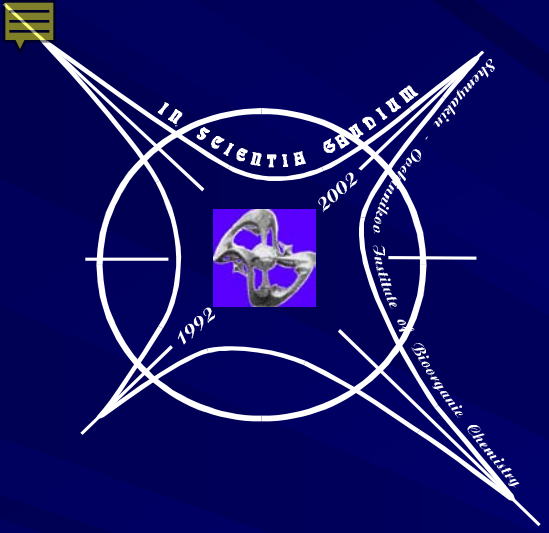
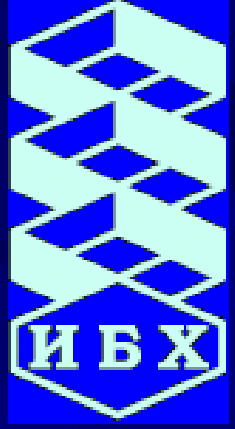




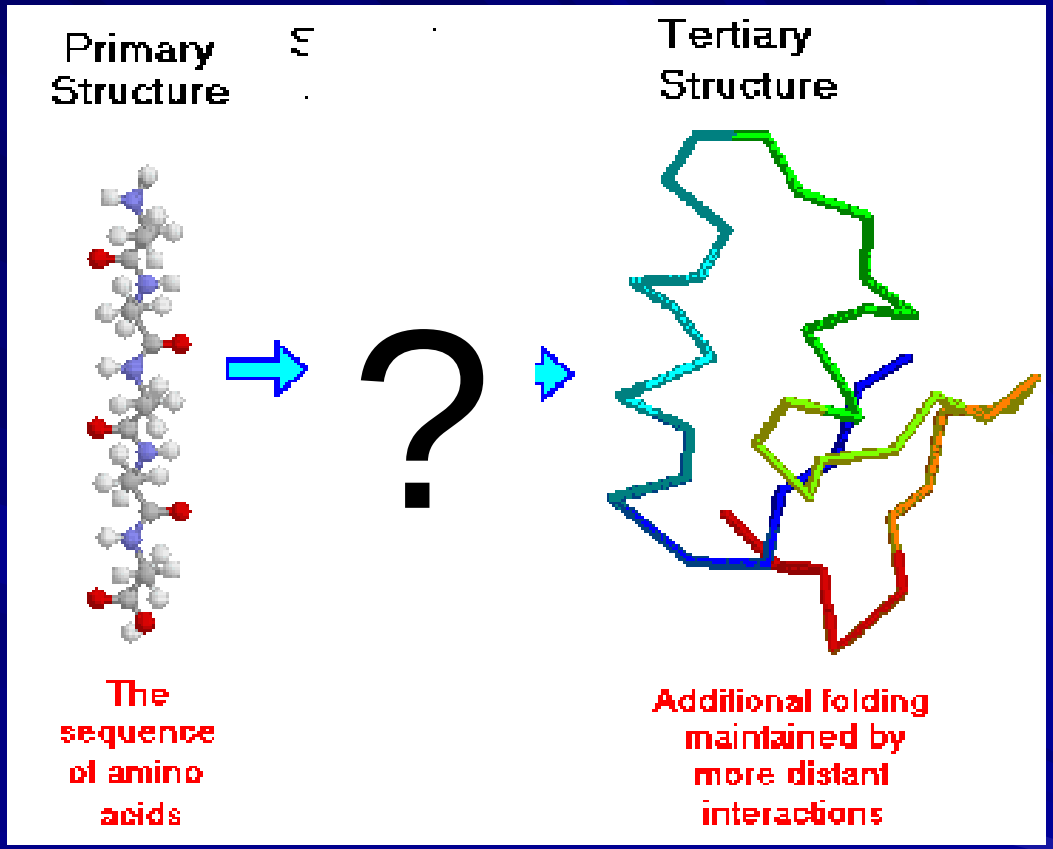
A new paradigm of protein structural organization

Nekrasov A.N., Anashkina A.A., Zinchenko A.A.

Moscow, Russia

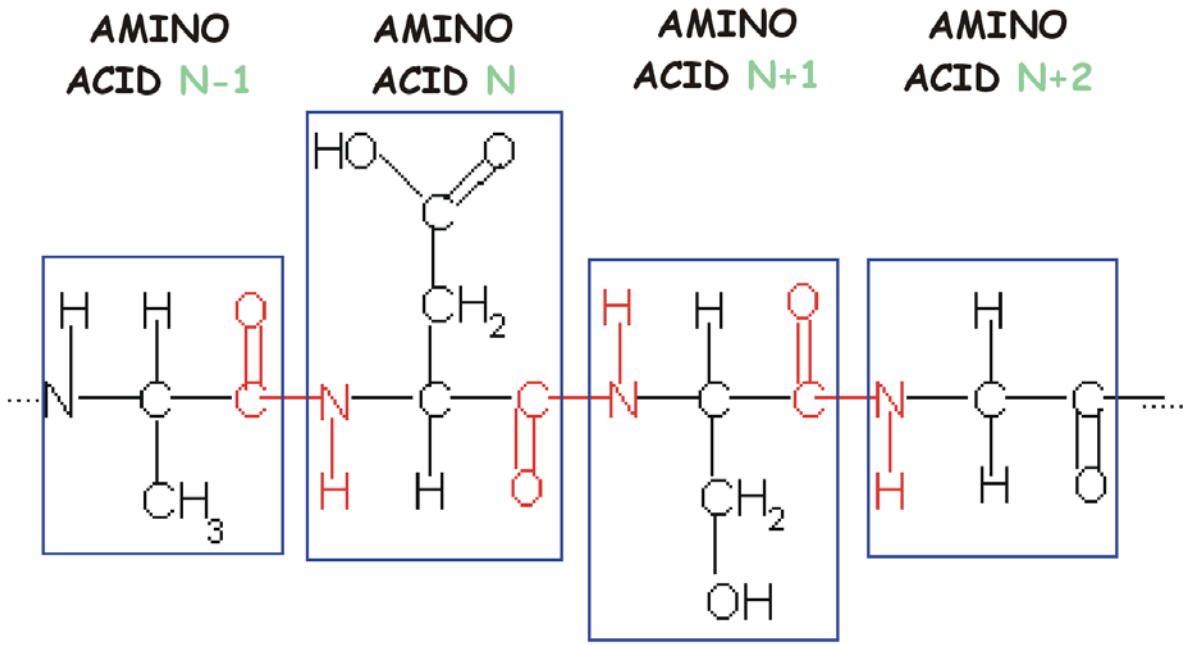
Proteins: structure and function

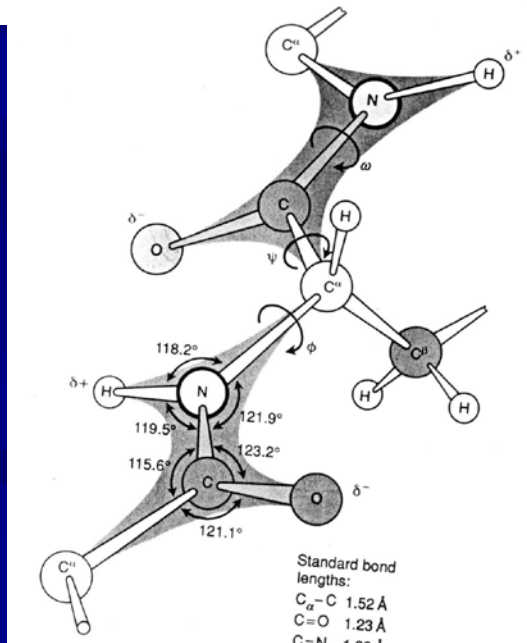
- Enzymes
- Structural
- Transport
- Motor
- Storage
- Signaling
- Receptors
- Gene regulation

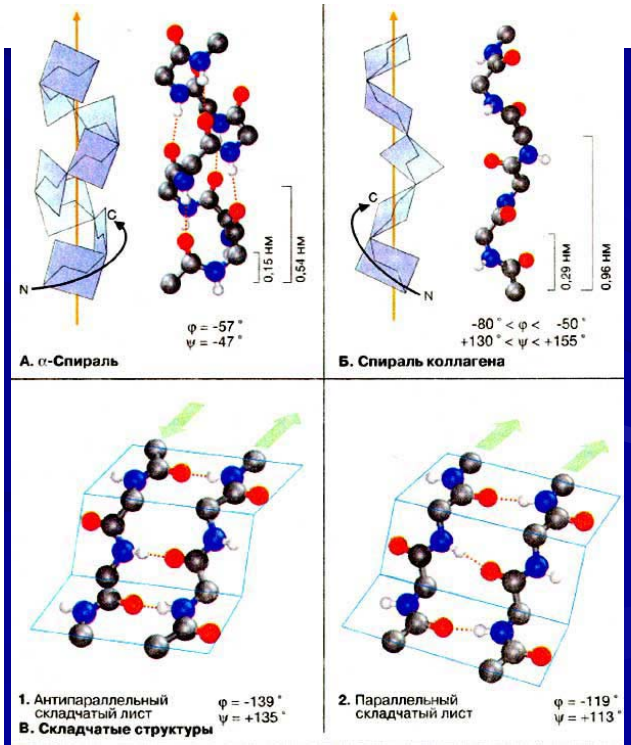


"THE SPATIAL STRUCTURE AND FUNCTIONAL PROPERTIES OF PROTEINS ARE COMPLETELY CODED IN THEIR SEQUENCE" (Anfinsen, 1961)

Linus Pauling







Experimental Data

- Protein structures

>90 000

- Protein sequences

>20 000 000

"… the information for the …. assumption of the native secondary and tertiary structures, is contained in the amino acid sequence itself."

Anfinsen, B. C., Haber, E., Sela, M., and White, Jr, F. H.

"The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain". *Biochemistry*, (1961) v. 47, pp.1309-1314

VS

"…proteins are slightly corrected random heteropolymers"

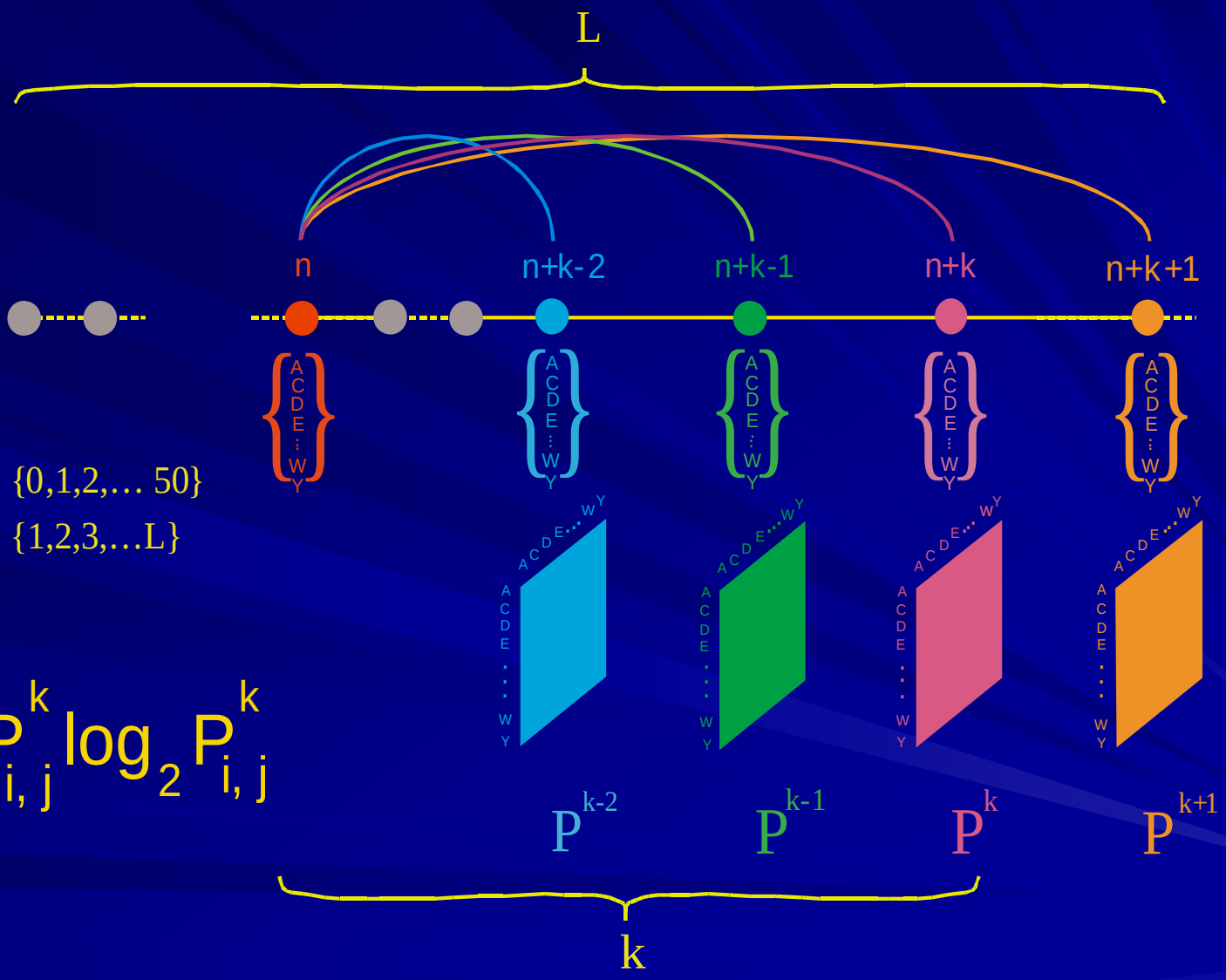
Ptitsin O.B., Vol'kenshtein M.V. J.Biomol. Struct. Dyn. (1986) v.4(1), pp.137-156

DATA FOR THE ANALYSIS

DATABASE	PROTEINS
NRDB (release 3.0)	192519
NRDB (release 6.0)	252926
NRDB (release 9.0)	534938

