NUCLEOTIDE BINDING ENZYMES

The Rossmann fold

Relationship between sequence, structure and function.

Anna Casas, Júlia Gasull and Nerea Vega

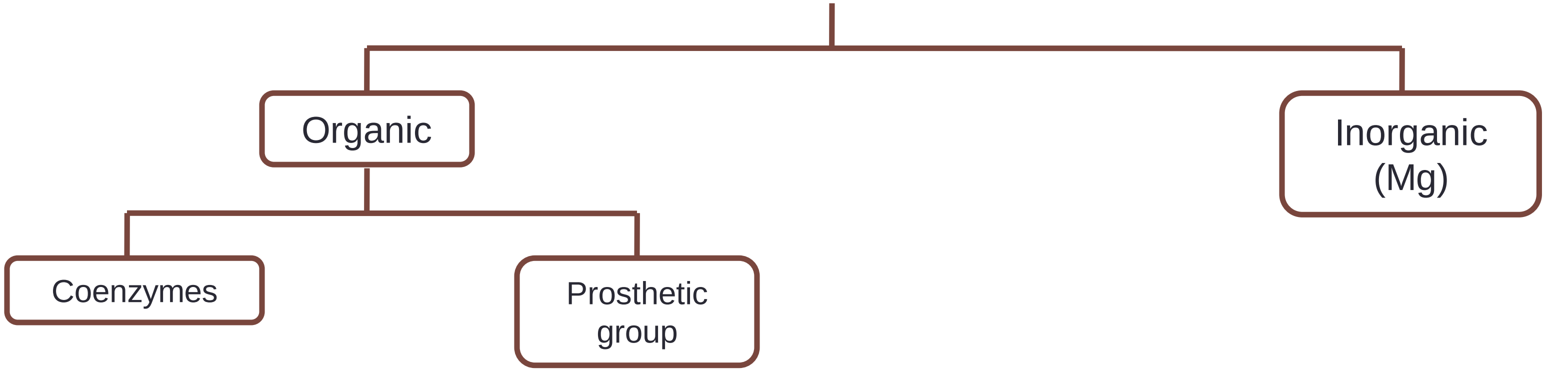
Index

INTRODUCTION

Adenine-nucleotides

All organisms posses small molecular weight co-factors with a crucial role in several metabolic and regulatory pathways.

Co-factor: non-protein chemical compound that is bound to a protein and is required for the protein’s biological activity.



Organic

Inorganic (Mg)

Coenzymes

Prosthetic group

NAD and FAD serve as a cofactors in many essential biologic processes, such as glycolysis (NAD) and the citric acid cycle (FAD and NAD)

Schulz, G.E., Schirmer, R.H. Princiiples of protein structure

Index

The Rossmann Fold

Introduction to Rossmann fold

Rossmann et. al. (1974) described two **β-α-β-α-β** units forming a **six-stranded parallel β-sheet** flanked by **four α-helices** in the structure of some dinucleotide-binding proteins.

⇩

ONE ROSSMANN FOLD UNIT = **β-α-β-α-β**.

