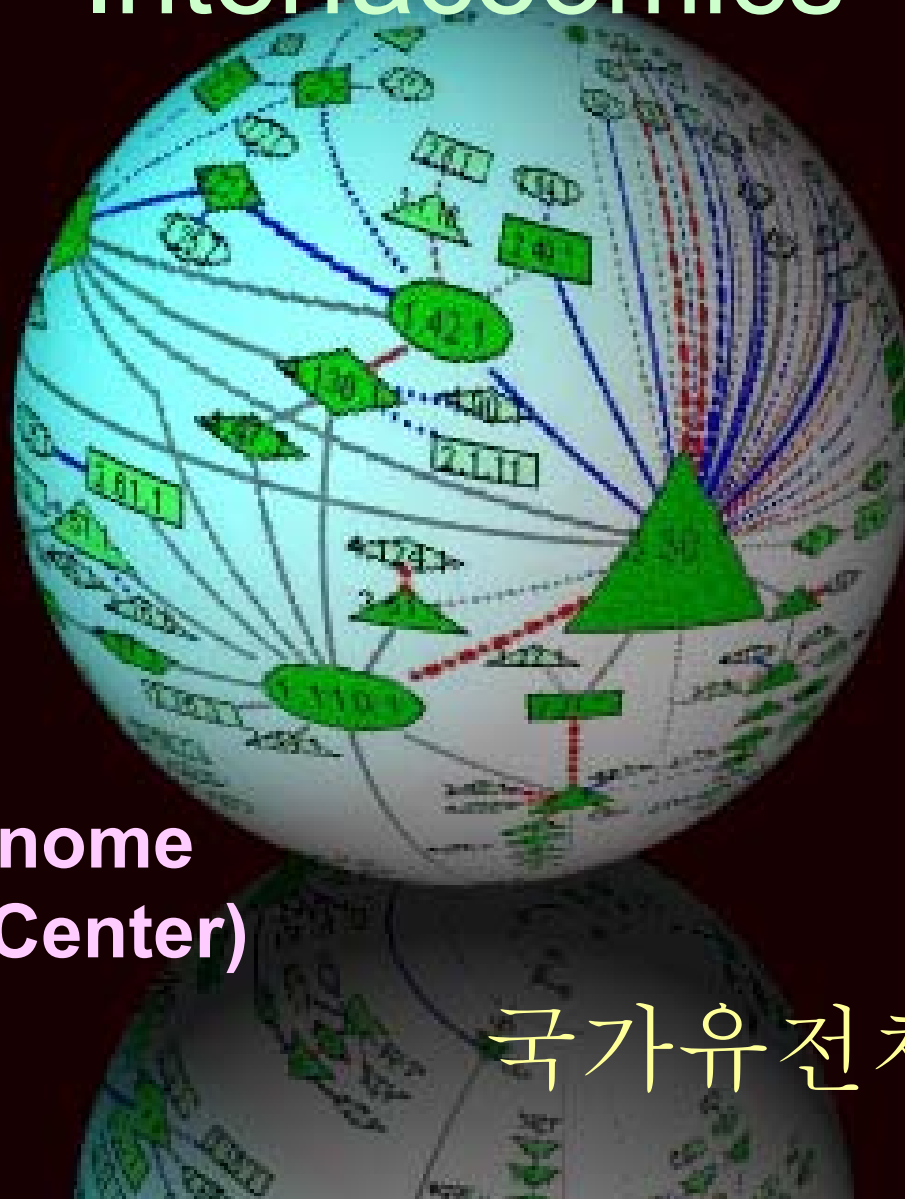


Protein Interactomics & Interfaceomics



박종화

Jong Bhak


NGIC

(National Genome
Information Center)

KRIBB

국가유전체정보센터

Acknowledgement

- People who do science because science is interesting and fun.
- People who support science and technology by paying tax.
- MOST of Korea which funds NGIC. 
- KRIBB for funding and support.



한국생명공학연구원 대전광역시 유성구 어은동 52번지 305-333
Korea Research Institute of Bioscience and Biotechnology(KRIBB)
Copy right by Korea Research Institute of Bioscience and Biotechnology (KRIBB)



Content

- Bioinformatics
- Protein Informatics
- Protein Interactomics & Interfaceomics

Brief History of Bioinformatics

- **Darwin:** A theoretical biologist
- **Mendel:** A theoretical prediction and validation
- **Perutz and Kendrew:** structural biologists
- **Crick and Watson:** DNA modellers
- **Crick, Brenner,** : Codon
- **Sanger:** DNA sequencing: Genomics
- **Sanger:** Protein sequencing
- **Sanger:** **Proteomics**
- **Lesk:** Visualization of proteins
- **Needleman & Wunch:** Computer algorithms
- **Southern:** Hybridization → Functional Genomics
- **Tim Berners-Lee** 1990: HTTPD → Internet

Darwin (evolution)
Mendel (genetic analysis)

DNA modelling (Watson & Crick, 1953)

Hemoglobin과 Myoglobin의 structure 비교 (Max Perutz, John Kentreu)

: structure 와 sequence 양대영역 개념정리

structure

sequence

Bio-Computation Methodology
(Chris Sander, Arther Lesk, etc)

Dynamic program을 이용한
Sequence comparison app module
(Niedleman & Bunsche)

DNA → codon → anticodon
→ peptide 개념정리

X-174 genome (F.sanger)
: full genome sequencing

DB construction (Gen Bank, PIR, ...)

DNA chip &
Microarray technology

Southern blot
Hybridization methology

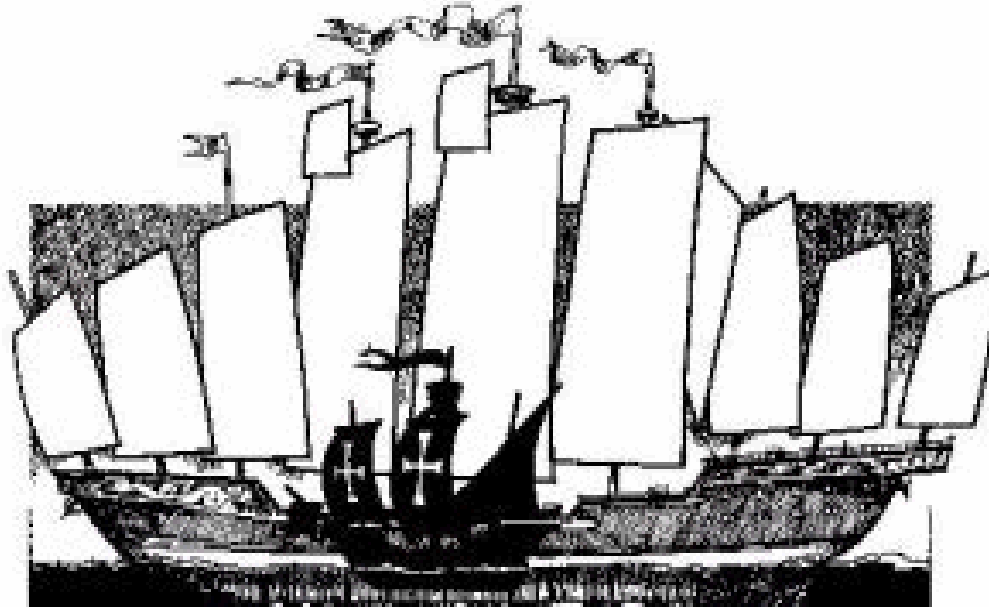
Bioinformatics Expanded

- Structural genomics
- Comparative genomics
- Sequencing
- Functional genomics
- SNP
- Proteomics (Mass spec. protein chip)
- DataBases
- Computational analysis Methodology

Functions

Computer → INTERNET

Bioinformatics People ?