J Biomol Struct Dyn. v. 20, p. 84

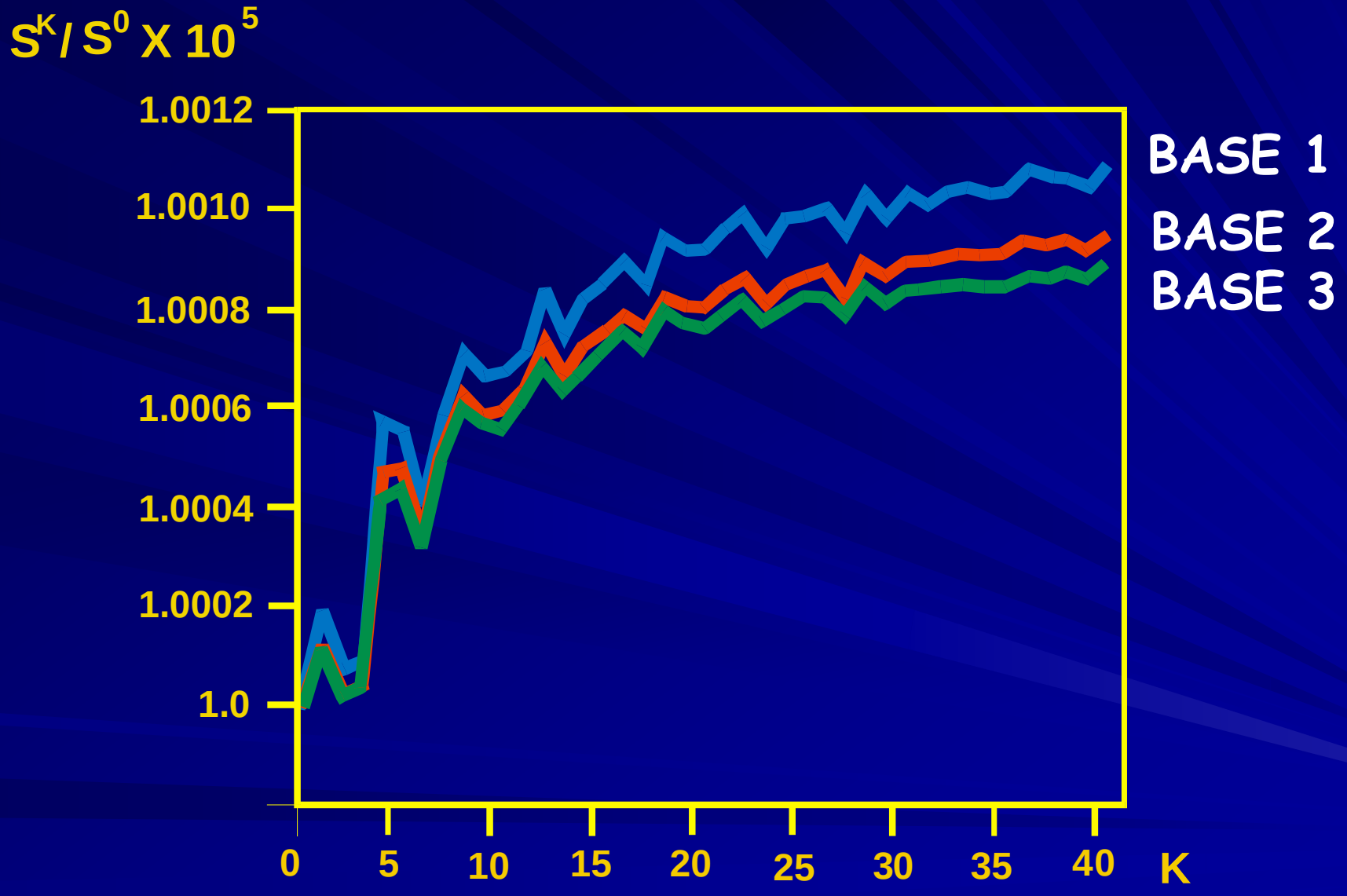
METHOD OF PROTEIN SEQUENCE ANALYSIS



$$S^{k} = -\sum_{i=1,20}^{20}\sum_{j=1,20}^{20} P_{i,j}^{k} \log_2 P_{i,j}^{k}$$

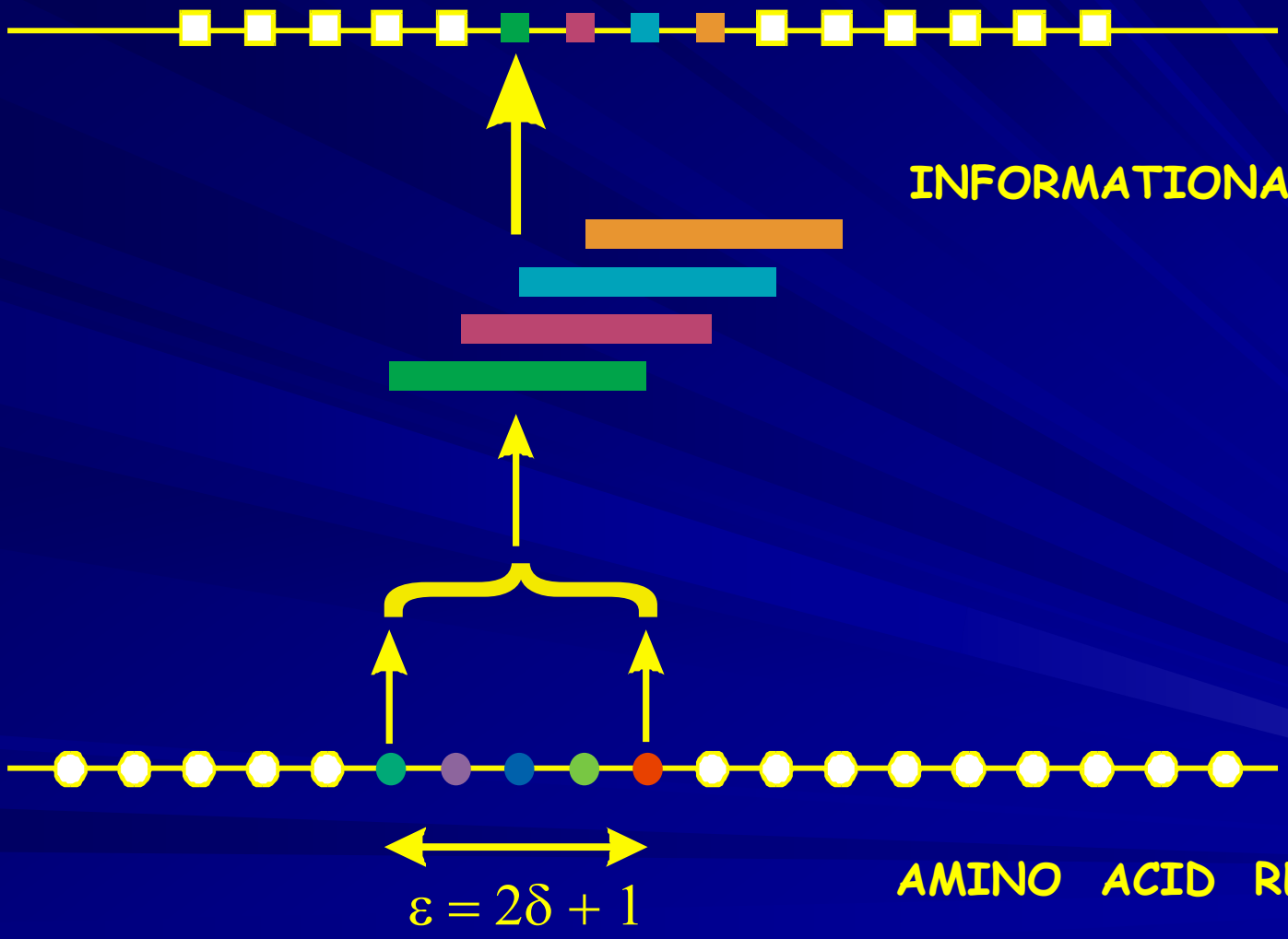
J Biomol Struct Dyn. v. 20, p. 84

POSITION SHANNON'S ENTROPY

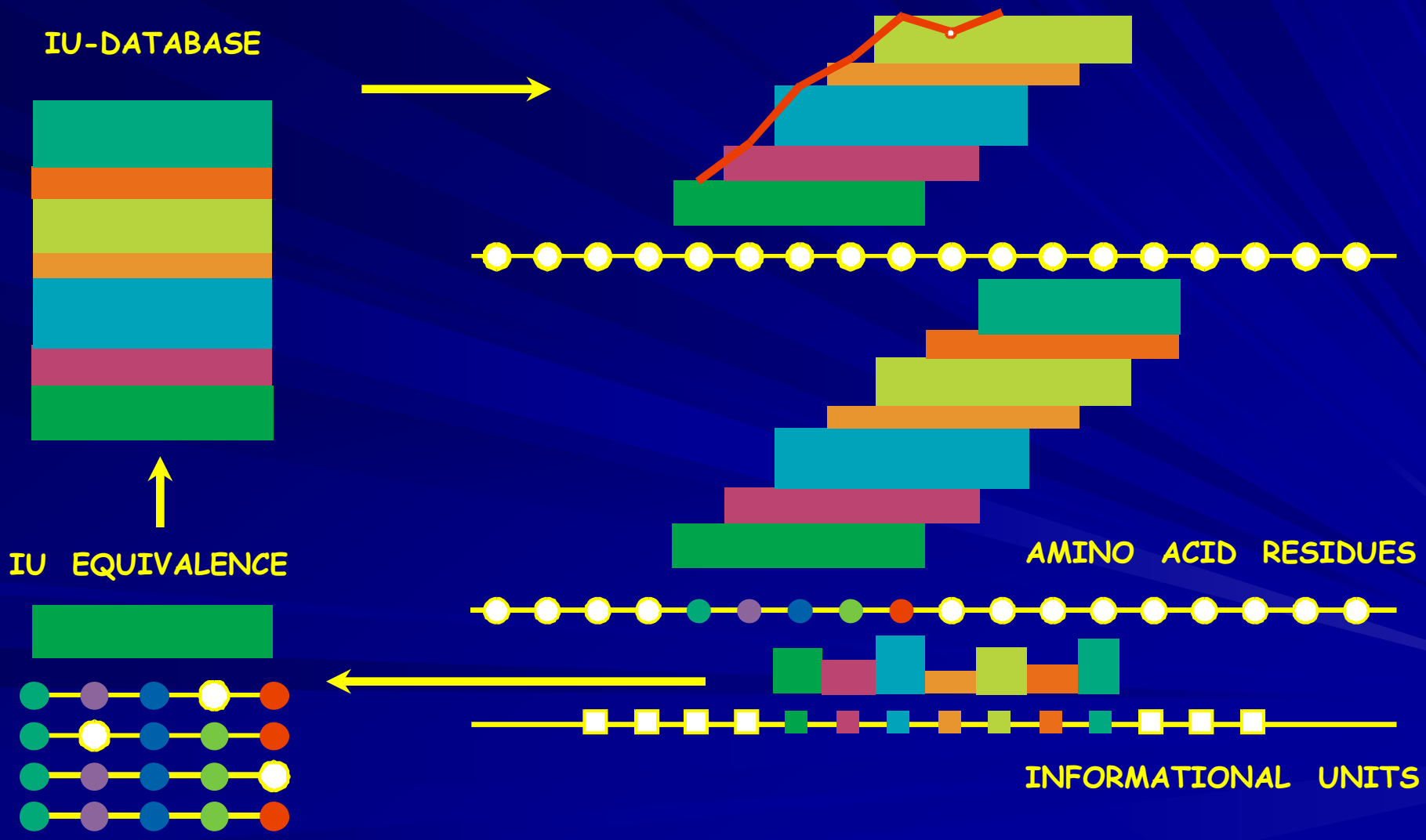


J Biomol Struct Dyn. v. 20, p. 84

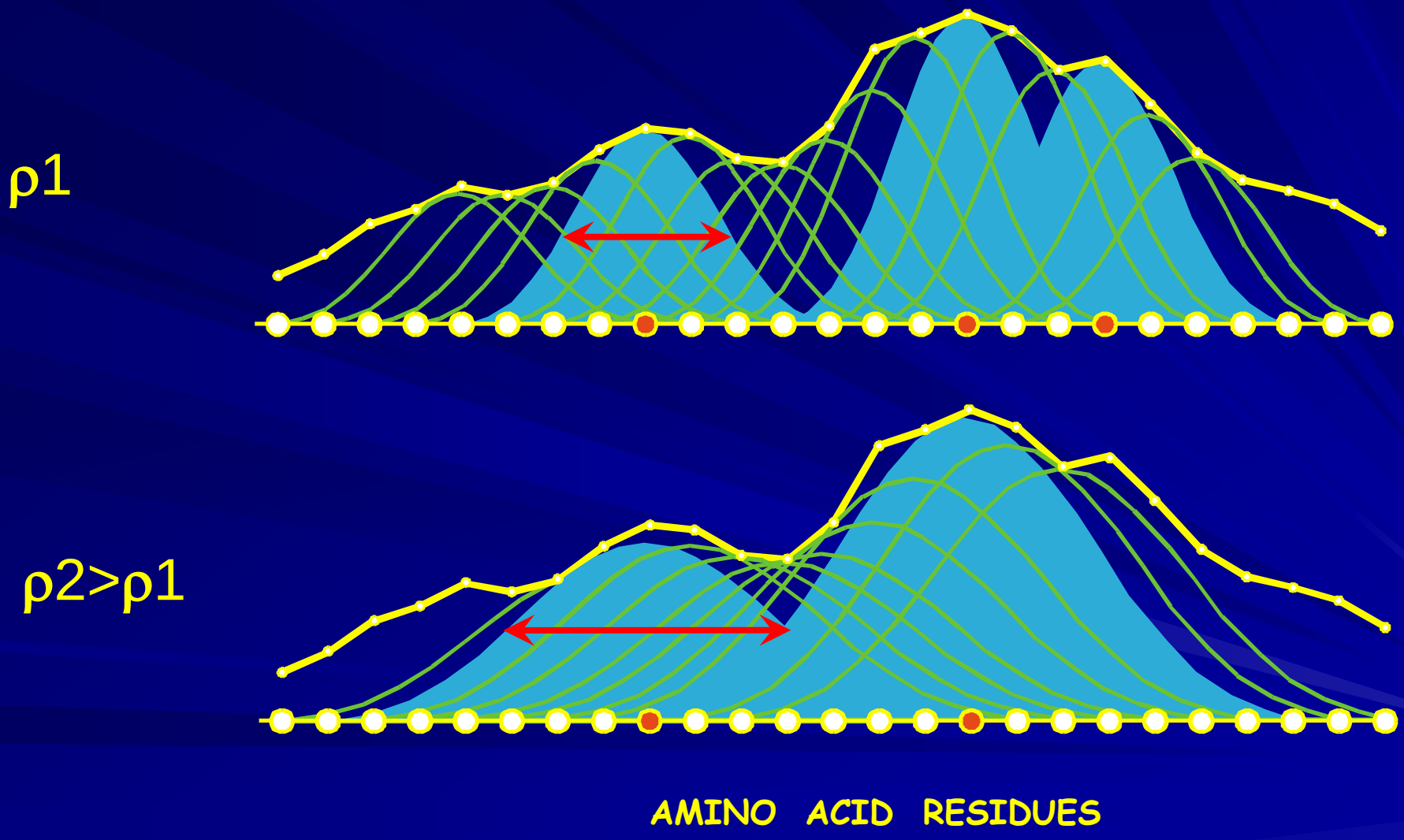
REPRESENTATION OF SEQUENCE AS A SET OF INFORMATIONAL UNITS



COMPUTATION OF INFORMATIONAL UNITS POPULATION PROFILE OF PROTEINS

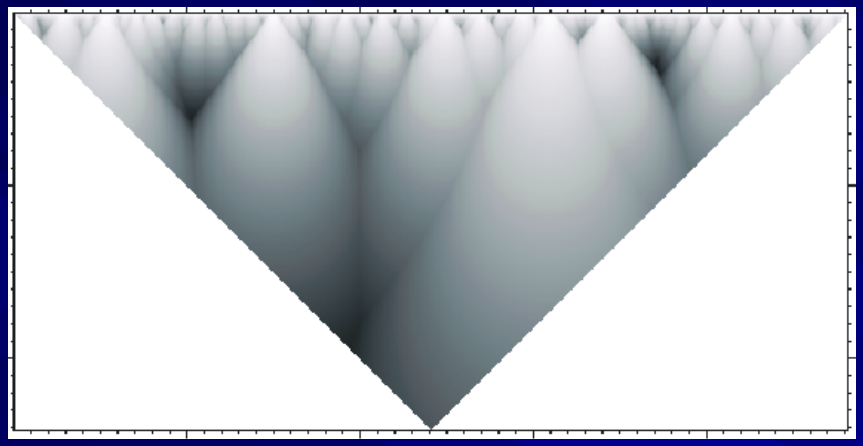


GAUSSIAN APPROXIMATION OF POPULATION PROFILE OF PROTEINS

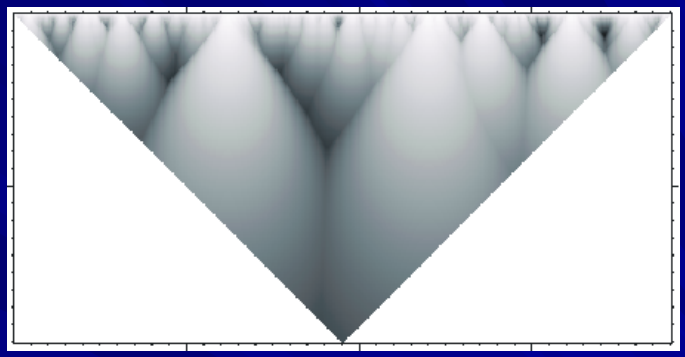


GRAPHICAL PRESENTATION OF PROTEIN INFORMATIONAL STRUCTURE

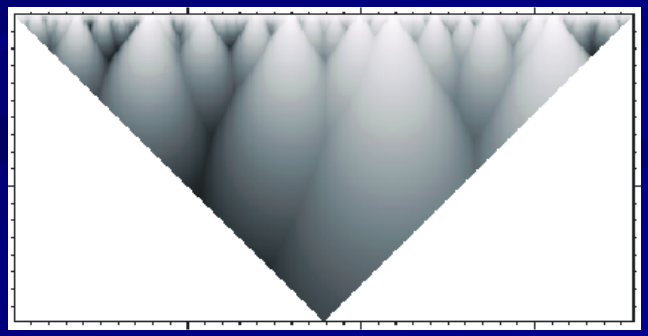
1AC5A



1BJWA

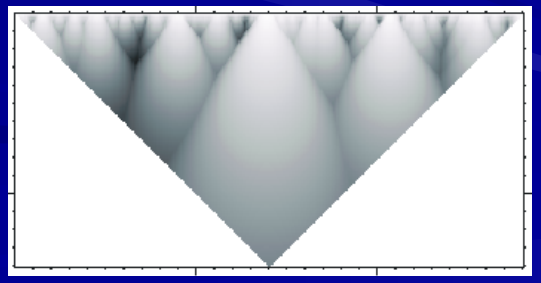


1AO5A



L
(sequence)

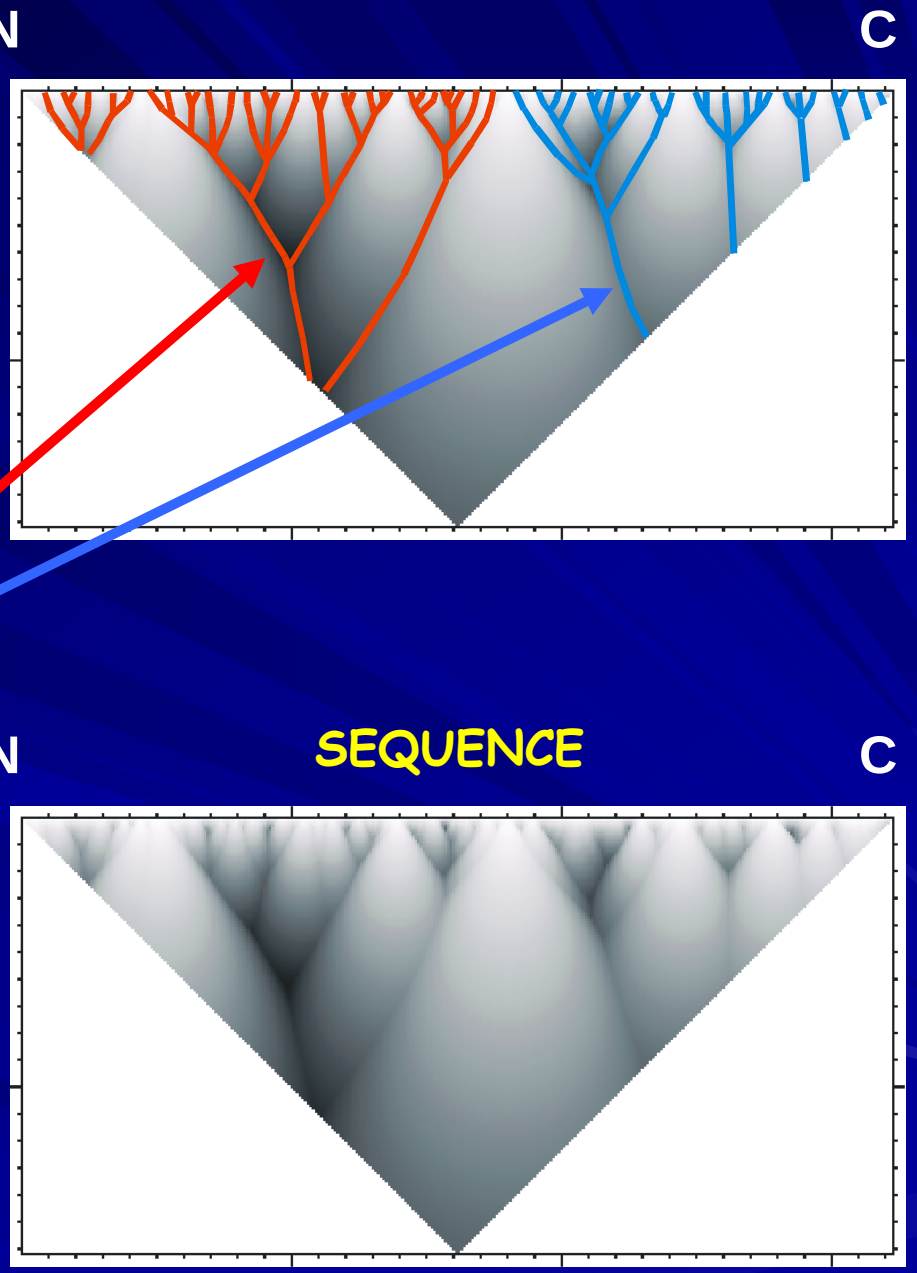
N C



ρ/2

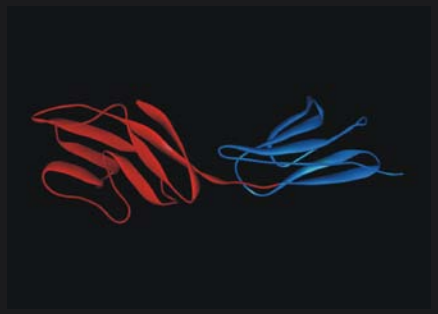
ANIS - **AN**alysis of **I**nformational **S**tructure method

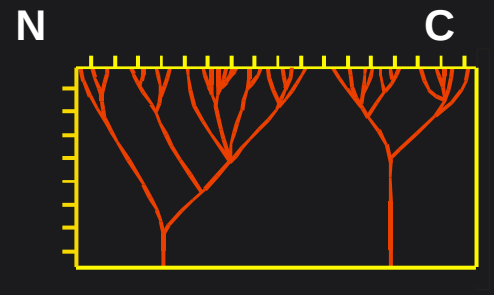
ELIS - High Rank **EL**ements of **I**nformational **S**tructure



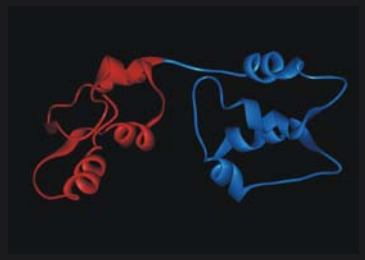
J Biomol Struct Dyn. v. 21, p. 615

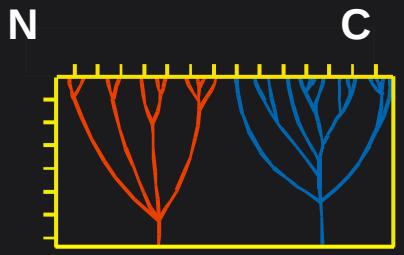
HIGH RANK ELIS RELATIONS WITH RESPECT TO STRUCTURAL DOMAINS



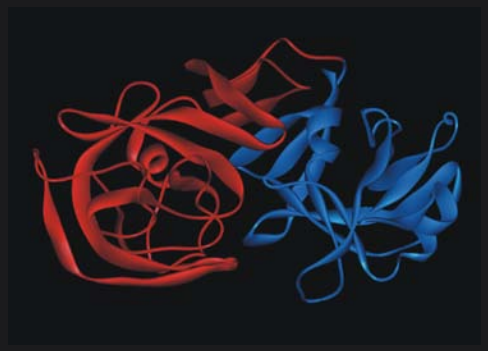


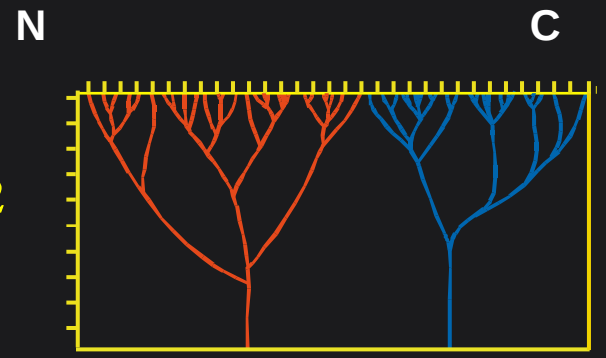
T-Lymphocyte Adhesion Glycoprotein (1HNG-A.PDB)





Myosin Regulatory Domain (1SCM-C.PDB)