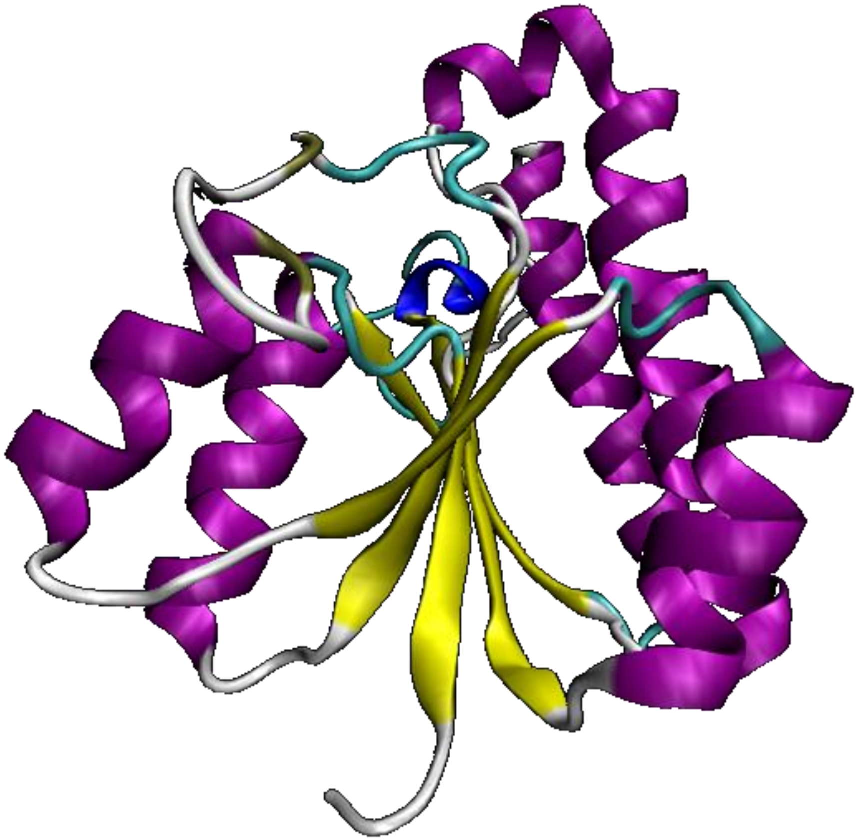


Fig.1 Dinucleotide-binding motif (Rossman fold)

Schulz, G.E., Schirmer, R.H. Princiiples of protein structure

The Rossmann Fold

β-α-β motif

Two adjacent β-strands in the aminoacid sequence are joined by an α-helix at opposite edges to form a **parallel β-strand** in the structure.

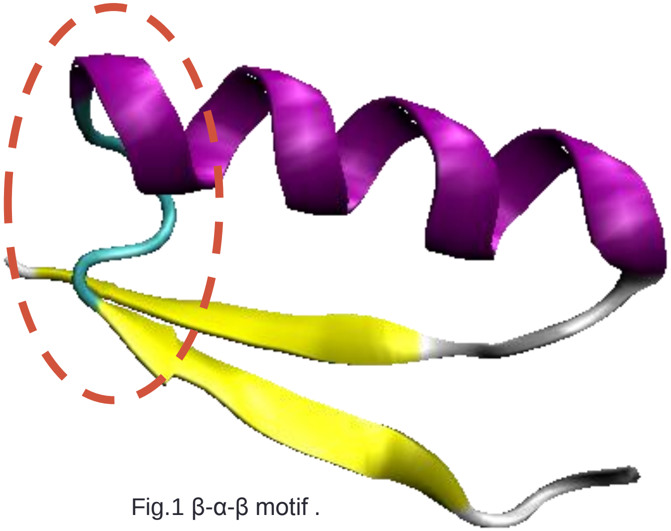
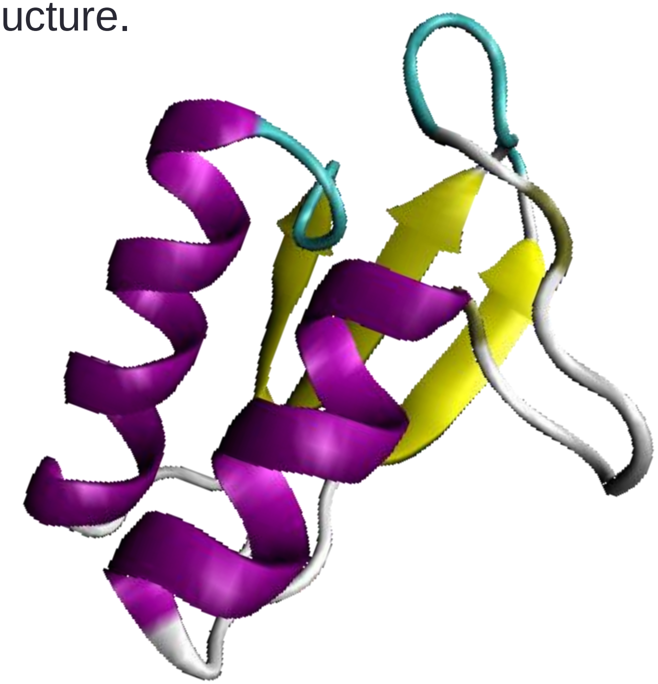


Fig.1 β-α-β motif .



