

- Introduction
- Stockholm Bioinformatics Center
- Functional annotations from domain architecture
- Expanding domain annotations
- Evolution of multi-domain proteins
- Personal

Gunnar von Heijne, SU, director

Arne Elofsson, SU, vice director	Jens Lagergren, KTH
----------------------------------	---------------------

Bob MacCallum, SU	Bengt Sennblad, KI
-------------------	--------------------

Lars Arvestad, KTH	Erik Lindahl, SU
--------------------	------------------

<i>Per Kraulis, SU</i>	<i>Bengt Persson, KI</i>
------------------------	--------------------------

<i>David Liberles, SU</i>	<i>Jesper Tegner, KTH</i>
---------------------------	---------------------------

- Funded with a 40 MSEK grant in 1999
- Research Center
- Currently seven groups
- Research areas:
 - Protein Structure (AE,BM,EL)
 - Sequence analysis (GvH,AE,BP)
 - Molecular evolution (JL,DL,BS)
 - System Biology (PK,JT)

- SBC and PDC
- Linux based
- 350 (single CPU) nodes
- AFS+Kerberos
- EASY batch queuing system

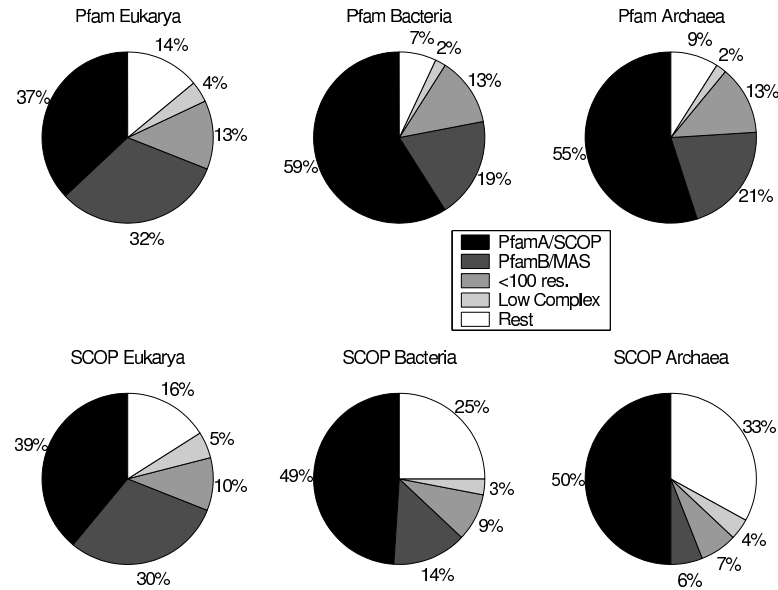
Domain family databases

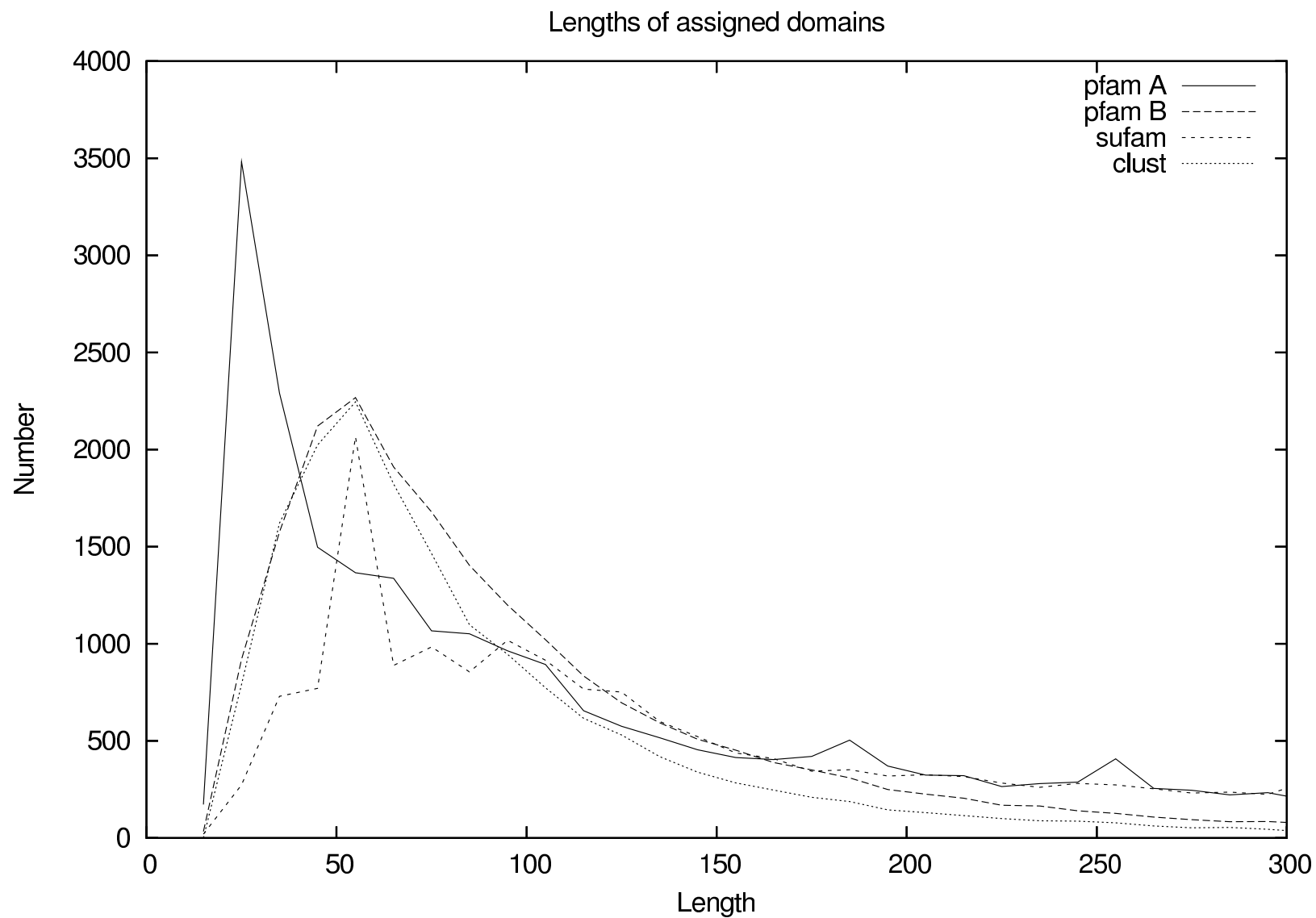
- SCOP and CATH *manual* from structures.
- FSSP *automatic* from structures.
- Pfam and Smart *manual* from sequences.
- PRODOM *automatic* from sequences.

