Proteins, remote homologies, …

Andrew Torda, ZBH

• What do we do in Bundesstrasse?
  • numerical methods
    • simulations - proteins, RNA, evolution
    • design of molecules – RNA
    • structure prediction – RNA, proteins
    • MD simulations (DNA + protein)

• Today mostly proteins
  • bit of RNA
RNA – design and structure

- Unusual spaces and dynamics
- Design molecules
  - useful structure (ribozyme, riboswitch, …)
  - find a new sequence which preserves structure
- energy functions? Literature conventions + tricks
- particles can be A, C, G, U
  - search space = $4 \times 4 \times 4 \ldots = 4^N$
  - discrete problem? maybe not
RNA design

- particles can be of mixed types
- treat as coordinates
- dynamics in this space
- will it work?

Stefan Bienert
Marco Mathies
Sequence Dynamics

- Classic Newtonian dynamics / fictitious space
- Energy functions
  - literature +
  - prevent mis-folding
- Toy system
- read out sequence at end
- real tests..
- proteins
Proteins – prediction and similarities

• goals – structure predictions, similarities

Similarities – what is my protein related to?
• sequences
  • what function do I have?
  • what structure do I have?

• structures
  • able to find more remote similarities
  • hopeless running time (NP complete)

• sequence changes faster than structure
Structure / sequence similarity

- TLR / toll-like receptors

1fyv

Gundolf Schenk
Thomas Margraf
Structure / sequence similarity

- TLR / toll-like receptors

48% sequence id
1o77

another TLR
Structure / sequence similarity

- TLR / toll-like receptors

6% sequence id
2qxy

response regulator
T. Maritima
Structure / sequence similarity

- TLR / toll-like receptors

9 % sequence id
1e5d

bacterial oxidoreductase
Structure / sequence similarity

- TLR / toll-like receptors

7 % sequence id
1ja0

rat oxidoreductase

- 100’s more examples

- how are they calculated?
Calculating alignments

- coordinates to vectors of structure properties
- fill score matrix
- find best path
Calculating alignments

- if one can do pairs of proteins swiftly..
Structural Phylogeny

- a bigger alignment
Structural Phylogeny
Peptidases
Peptidases
Peptidases
Peptidases
Peptidases

- includes TIM barrel
Peptidases
Methodology

- mostly classic phylogeny – some unique features
  - numerical approach
    - nodes are imaginary structures

- structure comparing?

  cardigan.zbh.uni-hamburg.de/salami/

- who cares?
Kinases

- kinases
  - structure based methods
  - accurate alignments
  - active / conserved residues
  - reliable classification

- needs structure?
  - mostly
  - interplay of sequence and structure
Sequence versus structure

• close homology
  • use sequence – models, classification, function
  • easy

• remote homology
  • more speculative
  • needs structural information
From classification to prediction

- previously
  - vectors of structural properties
- now
  - mix sequence and structure properties
- result
  - from known (sequence)
  - to unknown (structure)
- via known structures (threading)
  
  www.zbh.uni-hamburg.de/wurst/
- completely new (Monte Carlo like methods)