The **loop** connecting C-end of the β-strand with N-end of the α-helix often have conserved amino acid sequence in homologous proteins and is involved in forming the **active site.**

Right-handed.

Schulz, G.E., Schirmer, R.H. Princiiples of protein structure

The Rossmann Fold

Dinucleotide-binding fold

Rossman fold is a super-secondary structural open sheet domain composed of alternating α-helices and β-strands along the backbone. The β-strands are therefore mostly *parallel*.

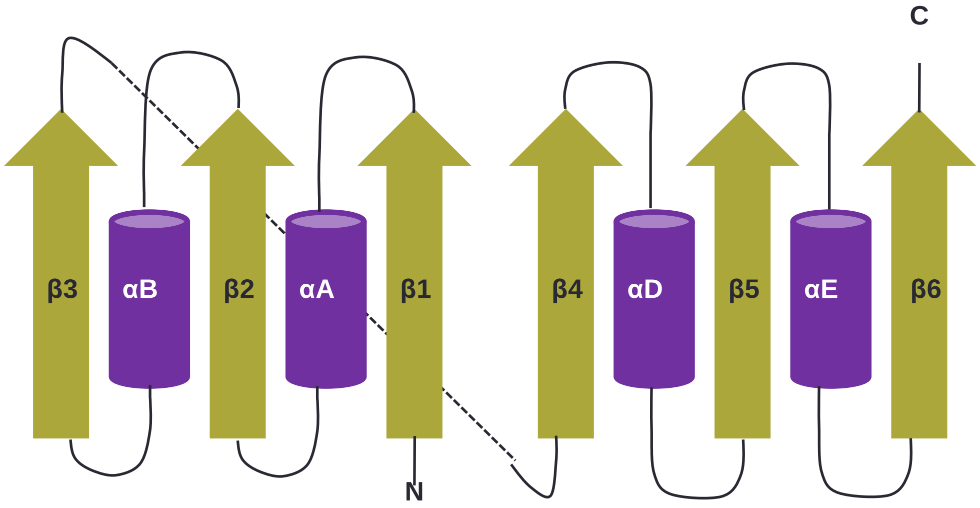


Fig.2 Topological switch diagram.

The **active site** is easy to find in a topological diagram

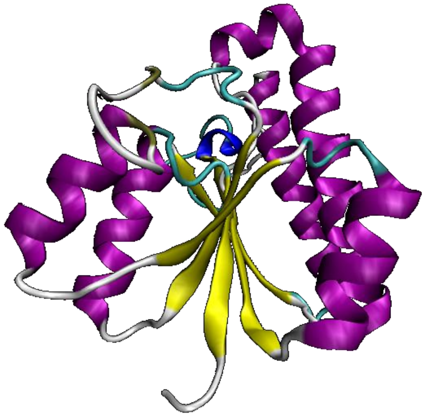


Fig.3Dinucleotide-binding motif (Rossman fold)