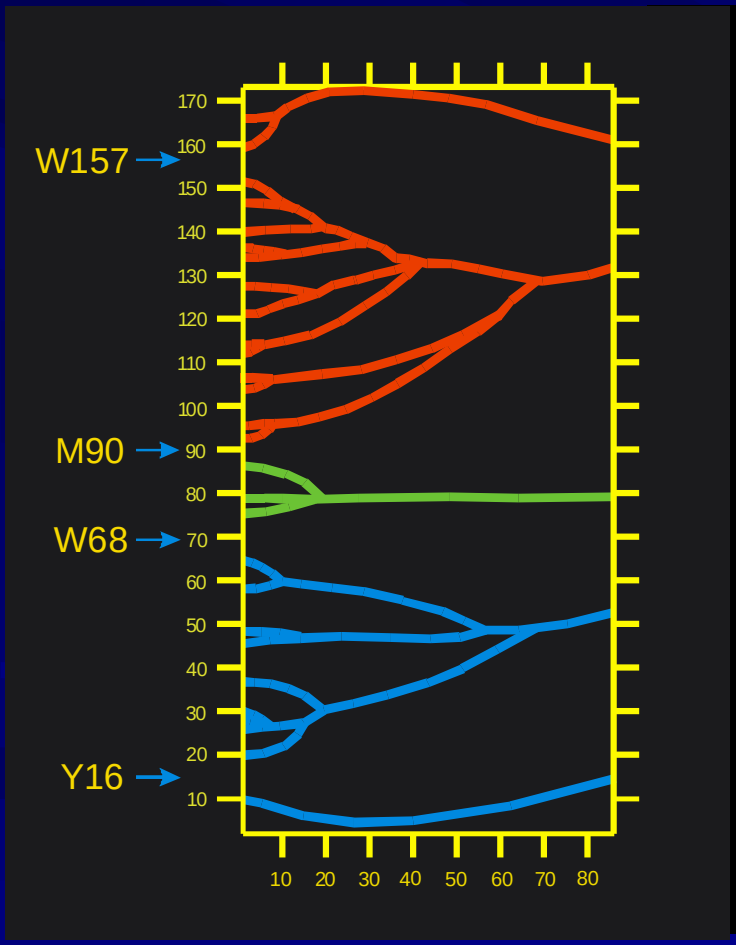


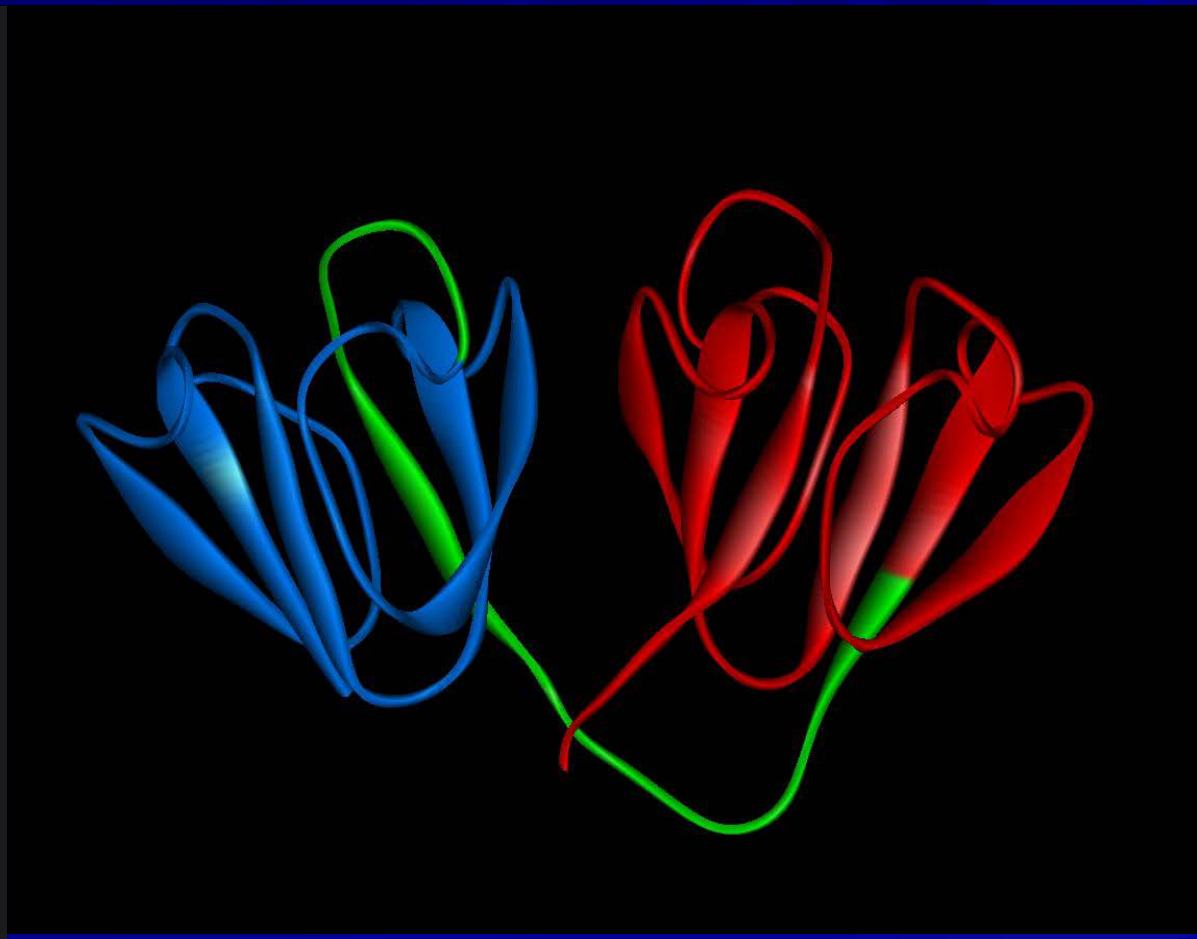


Pepsin (1AM5.PDB)

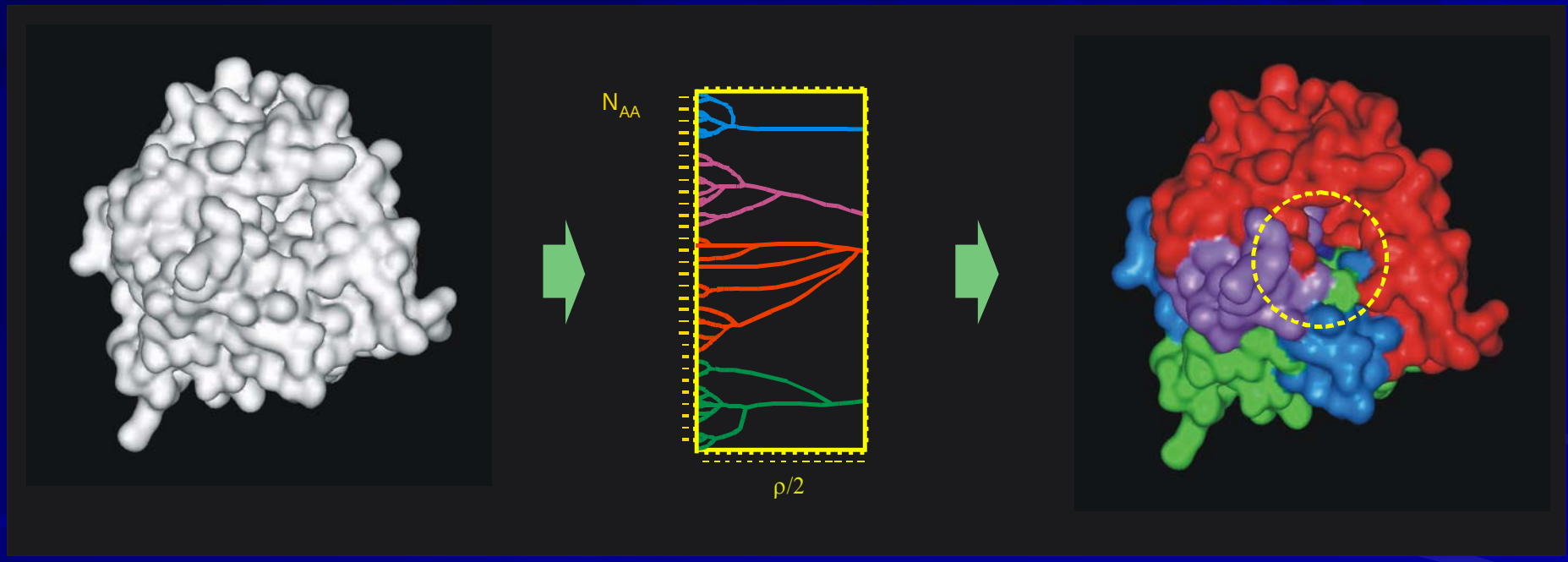
HIGH RANK ELIS RELATIONS WITH RESPECT TO STRUCTURAL DOMAINS

γ-CRYSTALLIN (1GCR.PDB)

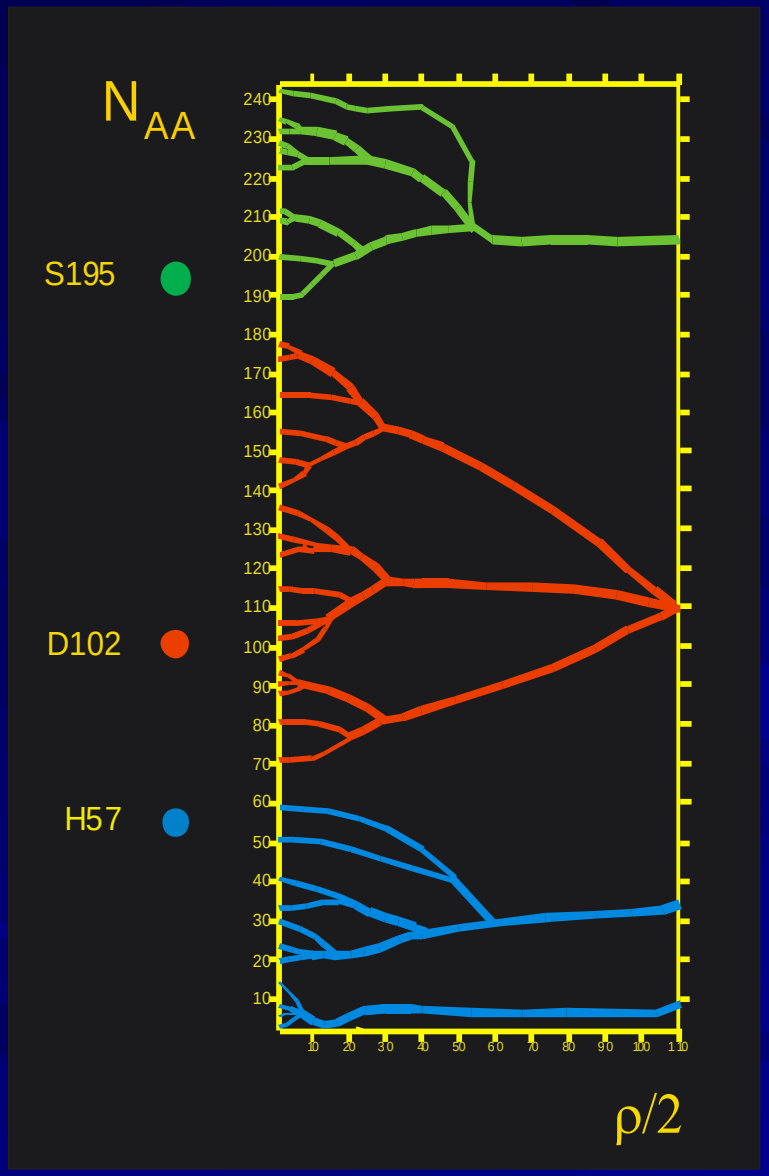




CATALYTIC SITES IDENTIFICATION METHOD

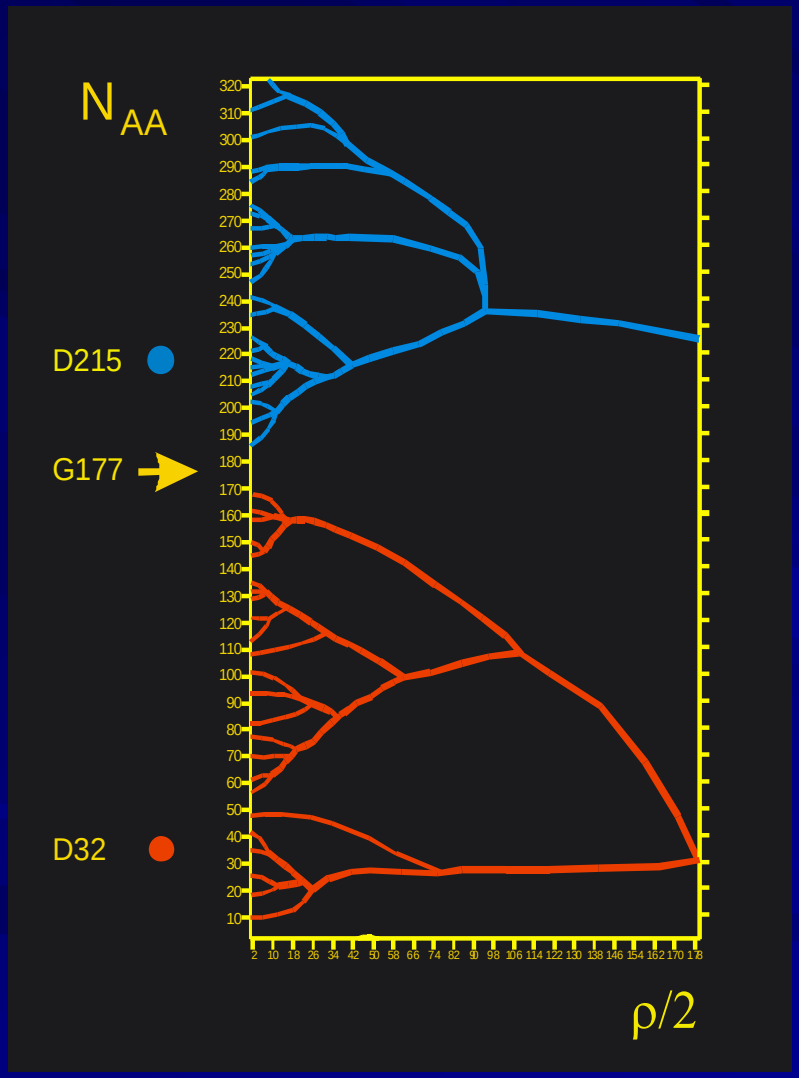


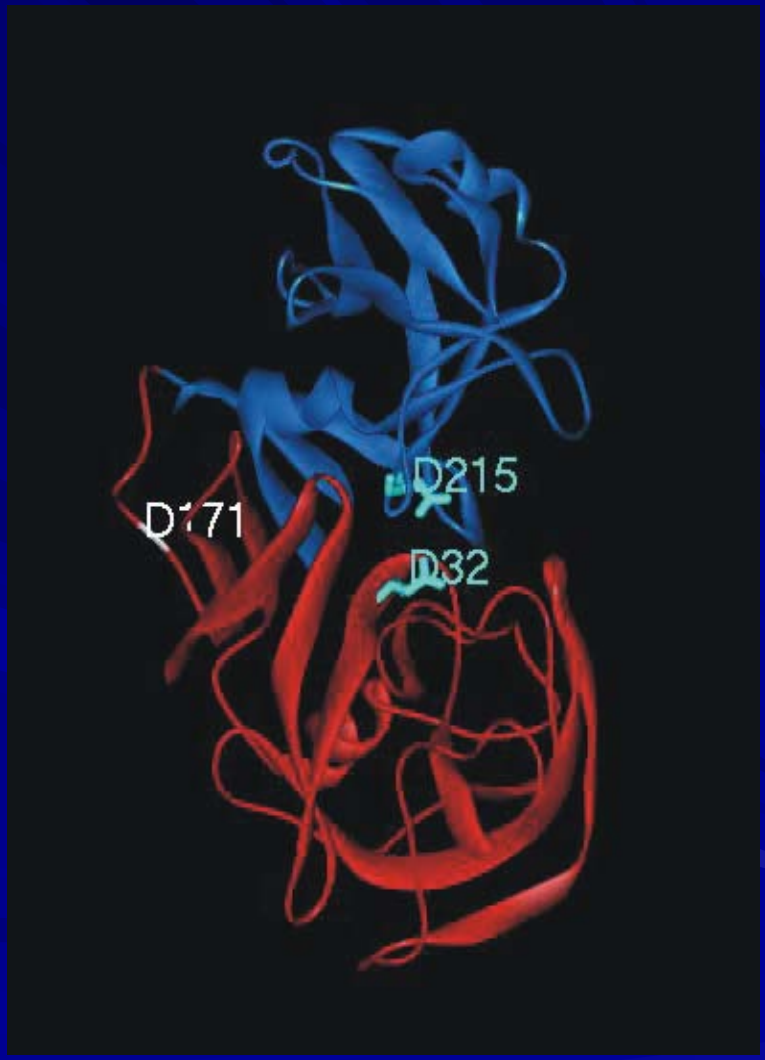
TRYPSIN FROM BOS TAURUS (EC: 3.4.21.4) 1CO7.PDB



J Biomol Struct Dyn. (2008) v. 25, p. 554

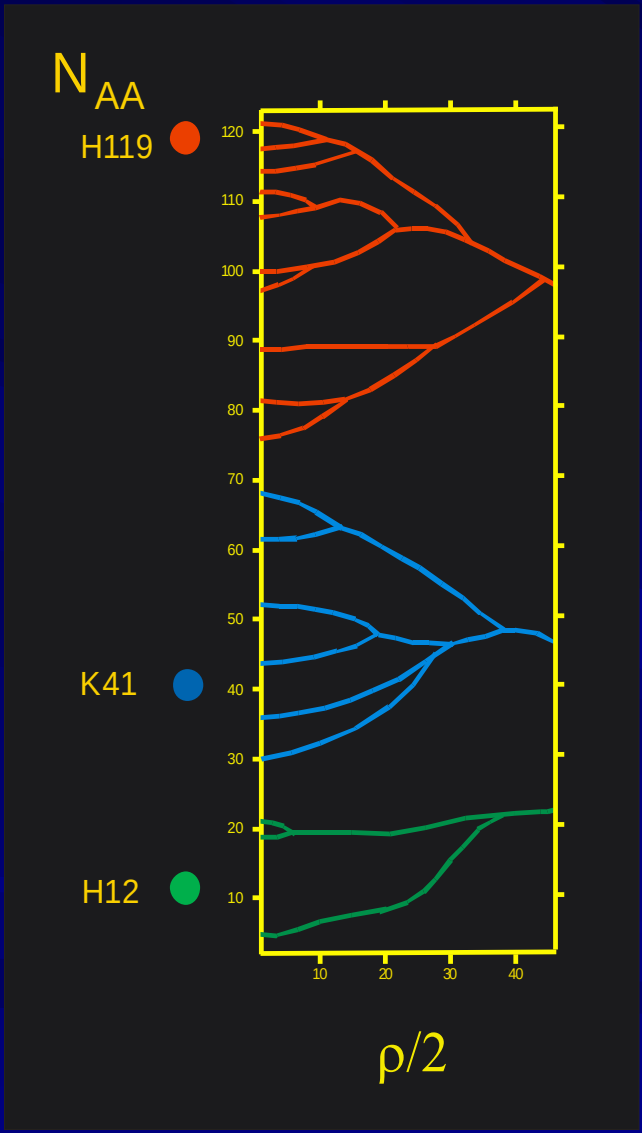
PEPSIN FROM GADUS MORHUA (EC: 3.4.23.1) 1AM5.PDB