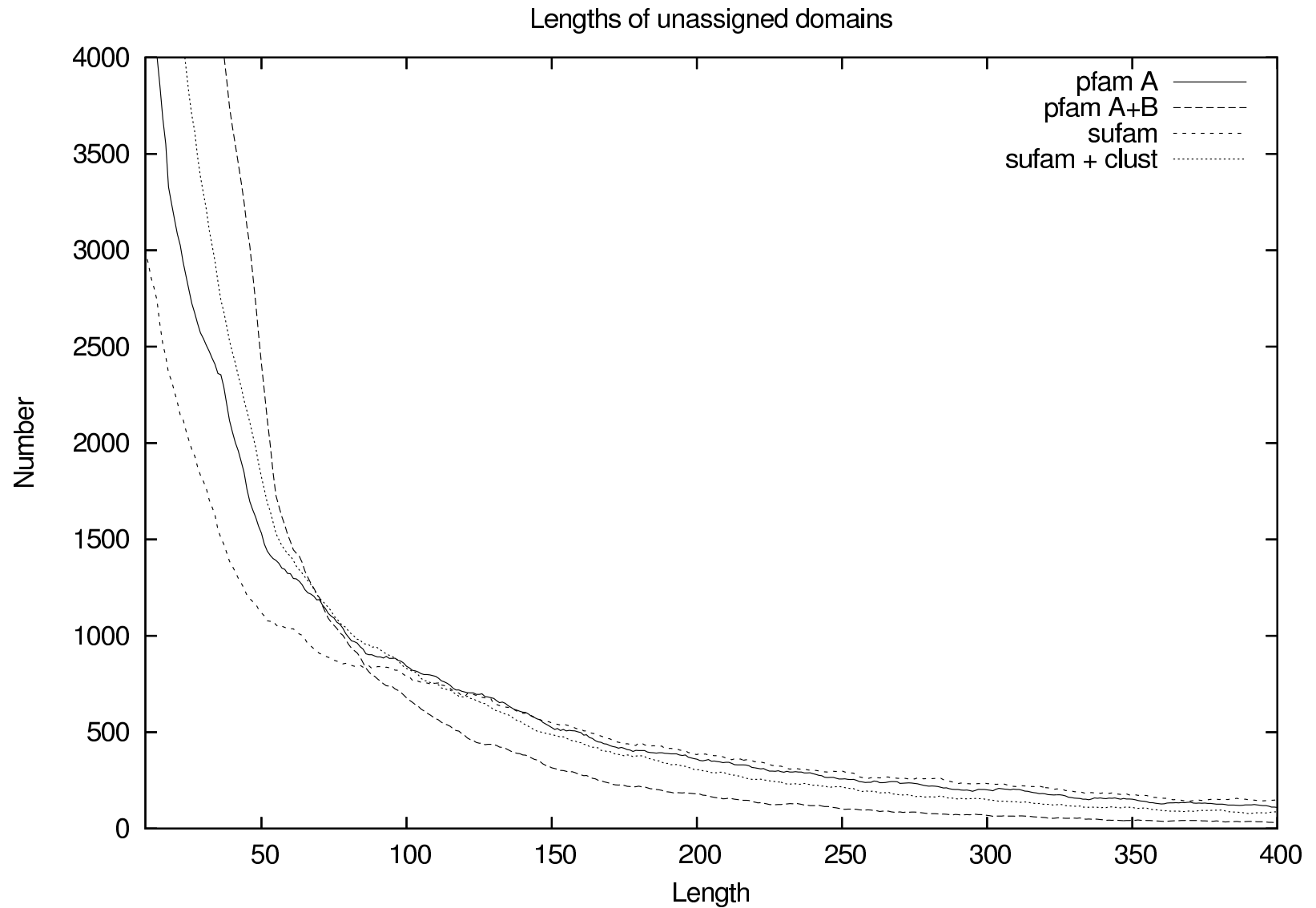


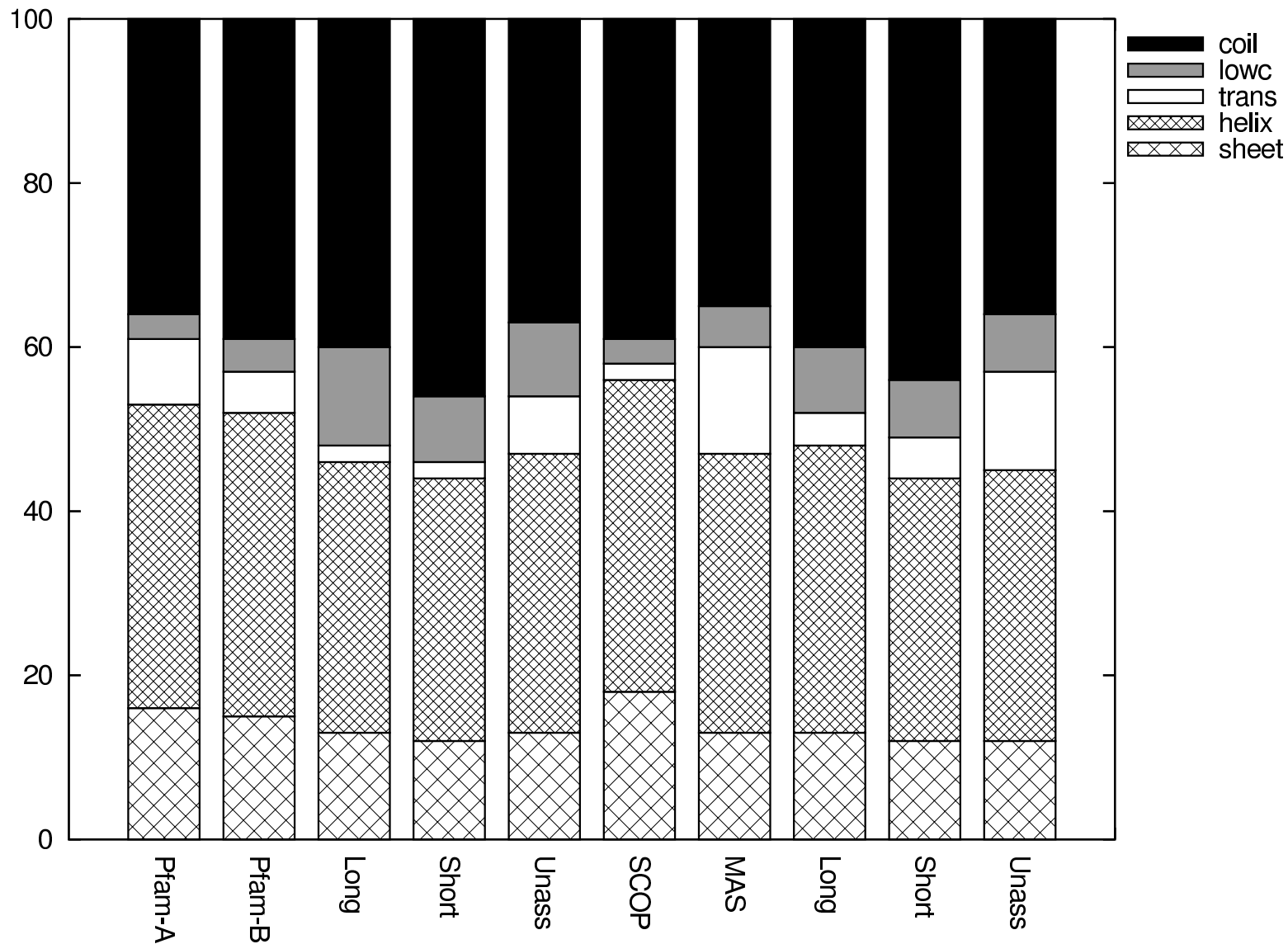
- Phylogenetic Profiles (Gaasterland, 1998, Eisenberg, Bork)
 - Inclusion of phylogenetic specie information (Liberles, 2002)
 - Inverted profiles
- Gene Fusion (Enright, Marcotte, Bork)
 - Phylogeny of multi-domain proteins