Carl Branden, John Tooze. Introduction to protein strructure

The Rossmann Fold

NAD-binding Rossmann fold

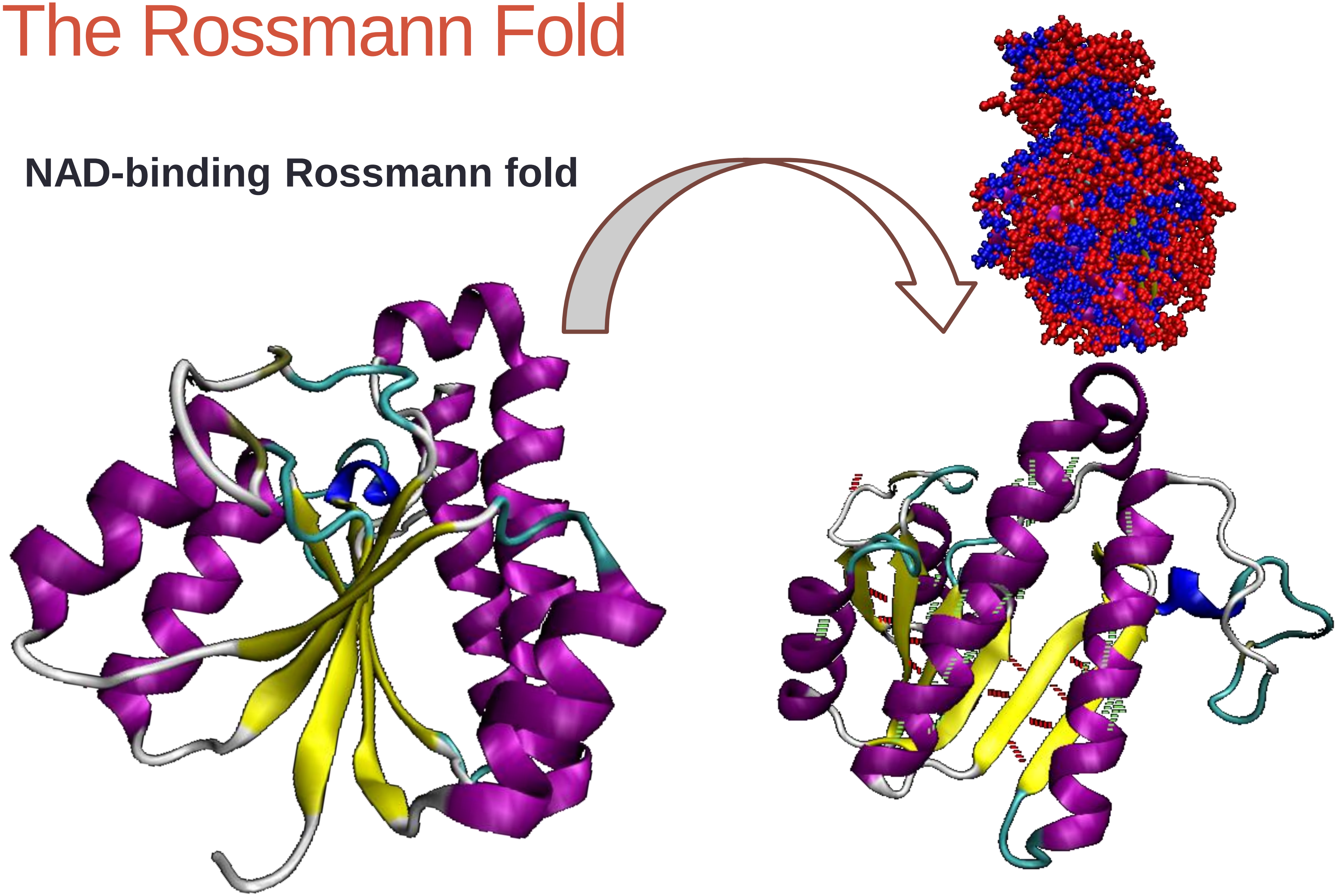
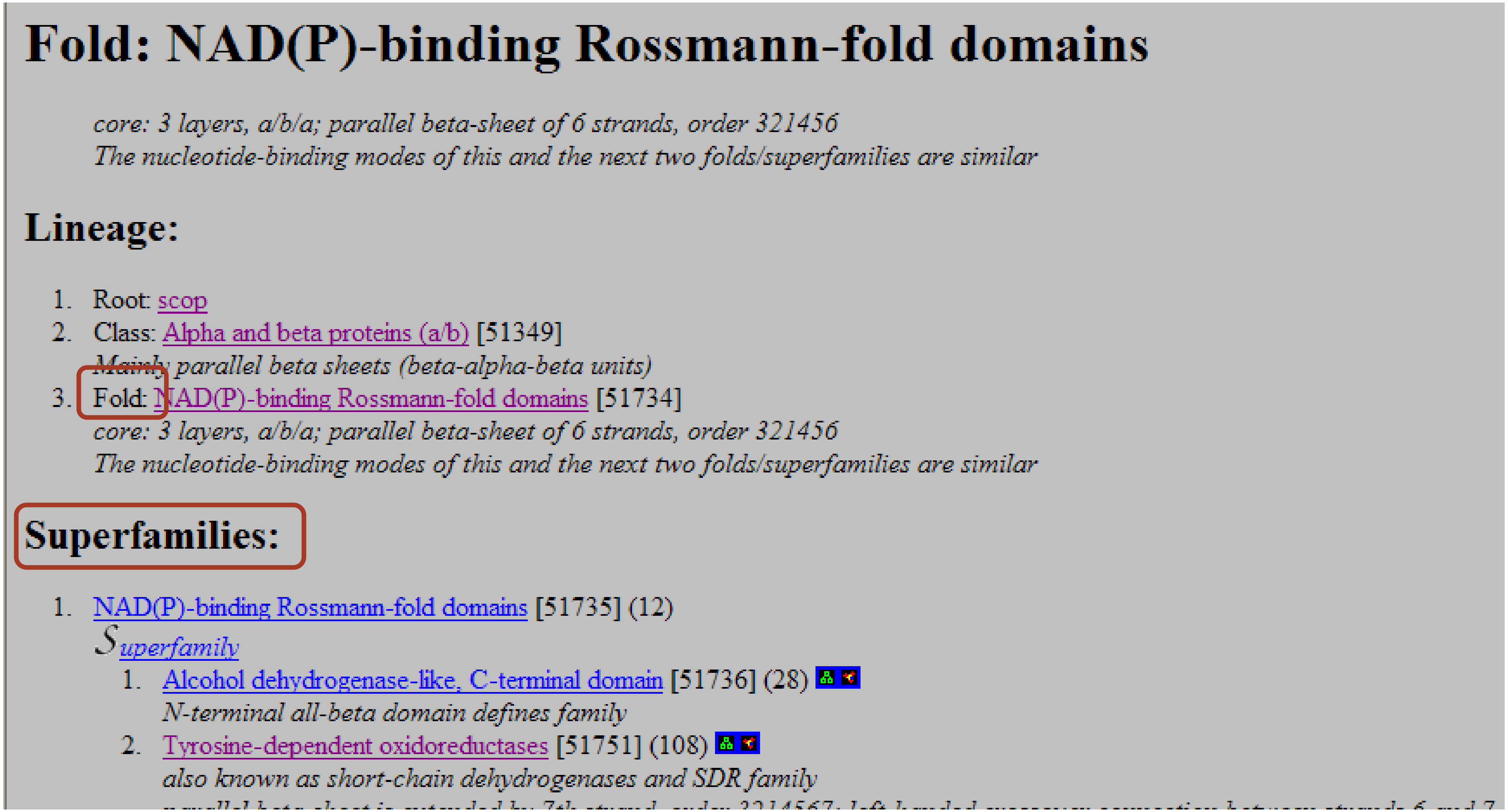


