Protein Networks

Proteins are the building blocks of cells. By interacting they form complex networks, which execute most of the cellular functions.

Models for protein networks range between two extremes:
- Differential equations \( \frac{dQ_{ij}}{dt} = k_{ii}(Q_i) - k_{ij}(Q_j) \)
- Interaction graphs

We introduce protein hypernetworks [3], a new model that stands in between.

Protein Hypernetworks

Rationale: The functional versatility of protein networks emerges from dependencies between the interactions, e.g. allosteric regulation, competition on binding domains, and post-translational modifications (e.g. phosphorylation).

**Protein Hypernetwork \([P, I, C]\):**
- Use a plain undirected graph \((P, I)\) for proteins and interactions.
- Add a set \(C\) of propositional logic constraints modeling interaction dependencies.

Minimal Network States

Under the given constraints \(C\), we are interested in all minimal sufficient network configurations that allow the occurrence of an interaction or protein \(q \in P \cup I\).

Minimal network states can be calculated by enumerating satisfying solutions for

\[ q \land \land C \]

using the propositional logic tableau calculus [4].

If all minimal network states of two interactions clash, they are not possible simultaneously:

Prediction of Protein Complexes

Protein complexes are expected to be dense subnetworks. Protein hypernetworks can be used to improve the prediction of protein complexes by taking constraints into account.

1. **Predict protein complexes on the plain graph**

2. **Calculate maximal combinations of minimal network states**

3. **Redecrypt refined complexes**

Results on the yeast protein network (4579 proteins, 12576 interactions [1], 458 constraints [2]):

Text-Mining for Interaction Dependencies

In contrast to protein interactions, that can be found in online databases like CYGD [1] or BioGrid [5], interaction dependencies are so far hidden as natural language statements in scientific publications. We propose a linear-time two-step algorithm to mine large collections of publications for interaction dependencies.

Given a text, e.g.

"... It is also possible that the interaction is indirect, via a Cbl-associated protein. Shishido and colleagues proposed that binding of Abl induces a conformational change in Cbl that allows binding of Src via its SH3 domain (Shishido et al., 2000); association of endogenous Src could explain the weak phosphorylation of Cbl occasionally induced by kinase-inactive Abl (Figure 2c)."

the algorithm performs the following two steps:

1. **Translate the text into a sequence of tokens.** Each token represents a relevant class of words.

2. **On the sequence of tokens, search for given regular expression patterns.** Matches hint to interaction dependencies.

For the human adhesome protein network [6] we extracted 144 matches from 16,419 publications. From these, we manually curated 41 interaction dependencies.

**References**