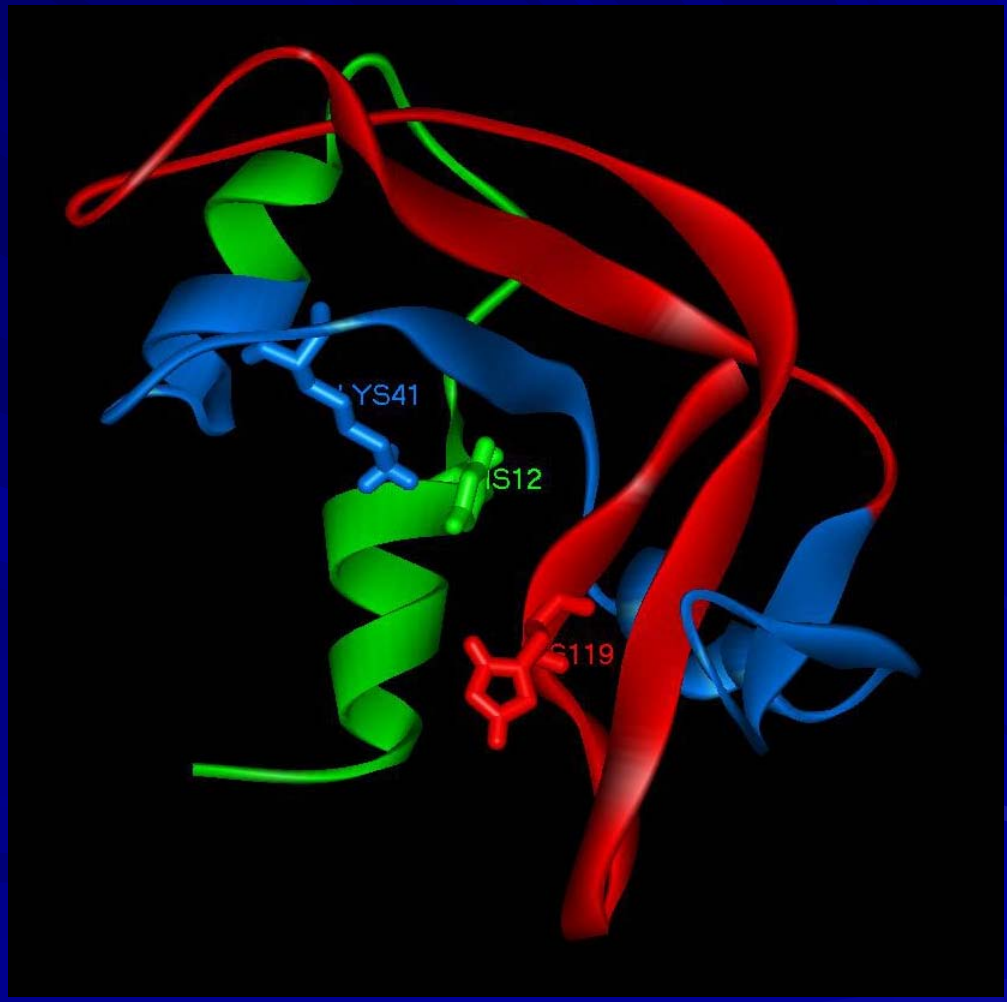




J Biomol Struct Dyn. (2008) v. 25, p. 554

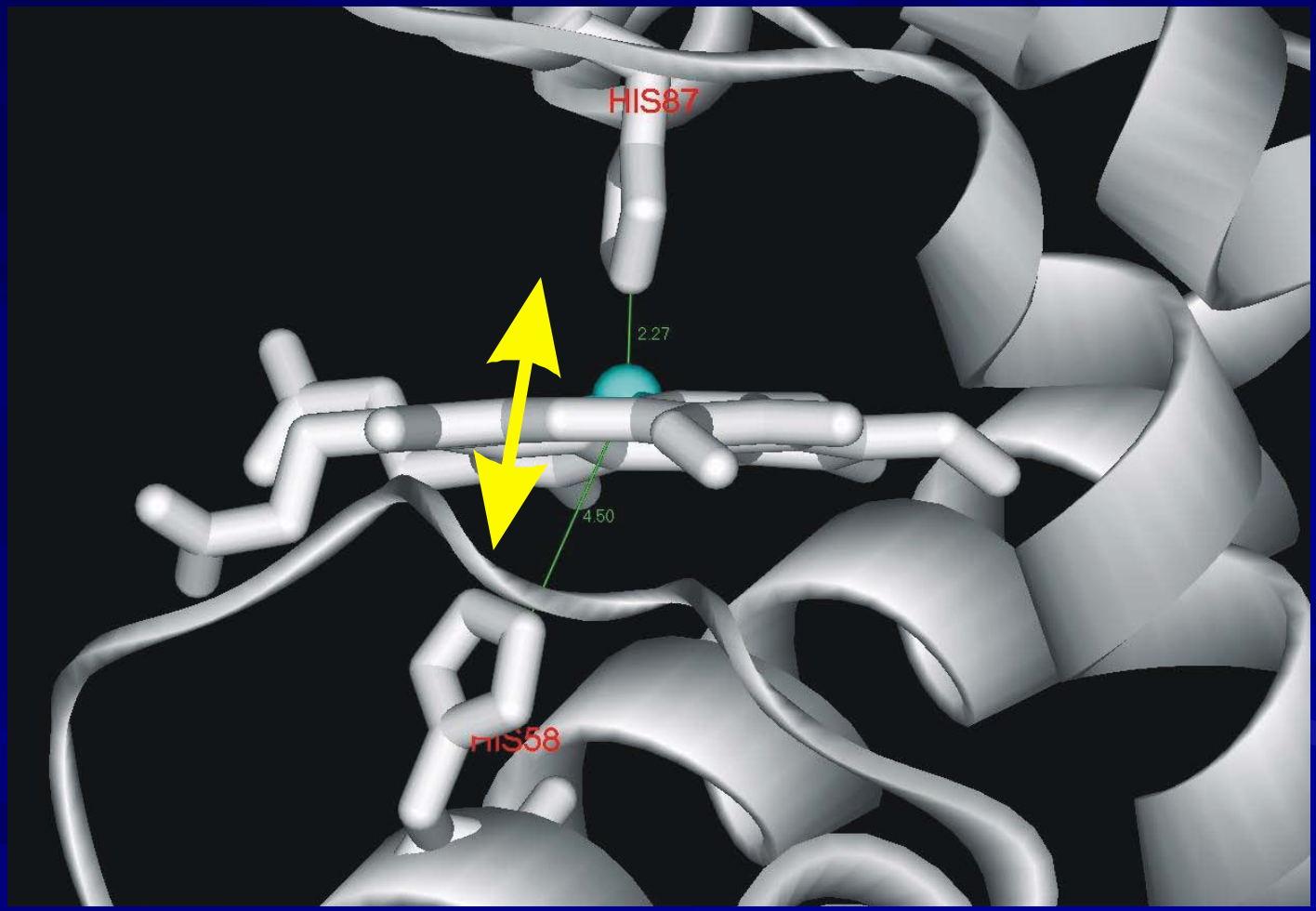
RNAse A FROM BOS TAURUS (EC: 3.1.27.5) 1AQP.PDB





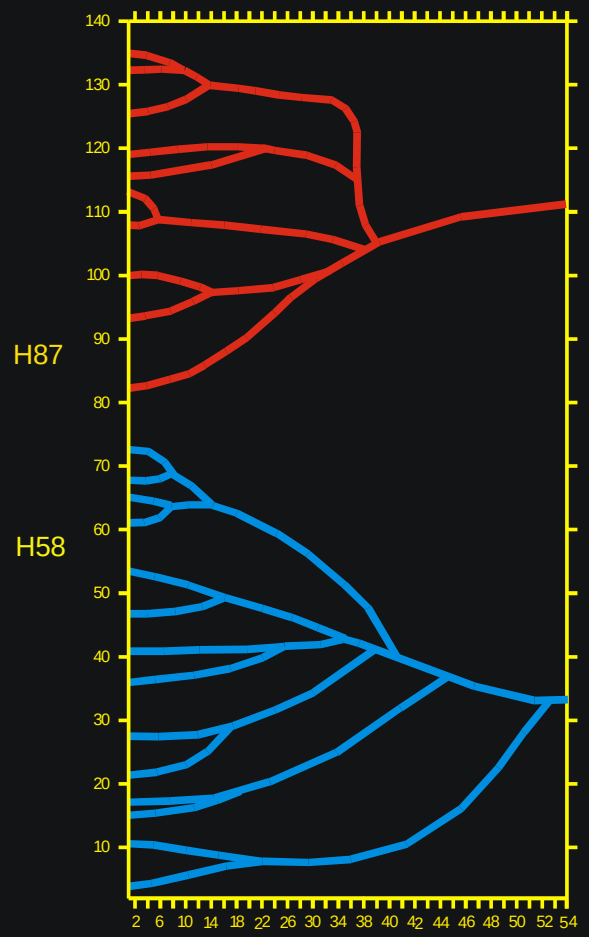
J Biomol Struct Dyn. (2008) v. 25, p. 554

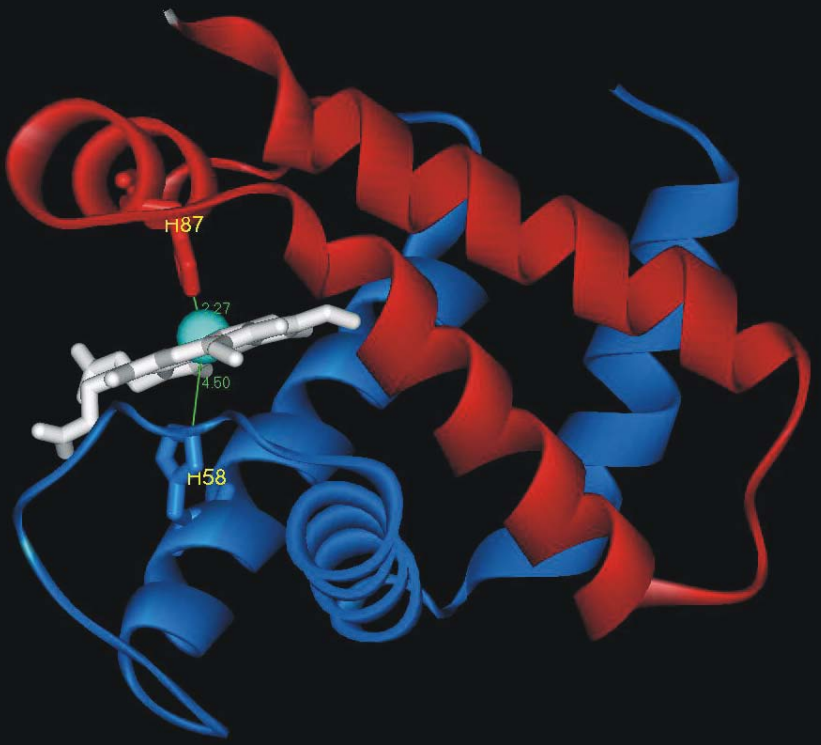
ACTIVE CENTRE OF HEMOGLOBIN β-CHAIN



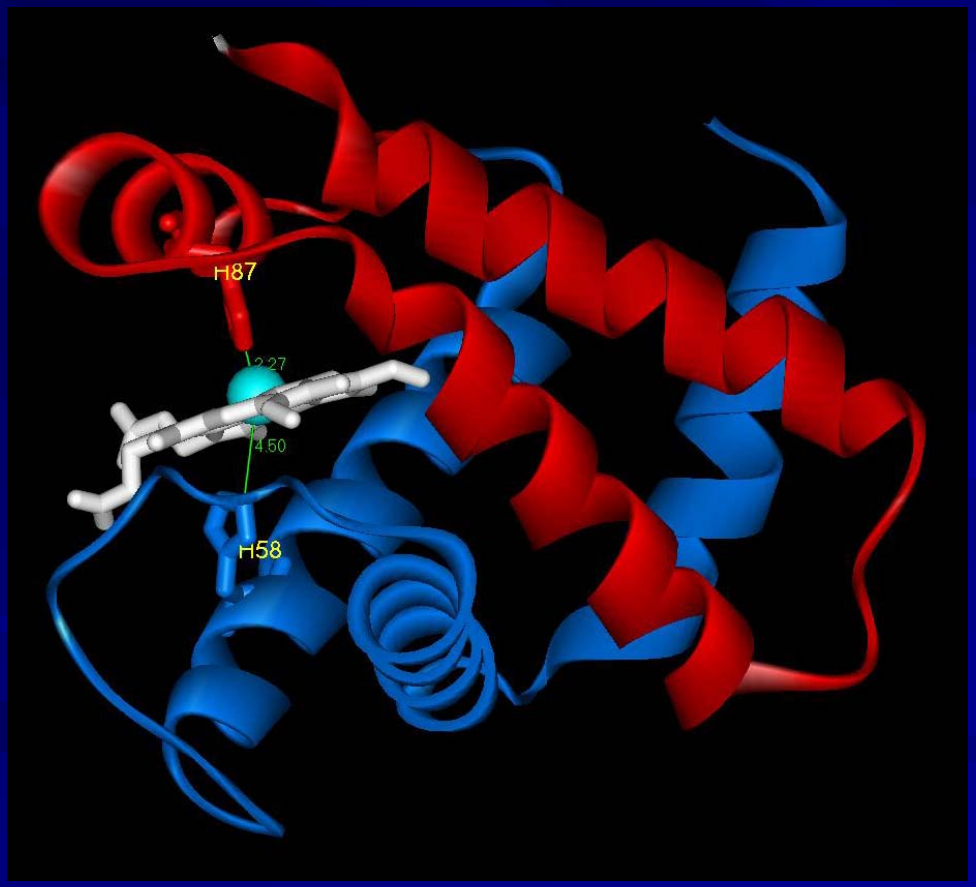
Paoli et. al. JMB (1996) v. 256, p. 775

HIGH RANK ELIS OF HEMOGLOBIN β-CHAIN (1BZ0.PDB)



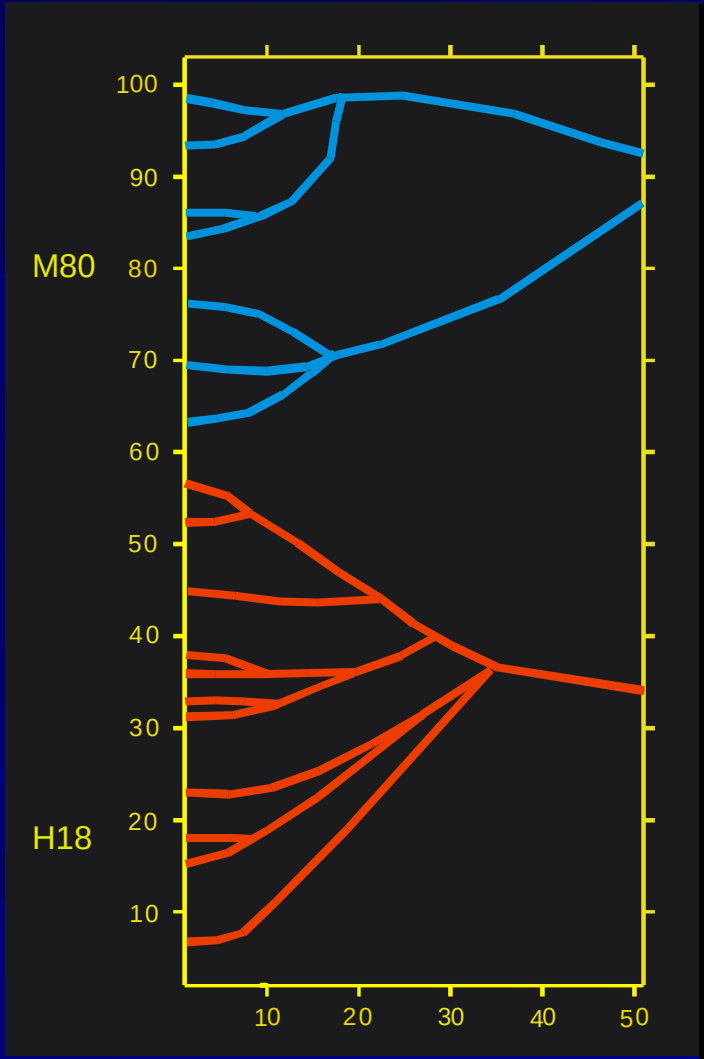


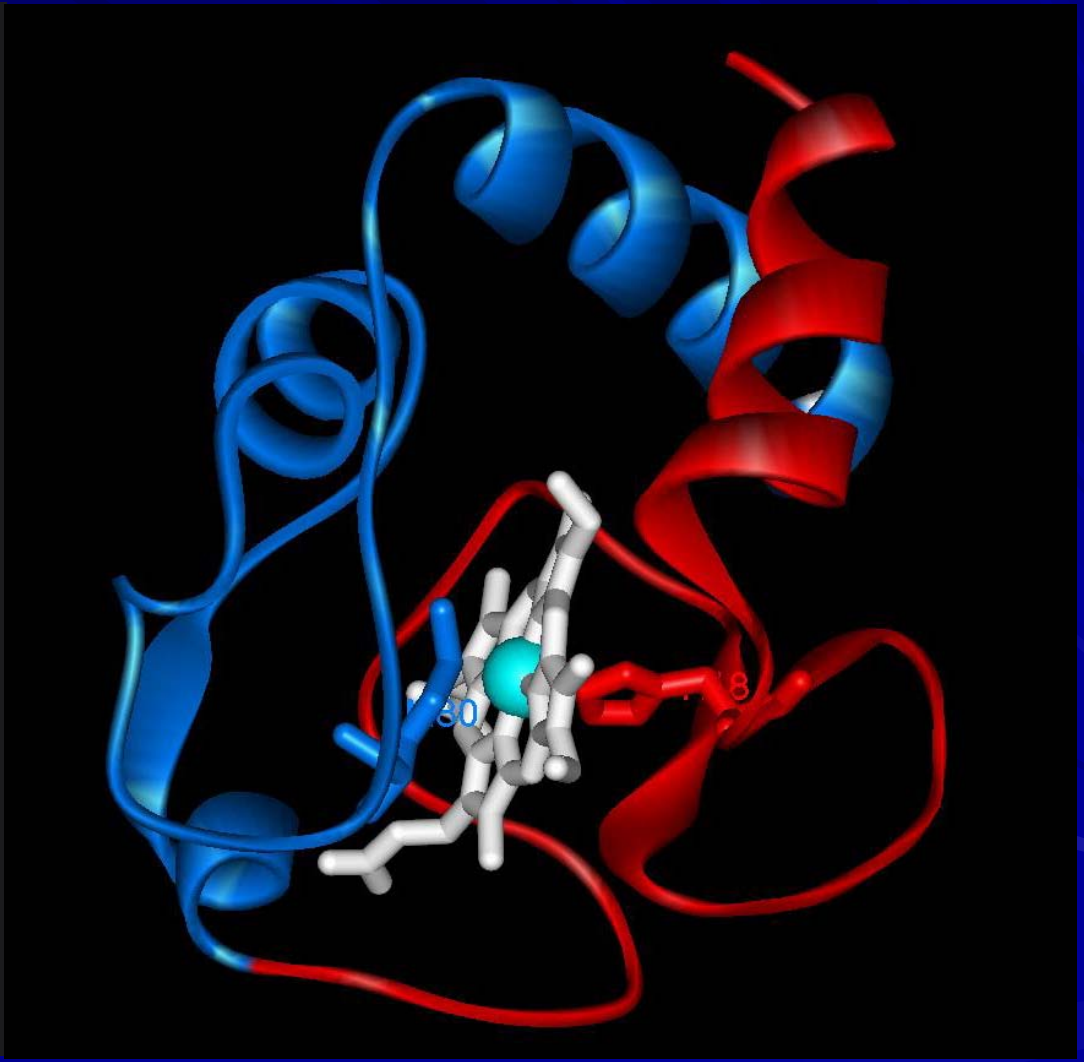
HEMOGLOBIN (1BZO.PDB)



CYTOCHROME C (1HRC.PDB)

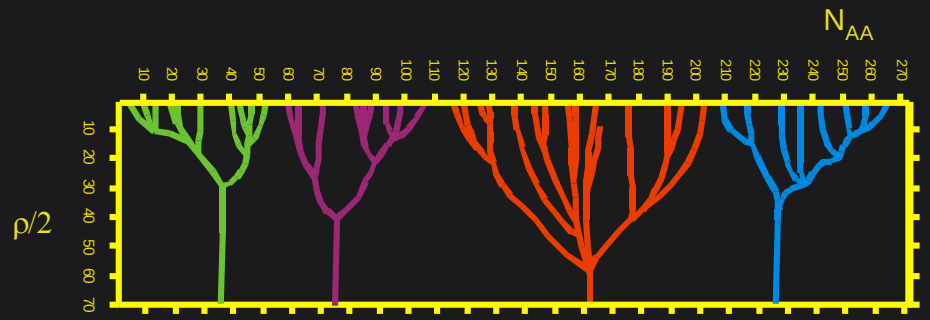
HIGH RANK ELIS OF CYTOCHROME C FROM EQUUS CABALLUS (1RHC.PDB)





METHOD OF NEW RECOMBINANT PROTEIN DESIGN

```
>PROTEIN  SEQUENCE
AQTVPYGIPL IKADKVQAQG FKGANVKVAV LDTGIQASHP DLNVVGGASF
VAGEAYNTDG NGHGTHVAGT VAALDNTTGV LGVAPSVSLY AVKVLNSSGS
GSYSGIVSGI EWATTNGMDV INMSLGGASG STAMKQAVDN AYARGVVVVA
AAGNSGNSGS TNTIGYPAKY DSVIAVGAVD SNSNRASFSS VGAELEVMAP
GAGVYSTYPT NTYATLNGTS MASPHVAGAA ALILSKHPNL SASQVRNRLS
STATYLGSSF YYGKGLINVE AAAQ*
```

