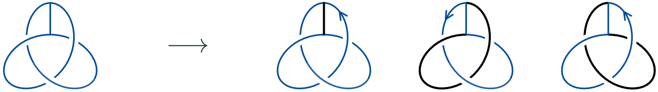
$$\mathcal{B} = \left\{ \Theta_1^{n_1} \Theta_2^{n_2} \cdots \Theta_k^{n_k} \mid \vec{n} \in \mathbb{N}_0^k \setminus \vec{0} \right\} \cup \{U\}.$$

It holds for a knot K with k bonds that

$$[K]_{\mathcal{B}} = \left(\frac{-lm}{1+l^2} \right)^{k-1} P(K')\, B,$$

where $B \in \mathcal{B}$ and $P(K')$ is the *HOMFLYPT polynomial* of K without bonds.

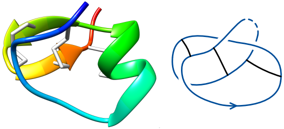
Example 1 (non-rigid case)

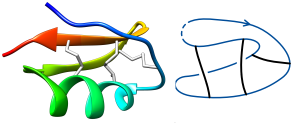


$\Big[$  $\Big]_{\mathcal{B}} = P\Big($  $\Big)\,\Theta = (l^{-2}m^2 - 2l^{-2} - l^{-4})\,\Theta$

$\Big[$  $\Big]_{\mathcal{B}} = P\Big($  $\Big)\,\Theta = \Theta$

Example 2 (non-rigid case)





CN29 toxin (Mexican Nayarit Scorpion)

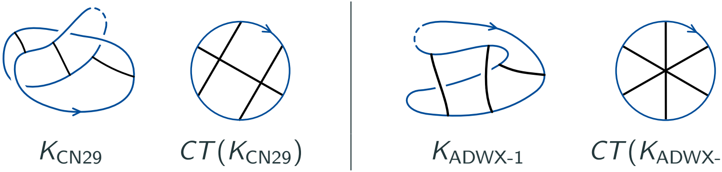
ADWX-1 toxin (Chinese scorpion)

$$[K_{\mathrm{CN29}}]_{\mathcal{B}} = \frac{l^2 m^2}{(1+l^2)^2}\,\Theta^3,$$

$$[K_{\mathrm{ADWX\text{-}1}}]_{\mathcal{B}} = \frac{l^2 m^2}{(1+l^2)^2}\,\Theta^3.$$

Expanding the variables / further work

Incorporate into the invariant information about the bonded knots' *CT (circuit topology)*.



K_{CN29} $\quad CT(K_{\text{CN29}})$ $\quad K_{\text{ADWX-1}}$ $\quad CT(K_{\text{ADWX-1}})$

Expanding the variables / further work

Generalize the *Yamada polynomial* $R : (\mathcal{G}_{\text{rig}} \subset \mathbb{R}^3) \to \mathbb{Z}[A^{\pm 1}]$ to bonded knots.

The yamada polynomial is defined by the following relations:

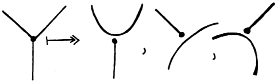
1. $R\Big($  $\Big) = AR\Big($  $\Big) + A^{-1}R\Big($  $\Big) + R\Big($  $\Big)$,
2. $R(G) = R(G - e) + R(G/e)$, e ni zanka,
3. $R(G_1 \sqcup G_2) = R(G_1)R(G_2)$,
4. $R(G_1 \vee G_2) = -R(G_1)R(G_2)$,
5. $R\Big($  $\Big) = -(-A - 1 - A^{-1})^n$,
6. $R(\emptyset) = 1$.

Remark: the R is an invariant of rigid-vertex graphs with max degree ≥ 4, but an invariant non-rigid-vertex graphs with max degree ≤ 3.

Expanding the variables / further work

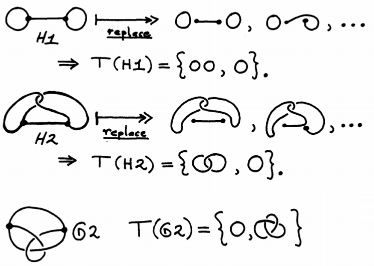
Generalize *Kauffman's T invariant*

Let $G \subset S^3$ be spatial graph. Consider the local replacements of a vertex:



Let $r(G)$ be the set of closed curves obtained by local replacements of all vertices.

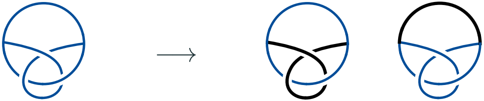
$$T(G) = \{r(G)\}_r$$



Expanding the variables / further work

We can expand T by counting the number of bonds lying on the closed components.

Consider coloring two different arcs of the Θ-curve $\Theta 3_1$:



Values of the extended invariant T' on these two bonded knots