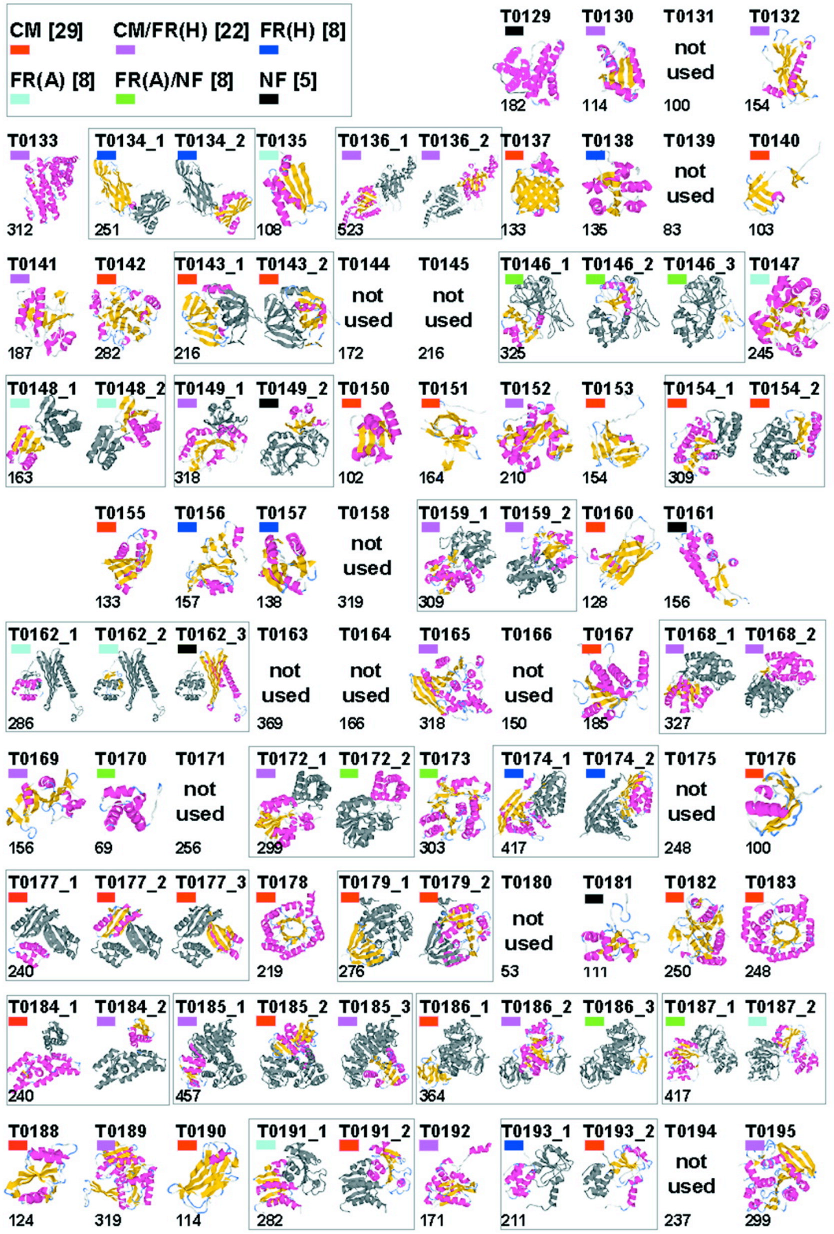
The Five Categories of CASP Targets

1. CM/E (Comparative Modeling / Easy) ← Structural homolog found by BLAST.
2. CM/H (Comparative Modeling / Hard) ← structural homolog found by 5 rounds of PSI-BLAST.
3. FR/H (Fold Recognition / Homology) ← Structural comparison to PDB finds a structure found by PSI-BLAST.
4. FR/A (Fold Recognition / Analogy) ← Finds a similar structure, no evidence of sequence homology.
5. NF (New Fold) ← nothing “similar” in the PDB

CASP

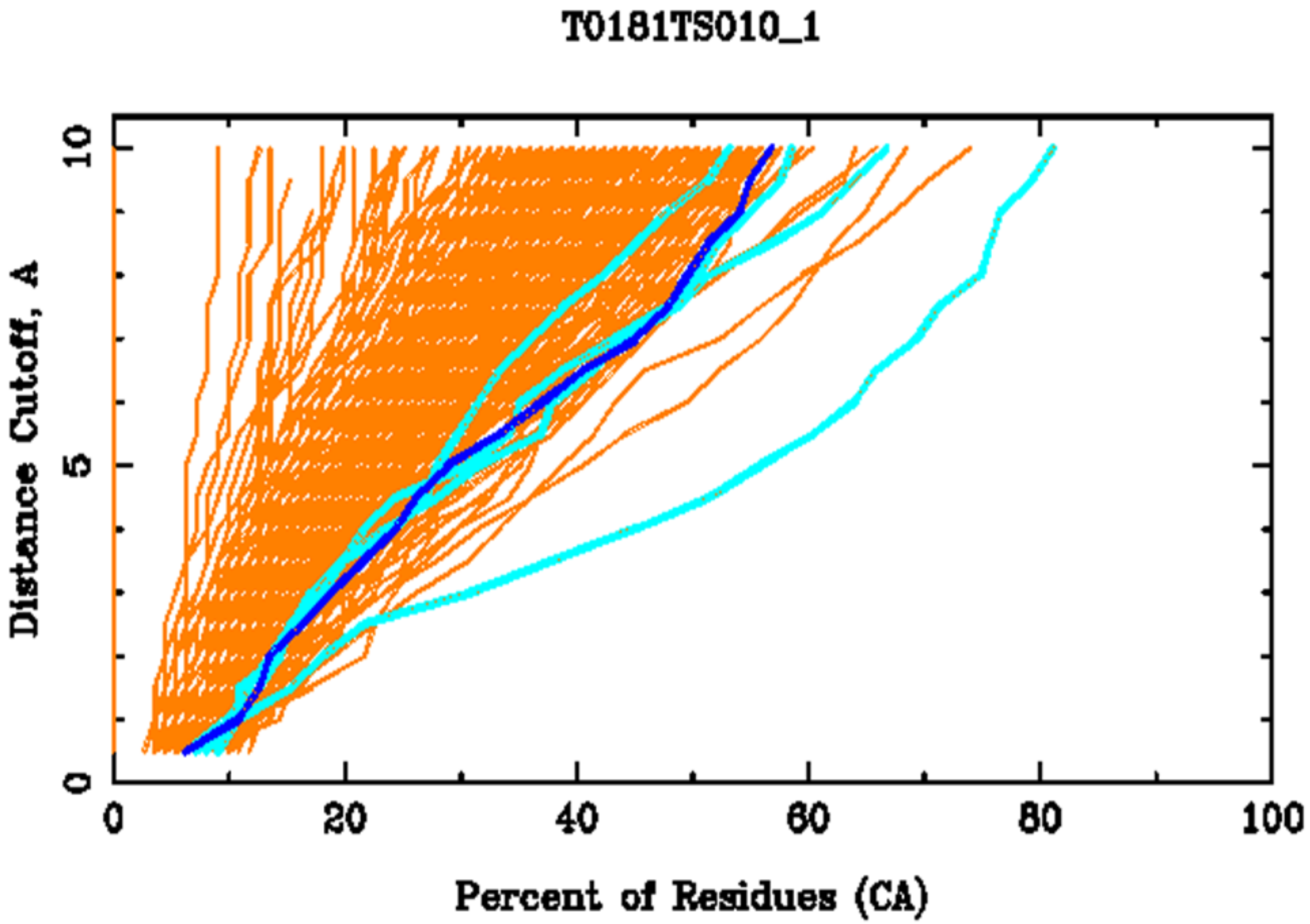


CM [29]
CM/FR(H) [22]
FR(H) [8]
FR(A) [8]
FR(A)/NF [8]
NF [5]
T0129 182
T0130 114
T0131 not used 100
T0132 154
T0133 312
T0134_1 T0134_2 251
T0135 108
T0136_1 T0136_2 523
T0137 133
T0138 135
T0139 not used 83
T0140 103
T0141 187
T0142 282
T0143_1 T0143_2 216
T0144 not used 172
T0145 not used 216
T0146_1 T0146_2 T0146_3 325
T0147 245
T0148_1 T0148_2 163
T0149_1 T0149_2 318
T0150 102
T0151 164
T0152 210
T0153 154
T0154_1 T0154_2 309
T0155 133
T0156 157
T0157 138
T0158 not used 319
T0159_1 T0159_2 309
T0160 128
T0161 156
T0162_1 T0162_2 T0162_3 286
T0163 not used 369
T0164 not used 166
T0165 318
T0166 not used 150
T0167 185
T0168_1 T0168_2 327
T0169 156
T0170 69
T0171 not used 256
T0172_1 T0172_2 299
T0173 303
T0174_1 T0174_2 417
T0175 not used 248
T0176 100
T0177_1 T0177_2 T0177_3 240
T0178 219
T0179_1 T0179_2 276
T0180 not used 53
T0181 111
T0182 250
T0183 248
T0184_1 T0184_2 240
T0185_1 T0185_2 T0185_3 457
T0186_1 T0186_2 T0186_3 364
T0187_1 T0187_2 417
T0188 124
T0189 319
T0190 114
T0191_1 T0191_2 282
T0192 171
T0193_1 T0193_2 211
T0194 not used 237
T0195 299

