Bonded knots

A topological model for knotted proteins

Boštjan Gabrovšek

8ECM, MS-11 Low-dimensional Topology

21-22 June 2021, Portorož, Slovenia
History of protein knots

- 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 UCHL3 contains the knot $5_2$
Protein Tp0642, deepest knot found up to date (Lim, Jackson, 2015)
Questions/Problems

- Why are proteins knotted (evolutionary advantages)?
- How do protein form knots in microbiological processes?
- How do we distinguish/classify/analyse such structures?

Hypothesised (biological) advantages of knotted proteins:
- increases thermal stability
- increases kinetic stability
- increases chemical stability
- prevention to being pulled into the proteasome
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  inequivalent rigid-vertex spatial graphs
```

*Rigid bonded knots* are easier to study, but *non-rigid knots* better reflect spatial isotopy.
A (non-rigid) colored bonded knot is the triple $(K, B, c)$, where:

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

Two bonded 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:**

For two non-rigid bonded knot diagrams to represent isotopic knots, they must be connected through a finite sequence of moves I–V.

In order to study *rigid isotopy*, we replace move V by

**Theorem**

Two non-rigid bonded knot diagrams represent isotopic knots iff they are connected through a finite sequence of moves I–V.
Let $\mathcal{D}$ be the set of all colored bonded knot diagrams. 

*Rigid (colored) bonded knots* are equivalence classes

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

where $D_1 \sim D_2$ iff thy are connected through planar isotopy and a finite sequence of moves I–IV and RV.
The HOMFLYPT polynomial

The HOMFLYPT polynomial of classical knots

\[ P : \mathcal{L}(S^3) \rightarrow \mathbb{Z}[l^{\pm 1}, z^{\pm 1}] \]

is defined using skein relations:

\[ P(\bigcirc) = 1 \quad \text{and} \quad lP(\begin{array}{c}
\text{\includegraphics[width=0.1\textwidth]{diagram1.png}}
\end{array}) + l^{-1}P(\begin{array}{c}
\text{\includegraphics[width=0.1\textwidth]{diagram2.png}}
\end{array}) + mP(\begin{array}{c}
\text{\includegraphics[width=0.1\textwidth]{diagram3.png}}
\end{array}) = 0 \]

In the case of bonded knots \( P \) is not well defined:

\[ P(\begin{array}{c}
\text{\includegraphics[width=0.18\textwidth]{diagram4.png}}
\end{array}) = -l^2 P(\begin{array}{c}
\text{\includegraphics[width=0.18\textwidth]{diagram5.png}}
\end{array}) - lm P(\begin{array}{c}
\text{\includegraphics[width=0.18\textwidth]{diagram6.png}}
\end{array}) \]

We will form an \( R \)-module in which it holds

\[ \begin{bmatrix}
\text{\includegraphics[width=0.18\textwidth]{diagram7.png}}
\end{bmatrix} = -l^2 \begin{bmatrix}
\text{\includegraphics[width=0.18\textwidth]{diagram8.png}}
\end{bmatrix} - lm \begin{bmatrix}
\text{\includegraphics[width=0.18\textwidth]{diagram9.png}}
\end{bmatrix}. \]
The HOMFLYPT skein module of bonded knots

Let

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

\[
I \left[ \begin{array}{c} \includegraphics{image1} \\ \includegraphics{image2} \end{array} \right]_\mathcal{B} + l^{-1} \left[ \begin{array}{c} \includegraphics{image3} \\ \includegraphics{image4} \end{array} \right]_\mathcal{B} + m \left[ \begin{array}{c} \includegraphics{image5} \\ \includegraphics{image6} \end{array} \right]_\mathcal{B}
\]

The non-rigid HOMFLYPT skein module is the quotient module

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

By taking \( \bar{\mathcal{L}} \) to be the set of rigid bonded knots, we similarly define the rigid HOMFLYPT skein module \( \bar{\mathcal{H}}(\mathcal{R}, l, m) \)
We define the following *elementary bonded knots* with color $i$:

$$
\Theta_i = \begin{array}{c}
\includegraphics[width=0.1\textwidth]{theta_i.png}
\end{array},
\bar{\Theta}_i = \begin{array}{c}
\includegraphics[width=0.1\textwidth]{theta_bar_i.png}
\end{array},
H_i = \begin{array}{c}
\includegraphics[width=0.08\textwidth]{h_i.png}
\end{array},
\bar{H}_i = \begin{array}{c}
\includegraphics[width=0.08\textwidth]{h_bar_i.png}
\end{array},
$$

### Theorems (G., 2020)

1. The H.S.M. of *rigid* bonded knots $\mathcal{H}$ is freely generated by

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

2. The H.S.M. of *non-rigid* bonded knots $\mathcal{H}$ is freely generated by

$$
\mathcal{B} = \left\{ \prod_{i=1}^{k} \Theta_i^{n_i} \mid \vec{n} \in \mathbb{N}_0^k \setminus \vec{0} \right\} \cup \{U\}.
$$

E.g. $\Theta_1^3 \bar{\Theta}_1 H_2 \bar{H}_3^2 = \begin{array}{c}
\includegraphics[width=0.4\textwidth]{example.png}
\end{array}$
Ideas of proof (generating set)

First, we show that $B$ 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 $\Theta$’s and $H$’s,
3. repeat the process until no bonds left.

Using the HOMFLYPT relation, we can compute:

\[
(l^2 + lm + 1)(l^2 - lm + 1) = l^2m^2 \left( \frac{\cdot H_i +}{\cdot \Theta_i} \right) + \frac{l^3m^3}{1+l^2} \left( \frac{\cdot \Theta_i +}{\cdot H_i} \right).
\]

and

\[
(l^2 + lm + 1)(l^2 - lm + 1) = l^2m^2 \left( \frac{\cdot \bar{H}_i +}{\cdot \bar{\Theta}_i} \right) + \frac{l^3m^3}{1+l^2} \left( \frac{\cdot \bar{\Theta}_i +}{\cdot \bar{H}_i} \right).
\]
We can associate three bonded knots to the theta-curve $\Theta 3_1$.

$$\left[ \begin{array}{c} \includegraphics[width=2cm]{knot1} \end{array} \right]_B = (l^{-2}m^2 - 2l^{-2} - l^{-4}) \Theta$$

$$\left[ \begin{array}{c} \includegraphics[width=2cm]{knot2} \end{array} \right]_B = \Theta$$
Example (rigid case)

\[
K_{\text{CN29}} = \frac{1}{(1 + \ell^2)^2(\ell^2 + ml + 1)^2(\ell^2 - ml + 1)^2} \left( \ell^6 m^4 (-1 - 3\ell^2 - 3\ell^4 - \ell^6 + \ell^2 m^2 + 2\ell^4 m^2) \right) + \ell^5 m^3 (1 + 3\ell^2 + 3\ell^4 + \ell^6 - m^2 - 6\ell^2 m^2 - 6\ell^4 m^2 - \ell^6 m^2 + \ell^2 m^4 + 3\ell^4 m^4) + \ell^7 m^5 (-1 - \ell^2 + \ell^2 m^2) + \ell^6 m^4 (-1 - 2\ell^2 + \ell^2 m^2) + \ell^6 m^4 (-1 - 2\ell^2 - \ell^4 - m^2 - \ell^2 m^2 + \ell^4 m^2 + \ell^2 m^4) + \ell^7 m^5 (-1 - 3\ell^2 - 2\ell^4 + \ell^2 m^2 + \ell^4 m^2)
\]

\[
K_{\text{ADWX-1}} = \frac{1}{(1 + \ell^2)^2(\ell^2 + ml + 1)^2(\ell^2 - ml + 1)^2} \left( \ell^6 m^4 (-1 - 2\ell^2 - \ell^4 + \ell^4 m^2) \right) + \ell^7 m^5 (-4 - 4\ell^2 + 2\ell^2 m^2) + \ell^7 m^5 (-1 + \ell^4) + \ell^7 m^5 (-1 + \ell^4) + \ell^7 m^5 (-2 - 2\ell^2 + \ell^2 m^2) + \ell^4 m^4 (1 + 2\ell^2 + \ell^4 - 2\ell^2 m^2 - 3\ell^4 m^2 + \ell^4 m^4) + \ell^6 m^4 (-2 - 4\ell^2 - 2\ell^4 + 2\ell^4 m^2)
\]
Example (rigid case)

\[
K_{CN29} = \frac{1}{(1 + l^2)^2(i^2 + ml + 1)^2(i^2 - ml + 1)^2} \left( i^6 m^4(-1 - 3l^2 - 3l^4 - l^6 + l^2 m^2 + 2l^4 m^2) \right)
\]

\[
= \frac{1}{(1 + l^2)^2(i^2 + ml + 1)^2(i^2 - ml + 1)^2} \left( i^5 m^3(1 + 3l^2 + 3l^4 + l^6 - m^2 - 6l^2 m^2 - 6l^4 m^2 - l^6 m^2 + l^2 m^4 + 3l^4 m^4) \right)
\]

\[
+ i^7 m^5(-1 - l^2 + l^2 m^2) \left( i^6 m^5(-1 - 2l^2 + l^2 m^2) \right)
\]

\[
+ i^6 m^4(-1 - 2l^2 - l^4 - m^2 - l^2 m^2 + l^4 m^2 + l^2 m^4) \left( i^5 m^5(-1 - 3l^2 - 2l^4 + l^2 m^2 + l^4 m^2) \right)
\]

\[
K_{ADWX-1} = \frac{1}{(1 + l^2)^2(i^2 + ml + 1)^2(i^2 - ml + 1)^2} \left( i^6 m^4(-1 - 2l^2 - l^4 + l^4 m^2) \right)
\]

\[
= \frac{1}{(1 + l^2)^2(i^2 + ml + 1)^2(i^2 - ml + 1)^2} \left( i^7 m^5(-4 - 4l^2 + 2l^2 m^2) \right)
\]

\[
+ i^7 m^5(-1 + l^4) \left( i^7 m^5(-2 - 2l^2 + l^2 m^2) \right)
\]

\[
+ i^4 m^4(1 + 2l^2 + l^4 - 2l^2 m^2 - 3l^4 m^2 + l^4 m^4) \left( i^6 m^4(-2 - 4l^2 - 2l^4 + 2l^4 m^2) \right)
\]

Thank you!