- **Explorers** of 1300-1400
- **Charting** the unknown territory of **Biological Life.**



Gangnido 1402



World Traffic Map



Long Definition of Bioinformatics

- Bioinformatics is a **discipline of science** that analyses, seeks understanding and models the **whole life** as an **information processing phenomenon** utilizing **energy** with methods from **philosophy, mathematics and computer science** using **biological experimental data.**

-- Jong Bhak

A short definition of Bioinformatics

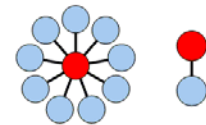
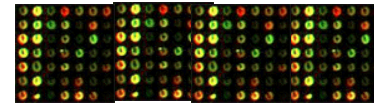
- **Biology is bioinformatics and bioinformatics is biology.**

-- Jong Bhak

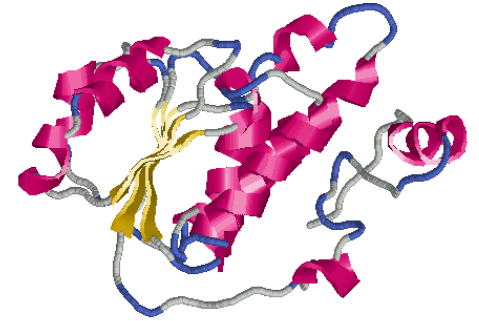
Research Domains of Bioinformatics

- Sequence
- Structure
- Expression
- **Interaction**
- Function
- Literature

```
>M19217  
TTTTTTGGCGTGGCGGCGTGGCTGGCGGT  
TCGGCAGTGAGTGCAAGGATCACCATGATT  
CTTCAGAGACTCTCAGGTTGTCCTCTGCT
```

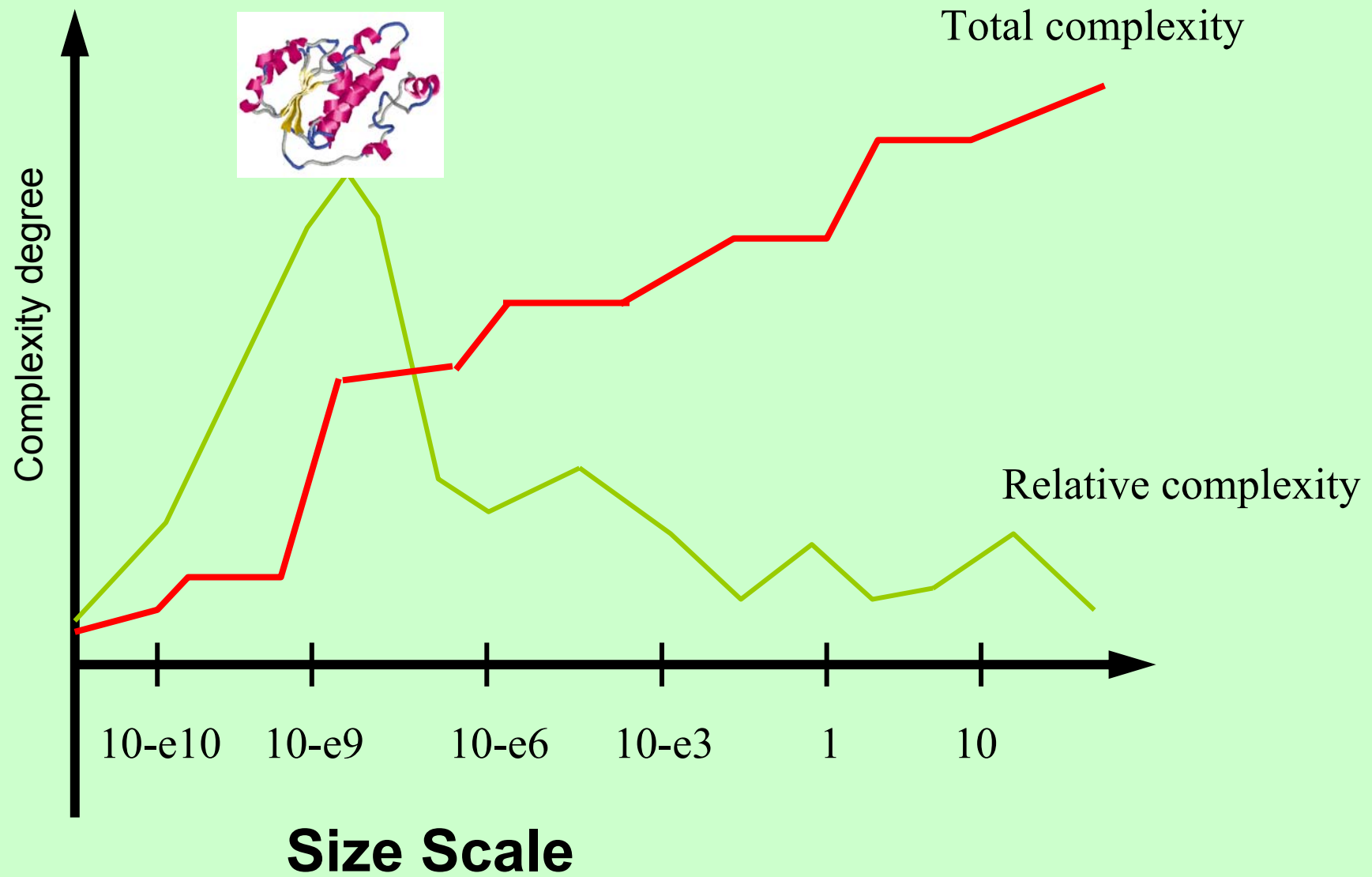


What is Protein?



- The most informative object in the biological universe.
- Protein level is efficient to work with.
 - ➔ a Naturally distinct unit. So, favoured by bioinformatic computing.

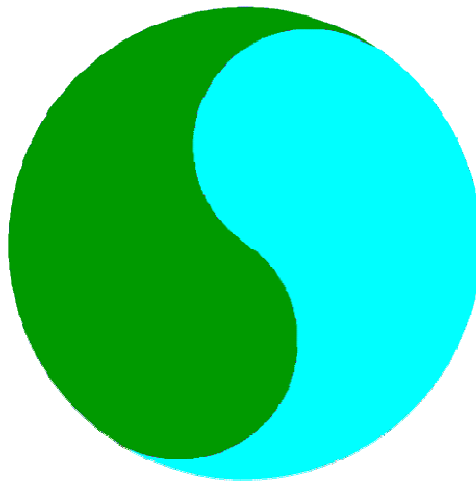
BioComplexity Density Chart



Why Interactome and Interactomics?

- Context in Biology is always the problem of ‘interactions’

Dan Bolser & Jong Bhak, *Genome informatics*



Why **Protein** Intearctome ?

- **Protein** Intearctome can give us the most valuable insights and information about biological functions in cells.

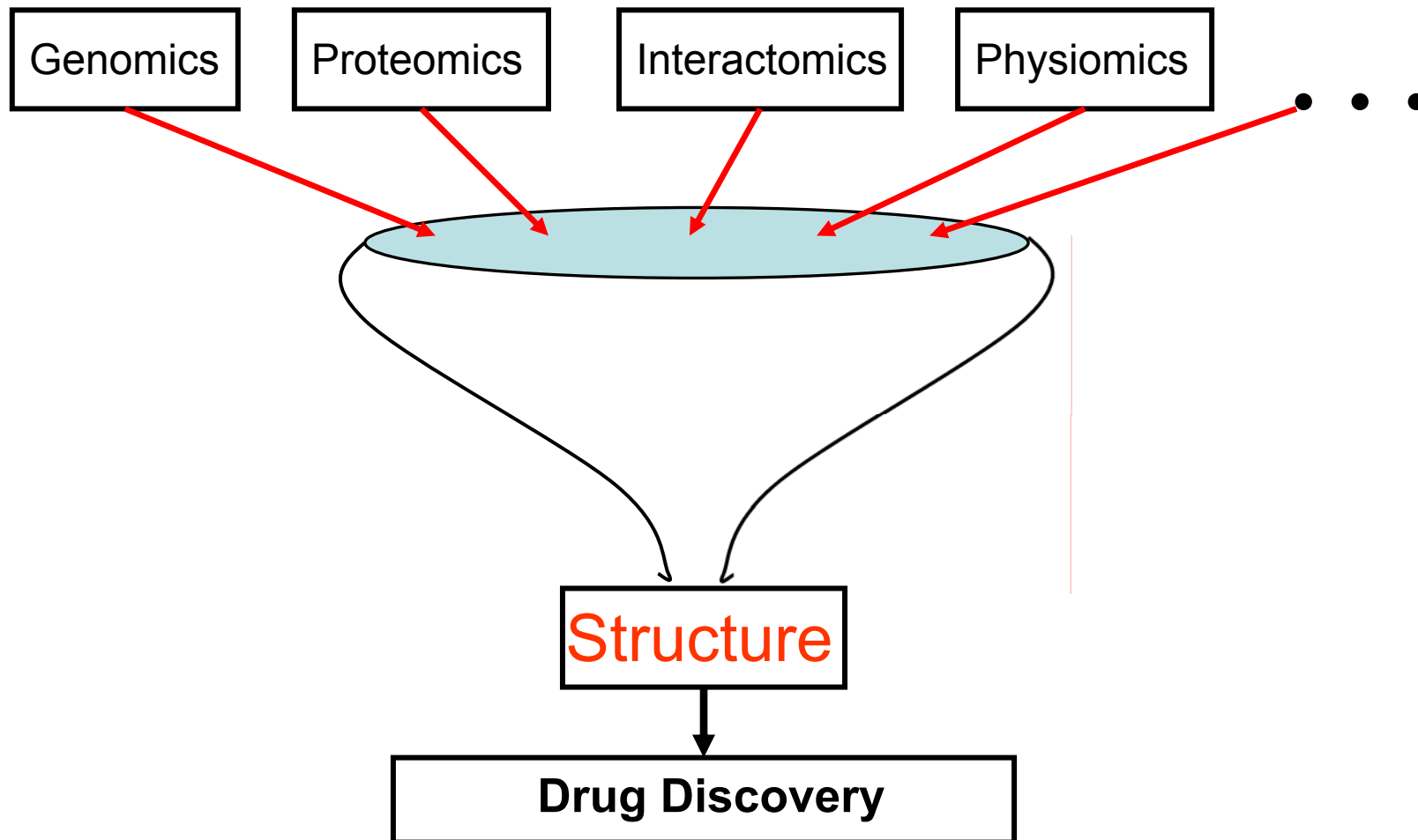
(it is the best map you can have in biology)

