Introduction:
Video games for structure prediction and design

CS/CME/Biophys/BMI 371

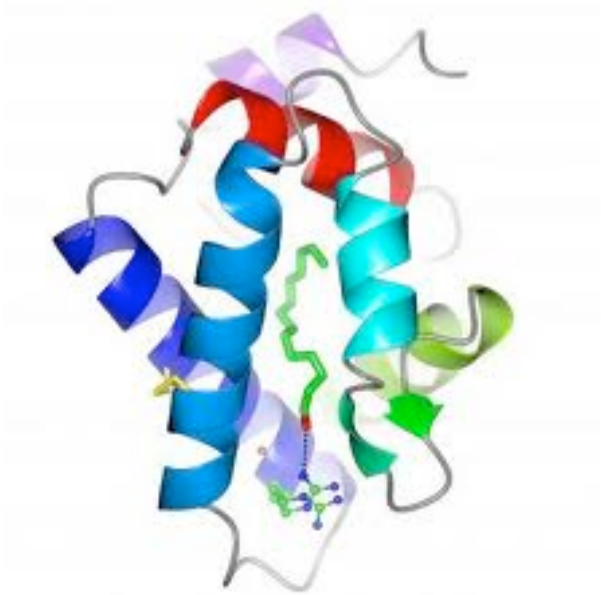
Jan. 16, 2018

Ron Dror

Protein Structure Prediction

- Goal: given the amino acid sequence of a protein, predict its three-dimensional structure
  - Proteins are constantly in motion. We want the average structure, which is roughly what's measured experimentally.
- Protein structure prediction is usually done by template-based modeling (also known as homology modeling), which requires a protein with similar sequence whose structure is known
  - If there isn't one, we need to do *ab initio* prediction, which is harder

SVYDAAAQLTADVKKDLRDSW
KVIGSDKKGNGVALMTTLFAD
NQETIGYFKRLGNVSQGMAND
KLRGHSITLMYALQNFIDQLD
NPDSLDLVCS…….

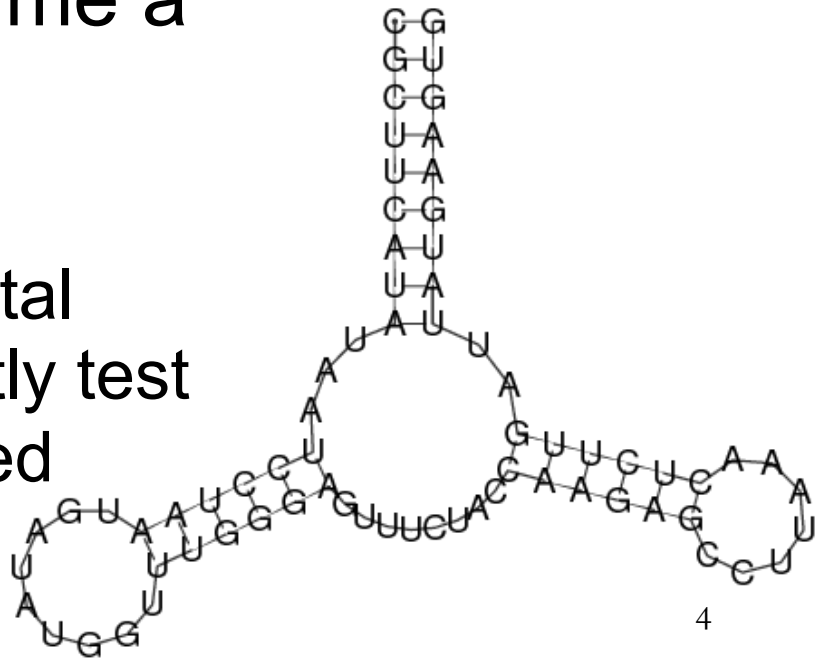


FoldIt: Human-assisted structure prediction

- Rosetta, a popular software package for *ab initio* structure prediction, has two key components:
  - An energy function, which assigns an energy to each possible structure
    - This energy function is inherently an approximation
  - An optimization procedure that searches for global minimum of the energy function
    - Existing optimization procedures are imperfect
- FoldIt lets humans help with the optimization process
  - FoldIt players have discovered clever structural optimization tricks that can be coded on a computer

RNA design

- For RNA, “secondary structure” — that is, how the bases pair — is of primary interest
- *Predicting* this computationally is usually a tractable problem
- *Designing* RNAs that will assume a desired secondary structure, however, is still tricky
  - A recently developed experimental procedure allows one to efficiently test whether one achieved the desired structure



http://www.tbi.univie.ac.at/~pkerp/forgi/_images/1y26_ss.png

EteRNA: Human-assisted RNA design

- Goal: design RNA with specified secondary structure
- Players who reach a certain level get to contribute to designs that are tested experimentally on a periodic basis
- Feedback from these experiments helps players improve their designs
- Players have discovered new rules for RNA design

Papers for Thursday

- One paper introduces FoldIt
- One covers EteRNA (developed at Stanford and Carnegie Mellon)
- One describes use of FoldIt (as well as a separate classroom project) to “refine” an experimental structure
  - Here the starting point is an “electron density map” from x-ray crystallography

- For background on protein structure prediction and crystallography, see slides from CS/CME/BioE/ Biophys/BMI 279:
  - http://web.stanford.edu/class/cs279/lectures/lecture5.pdf
  - http://web.stanford.edu/class/cs279/lectures/lecture14.pdf

Reading critically

- It's tempting to imagine that crowdsourcing will take over all of computational structural biology, but at least so far, its reach has been limited
  - Why?
- When reading these papers (and all the others in this course), try to identify the limitations