Fig.4 Rossmann fold (1EK5)

Carl Branden, John Tooze. Introduction to protein strructure

The Rossmann Fold

SCOP classification



Fold: NAD(P)-binding Rossmann-fold domains

core: 3 layers, a/b/a; parallel beta-sheet of 6 strands, order 321456
The nucleotide-binding modes of this and the next two folds/superfamilies are similar

Lineage:

1. Root: scop
2. Class: Alpha and beta proteins (a/b) [51349]
   Mainly parallel beta sheets (beta-alpha-beta units)
3. Fold: NAD(P)-binding Rossmann-fold domains [51734]
   core: 3 layers, a/b/a; parallel beta-sheet of 6 strands, order 321456
   The nucleotide-binding modes of this and the next two folds/superfamilies are similar

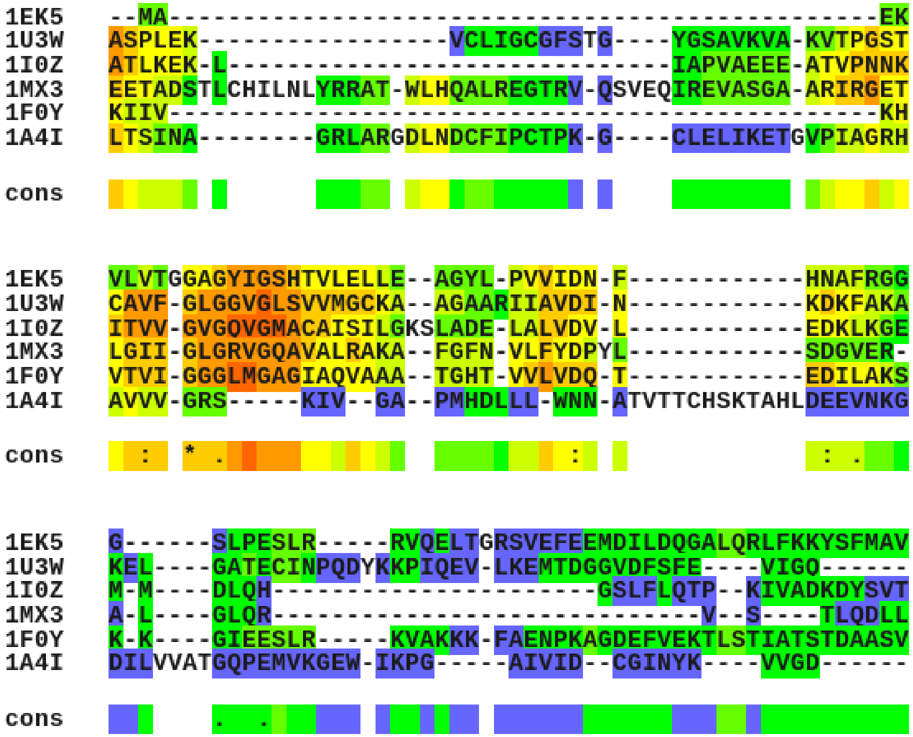
Superfamilies:

1. NAD(P)-binding Rossmann-fold domains [51735] (12)
   Superfamily
   1. Alcohol dehydrogenase-like, C-terminal domain [51736] (28)
      N-terminal all-beta domain defines family
   2. Tyrosine-dependent oxidoreductases [51751] (108)
      also known as short-chain dehydrogenases and SDR family

The Rossmann Fold

Rossmann fold

Although dinucleotide-binding domains show very low overall sequence homology, large portions of their proteins backbones superimpose very well due to Rossman fold.