<http://interactome.org>

Why Structural Interactomics ?

- Most definite.
- Most informative
- Most accurate
- Directly connected to drug discovery.
- **Structural** → certain well-known forces and physical rules are considered.
- Most fun and beautiful to work with ?


Funnel of Biological Drug Discovery



The unit of Interaction

Protein Structural Domains

Structural Classification of Proteins































Class: All alpha proteins

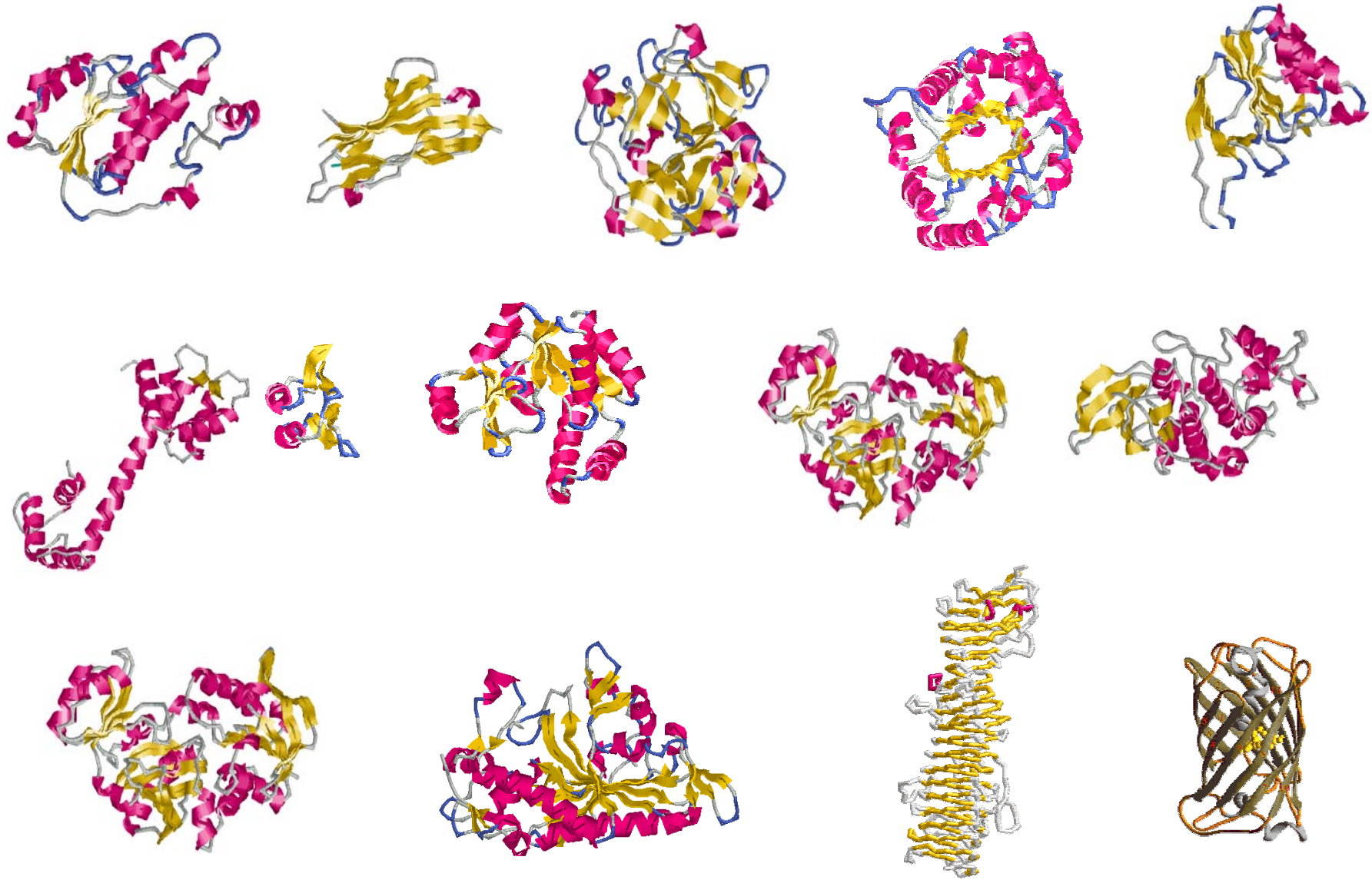
Lineage:

1. Root: [scop](#)
2. Class: [All alpha proteins](#)

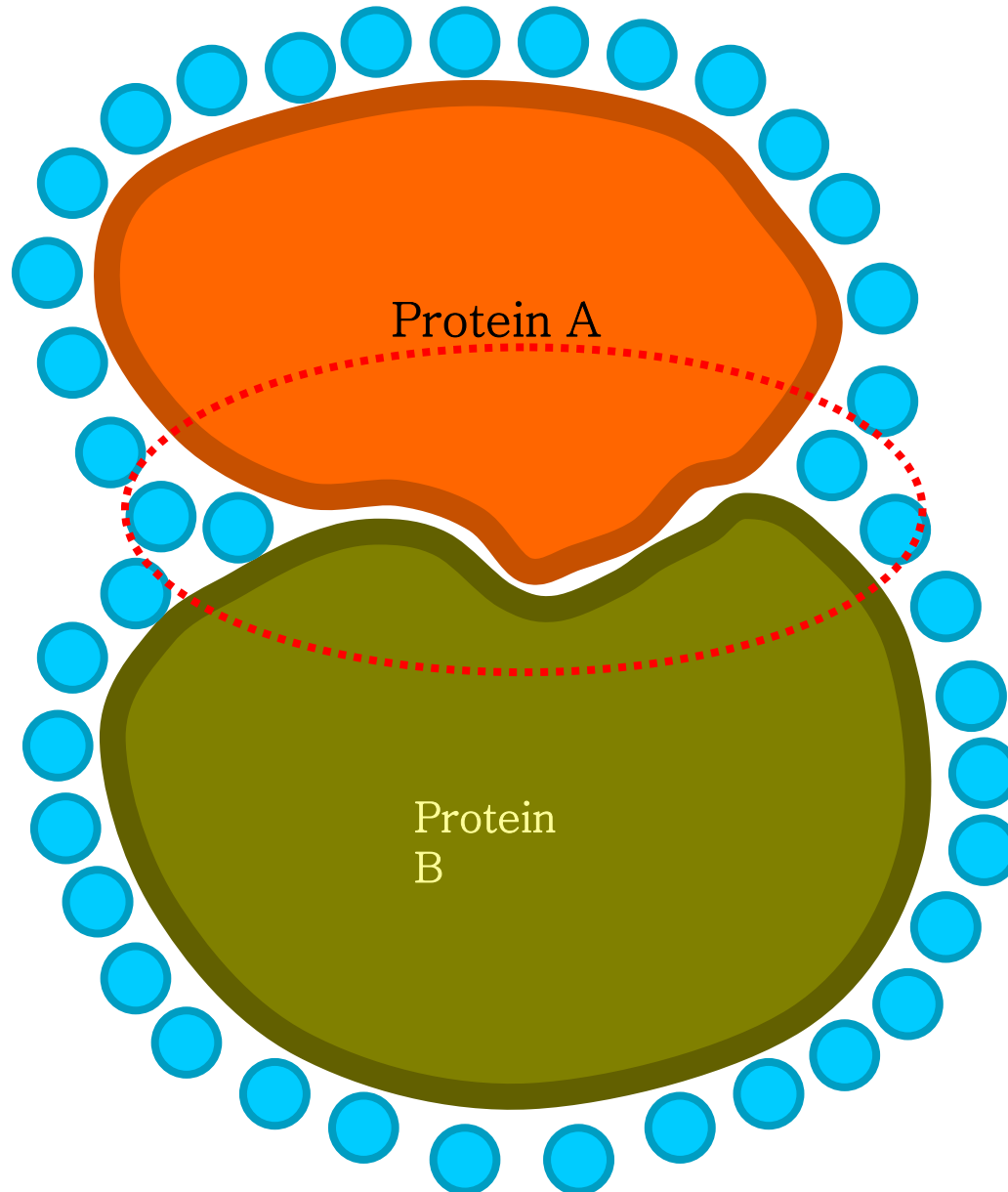
Folds:

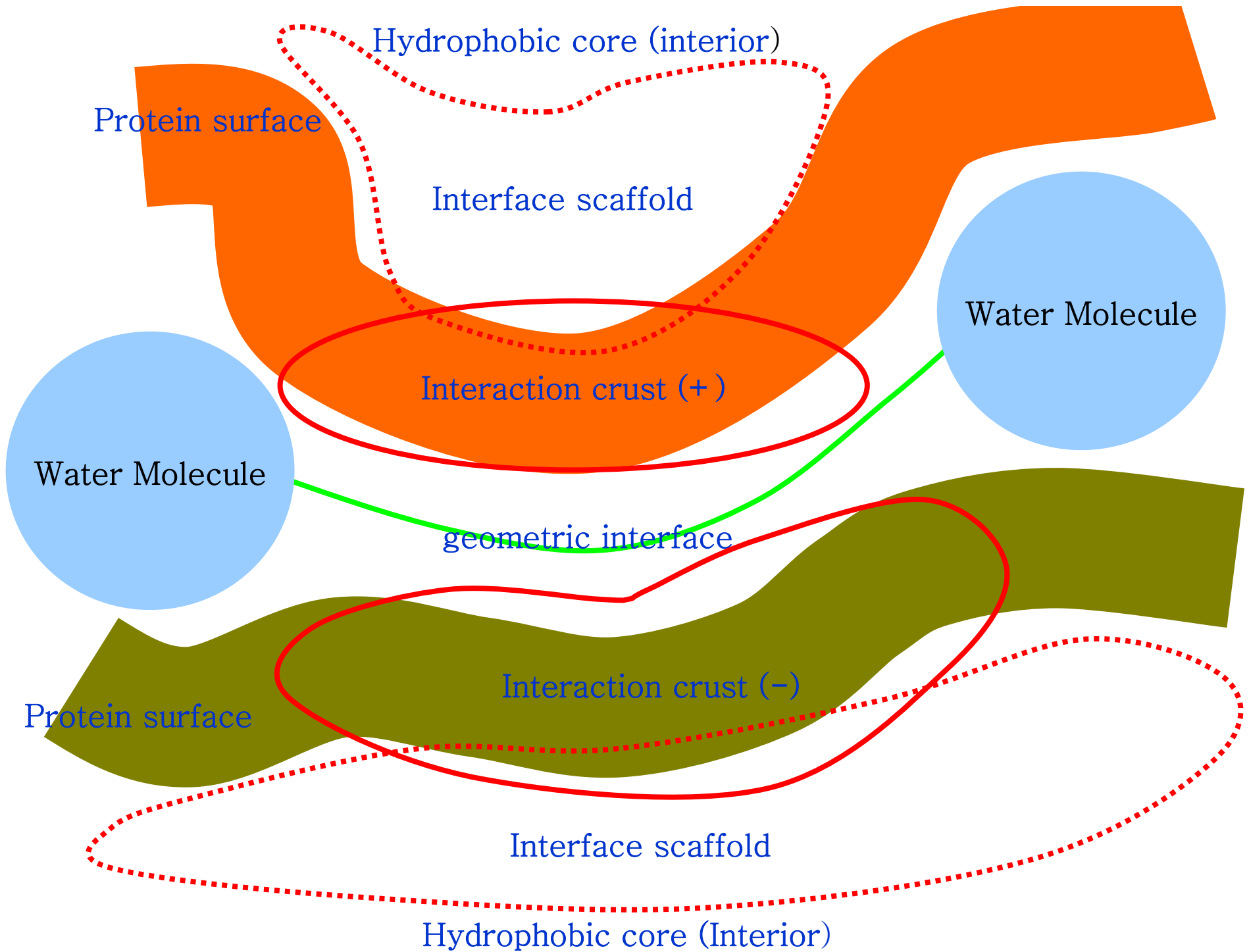
1. [Globin-like](#) (2)
core: 6 helices; folded leaf, partly opened
 1. [Globin-like](#) (4)  
 2. [alpha-helical ferredoxin](#) (2)  
contains two Fe4-S4 clusters
2. [Long alpha-hairpin](#) (11)
2 helices; antiparallel hairpin, left-handed twist
 1. [GreA transcript cleavage protein, N-terminal domain](#) (1)  
 2. [Ribosomal protein L29 \(L29p\)](#) (1)  
 3. [Chaperone J-domain](#) (1)  
 4. [Theta subunit of DNA polymerase III](#) (1)  
fold is similar to the chaperone J-domain ?
 5. [Prefoldin](#) (1)  
 6. [Effector domain of the protein kinase pkn/prk1](#) (1)  
 7. [tRNA-binding arm](#) (4)  
formerly a class II aminoacyl-tRNA synthetase N-domain
 8. [Eukaryotic DNA topoisomerase I, dispensable insert domain](#) (1)  
 9. [C-terminal UvrC-binding domain of UvrB](#) (1)  
 10. [Epsilon subunit of F1F0-ATP synthase C-terminal domain](#) (1)  
 11. [Fe,Mn superoxide dismutase \(SOD\), N-terminal domain](#) (1)  
3. [Type I dockerin domain](#) (1)
tandem repeat of two calcium-binding loop-helix motifs, distinct from the EF-hand
 1. [Type I dockerin domain](#) (1)  

What are Structural Domains ?