- Clustering unassigned regions
- Analyzing unassigned ORPHan regions

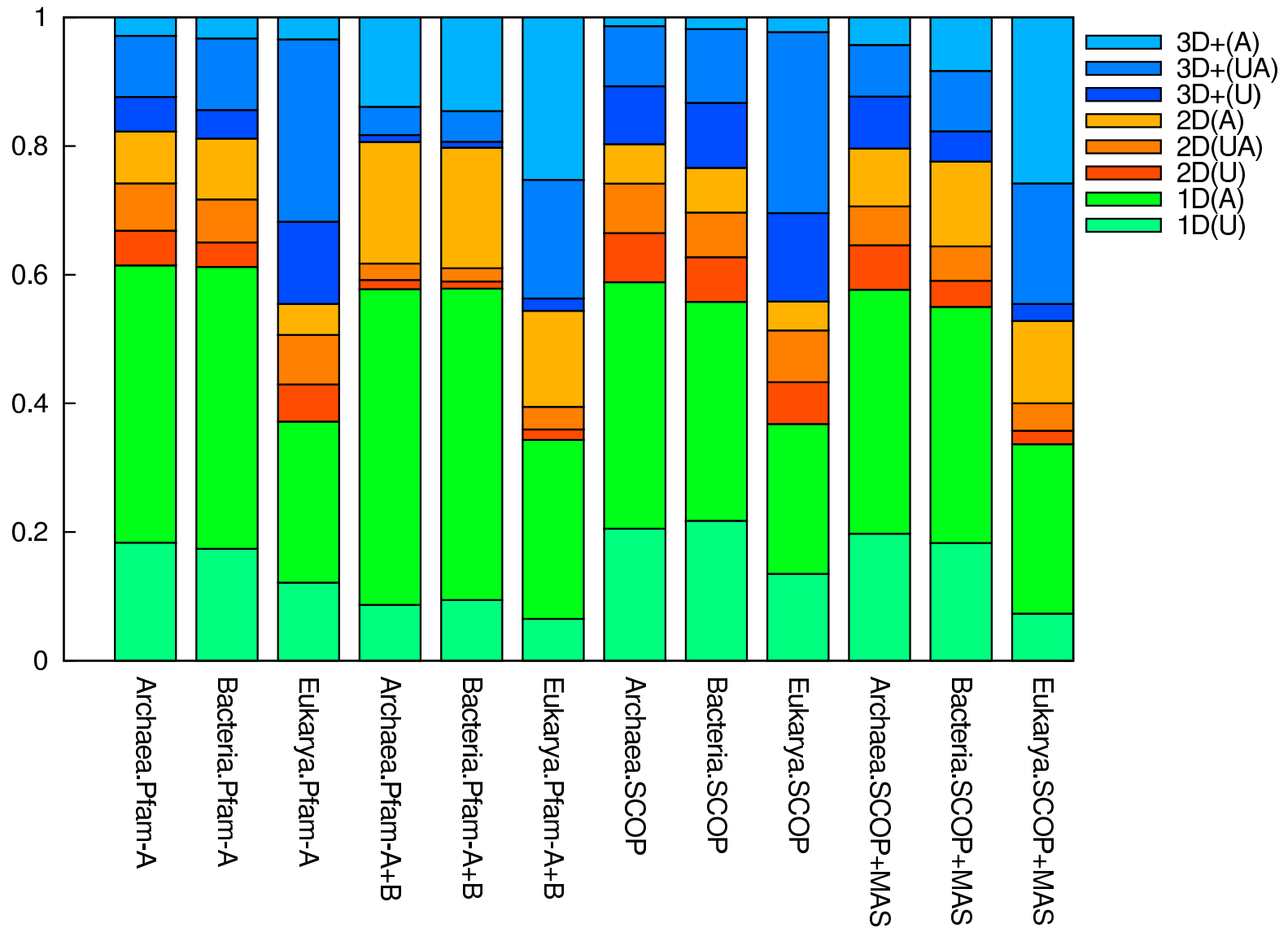




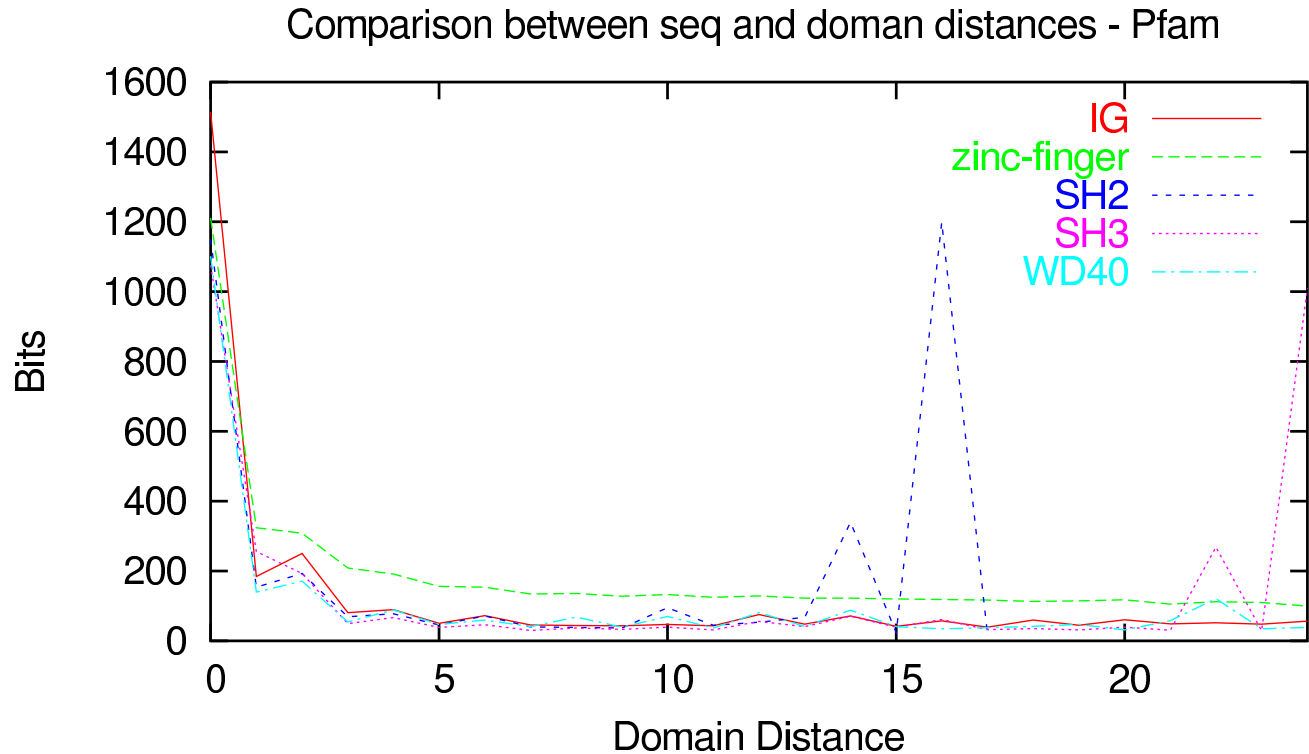