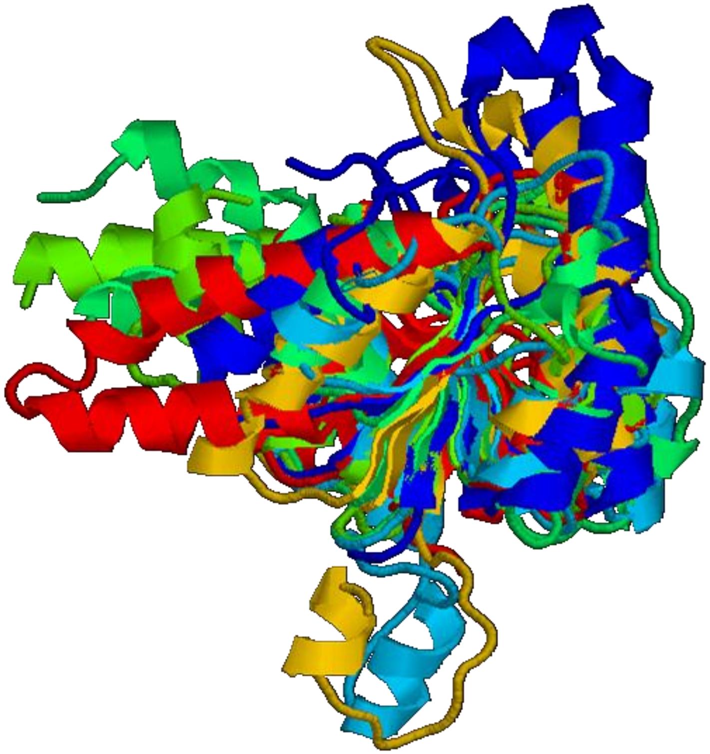


```
1EK5    --MA------------------------------------------------EK
1U3W    ASPLEK-----------------VCLIGCGFSTG----YGSAVKVA-KVTPGST
1I0Z    ATLKEK-L------------------------------IAPVAEEE-ATVPNNK
1MX3    EETADSTLCHILNLYRRAT-WLHQALREGTRV-QSVEQIREVASGA-ARIRGET
1F0Y    KIIV------------------------------------------------KH
1A4I    LTSINA--------GRLARGDLNDCFIPCTPK-G----CLELIKETGVPIAGRH

cons

1EK5    VLVTGGAGYIGSHTVLELLE--AGYL-PVVIDN-F------------HNAFRGG
1U3W    CAVF-GLGGVGLSVVMGCKA--AGAARIIAVDI-N------------KDKFAKA
1I0Z    ITVV-GVGQVGMACAISILGKSLADE-LALVDV-L------------EDKLKGE
1MX3    LGII-GLGRVGQAVALRAKA--FGFN-VLFYDPYL------------SDGVER-
1F0Y    VTVI-GGGLMGAGIAQVAAA--TGHT-VVLVDQ-T------------EDILAKS
1A4I    AVVV-GRS-----KIV--GA--PMHDLLL-WNN-ATVTTCHSKTAHLDEEVNKG

cons      :  * .                       :                : .

1EK5    G------SLPESLR-----RVQELTGRSVEFEEMDILDQGALQRLFKKYSFMAV
1U3W    KEL----GATECINPQDYKKPIQEV-LKEMTDGGVDFSFE----VIGQ------
1I0Z    M-M----DLQH----------------------GSLFLQTP--KIVADKDYSVT
1MX3    A-L----GLQR-----------------------------V--S----TLQDLL
1F0Y    K-K----GIEESLR-----KVAKKK-FAENPKAGDEFVEKTLSTIATSTDAASV
1A4I    DILVVATGQPEMVKGEW-IKPG-----AIVID--CGINYK----VVGD------

cons           .  .
```

The Rossmann Fold

Structure is more conserved than sequence

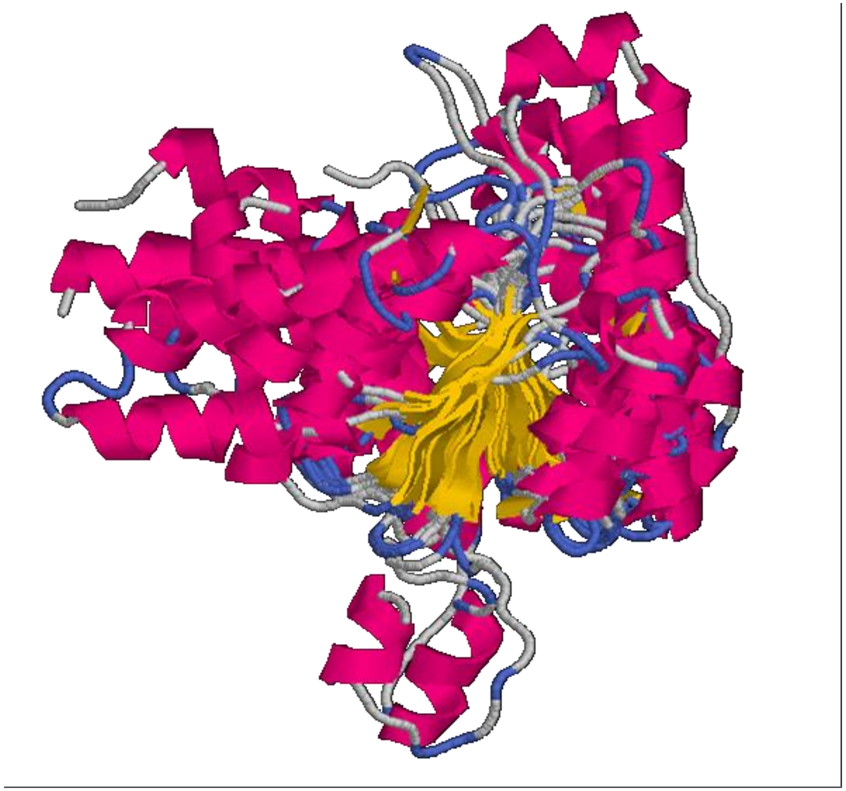
The degree of conservation of the three-dimensional structure is much higher than the degree of conservation of the amino acid sequence.