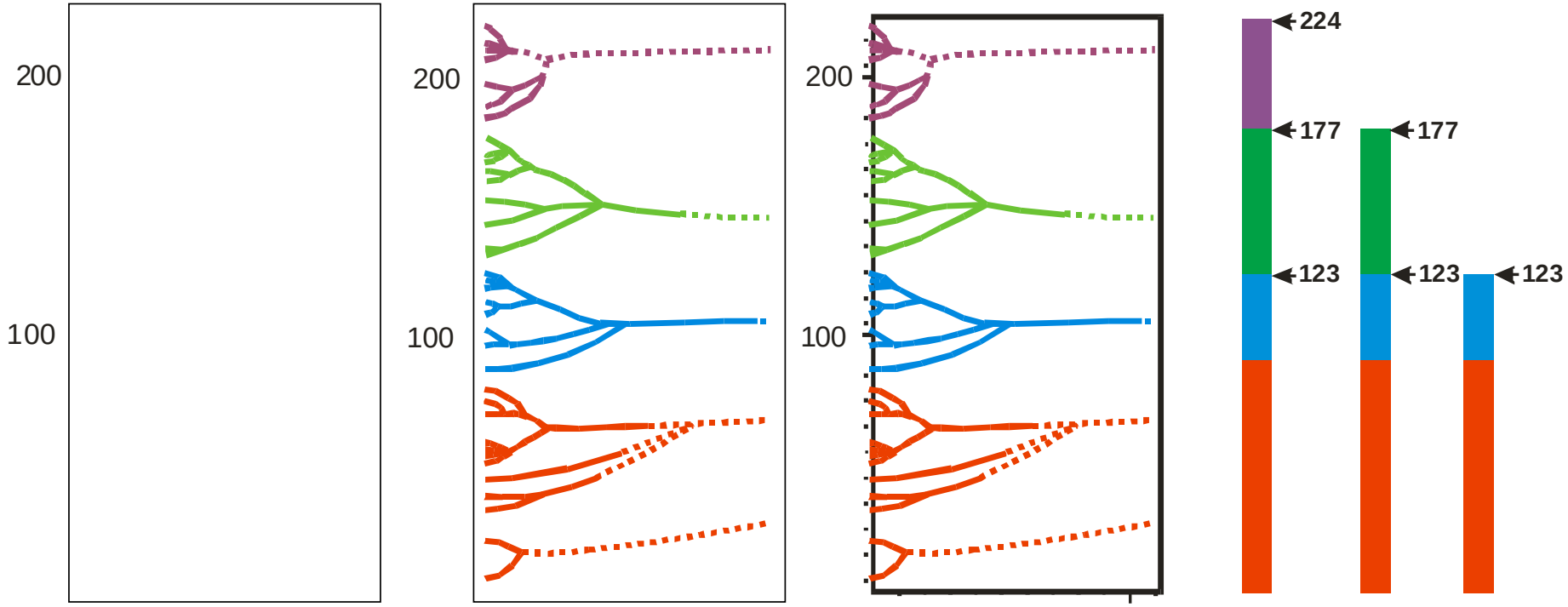



```
>PROTEIN  SEQUENCE
AQTVPYGIPL IKADKVQAQG FKGANVKVAV LDTGIQASHP DLNVVGGASF
VAGTNGMDVI NMSLGGASGS TAMKQAVDNA YARGVVVVAA AGNSGNSGST
NTIGYPAKYD SVIAVGAVDS NSNRASFSSV GAELEVMAPG AGVYSTYPTN
TYATLNGTSM ASPHVAGAAA LILSKHPNLS ASQVRNRLSS TATYLGSSFY
YGKGLINVEA AAQ*
```

```
>PROTEIN  SEQUENCE
AQTVPYGIPL IKADKVQAQG FKGANVKVAV LDTGIQASHP DLNVVGGASF
VAGEAYNTDG NGHGTHVAGT VAALDNTTGV LGVAPSVSLY AVKVLNSSGS
GSYSGIVSGI EWATTNGMDV INMSLGGASG STAMKQAVDN AYARGVVVVA
AAGNSGNSGS TNTIGYPAKY DSVIAVGAVD SNSNRASFSS VGAELEVMAP
GAGVYSTYPT NTYATLNGTS MASPHVAGAA ALILSKHPNL SASQVRNRLS
STATYLGSSF YYGKGLINVE AAAQ*
```

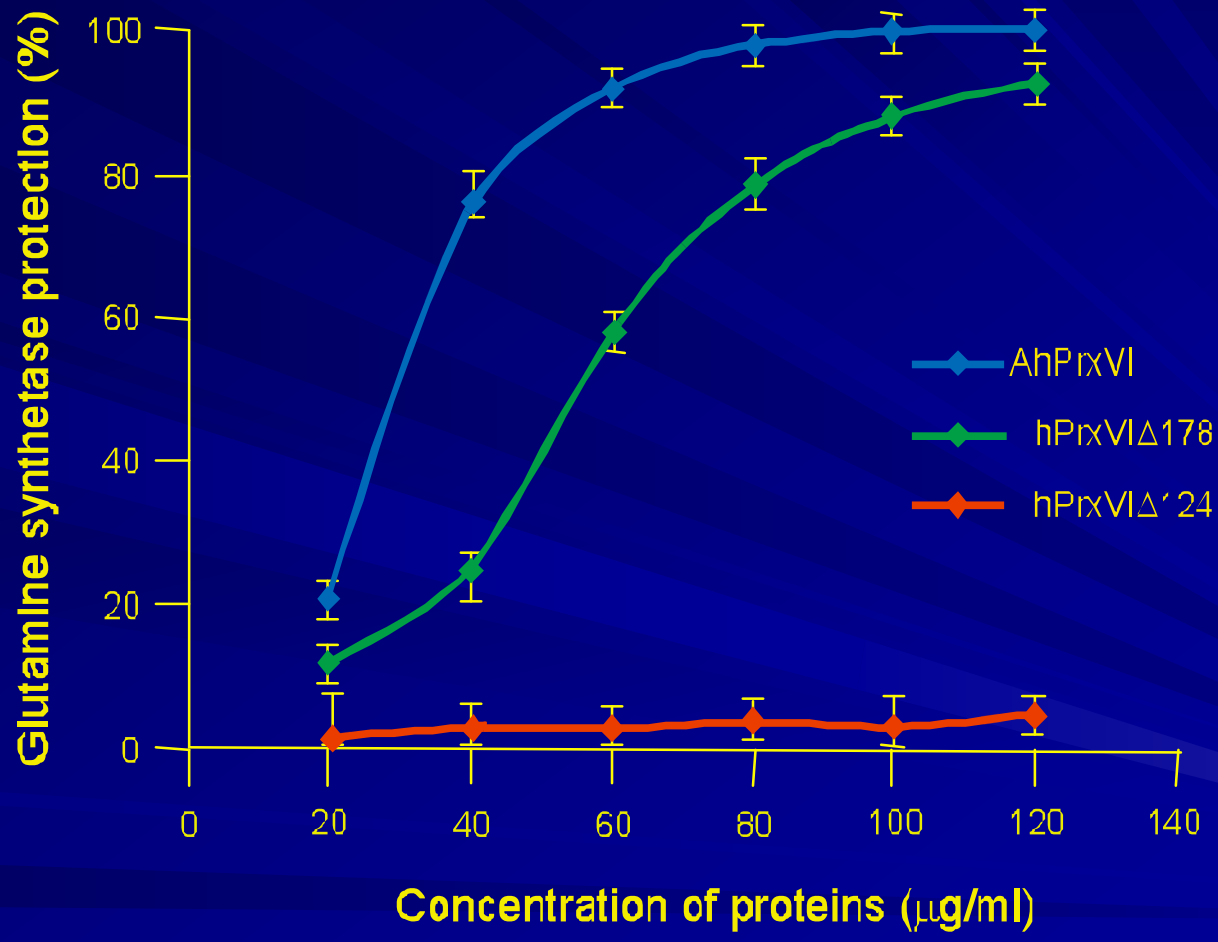
J Biomol Struct Dyn. v.24(5), p.455

TRUNCATED FORMS OF HUMAN 1-CYS PYROXIREDOXIN (PrxVI)



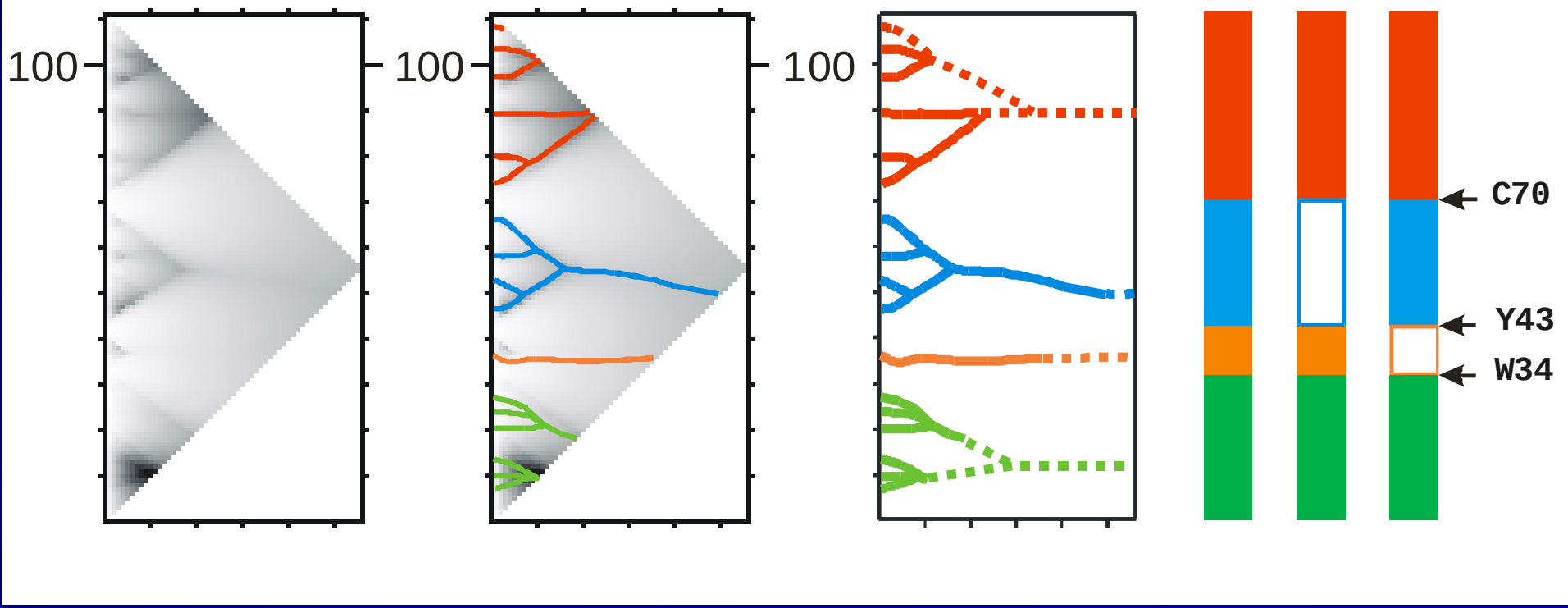
J Biomol Struct Dyn. v. 24, p. 455

PROTECTIVE ACTIVITIES of hPrxVI, hPrxVIΔ178 & hPrxVIΔ124

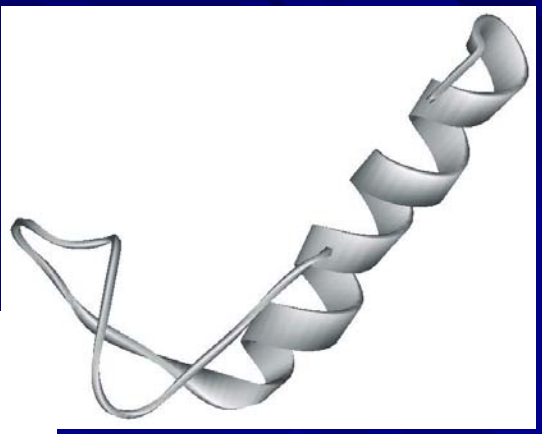


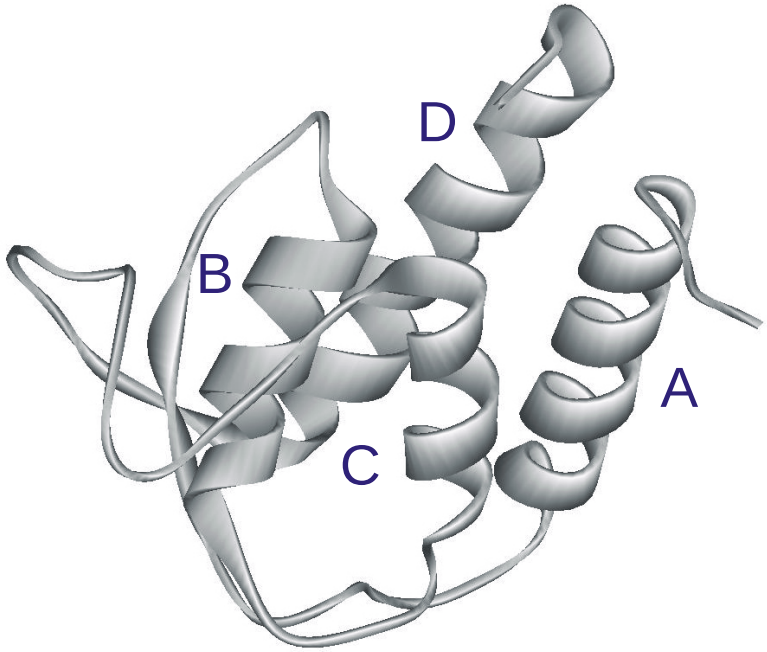
J Biomol Struct Dyn. v. 24, p. 455

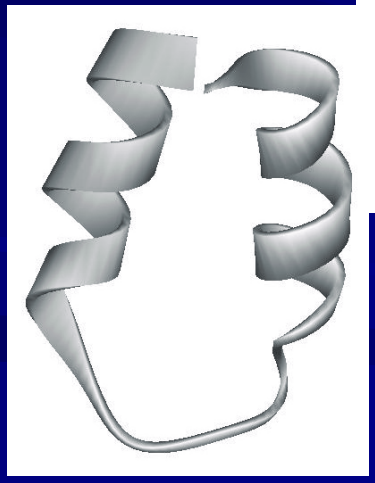
TRUNCATED FORMS OF INTERLEUKIN IL-13

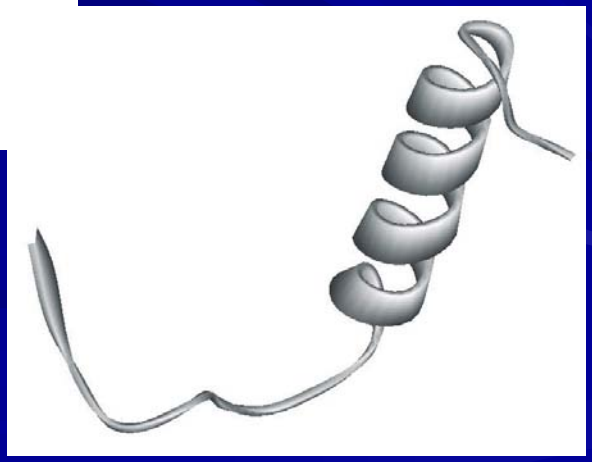


Biochemistry (Moscow), (2009) vol. 74, p. 493

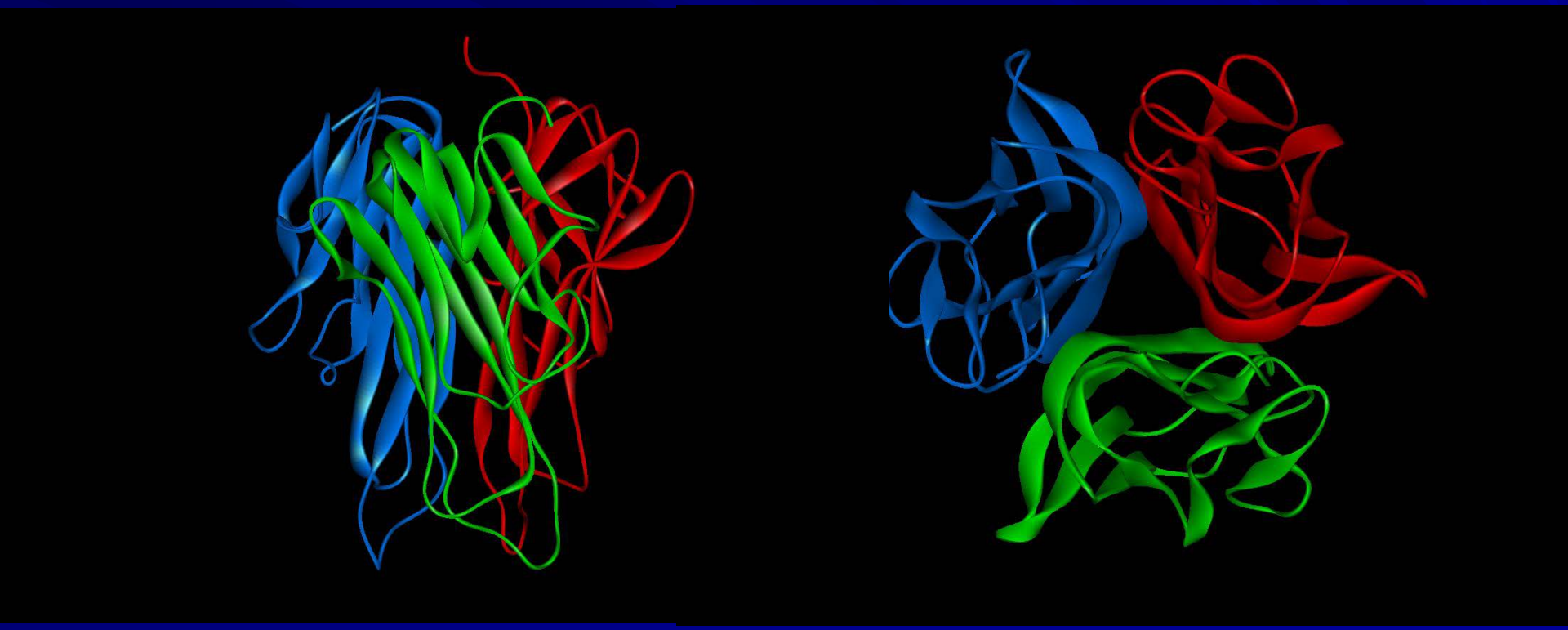



D
B
A
C



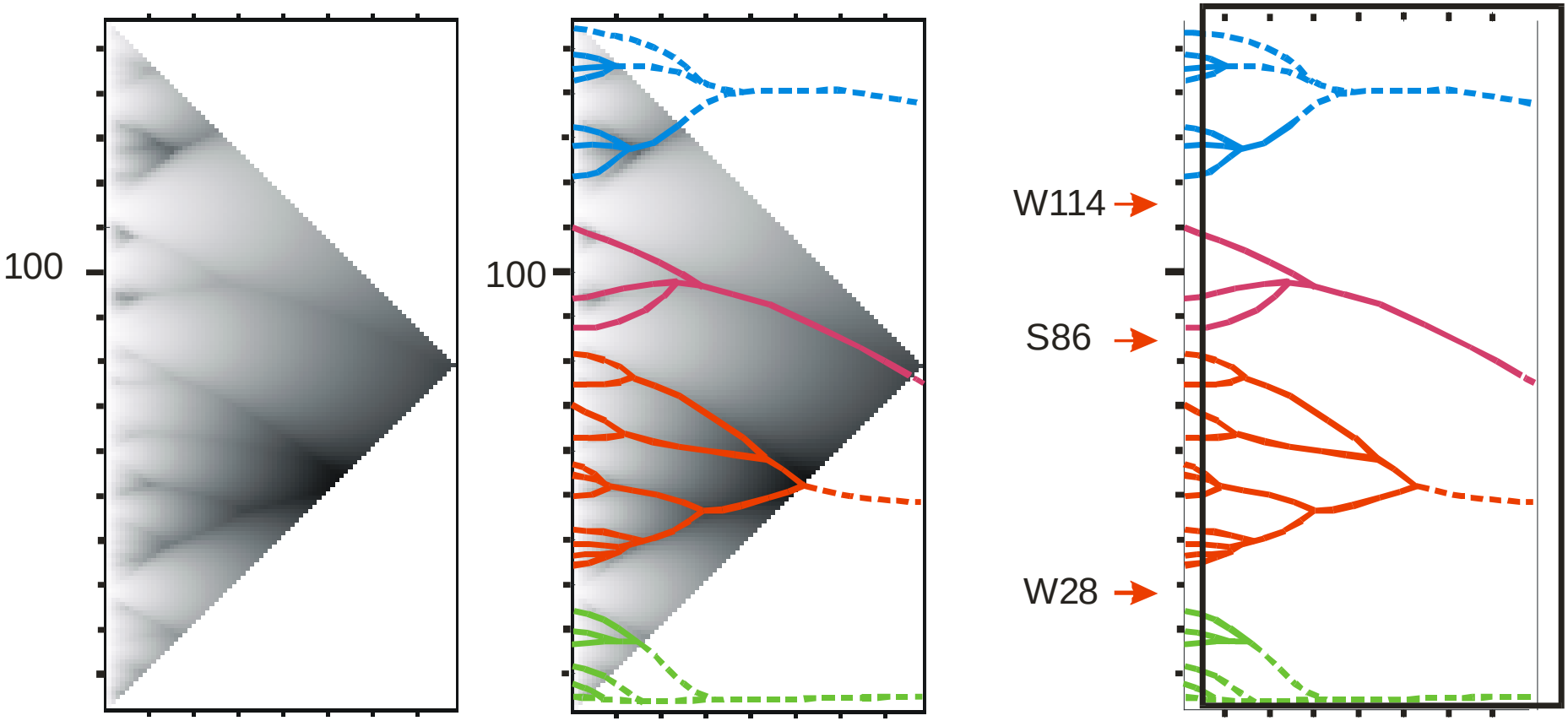


SPATIAL STRUCTURE OF TUMOUR NECROSIS FACTOR (TNF)