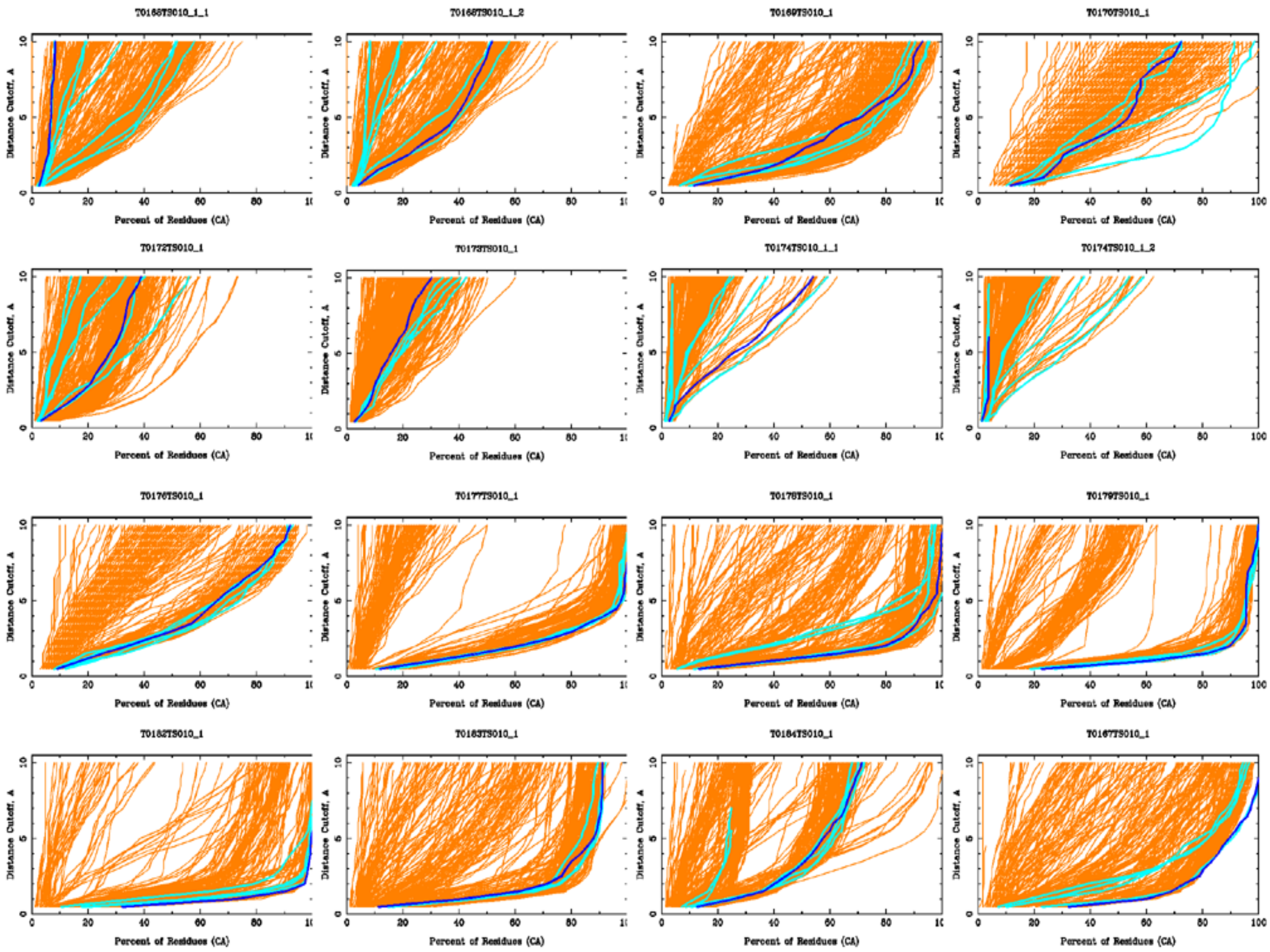
CASP Questions

1. Are the models produced similar to the corresponding experimental structure?
2. Is the mapping of the target sequence onto the proposed structure (i.e. the alignment) correct?
3. Have similar structures that a model can be based on been identified?
4. Are the details of the models correct?
5. Has there been progress from the earlier CASPs?
6. What methods are most effective?
7. Where can future effort be most productively focused?

Hubbard Plot



Hubbard Plots



http://www3.interscience.wiley.com/cgi-bin/jissue/106558999

Home / Chemistry / Biochemistry

Proteins: Structure, Function, and Genetics

Volume 53, Issue S6

Supplement: Fifth Meeting on the Critical Assessment of Techniques for Protein Structure Prediction

Copyright © 2003 Wiley-Liss, Inc., A Wiley Company

Save Title to My Profile   Set E-Mail Alert

Go to the homepage for this journal to access trials, sample copies, editorial and author information, news, and more.

Proteins: Structure, Function, and Genetics

TABLE OF CONTENTS

Volume 53, Issue S6, Pages 333-595 (2003)

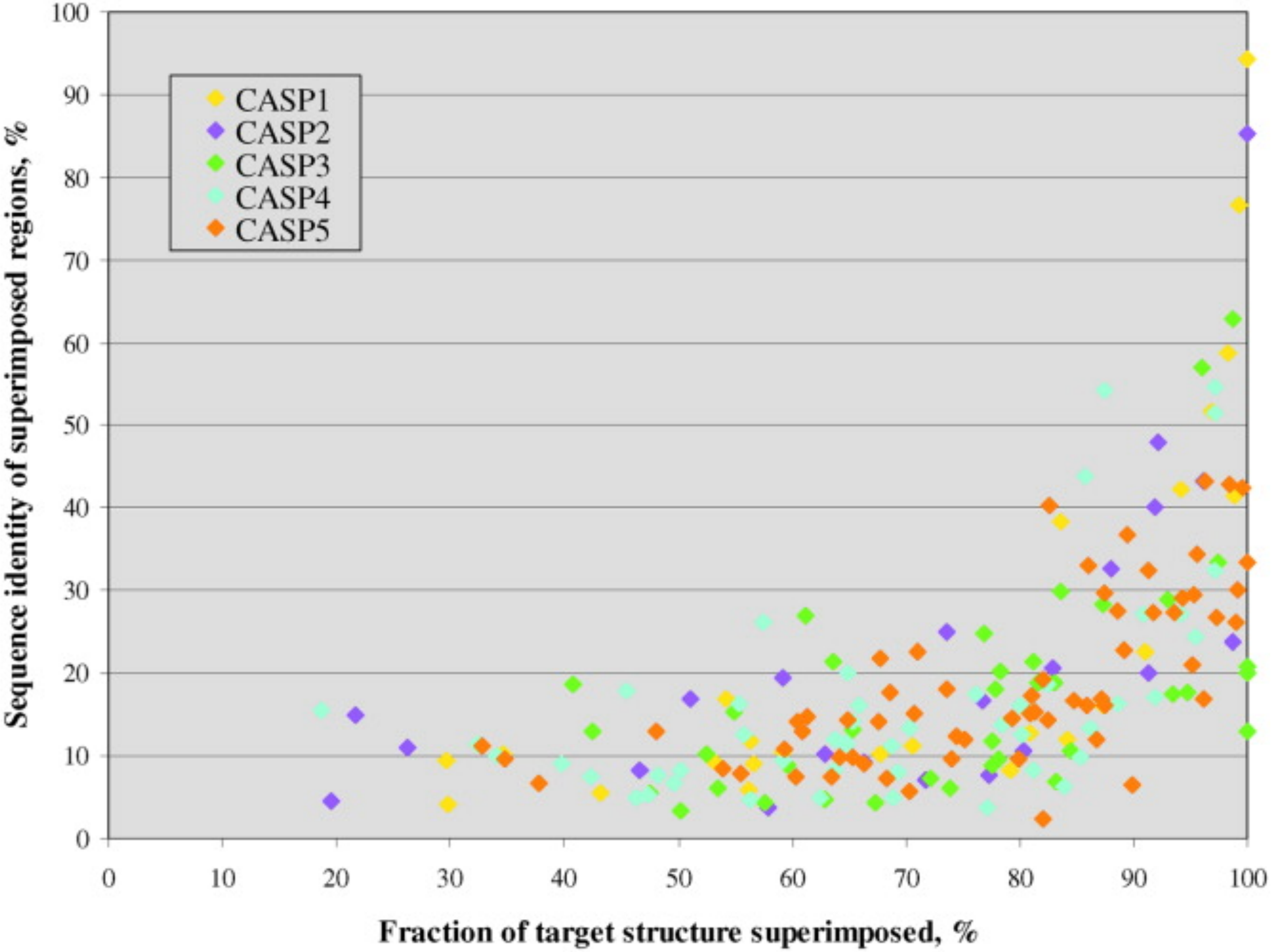
Editorial

Fifth Meeting on the Critical Assessment of Techniques for Protein Structure Prediction (p 333)
Eaton E. Lattman
Published Online: 15 Oct 2003
DOI: 10.1002/prot.10580
Abstract | Full Text: HTML , PDF (Size: 28K)
Save Article