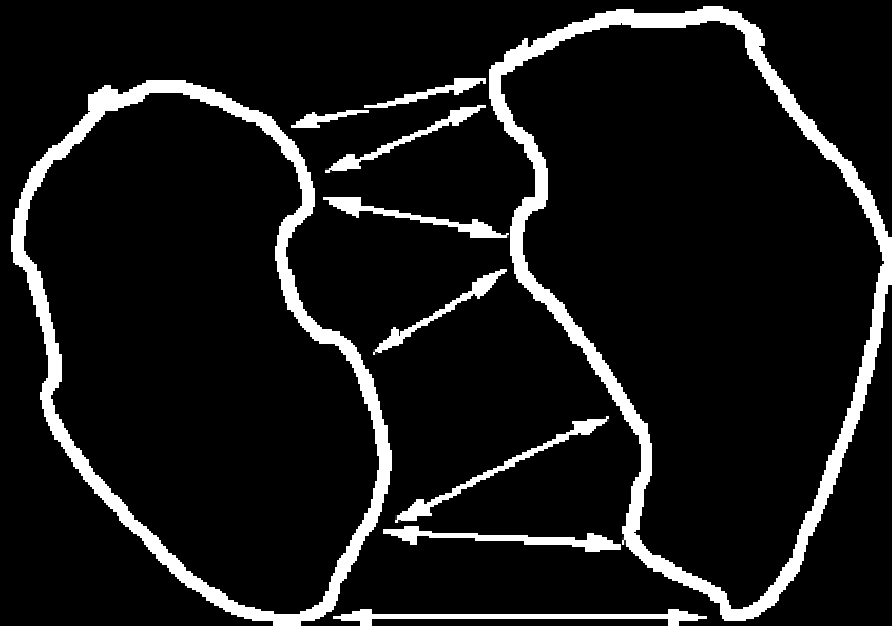
Interaction among domains





Detecting (defining) Domain Interaction

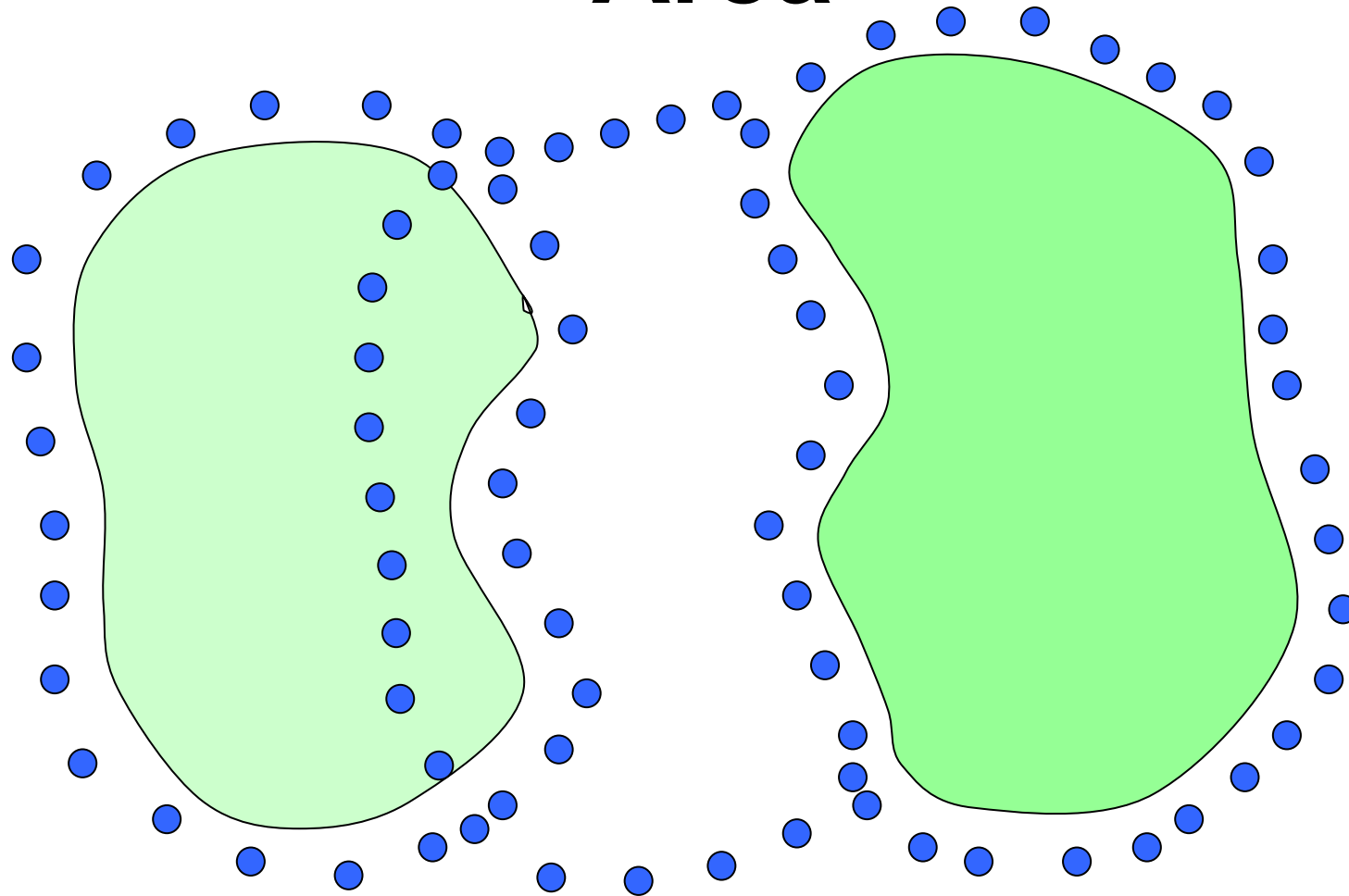
~5 Ångstrom?