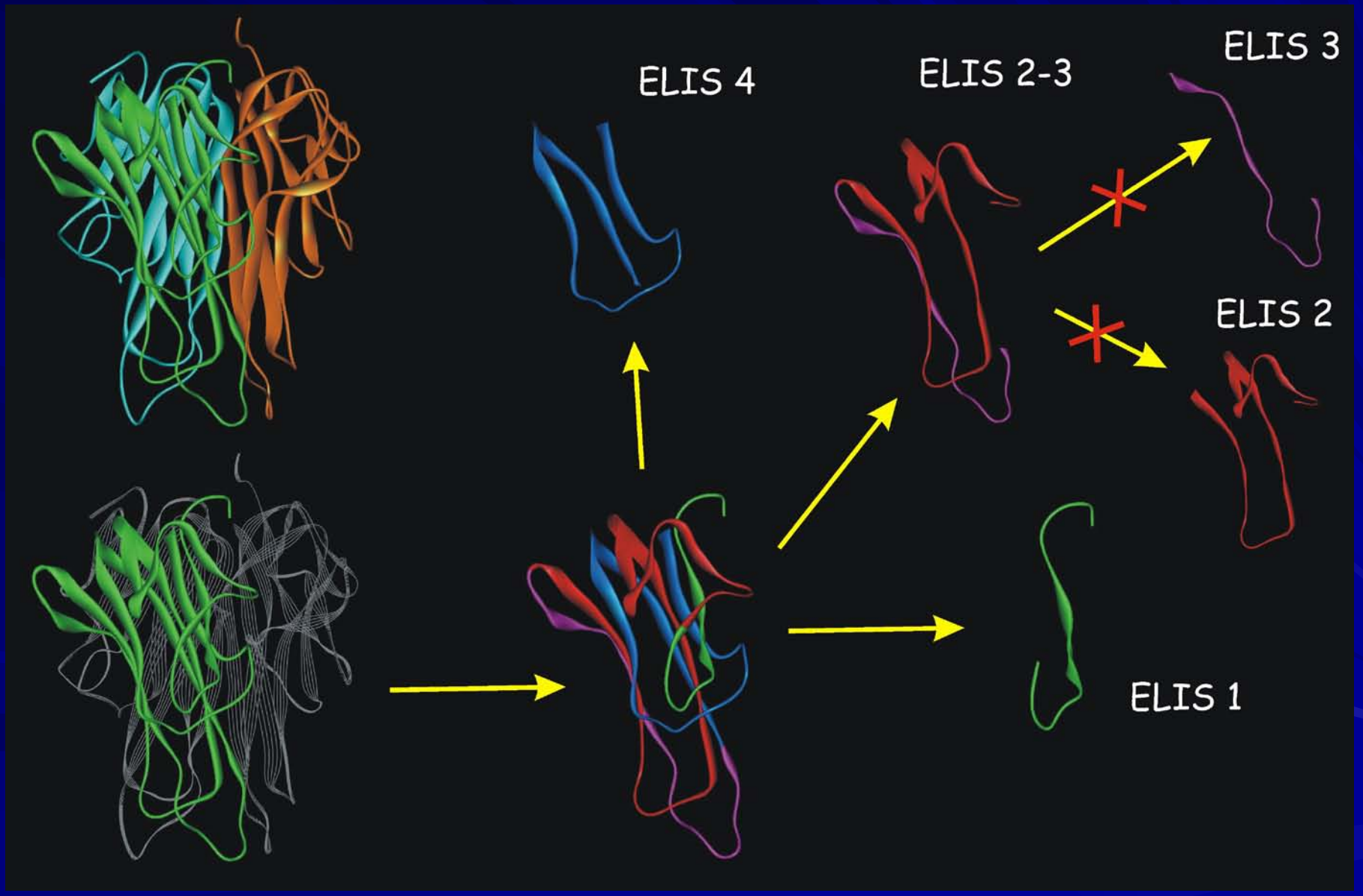
Bioorg Khim. (2010) vol. 36, p. 327 (RUS)

INFORMATIONAL STRUCTURE OF TNF



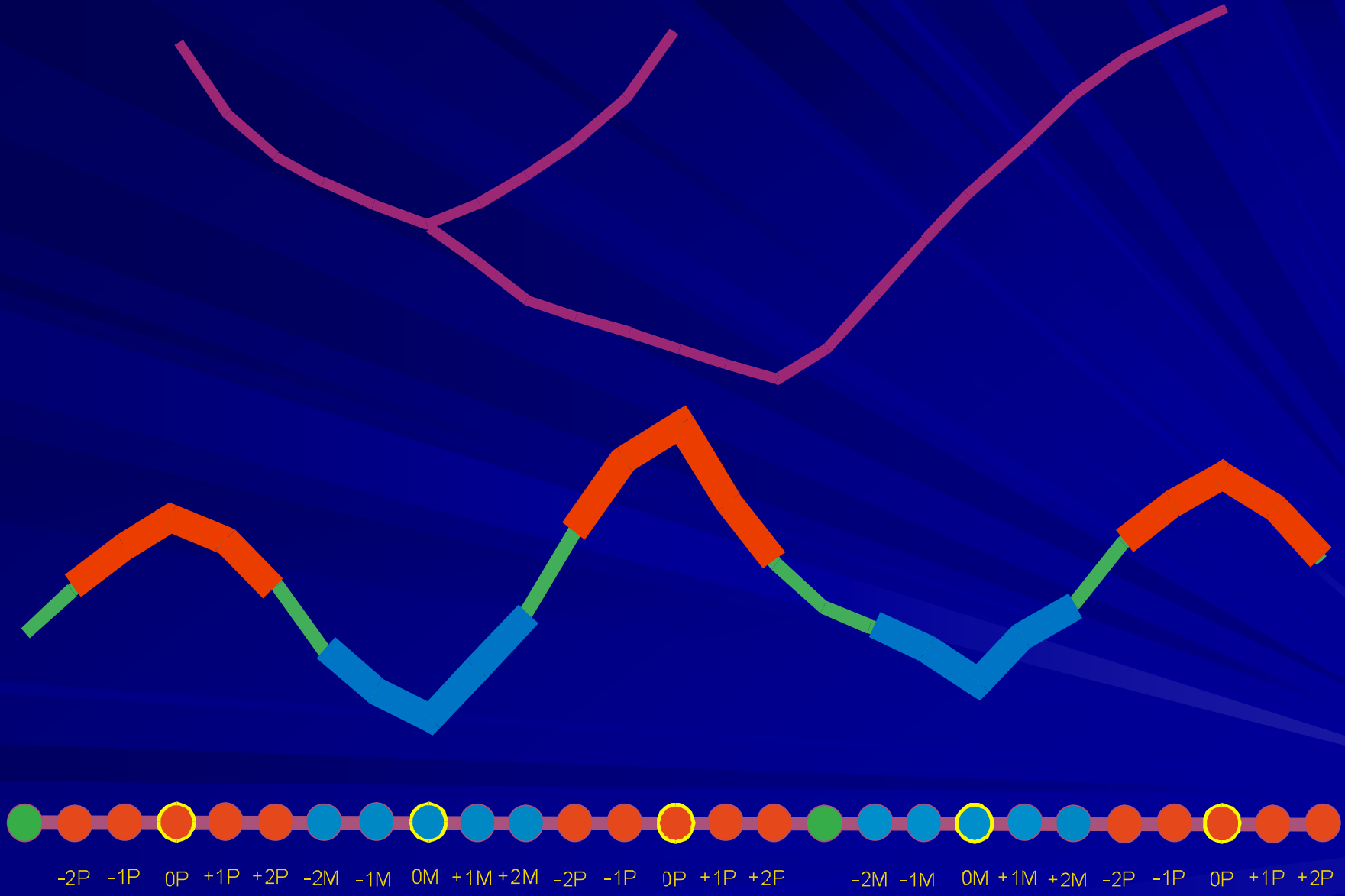
Bioorg Khim. (2010) vol. 36, p. 327 (RUS)

TNF ELIS FRAGMENTS

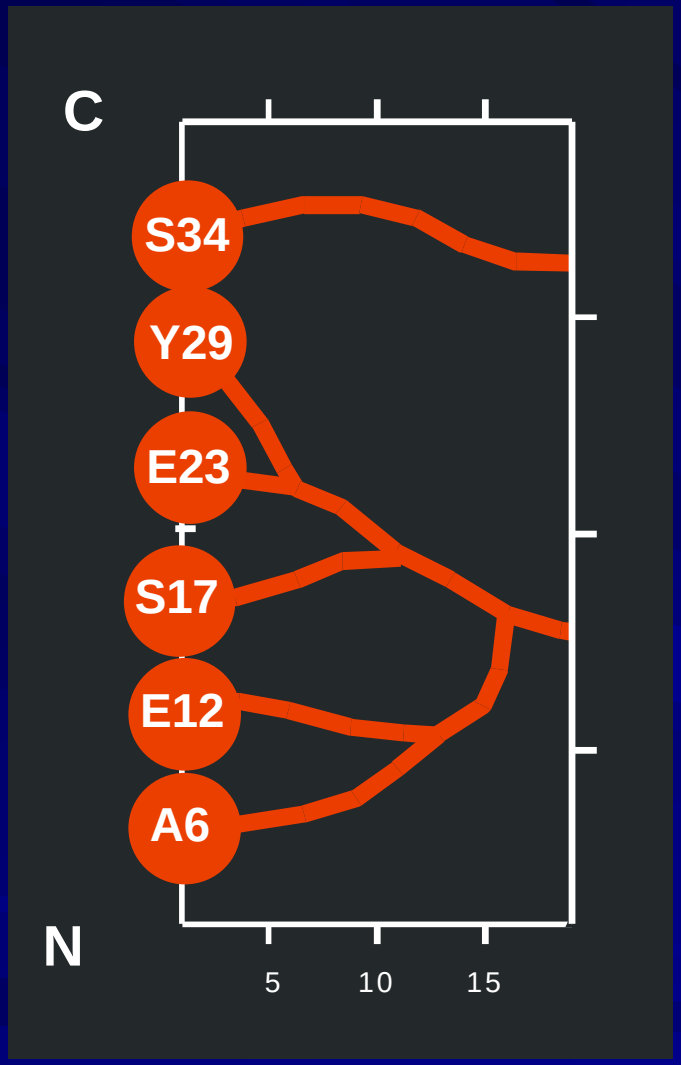


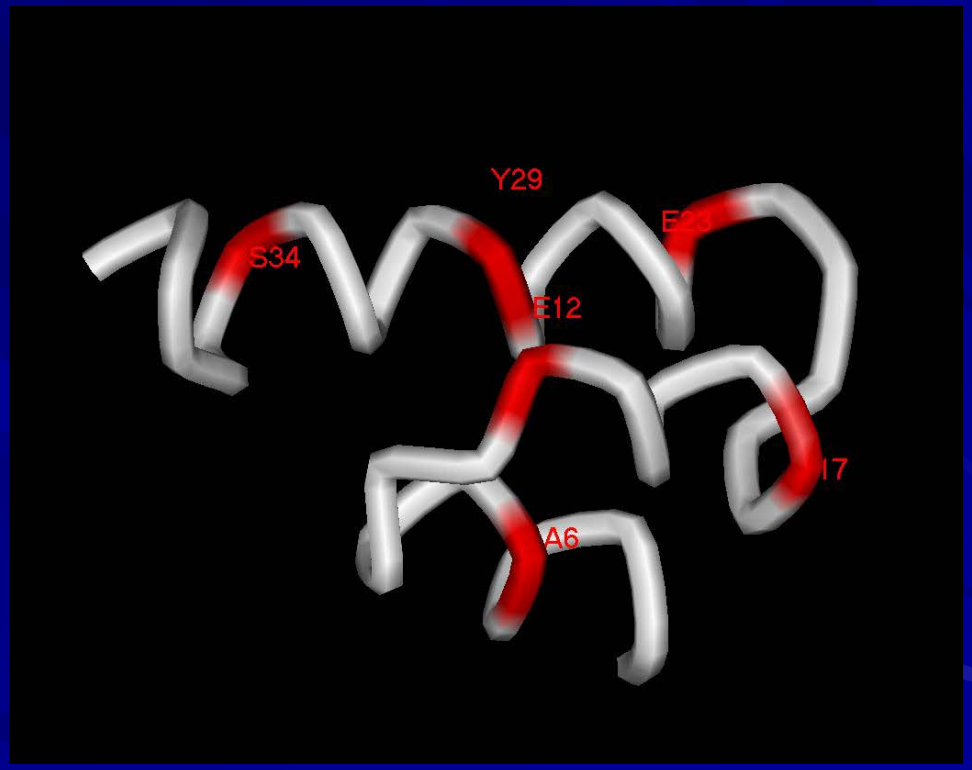
STRUCTURE FORMING AND STRUCTURE STABILAZING PENTAPEPTIDES

Protein Informational Profile



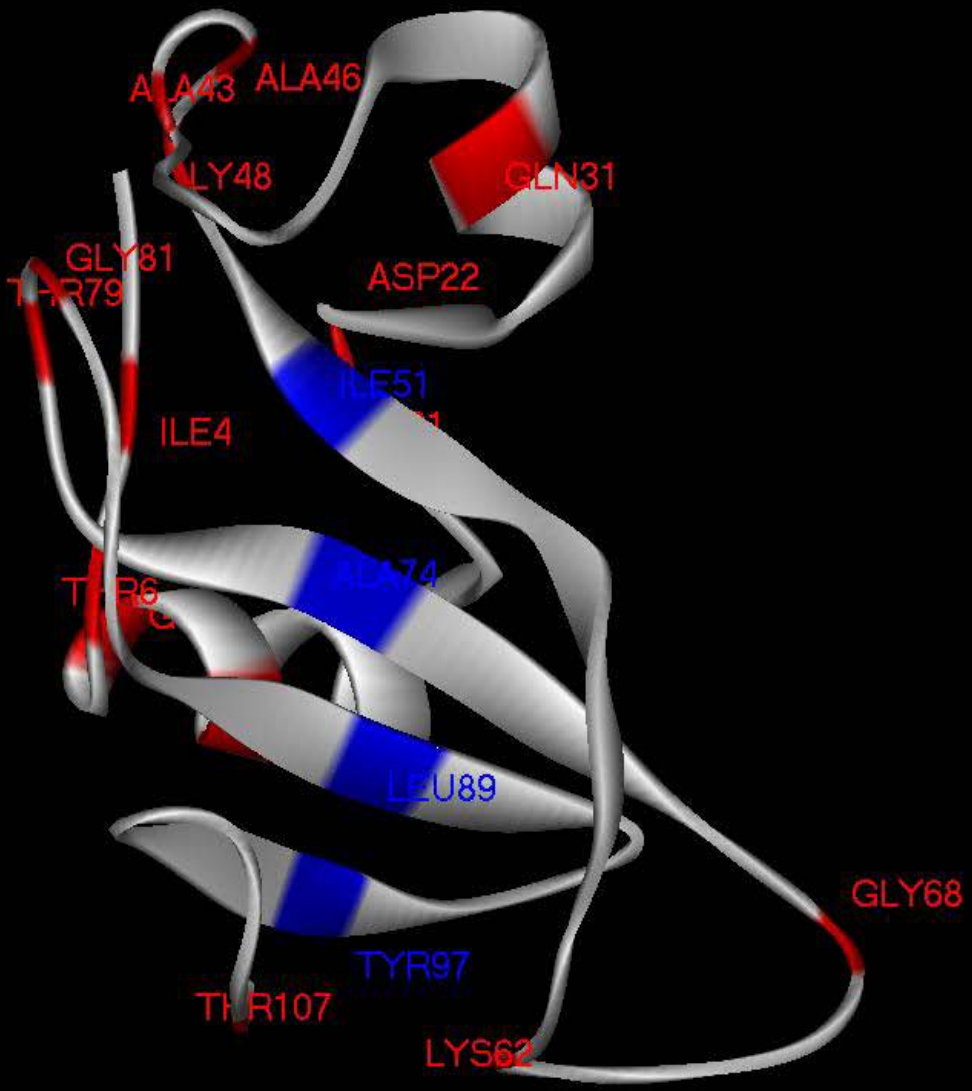
PHEROMONE ER-1 (2ERD.PDB)



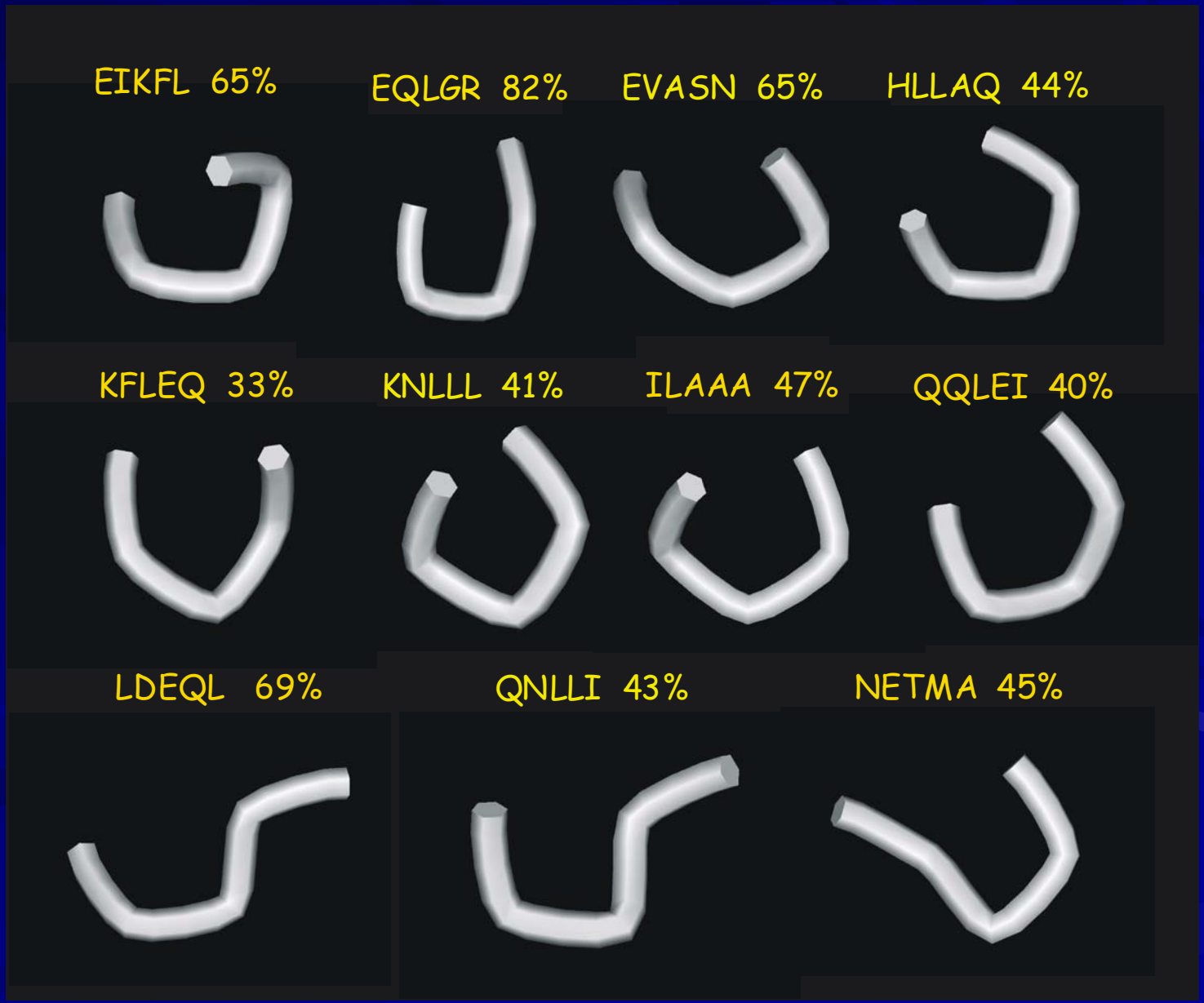


J Biomol Struct Dyn. vol. 28(1), p. 85.