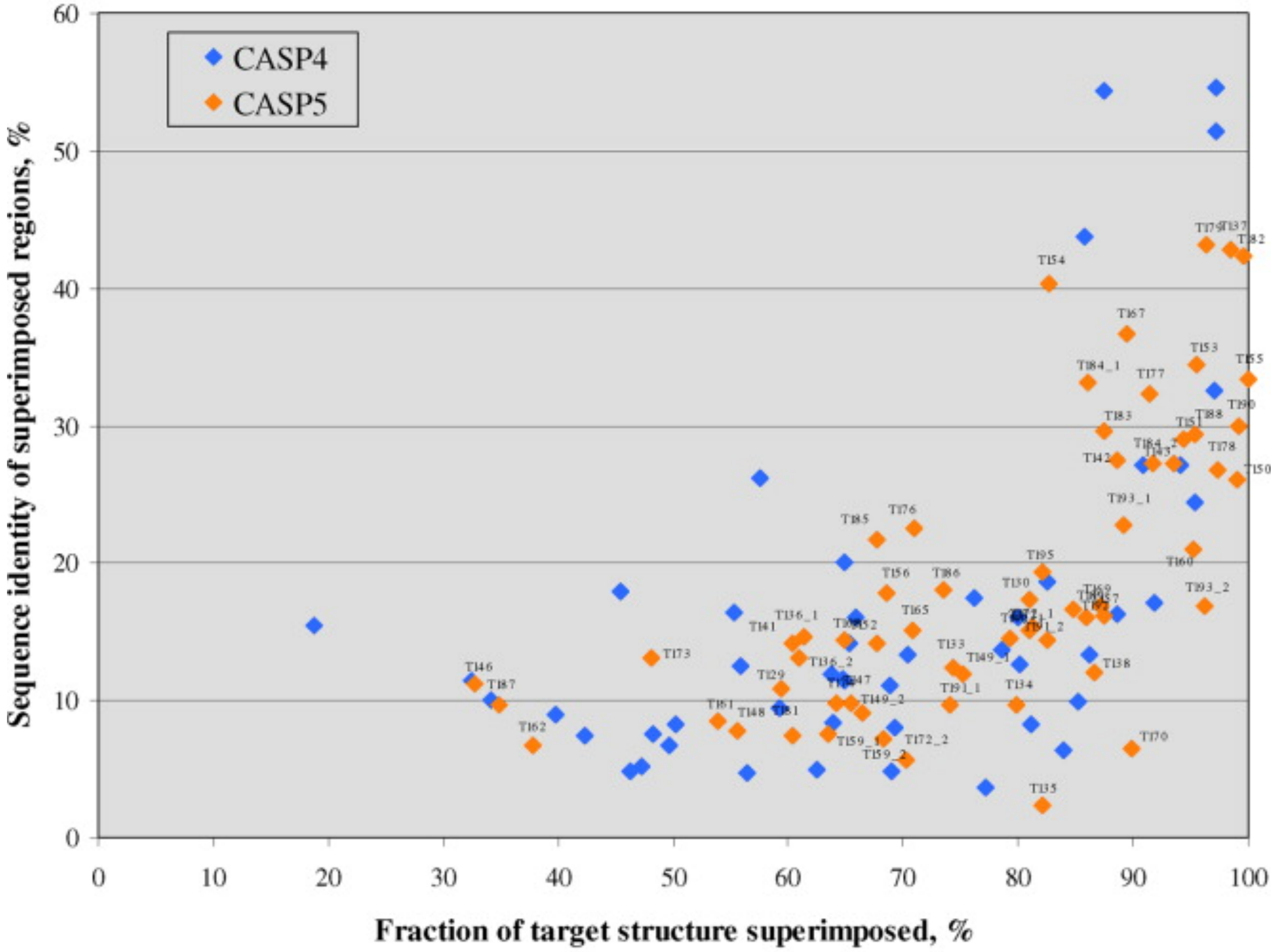
Introduction

Critical assessment of methods of protein structure prediction (CASP)-round V (p 334-339)
John Moult, Krzysztof Fidelis, Adam Zemla, Tim Hubbard
Published Online: 15 Oct 2003
DOI: 10.1002/prot.10556
Abstract | References | Full Text: HTML , PDF (Size: 71K)
Save Article

Distribution of Target Difficulty



Distribution of Target Difficulty

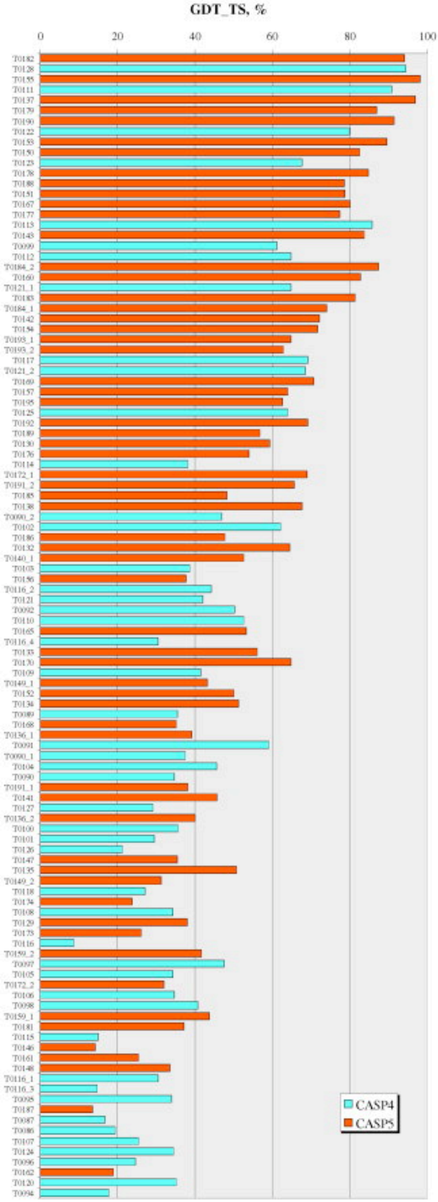


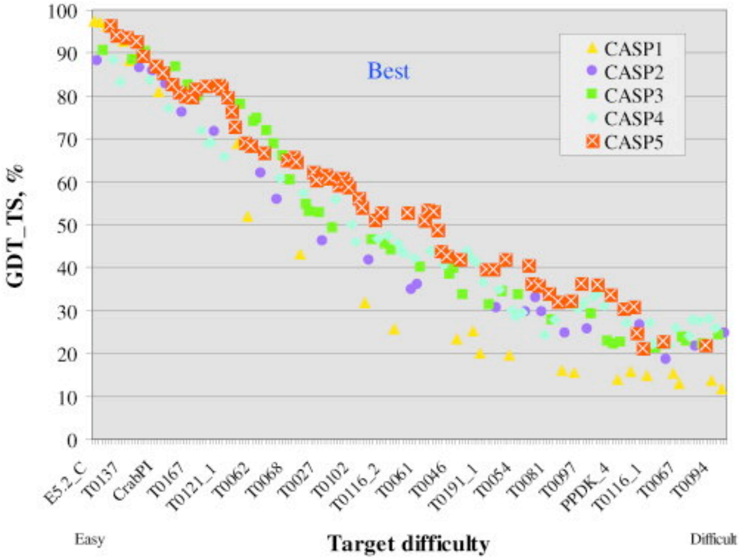
Overall Model Quality Assessment

Venclovas et al (2003): ”A large sample of possible structure superpositions of the model on the corresponding experimental structure is generated by superposing all sets of three, five, and seven consecutive Ca along the backbone (each peptide segment provides one superposition). Each of these initial superpositions is iteratively extended, including all residue pairs under a specified threshold in the next iteration, and continuing until there is no change in included residues. The procedure is conducted by using thresholds of 1, 2, 4, and 8 Å, and the superposition that includes the maximum number of residues, is selected for each threshold ... GDT_TS is then obtained by averaging over the four superposition scores for the different thresholds:

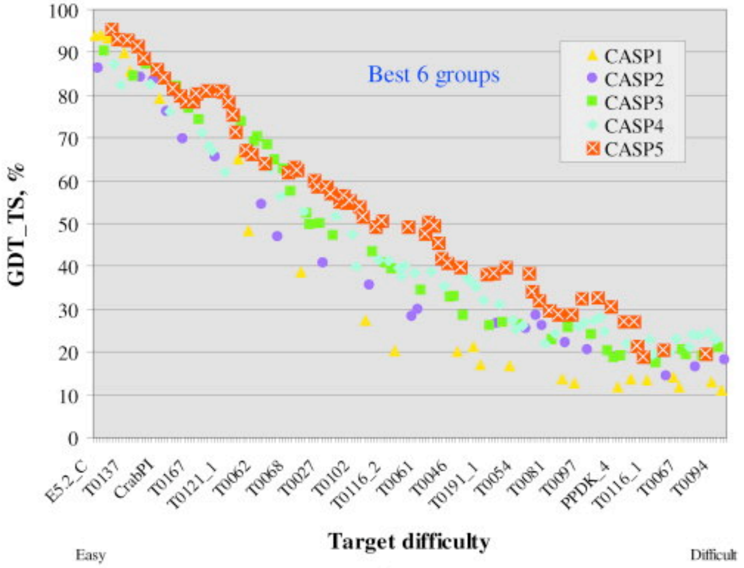
$$GDT\_TS = (N1+N2+N4+N8) / 4$$

CASP5 Progress





(A)



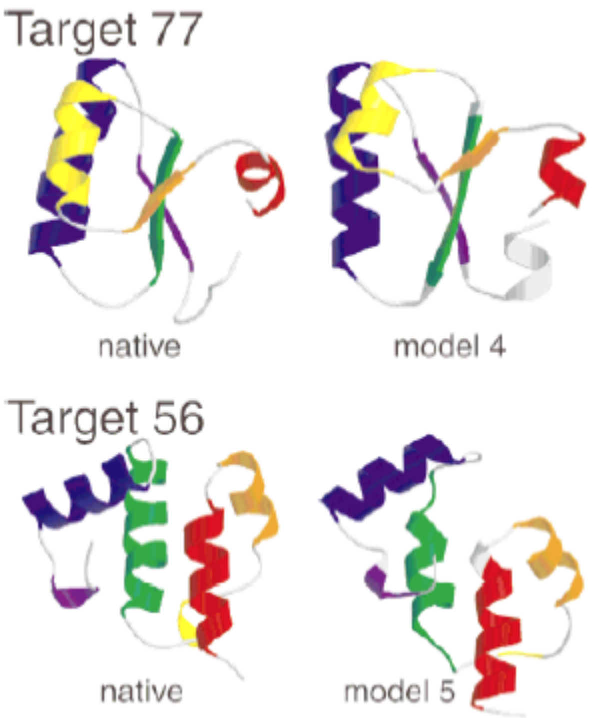
(B)