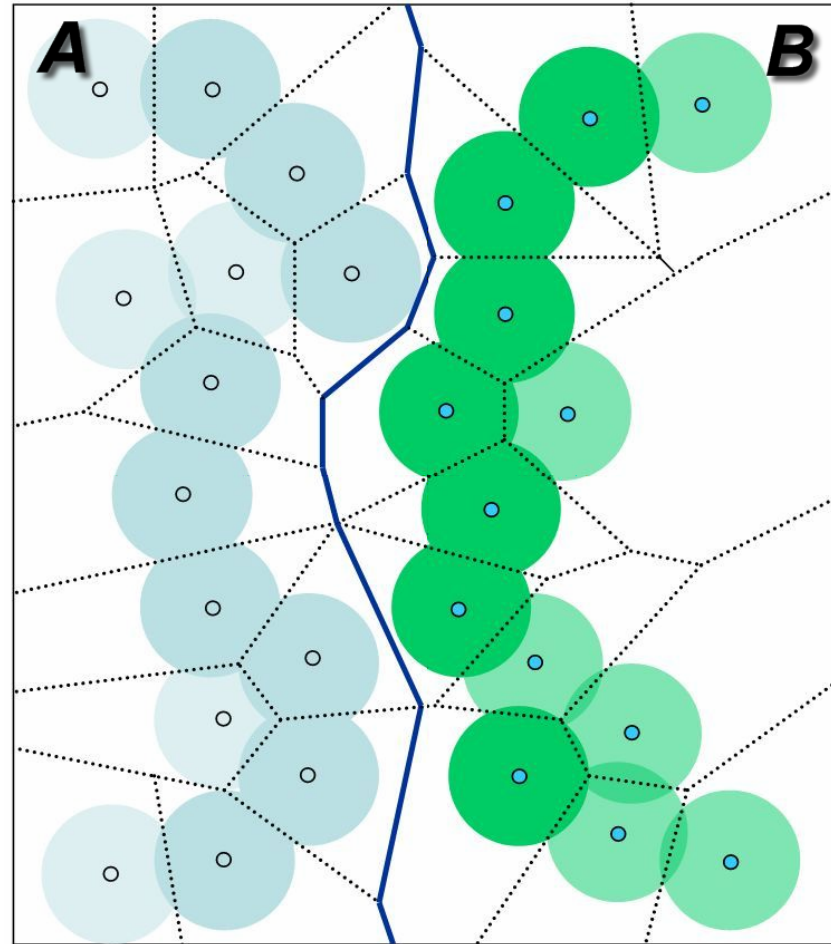
Domain A

Domain B

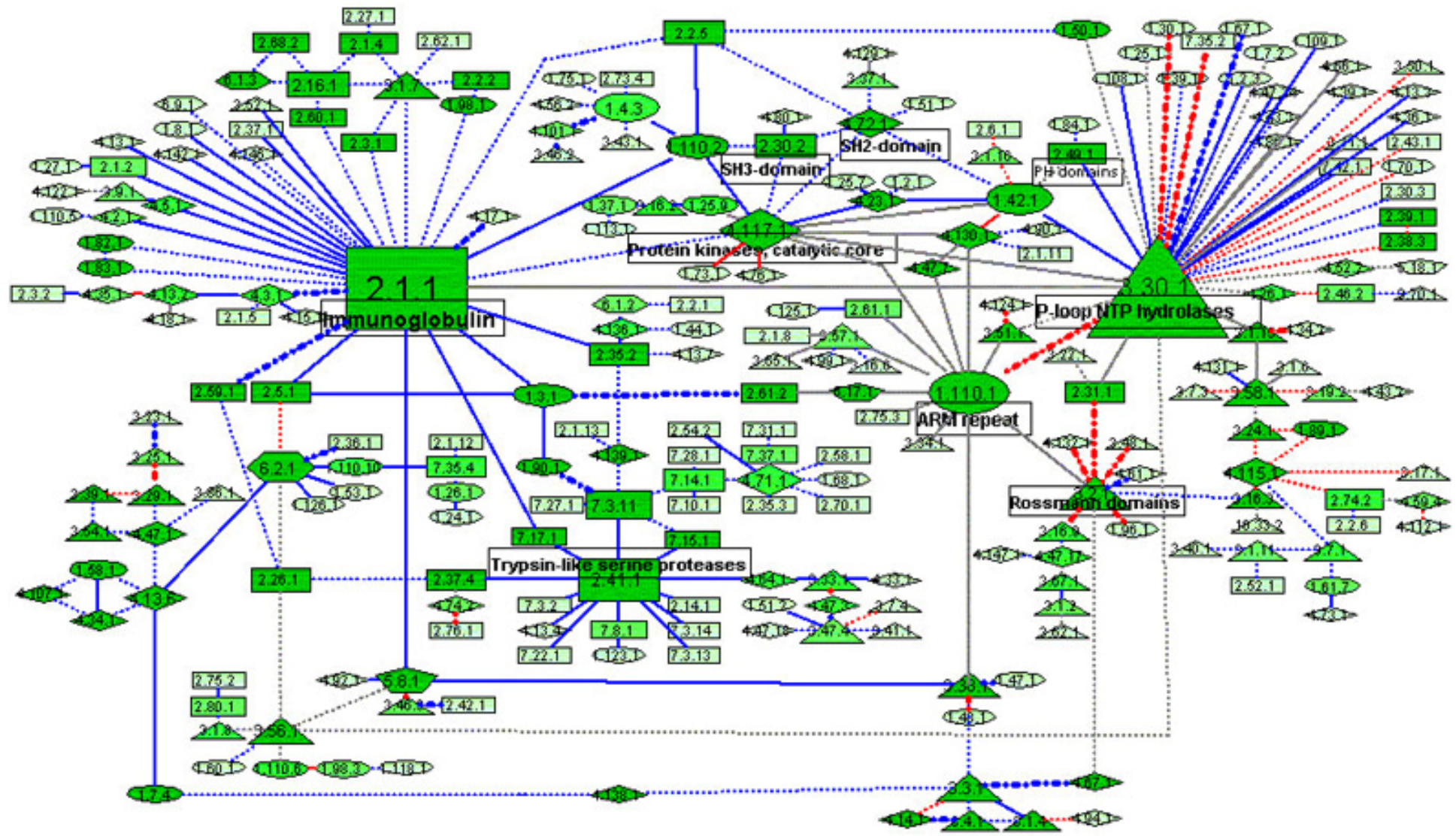
ASA: Accessible Surface Area



Voronoi Diagram



A protein domain interaction map



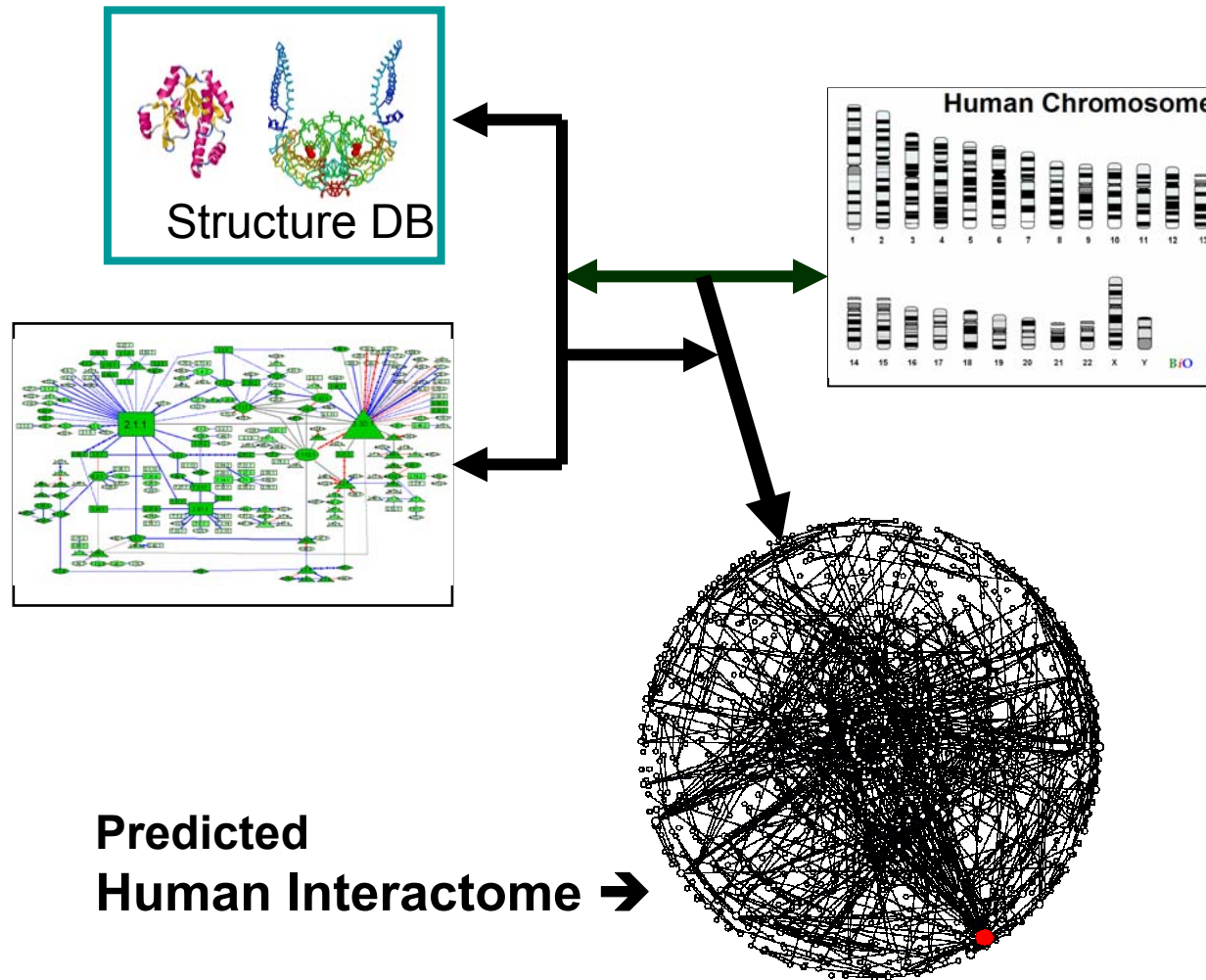
What can we do with Interactomes?

- The main academic goal of Interactomics
 - Mapping all the molecular interactions in cells

Computational Human Interactome

- Mapping all the **protein-protein** and **protein-ligand** interactions using computational and experimental methods.