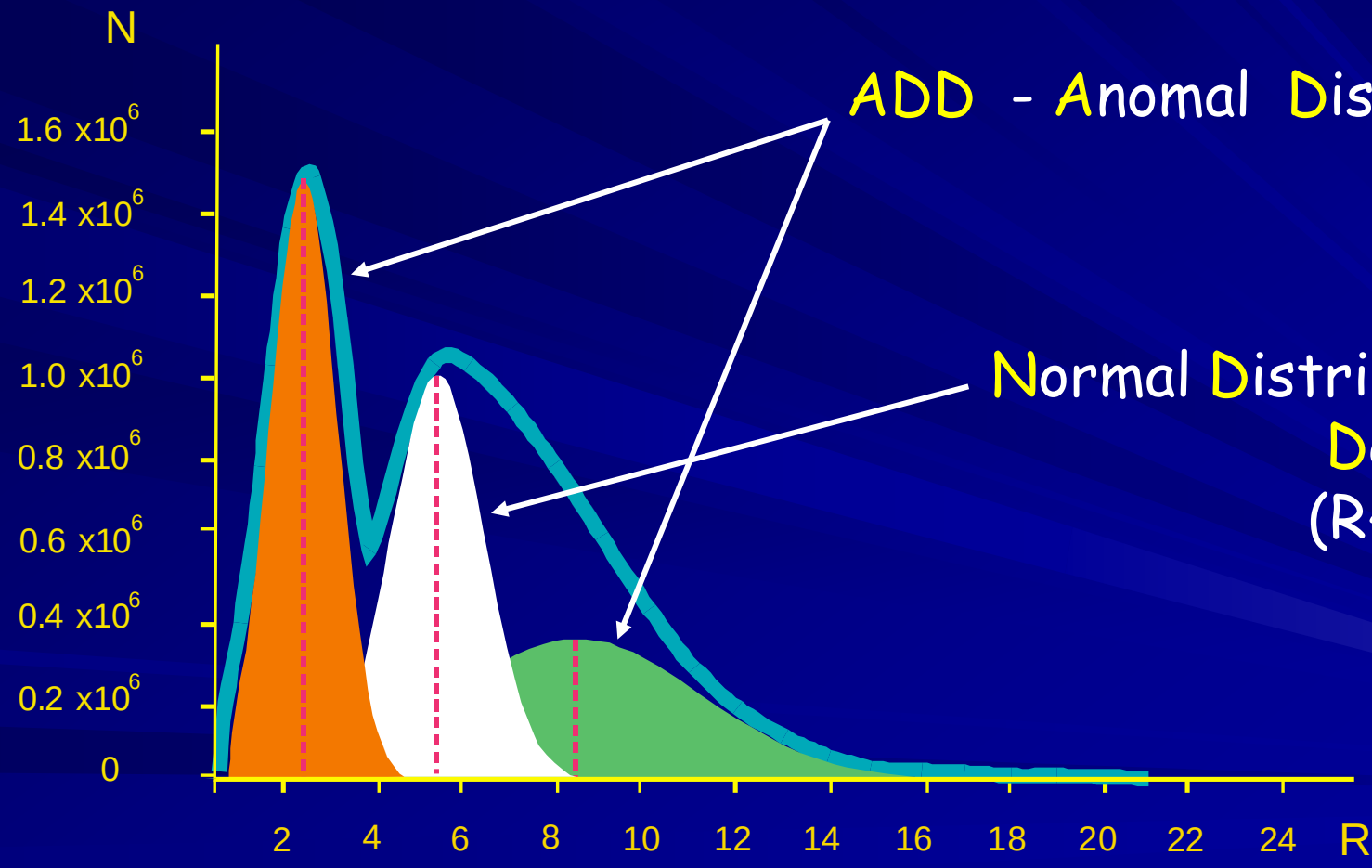
BARNASE (1BGS.PDB)



CONFORMATIONS OF LARGEST CLUSTERS OF FIRST RANK ELIS AFTER MD

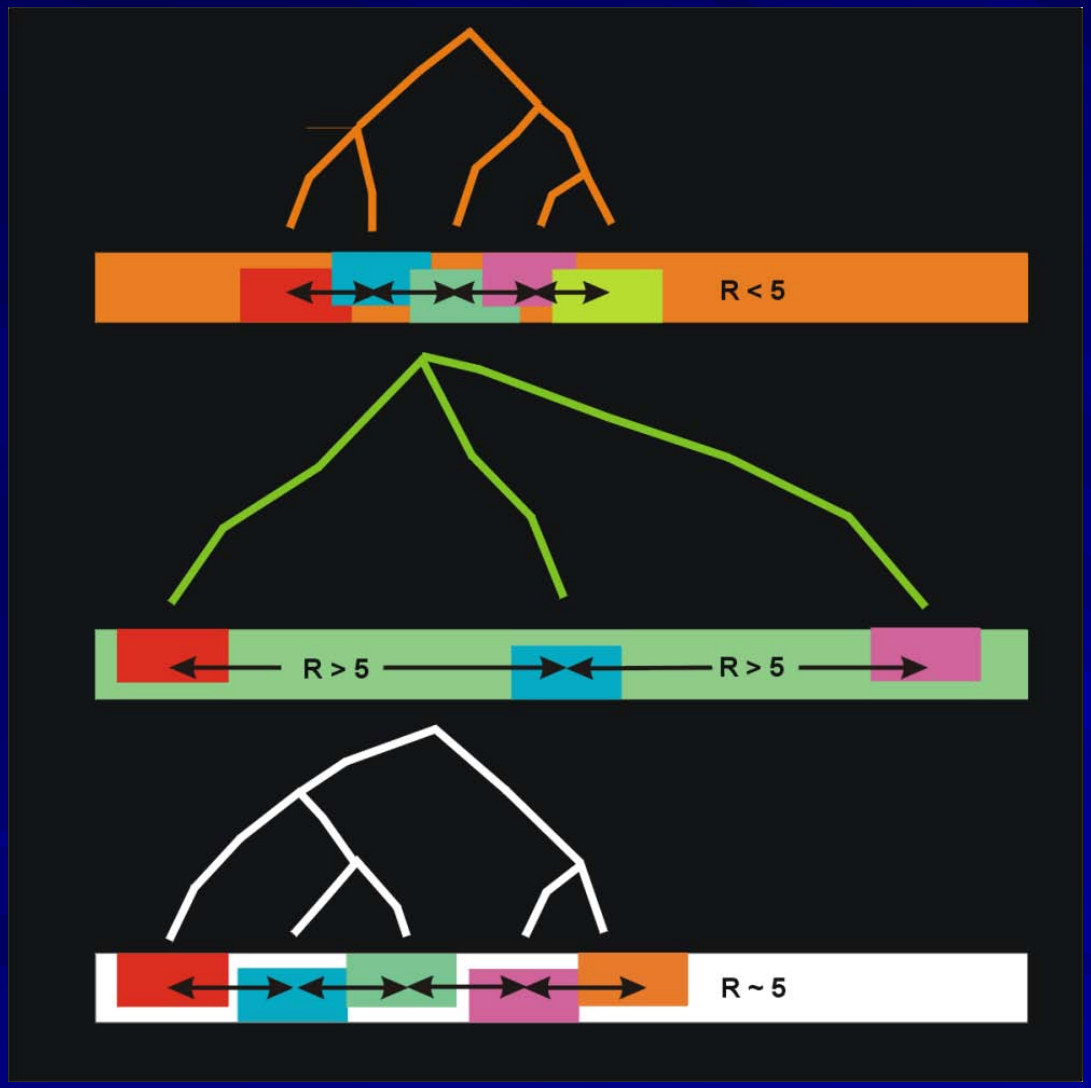


DISTRIBUTION OF FIRST RANK ELIS DISTANCE



J Biomol Struct Dyn. vol. 28(1), p. 85.

PROTEIN SITES CLASSIFICATION BY THE DENSITY OF FIRST RANK ELIS



ADD+

TOPOLOGICALLY CONSTRAINED SITE

ADD-

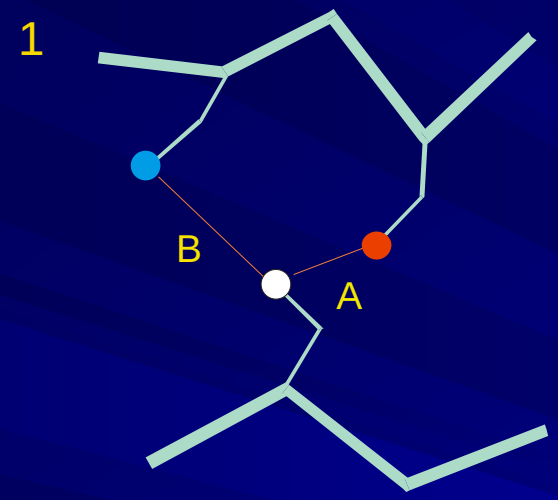
TOPOLOGICALLY VARIABLE SITE

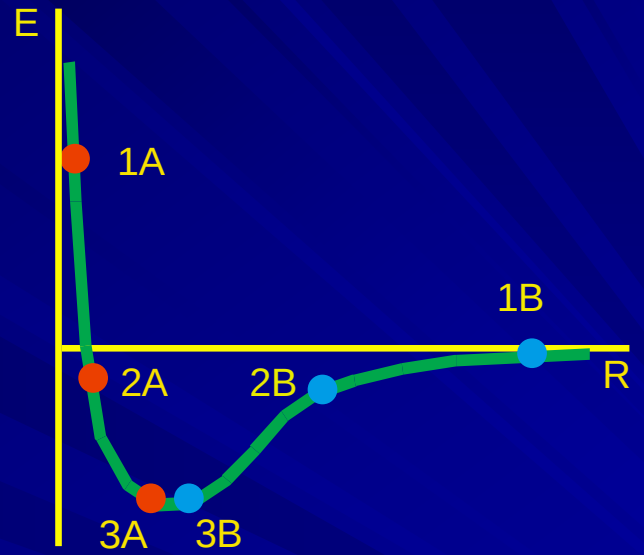
NORMAL

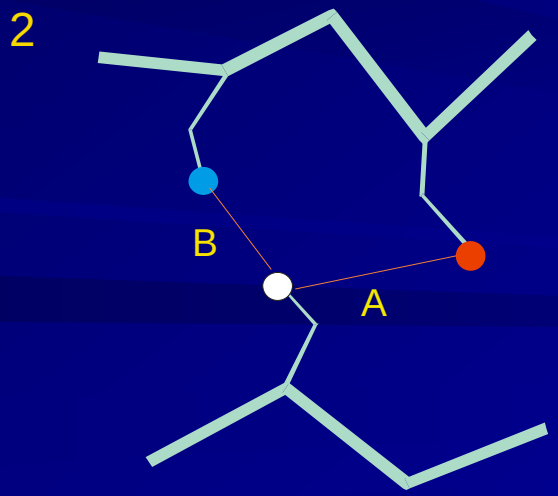
TOPOLOGICALLY STANDARD SITE

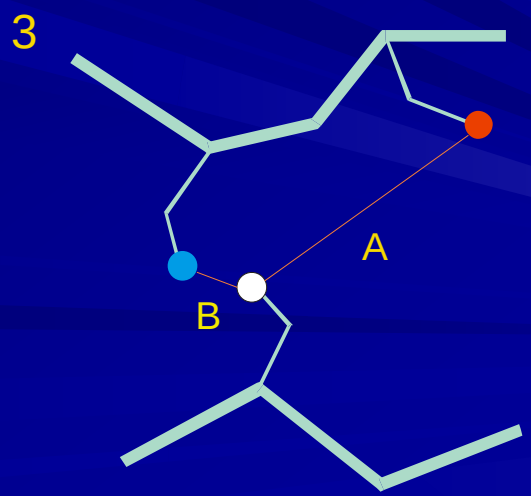
J Biomol Struct Dyn. vol. 28(1), p. 85.

ADAPTABLE CONFORMATIONAL CHANGES WHILE POLYPEPTIDE CHAINS INTERACTION









FACTORS THAT INFLUENCE ON INTERACTION BETWEEN POLYPEPTIDE CHAINS

- EXISTENCE OF FUNCTIONAL GROUPS OF AMINO ACID RESIDUES THAT PROVIDE INTERACTION

- ABILITY OF POLYPEPTIDE CHAIN TO CHANGE CONFORMATION IN ORDER TO PROVIDE OPTIMUM DISTANCE FOR INTERACTING FUNCTIONAL GROUPS

Conclusions

- 1. Proposed and statistically justified a new paradigm of the structural organization of proteins, according to which the basic element of a naturally occurring amino acid sequence if information unit.
- 2. On the basis of the paradigm developed a method of information structure analysis (ANIS) of proteins. The method allows to reveal hierarchically organized structural elements of the informational structure (ELIS) in the amino acid sequences.
- 3. Demonstrated examples of the ANIS method applications to explain the mechanism of proteins function. In works on protein engineering the method was used for the design of amino acid sequences with modified physico-chemical and functional characteristics, but retains the ability to self-organize.

Acknowledgements

- This work was partially supported by grant 12-04-01776-a of Russian Foundation for Basic Research

AUTHORS

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry RAS Moscow

Alexei N. Nekrasov (ANIS method, IU, ELIS, TNF, IL-13, PrxVI)

Alexei A. Zinchenko (IU, IS, ELIS)

Lada E. Petrovskaya (IL-13, TNF)

Sergey I. Rogov (IU)

Vitaly V. Radchenko (PrxVI)

Tatiana M. Shuvaeva (PrxVI)

Ljudmila N. Shingarova (IL-13, TNF)

Valery M. Lipkin

Mikhail P. Kirpichnikov

Institute of Cell Biophysics RAS Pushchino

Vladimir I. Novoselov (PrxVI)

Eugenyi E. Fesenko

Steklov Mathematical Institute RAS Moscow

Igor V. Volovich

Sergey V. Kozyrev (ANIS method)

Engelhardt Institute of Molecular Biology RAS, Moscow

Anastasia A. Anashkina (IU, ELIS)

THANKS FOR YOUR ATTENTION

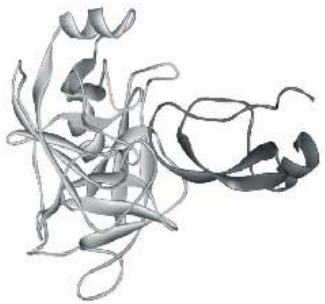
alexei_nekrasov@mail.ru

anastasya.anashkina@gmail.com

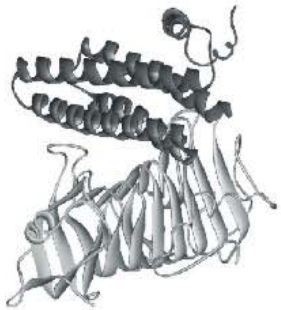
ОСОБЕННОСТИ ОРГАНИЗАЦИИ ИНТЕРФЕЙСОВ ВЗАИМОДЕЙСТВИЯ В БЕЛОК-БЕЛКОВЫХ КОМПЛЕКСАХ

SPATIAL STRUCTURES OF ENZYME-INHIBITOR COMPLEXES

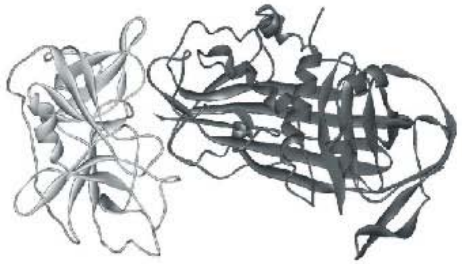
TRSP/BPTI
(3TGI.PDB)



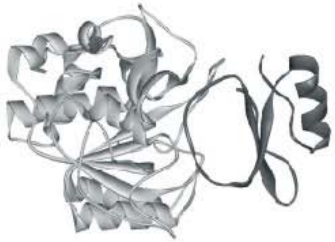
PME 1/PMEI
(1XG2.PDB)



PPE/α_1-PI
(2D26.PDB)

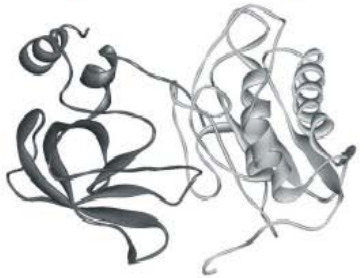


SUBT/CI-2A
(1A10.PDB)



RNase A/RI
(1DFJ.PDB)

MMP-3/TIMP-1
(1OO0.PDB)



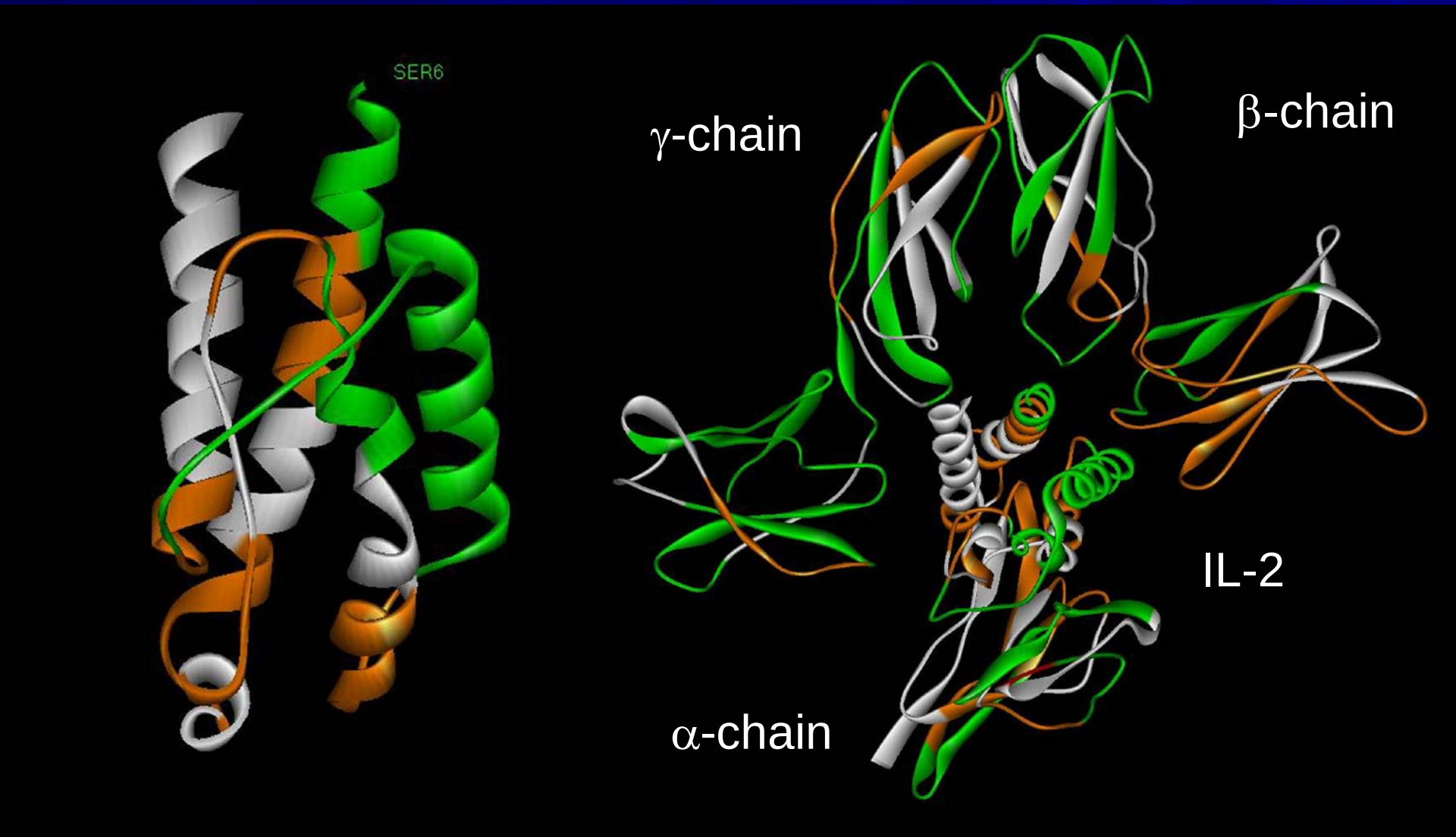
J Biomol Struct Dyn. vol. 28(1), p. 85.