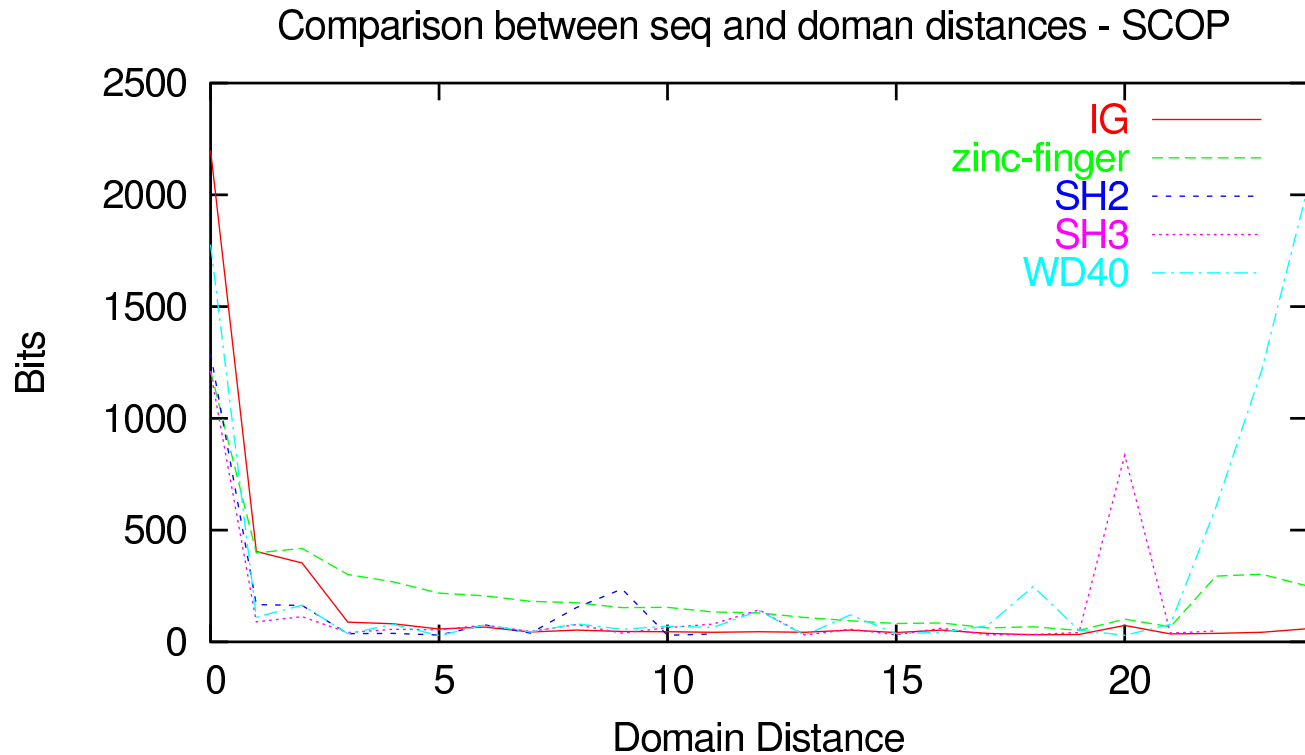
Fraction of multi-domain proteins



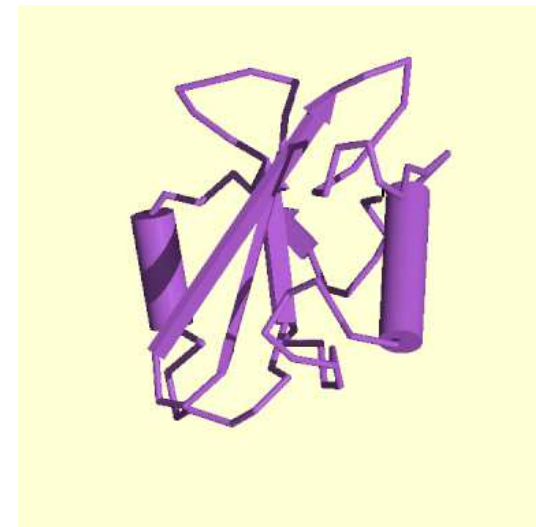
Proteins that are more similar in sequence are more similar in domain-distance



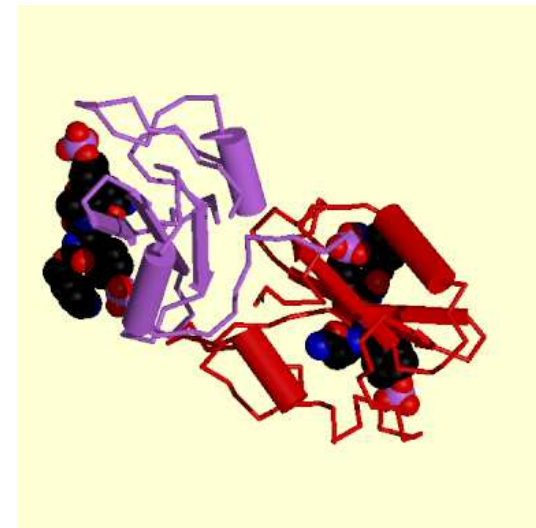
Proteins that are more similar in sequence are more similar in domain-distance