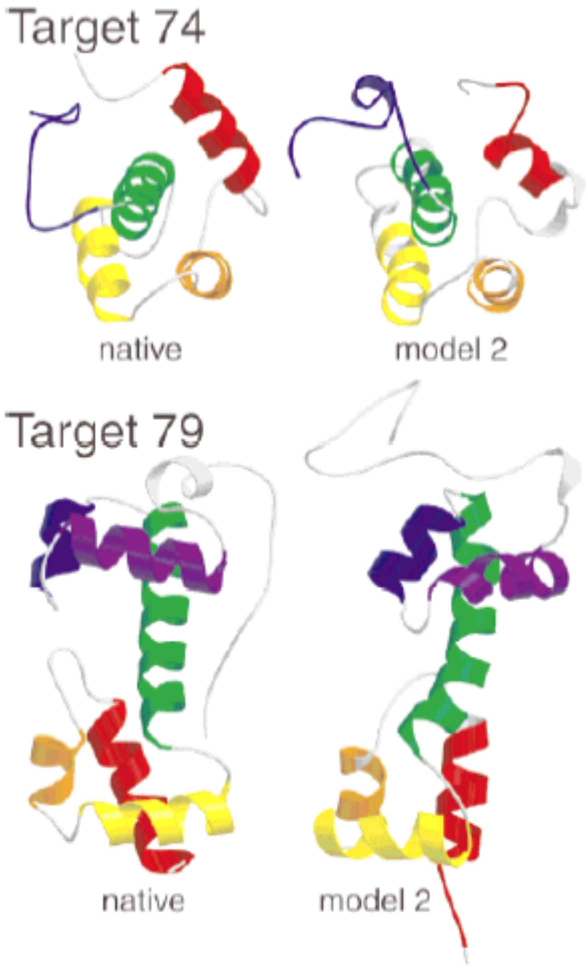
CASP Problem Areas and Bottlenecks

1. Alignment of a sequence onto a template fold.
2. Model refinement - improving accuracy of initial models.
3. Accurately modeling regions of insertion and deletion relative to a template structure.
4. Improved fold recognition, particularly for analogous, analogous/new fold targets.
5. Improved New Fold methods (for recognizing new folds).