IS TYPE CONTACTS REALIZED IN HYDROLASE-INGIBITOR COMPLEXES

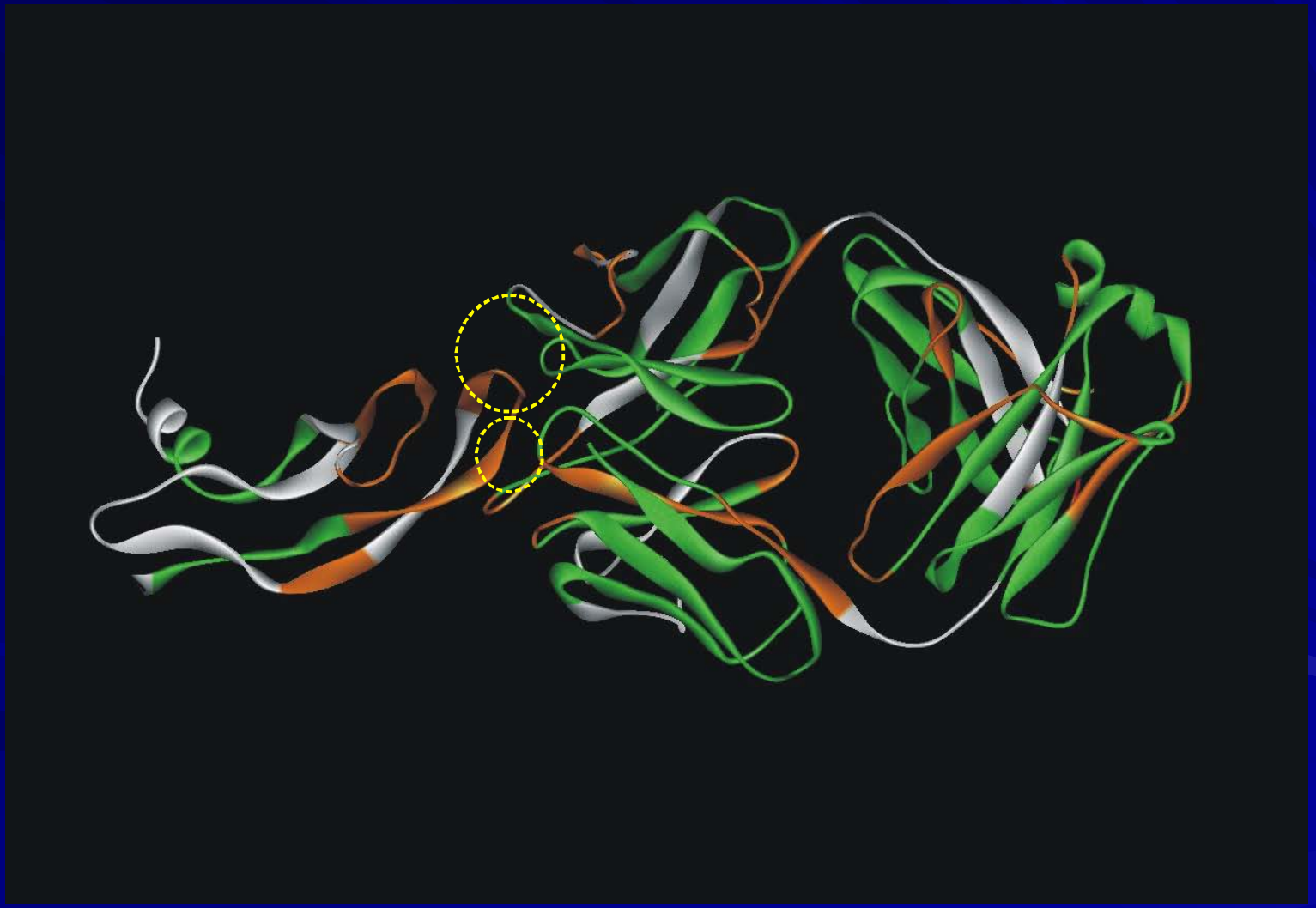
TYPE CONTACT	TRPS/BPTI		SUBT/CI2A		PME/PMEI		RNAse/RI		PPE/aPI		MMP3/TIMP		
ADD-/ADD+	93.6		39.0		33.8		31.4		14.3		9.6		
ADD-/NORM	0.0	95.2	33.7	72.7	31.0	85.9	23.5	96.1	14.3	92.9	46.9	62.8	84.2
ADD-/ADD-	1.6		0.0		21.1		41.2		64.3		6.3		
NORM/NORM	0.0		11.7		5.6		0.0		0.0		24.6		
ADD+/NORM	3.2	4.8	13.0	27.3	8.5	14.1	3.9	3.9	7.1	7.1	12.6	37.2	15.8
ADD+/ADD+	1.6		2.6		0.0		0.0		0.0		0.0		

J Biomol Struct Dyn. vol. 28(1), p. 85.

INTERLEUKIN-2 (IL-2) & RECEPTOR COMPLEX (2B5I.PDB)



VASCULAR ENDOTHELIAL GROWTH FACTOR IN COMPLEX WITH A NEUTRALIZING ANTIBODY (1BJ1.PDB)

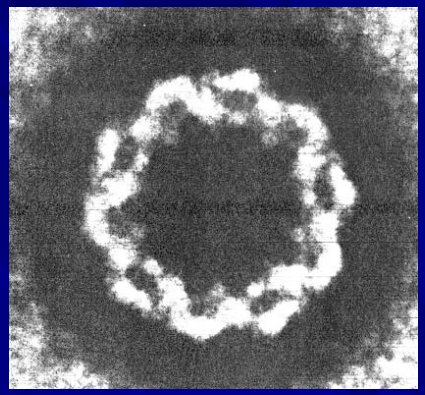


МЕТОД ПОЛУЧЕНИЯ РЕКОМБИНАНТНЫХ ВАКЦИН

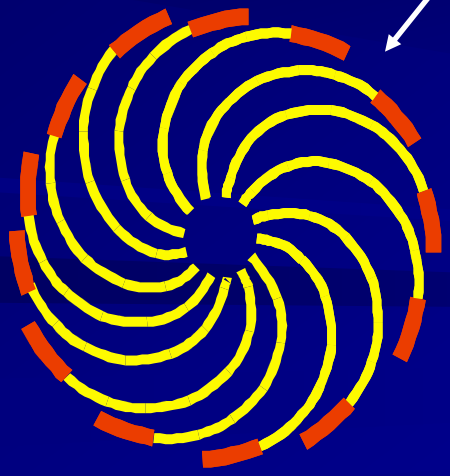
CREATION OF RECOMBINANT VACCINE

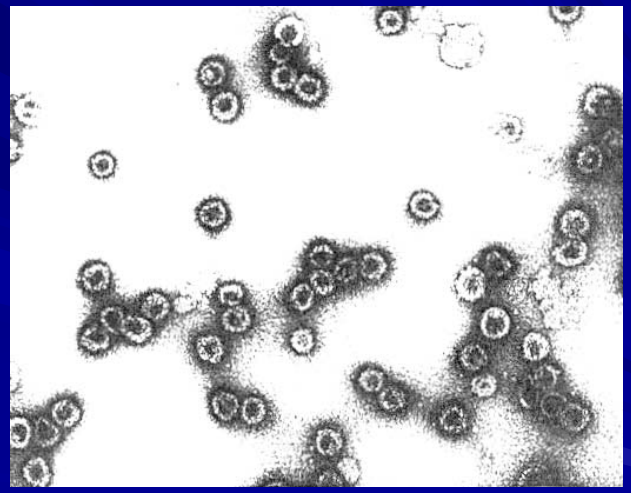
P1 protein from *Saccharomyces cerevsiae*

ADD+ site



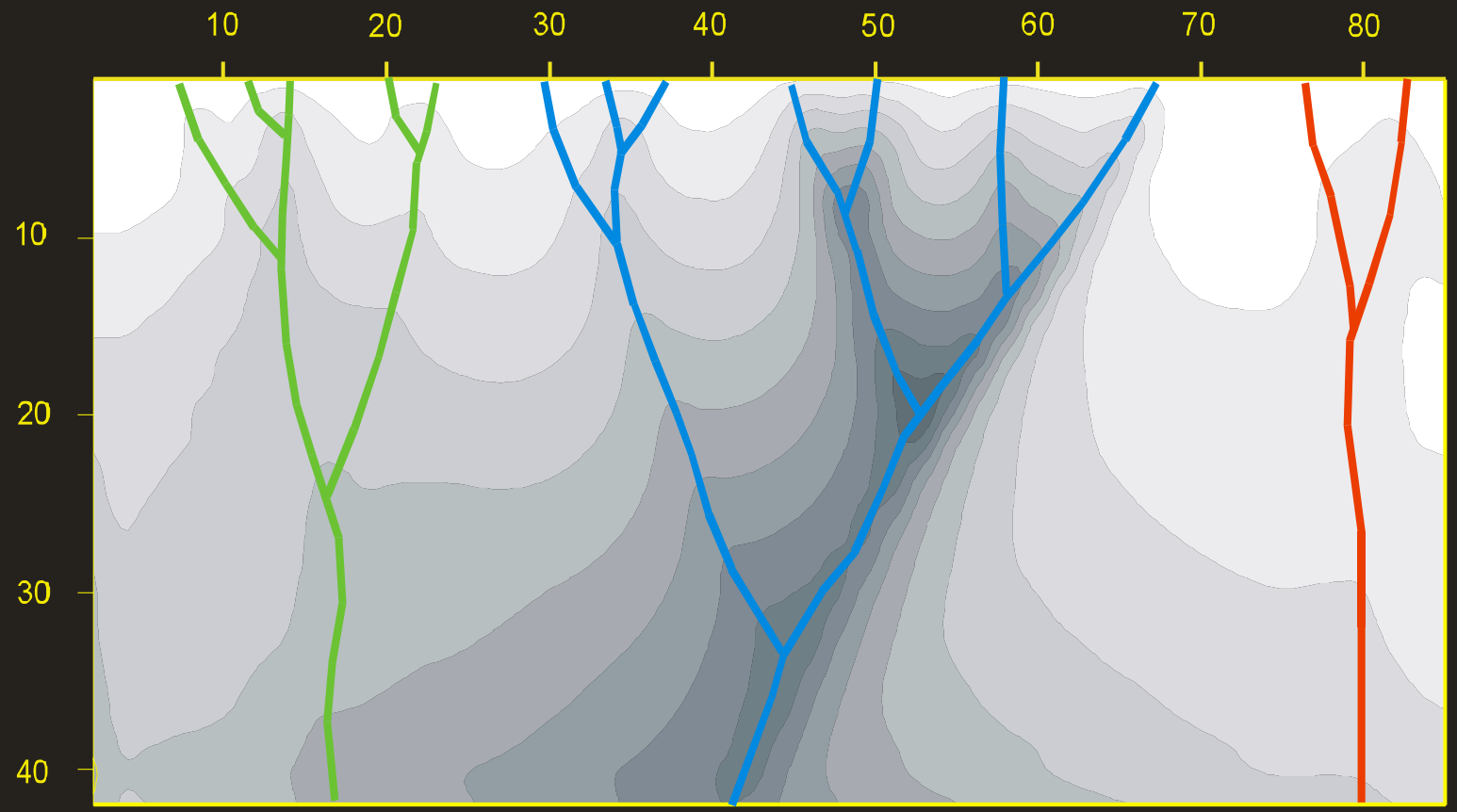
Virus-like particle





Vaccine v. 26, p. 2688

PROINSULINE INFORMATIONAL STRUCTURE



FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN

Beta- chain

C-peptide

Alpha-chain

DIRECTIONS OF POSSIBLE COLLABORATION

1. Development of tools for protein design and engineering tools.
2. Development of a method of hierarchical protein folding.
3. Computer modeling of spatial structure and investigate forming rules of protein subunits.
4. Design and modification of pharmacologically important proteins.
5. Research of protein structure organization and mechanisms of enzymes activity.
6. "Molecular tweezers" by fibronectin

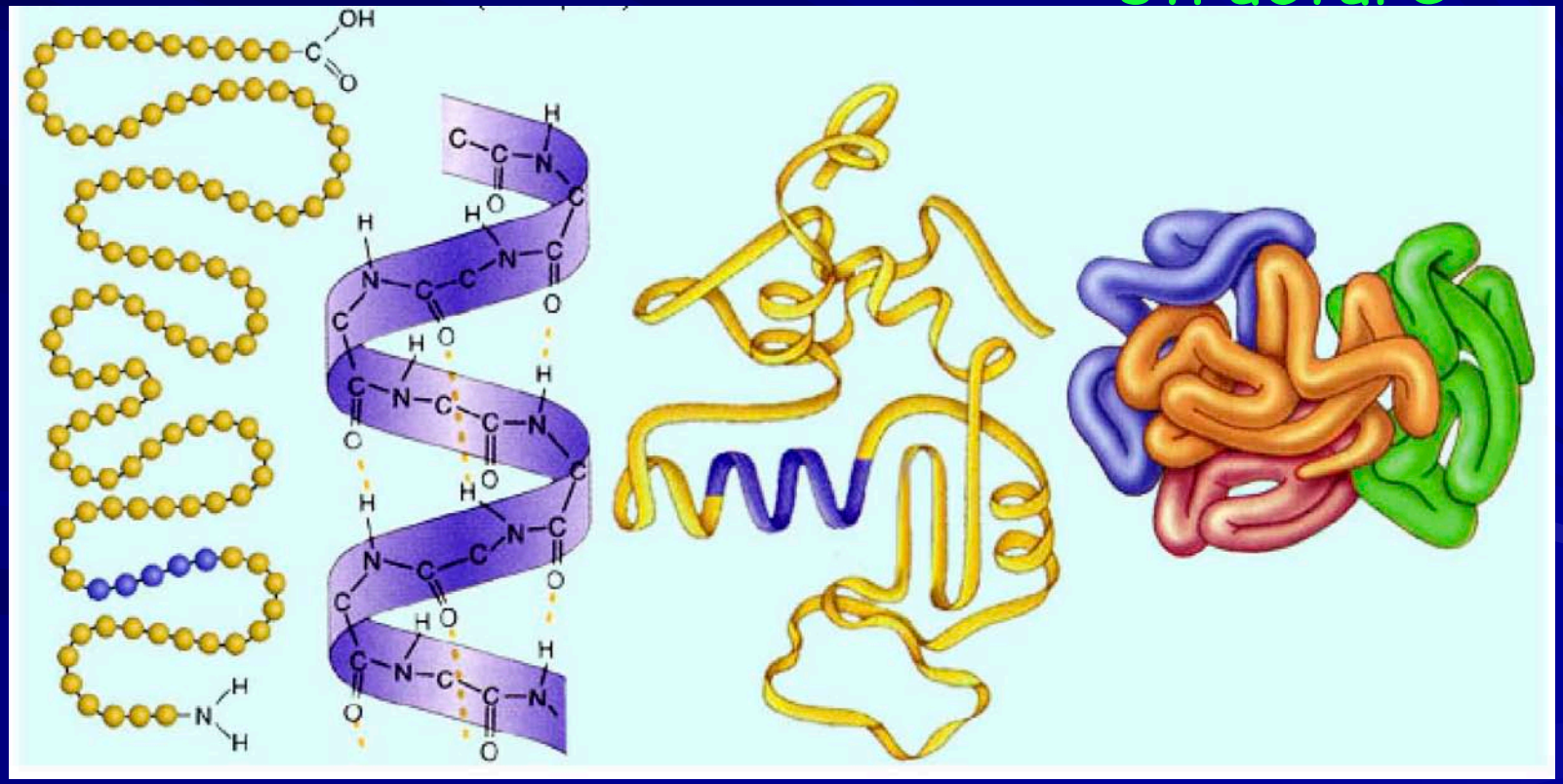
Classic view: levels of structural organization of proteins

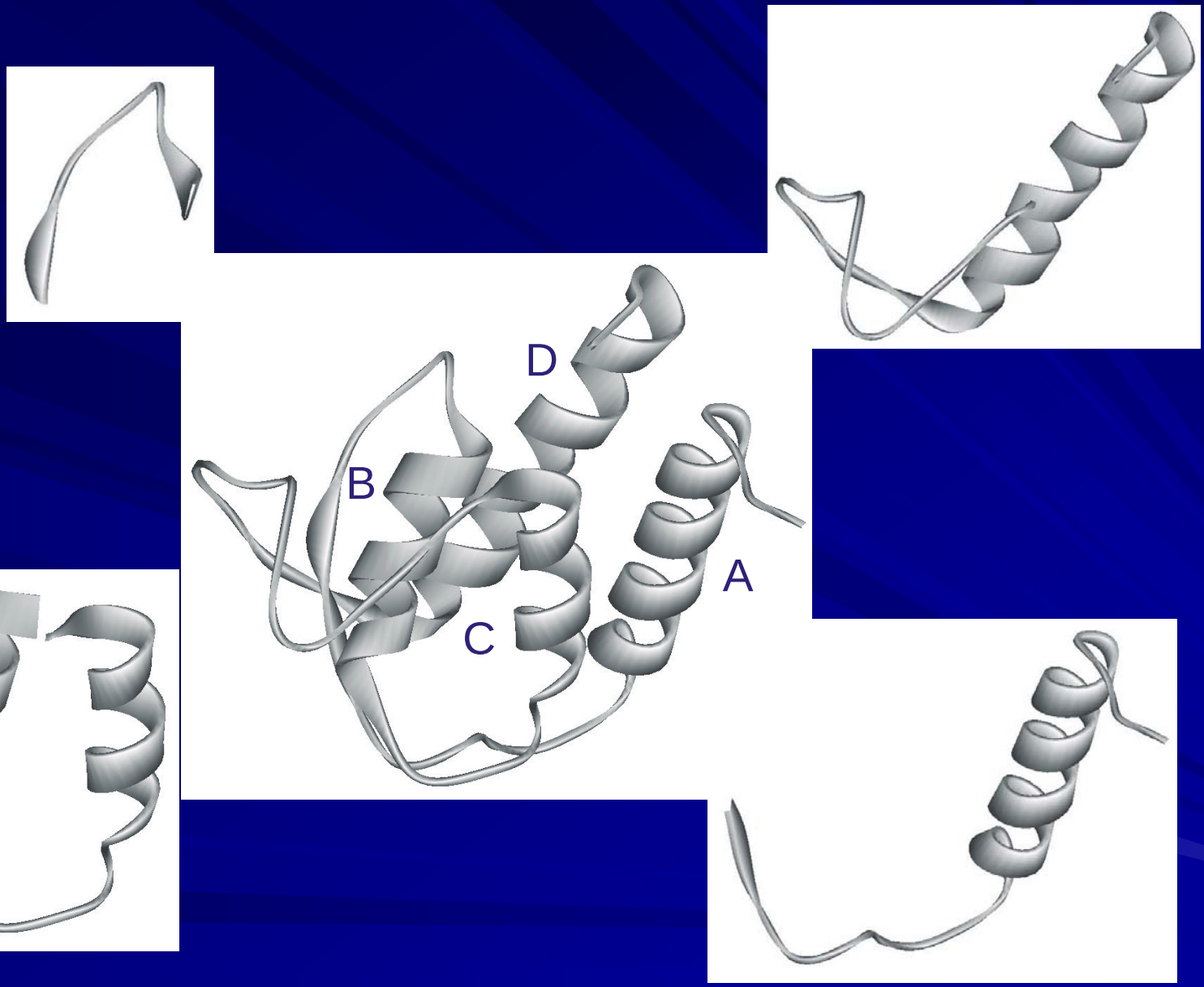
sequence

secondary structure

tertiary structure

quaternary structure





Biochemistry (Moscow), (2009) vol. 74, p. 493