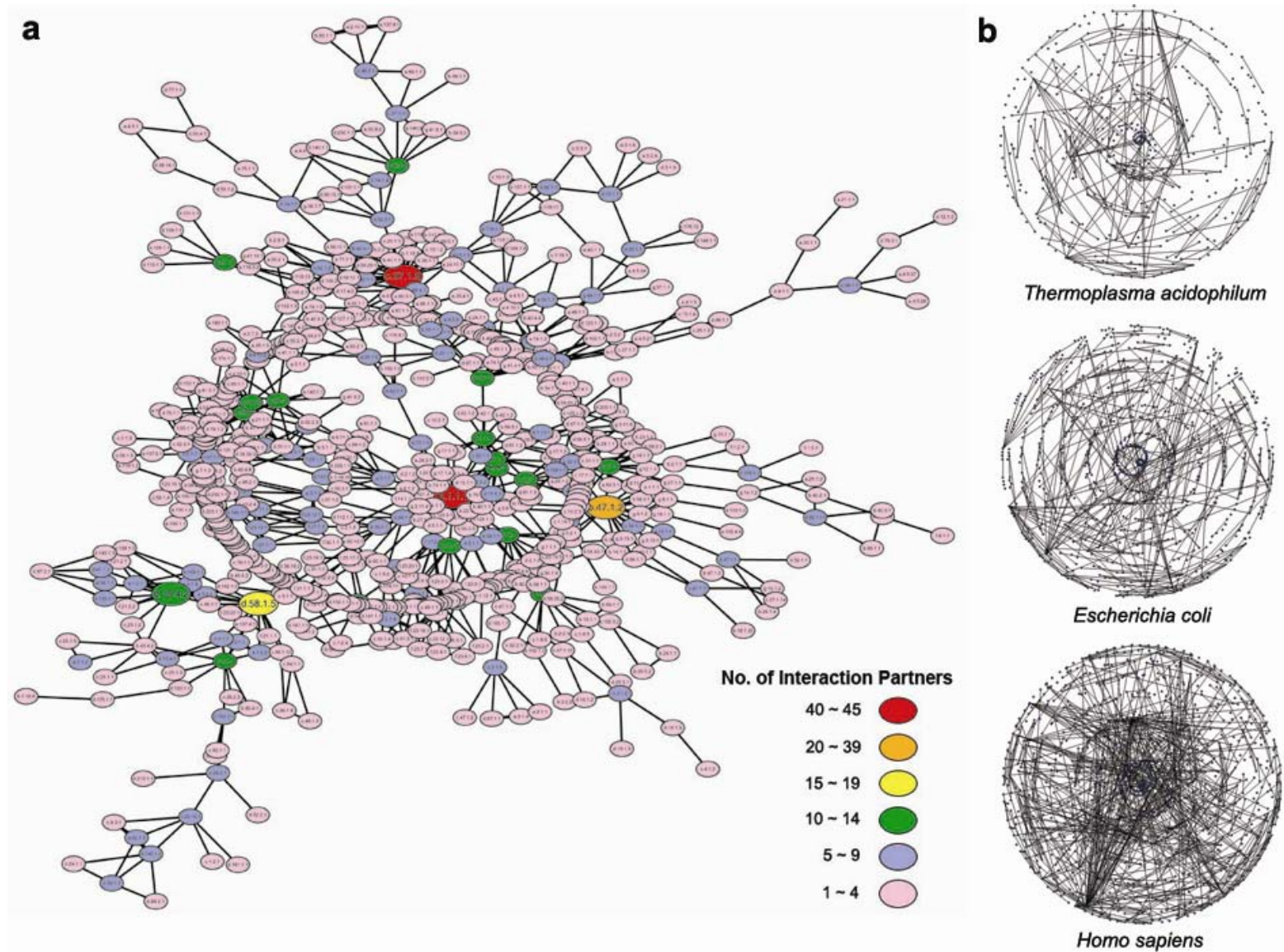
Practical Steps of Complete Human Interactome



Comparative Interactomics

- A large scale analysis of comparing interactomes of species to understand the evolution and function of cells.

Global view of protein family interaction networks for 146 genomes



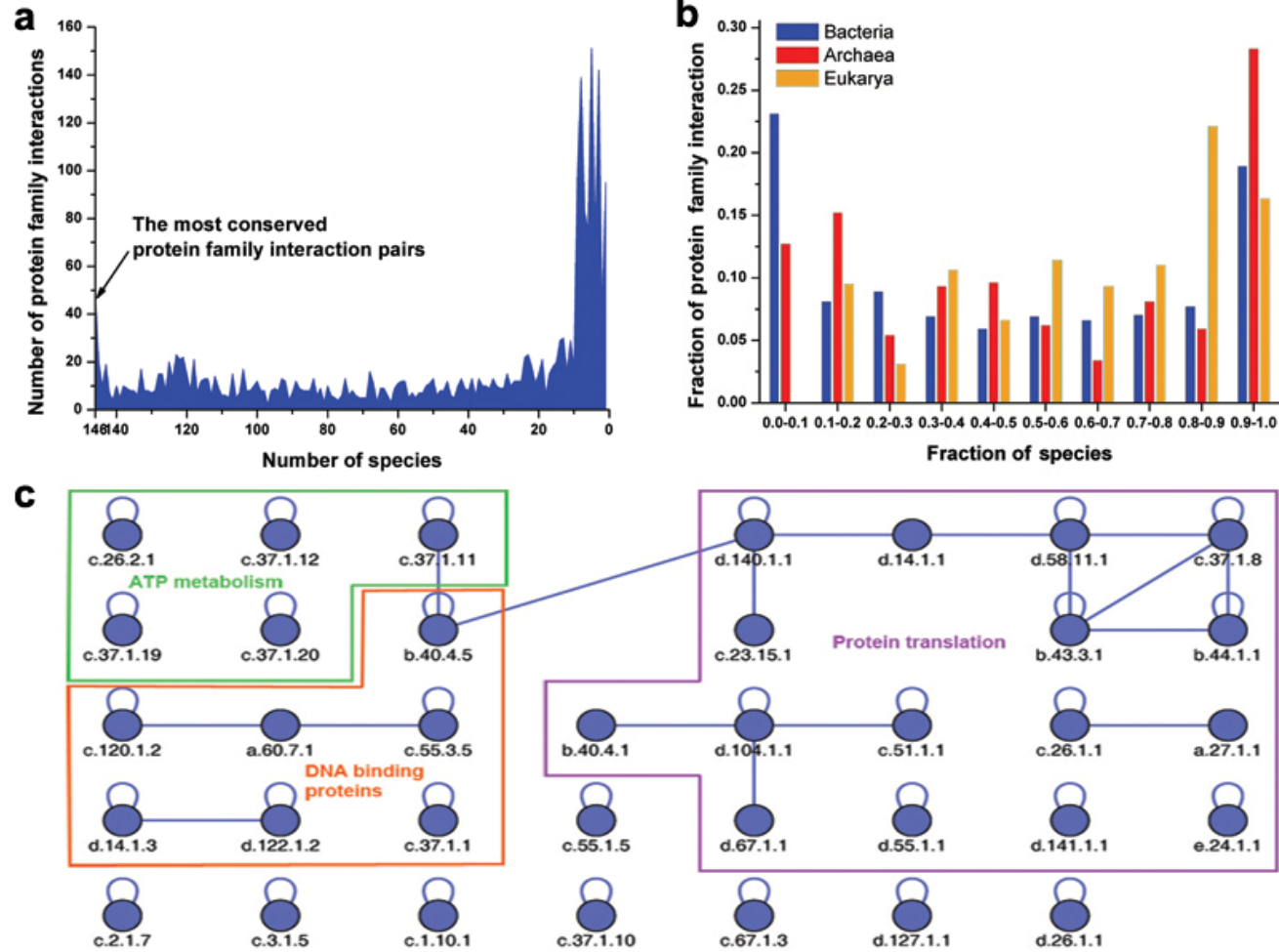
Comparative Interactomics

- We found that all 146 interactomes are scale-free networks, and they share a core protein-interactome comprising 36 protein families related to indispensable functions in a cell.
- *Daeui Park, Semin Lee, Dan Bolser, Michael Schroeder, Michael Lappe, Donghoon Oh, & Jong Bhak. Bioinformatics, 2005.*

Eukarya and Prokarya

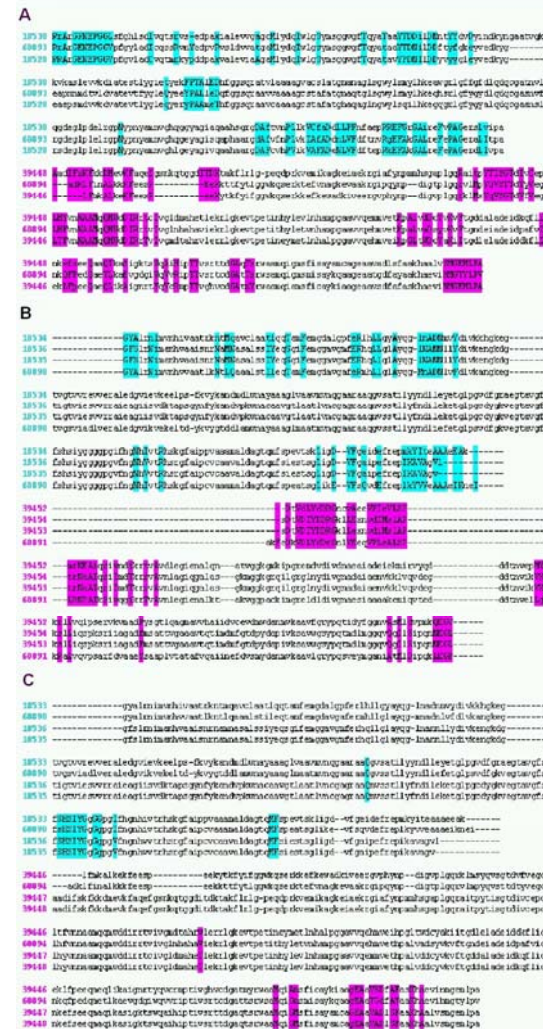
- We found two fundamental differences among prokaryotic and eukaryotic interactomes:
 - 1) eukarya had significantly more hub families than archaea and bacteria, and
 - 2) certain special hub families determined the shapes of the eukaryotic interactomes.