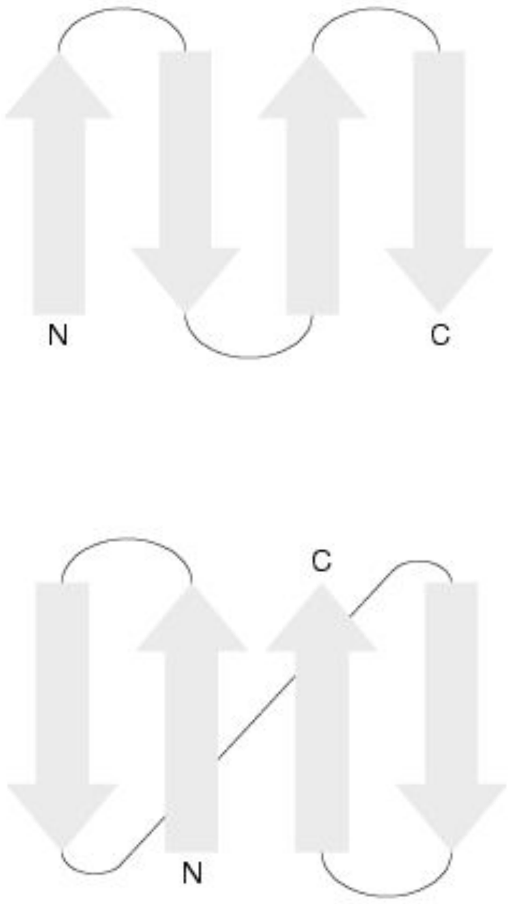
Always the same …

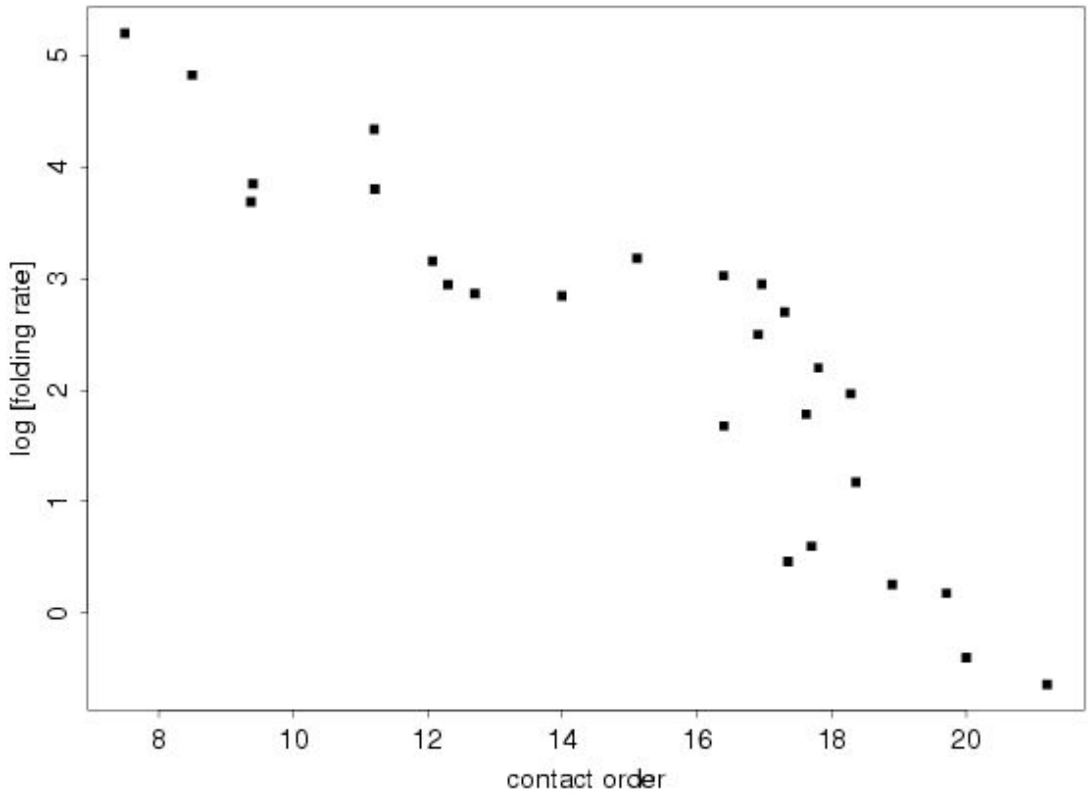
CASP3 Predictions



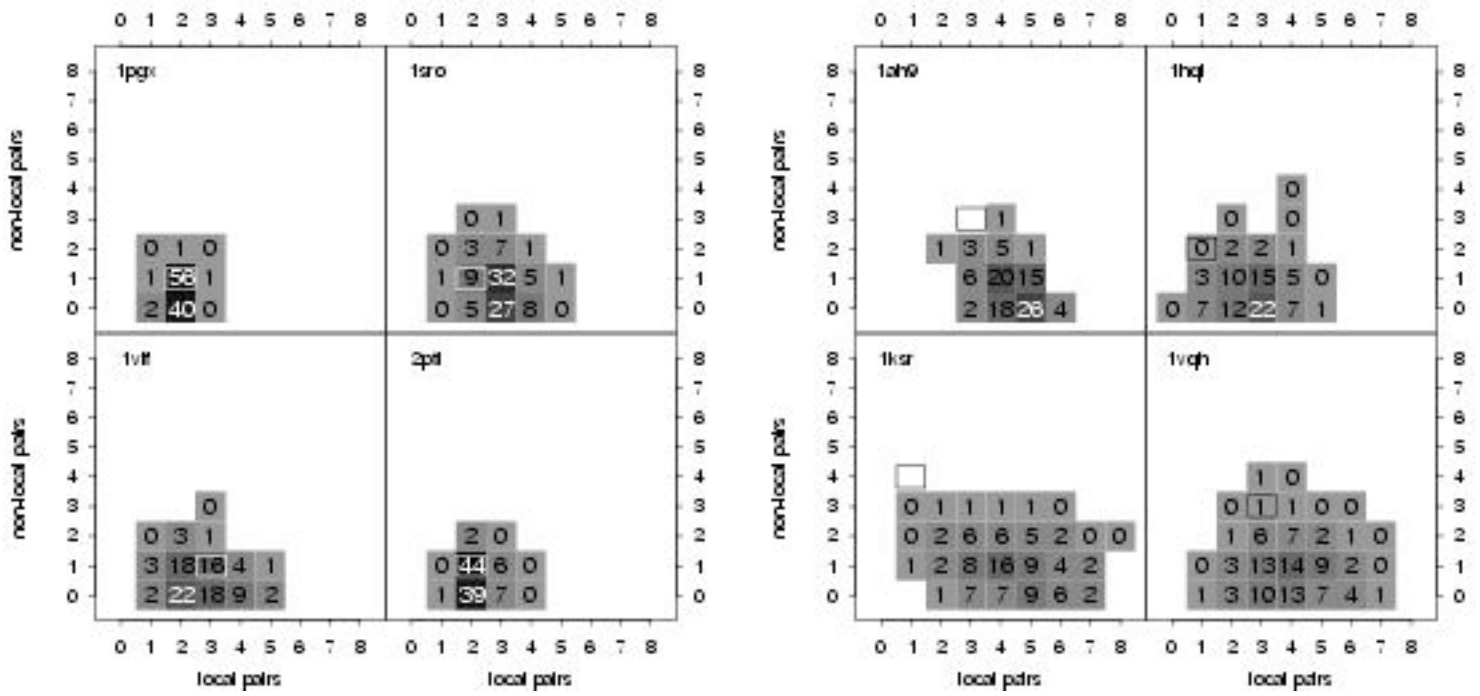


Contact Order

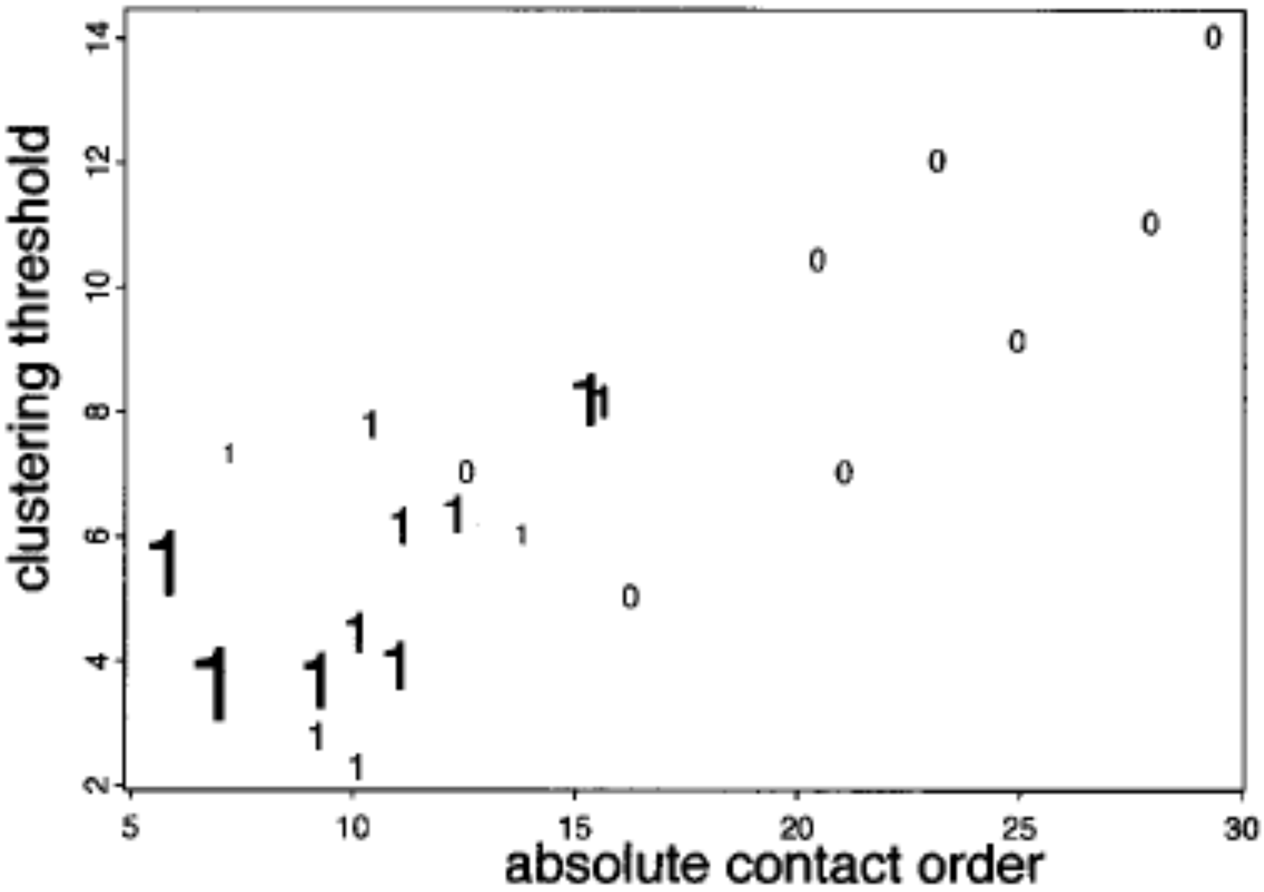




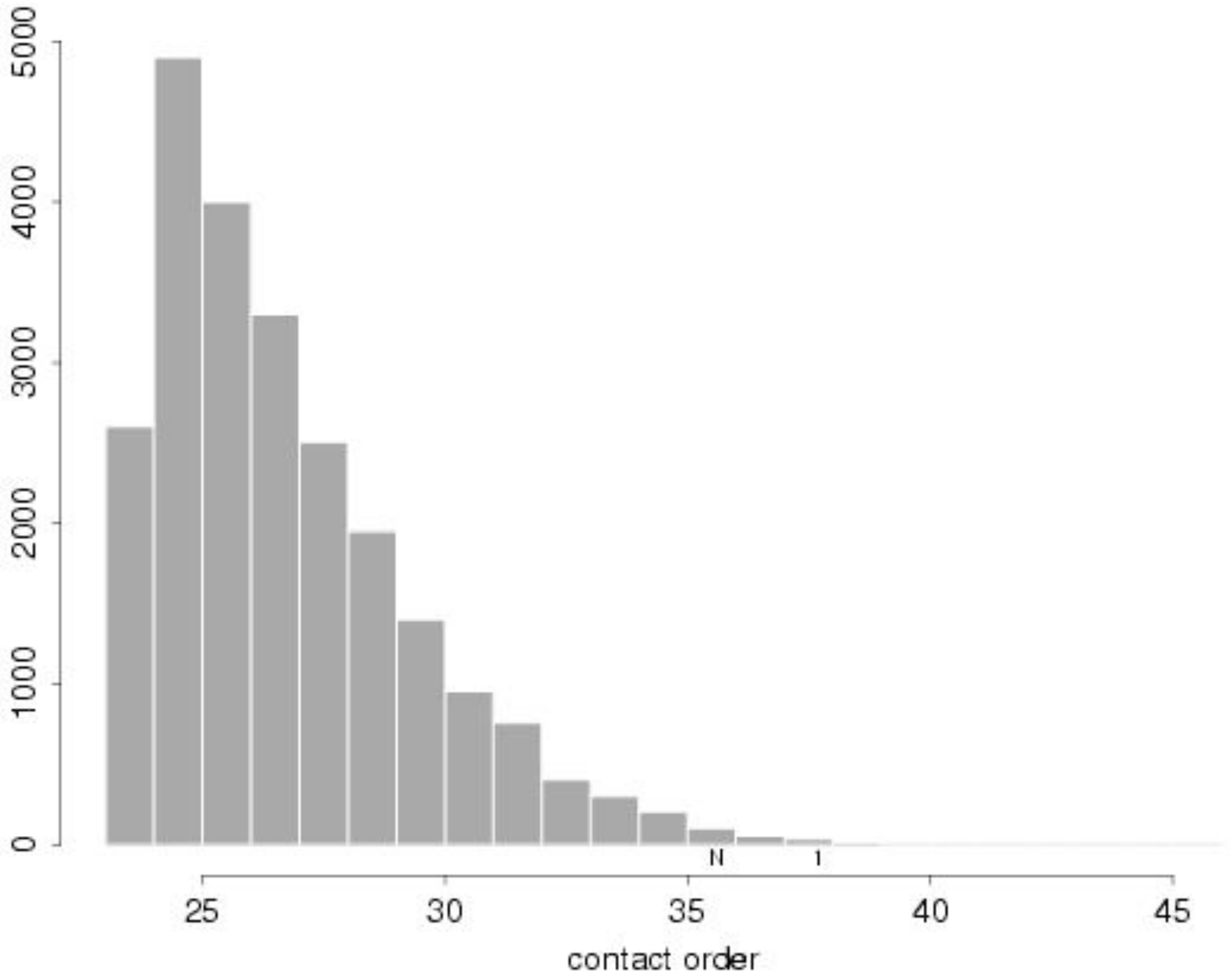
Contact Order



Clustering and Contact Order



Decoy Enrichment in CASP4

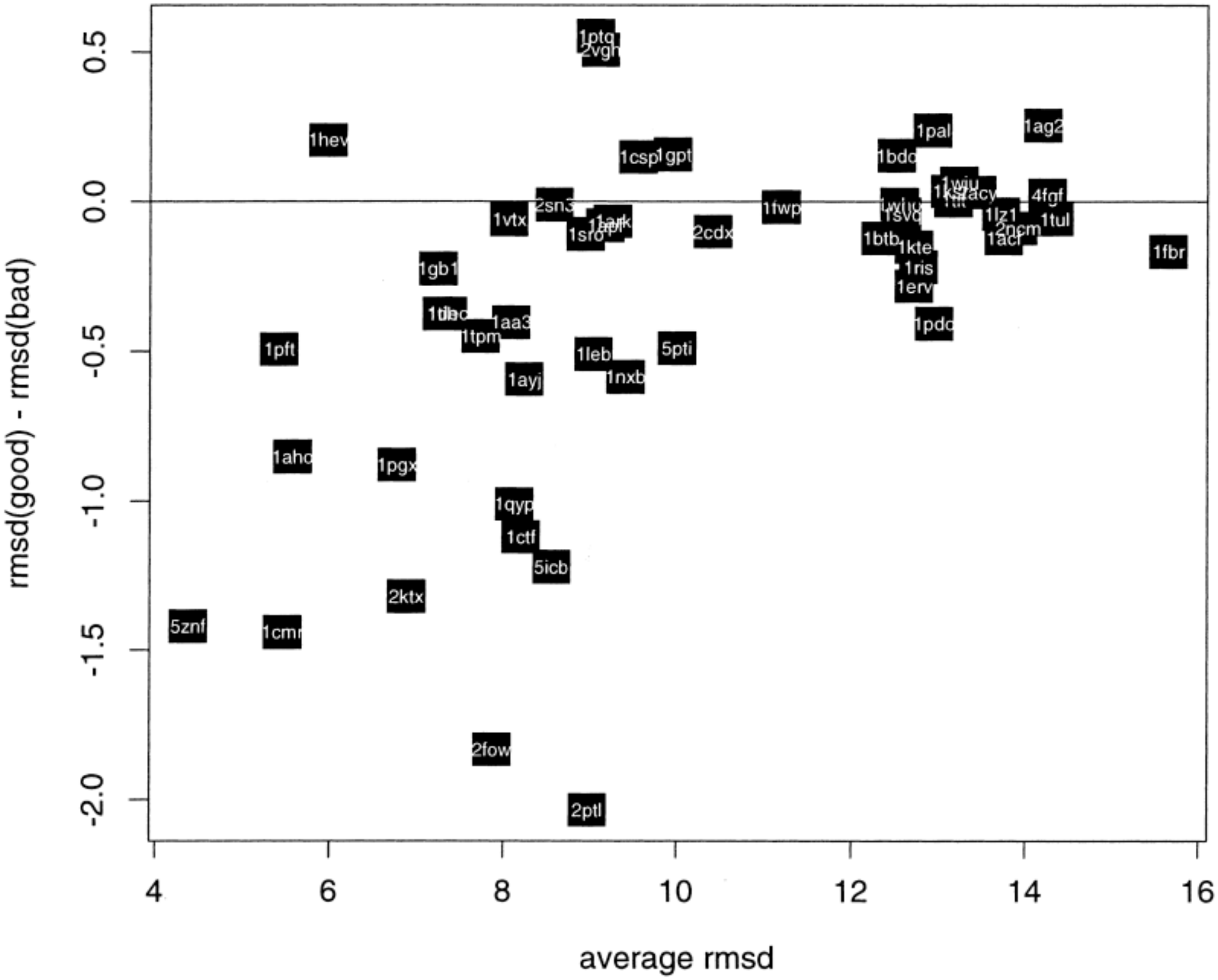


A Filter for Bad β-Sheets

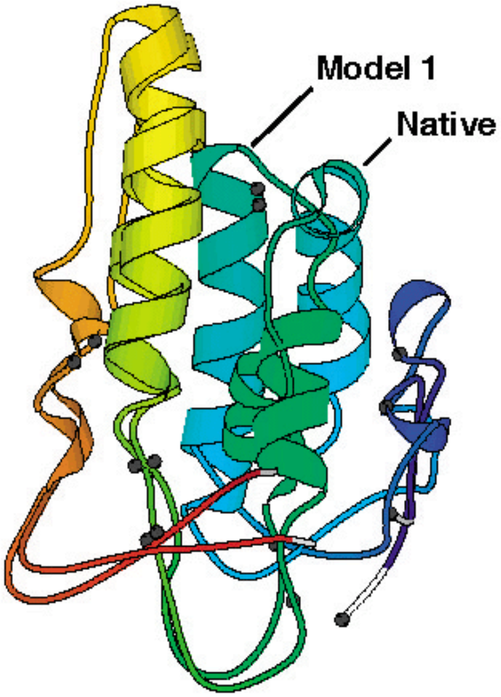
Many decoys do not have proper sheets. Filtering those out seems to enhance the rmsd distribution in the decoy set. Bad features we see in decoys include:

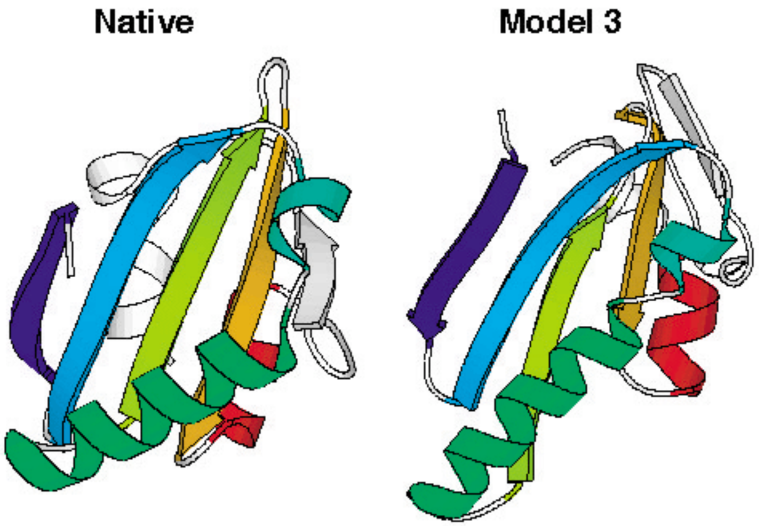
- No strands,
- Single strands,
- Too many neighbours,
- Single strand in sheets,
- Bad dot-product,
- False handedness,
- False sheet type (barrel),
- …

A Filter for Bad β-Sheets

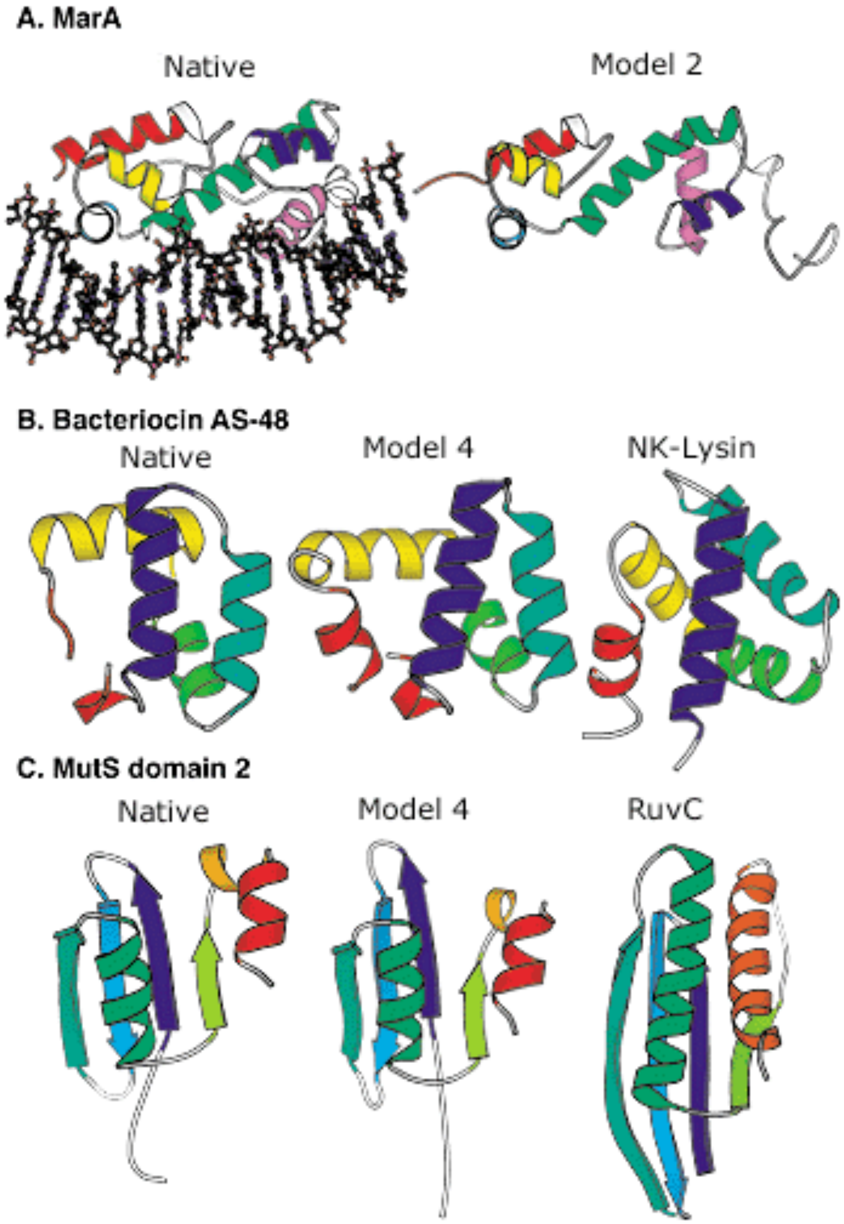


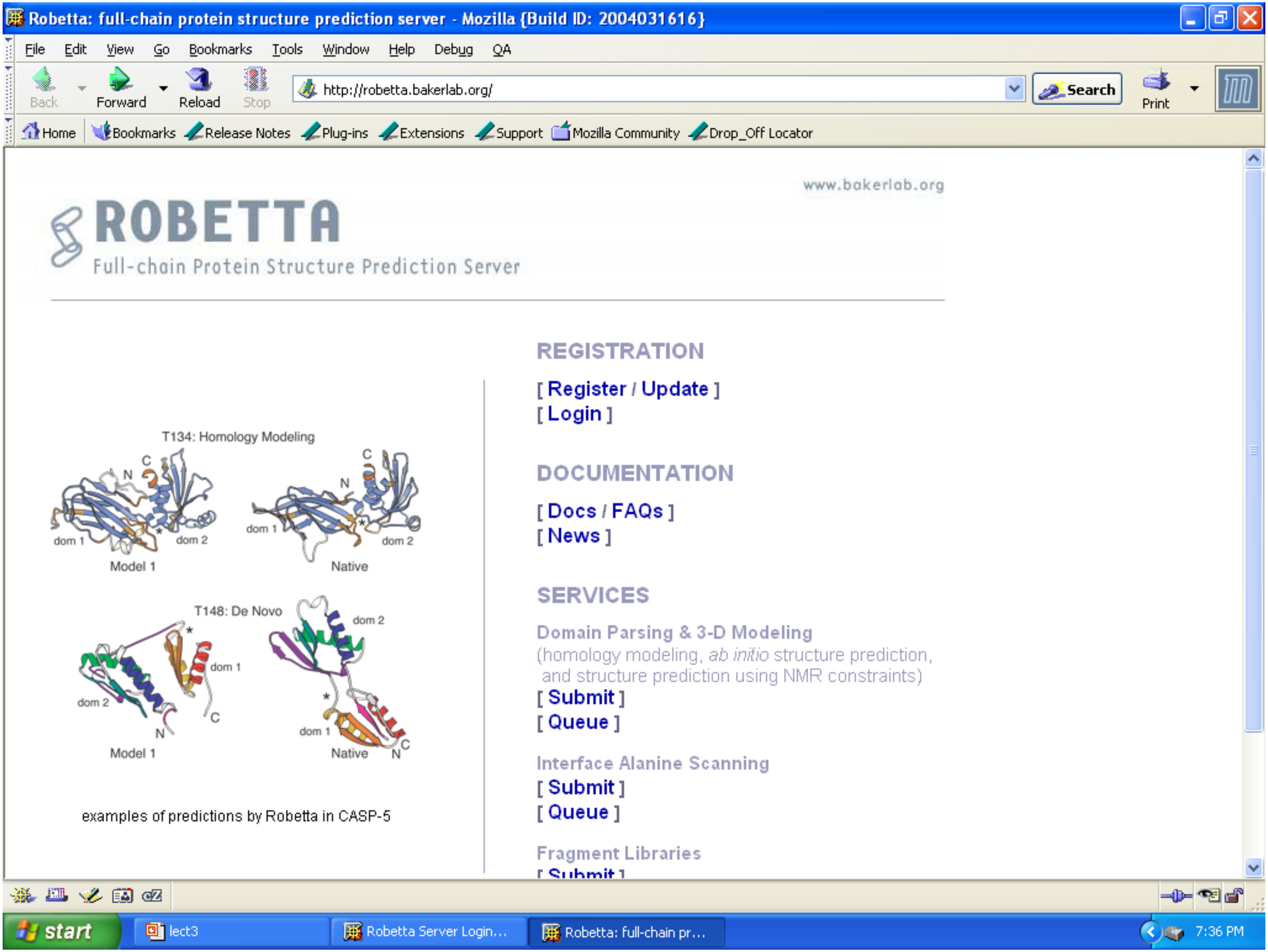
Rosetta in CASP4





CASP 4


A. MarA
Native
Model 2
B. Bacteriocin AS-48
Native
Model 4
NK-Lysin
C. MutS domain 2
Native
Model 4
RuvC



www.bakerlab.org

ROBETTA

Full-chain Protein Structure Prediction Server

examples of predictions by Robetta in CASP-5

REGISTRATION

[Register / Update]
[Login]