- protein domain of about 100 amino-acid residues
- SH2 domains function as regulatory modules of intracellular signaling cascades
- Interacts with high affinity to phosphotyrosine-containing target peptides
- 814 in Pfam 10.0

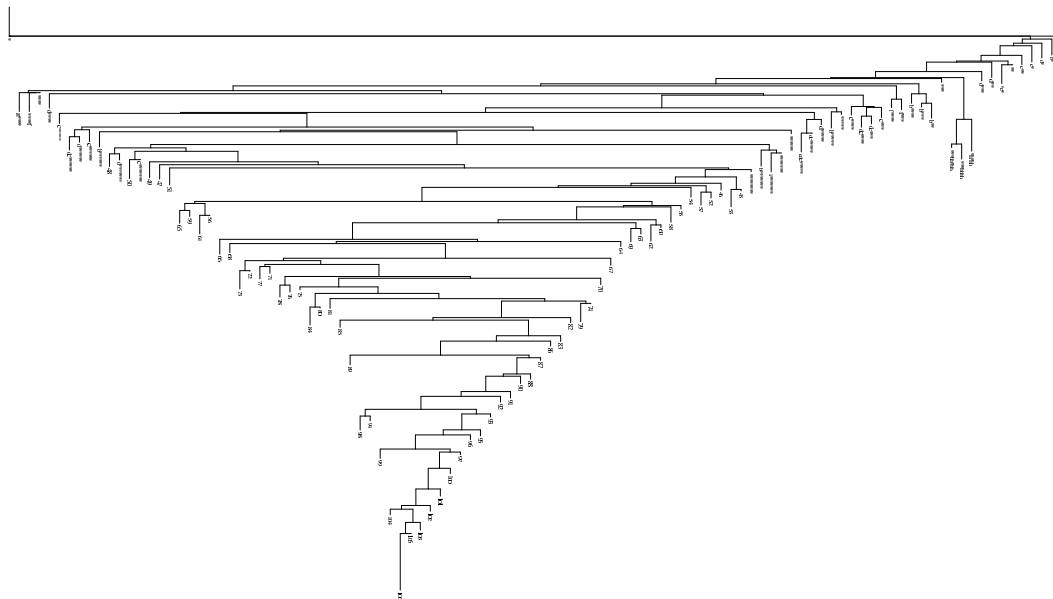


- small protein modules containing approximately 50 amino acid residues
- indicative of a protein involved in signal transduction related to cytoskeletal organization
- 1610 in Pfam 10.0



- nucleic acid-binding protein structures
- A zinc finger domain is composed of 25 to 30 amino-acid residues
- 2 conserved Cys and 2 conserved His residues in a C-2-C-12-H-3-H type motif.
- 23742 in Pfam 10.0





1. Methods for the reliable delineation of domain architecture (DA): Month 6
2. Methods for prediction of functional features from DA: Month 12
3. Results on benchmarks of methods for predicting interactions from DA:
Month 18
4. Improved methods for prediction of function and interactions from DA from a larger set of proteins with reliable annotations: Month 24
5. Function prediction from DA for non annotated protein regions: Month 36

1. WP1 Protein families and subfamilies
Improved alignments methods using profile–profile HMMs
2. WP2 Reliability scores for functional annotations
Extended experience in benchmarking
3. WP3 Functional annotations from domain architecture
See above
4. WP4 Structure-based prediction of function and interactions
Improved structure prediction methods
5. WP5 Meta-server, Context-based function prediction
Will provide methods
6. WP6 Meta-server, Integrated function prediction
Will provide methods

- Johannes Frey-Skött, Alternative splicing
- Åsa Björklund, Domains
- Diana Ekman, Domains