Core Interactome



Interfaceomics

- <http://interpare.net/>
- Analyzing the actual interaction interfaces of protein-protein interactions.
- Interfaceome: the whole set of interacting molecular pairs in cells.



Wankyung Kim, Dan M. Bolser and Jong Park, *Bioinformatics*, 2004, 20(7):1138-1150.

InterFacer





The screenshot shows a web browser window with the address bar containing <http://www.interfacer.org/>. The browser's address bar also shows other sites like zilla.org, mozillaZine, and mozdev.org. The main content area features the InterFacer logo, which consists of the word "Inter" in red and "Facer" in blue, with a red and blue background. To the right of the logo is the text "Front Page" and a search input field with a "Go" button. Below the logo and search field are several icons for navigation and utility, including a home icon, a search icon, a list icon, a refresh icon, a settings icon, a magnifying glass icon, a mail icon, a RSS icon, and a print icon. The main content area is titled "InterFacer #" and contains a paragraph of text: "InterFacer is a molecular viewer specialized to visualize domain-domain interfaces. Its advantage over previous programs is that it can retrieve domain classification information directly from the viewer utilizing protein structural classification database. It is a java application that can be used in any platform." Below this paragraph is a list of links: Downloads, Tutorial, Manual, InterFacerApplet, Installation, ScreenShots, VersionHistory, History, Links, Logo, FAQ, and Acknowledgement. At the bottom of the page, there are several icons: a home icon, a search icon, a list icon, a refresh icon, a settings icon, a magnifying glass icon, a mail icon, a RSS icon, a print icon, a W3C XHTML 1.0 icon, a W3C CSS icon, and a Powered by MoniWiki icon.

http://www.interfacer.org/

zilla.org mozillaZine mozdev.org



InterFacer Front Page

대문 | [찾기](#) | [목록](#) | [바뀐글](#) | [환경설정](#) |        

InterFacer

[InterFacer](#) is a molecular viewer specialized to visualize domain-domain interfaces. Its advantage over previous programs is that it can retrieve domain classification information directly from the viewer utilizing protein structural classification database. It is a java application that can be used in any platform.

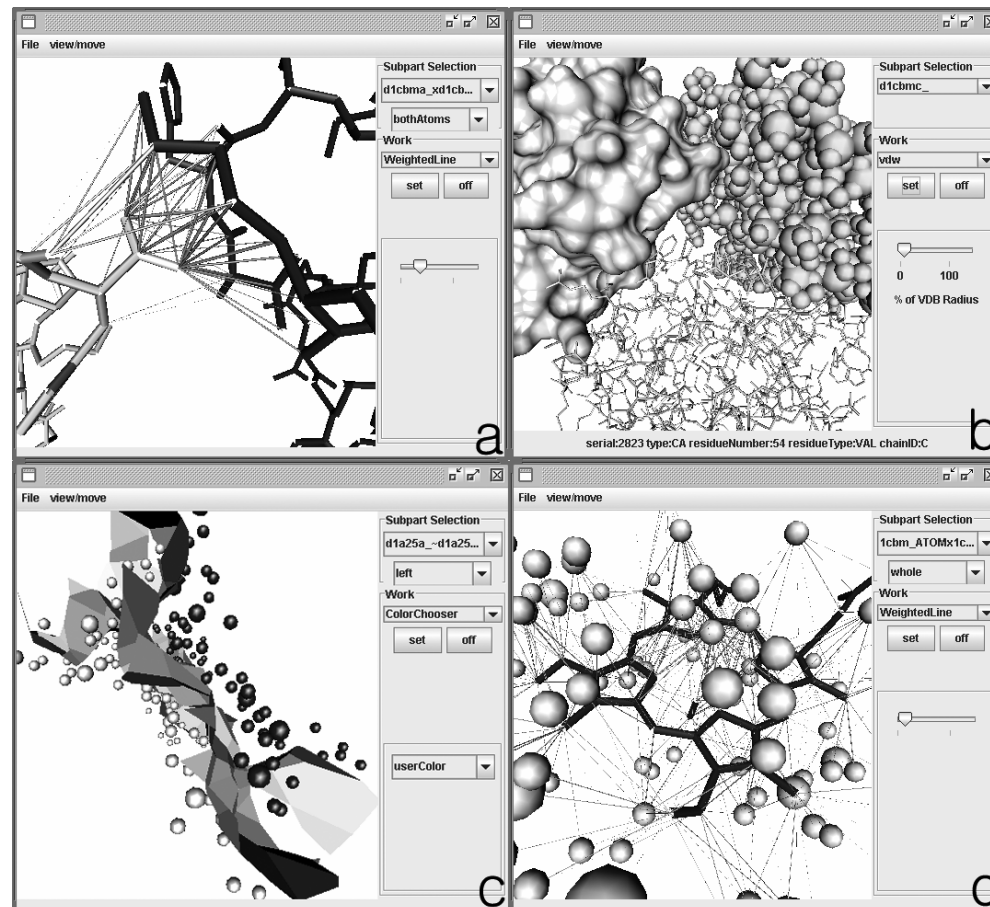
- [Downloads](#)
- [Tutorial](#)
- [Manual](#)
- [InterFacerApplet](#)
- [Installation](#)
- [ScreenShots](#)
- [VersionHistory](#)
- [History](#)
- [Links](#)
- [Logo](#)
- [FAQ](#)
- [Acknowledgement](#)

[고치기](#) | [찾기](#) | [쪽 지우기](#) | [비슷한 쪽](#) |   

Choi HanSol

InterFacer

- <http://www.interfacer.org/>



Graphic User Interface of InterFacer.

The image displays the InterFacer GUI, which is used for protein structure analysis. The central window shows a 3D representation of a protein structure. To the left, a menu lists various functions: openPDB, openPDBURL, openPDBID, openSCOP, openSCOPURL, custom domain, openEuclidianInterface, openEI_URL, openVoronoiInterface, save screen shot, and Calculate Euclidian Interface. Below this menu, there are two dialog boxes: 'Domain definition in SCOP for...' with a 'Custom domain' input field and '확인' (confirm) and '취소' (cancel) buttons; and a 'calculatelt' dialog box with dropdowns for 'd1a25a_', 'd1a25b_', and '5.0', and a 'calculatelt' button. On the right side, there are several control panels. The 'Subpart Selection' panel includes a dropdown for 'd1fpua_xd1fpub_', 'add' and 'remove' buttons, and a 'whole' dropdown. The 'Work' panel includes a 'CutePrint' dropdown, 'set' and 'off' buttons, and a list of options: whole, left, right, bothAtoms, leftAtoms, rightAtoms, bothInterface, and leftInterface. The 'ColorChooser' panel includes 'set' and 'off' buttons, a 'cpk' dropdown, and a slider for 'vbw' ranging from 0 to 100.

Goal of Interactomics & Interfaceomics

- Fastest Drug Discovery Possible

References

Large scale co-evolution analysis of Protein Structural Interlogues using the global Protein Structural Interactome Map(PSIMAP).

Wanky Kim, Dan M. Bolser and Jong Park, *Bioinformatics*, 2004, 20(7):1138-1150.

Visualisation and Graph-theoretic Analysis of a Large-scale Protein Structural Interactome

Dan M Bolser , Panos Dafas, Richard Harrington , Jong H Park and Michael Schroeder

BMC Bioinformatics 2003 4:45

Conservation of protein interaction network in evolution.

Park J, Bolser D.

Genome Informatics. 2001 ;12:135-40.

Mapping protein family interactions: intramolecular and intermolecular protein family interaction repertoires in the PDB and yeast.

Park J, Lappe M, Teichmann SA.

J Mol Biol. 2001 Mar 30;307(3):929-38.

Brief Introduction on NGIC

- 國家有傳體情保 (National Genome Information Center <http://ngic.re.kr/>)

NGIC was established with the vision to become the hub of Korean bioinformatics effort:

- 1) collects and distributes genomic and proteomic data,
- 2) provides bioinformatic databases and analysis platforms
- 3) promotes domestic and international research collaborations in bioinformatics.