DOCUMENTATION

[Docs / FAQs]
[News]

SERVICES

Domain Parsing & 3-D Modeling

(homology modeling, *ab initio* structure prediction, and structure prediction using NMR constraints)

[Submit]
[Queue]

Interface Alanine Scanning

[Submit]
[Queue]

Fragment Libraries

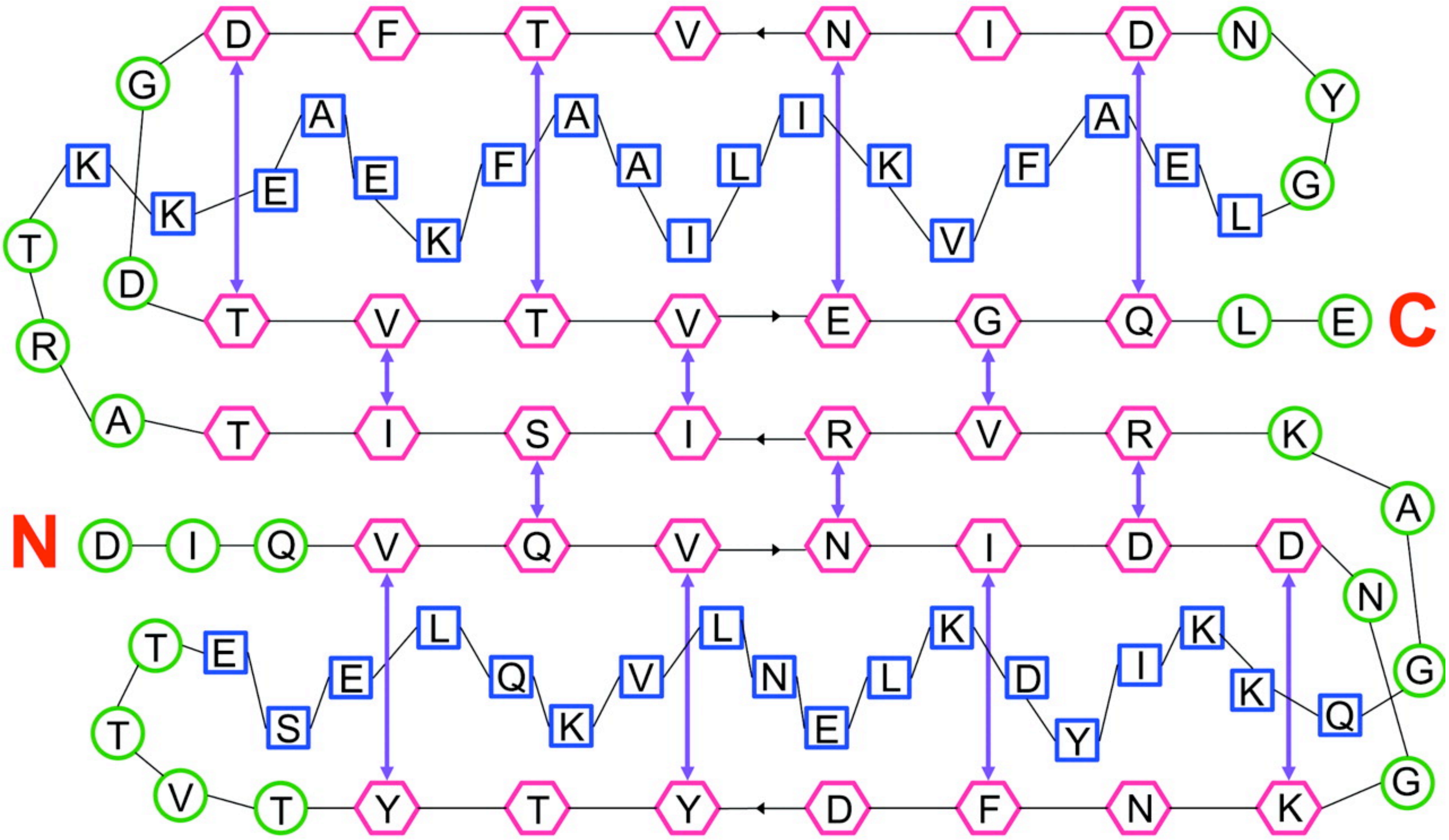
[Submit]

Applications and Other Uses of Rosetta

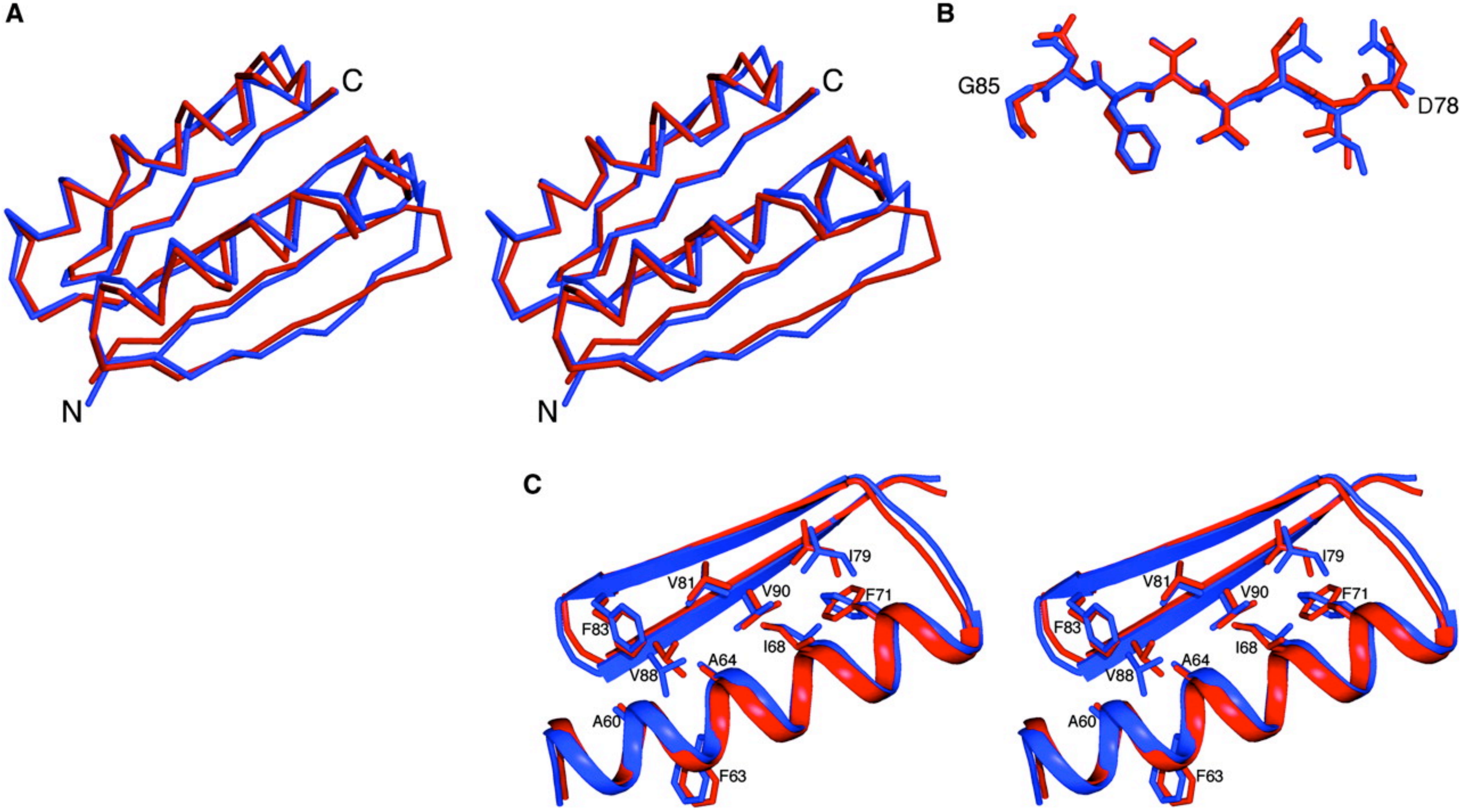
- Other uses of Rosetta:
  - Homology modeling.
  - Rosetta NMR.
  - Protein interactions (docking).

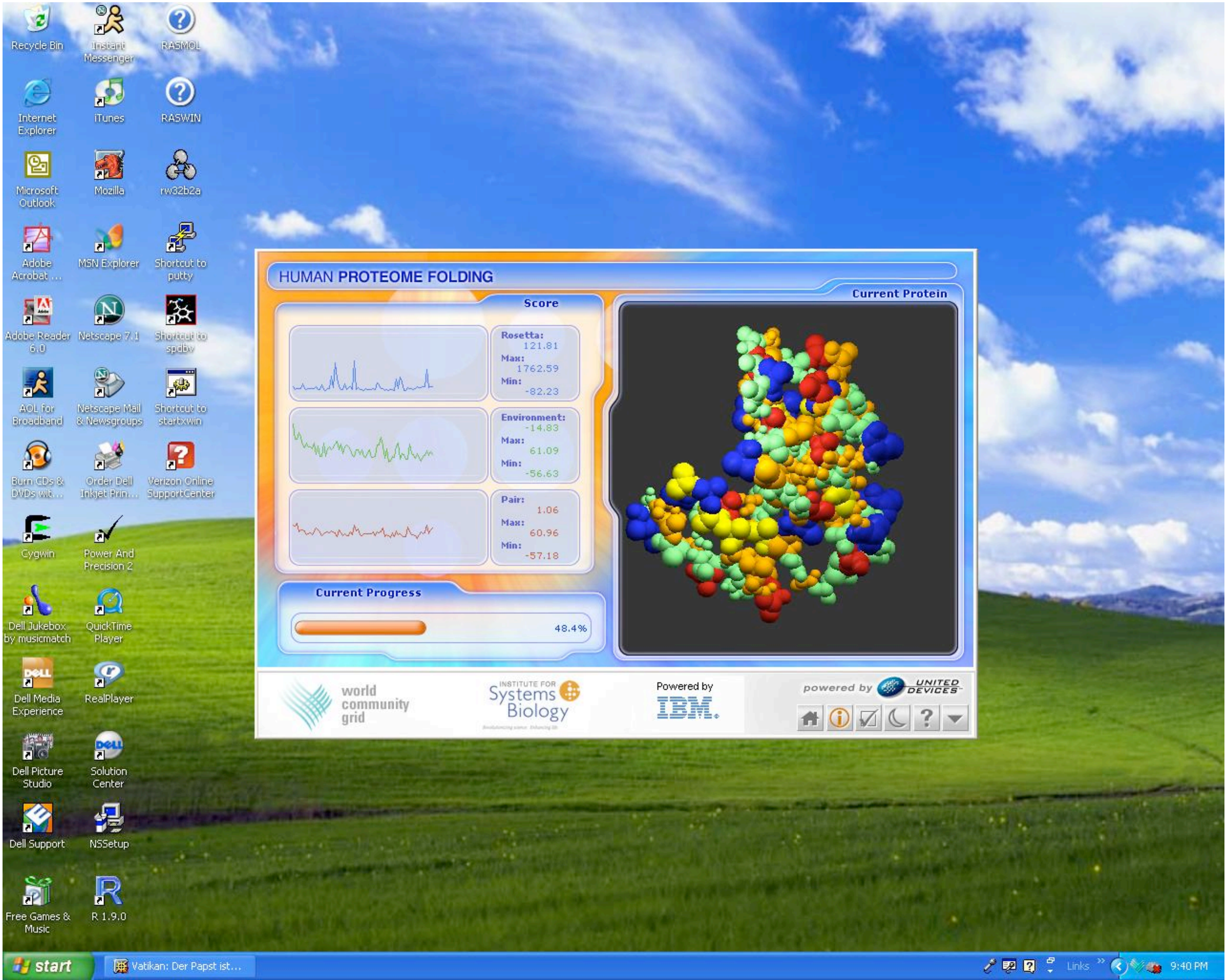
- Applications of Rosetta:
  - Functional annotation of genes.
  - Novel protein design.

Protein Design



Protein Design





HUMAN PROTEOME FOLDING

Score

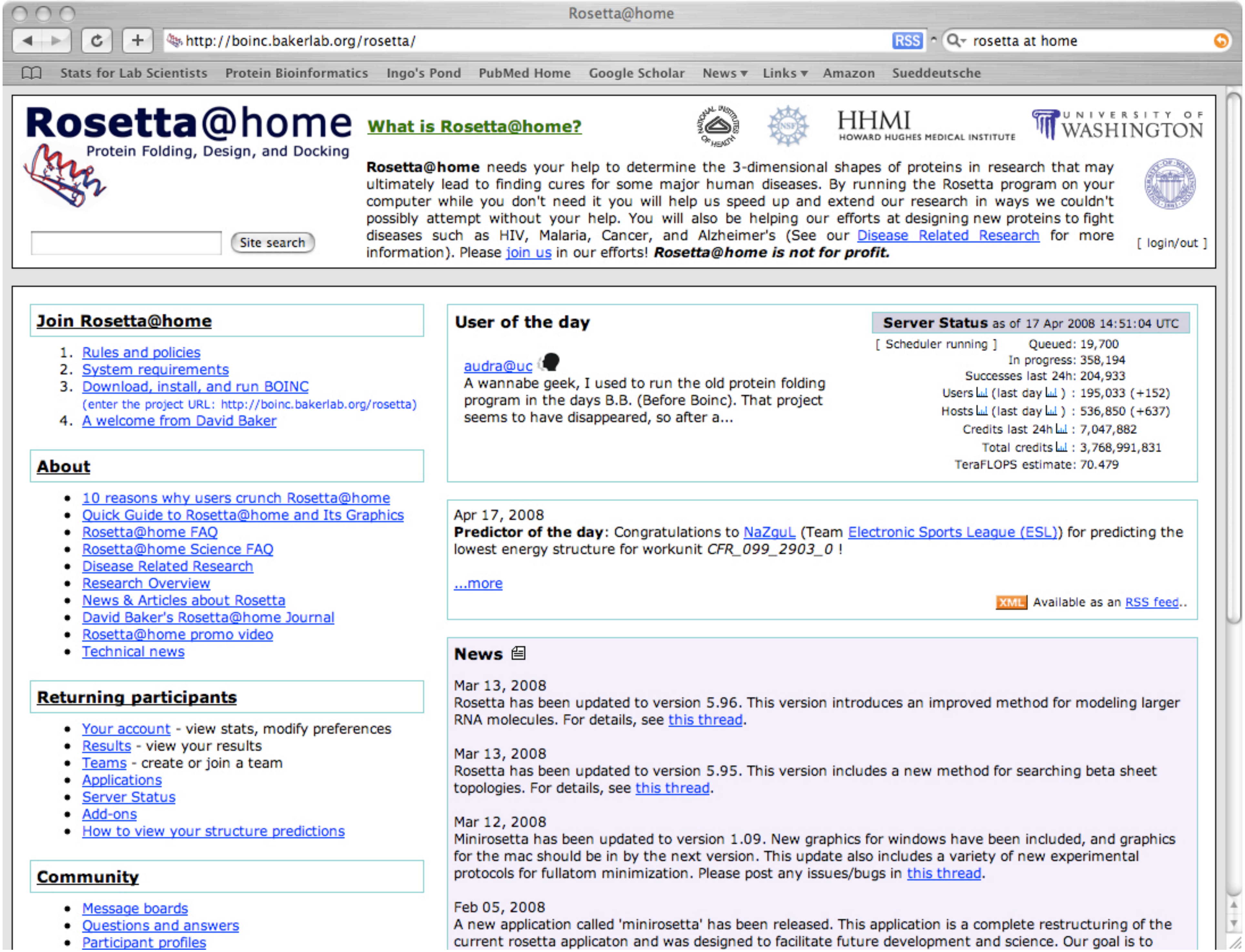
Rosetta: 121.81
Max: 1762.59
Min: -82.23

Environment: -14.83
Max: 61.09
Min: -56.63

Pair: 1.06
Max: 60.96
Min: -57.18

Current Progress 48.4%

Current Protein



Rosetta@home

Protein Folding, Design, and Docking

Site search

What is Rosetta@home?

Rosetta@home needs your help to determine the 3-dimensional shapes of proteins in research that may ultimately lead to finding cures for some major human diseases. By running the Rosetta program on your computer while you don't need it you will help us speed up and extend our research in ways we couldn't possibly attempt without your help. You will also be helping our efforts at designing new proteins to fight diseases such as HIV, Malaria, Cancer, and Alzheimer's (See our Disease Related Research for more information). Please join us in our efforts! ***Rosetta@home is not for profit.***

[login/out]

Join Rosetta@home

1. Rules and policies
2. System requirements
3. Download, install, and run BOINC
   (enter the project URL: http://boinc.bakerlab.org/rosetta)
4. A welcome from David Baker

About

- 10 reasons why users crunch Rosetta@home
- Quick Guide to Rosetta@home and Its Graphics
- Rosetta@home FAQ
- Rosetta@home Science FAQ
- Disease Related Research
- Research Overview
- News & Articles about Rosetta
- David Baker's Rosetta@home Journal
- Rosetta@home promo video
- Technical news

Returning participants

- Your account - view stats, modify preferences
- Results - view your results
- Teams - create or join a team
- Applications
- Server Status
- Add-ons
- How to view your structure predictions

Community

- Message boards
- Questions and answers
- Participant profiles

User of the day

audra@uc

A wannabe geek, I used to run the old protein folding program in the days B.B. (Before Boinc). That project seems to have disappeared, so after a...

Server Status as of 17 Apr 2008 14:51:04 UTC

[Scheduler running] Queued: 19,700

In progress: 358,194

Successes last 24h: 204,933

Users (last day) : 195,033 (+152)

Hosts (last day) : 536,850 (+637)

Credits last 24h : 7,047,882

Total credits : 3,768,991,831

TeraFLOPS estimate: 70.479

Apr 17, 2008

Predictor of the day: Congratulations to NaZquL (Team Electronic Sports League (ESL)) for predicting the lowest energy structure for workunit *CFR_099_2903_0* !

...more

XML Available as an RSS feed..

News

Mar 13, 2008

Rosetta has been updated to version 5.96. This version introduces an improved method for modeling larger RNA molecules. For details, see this thread.

Mar 13, 2008

Rosetta has been updated to version 5.95. This version includes a new method for searching beta sheet topologies. For details, see this thread.

Mar 12, 2008

Minirosetta has been updated to version 1.09. New graphics for windows have been included, and graphics for the mac should be in by the next version. This update also includes a variety of new experimental protocols for fullatom minimization. Please post any issues/bugs in this thread.

Feb 05, 2008

A new application called 'minirosetta' has been released. This application is a complete restructuring of the current rosetta applicaton and was designed to facilitate future development and science. Our goal is to