The Rossmann Fold

Structure is more conserved than sequence

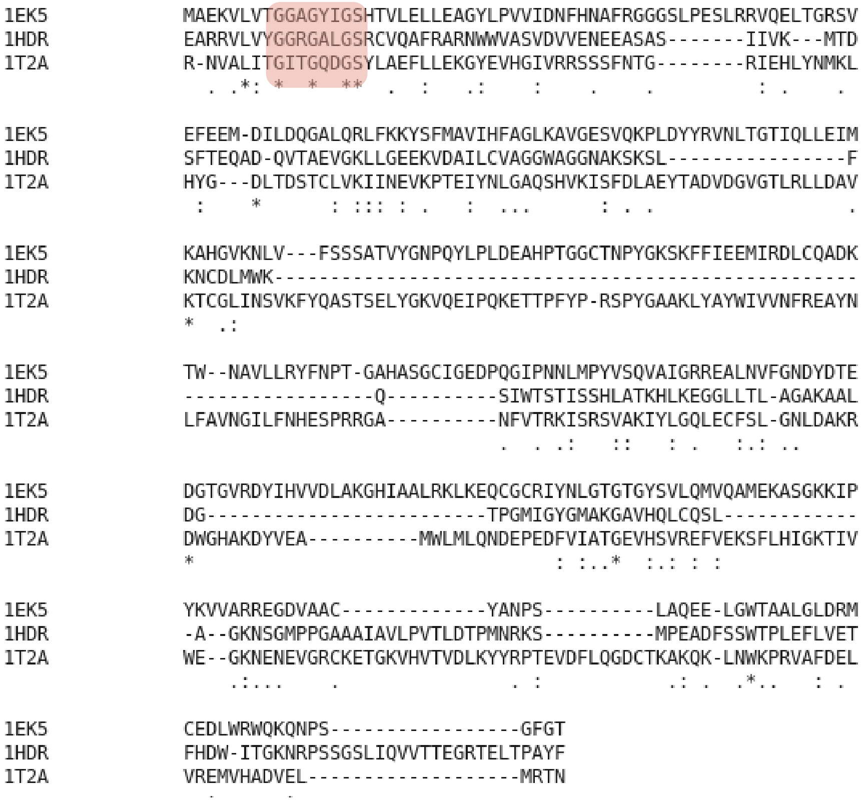
The degree of conservation of the three-dimensional structure is much higher than the degree of conservation of the amino acid sequence.



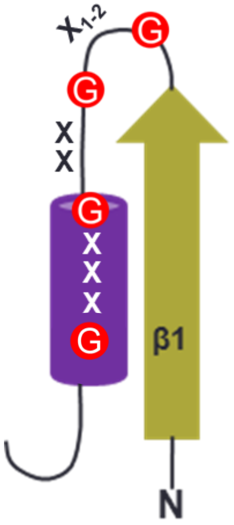
The Rossmann Fold

The Rossmann hypothesis

*A large number of proteins had incoporate the Rossmann fold as a consecuence of a **gene fusion** process. These proteins have adquired a new fuction*



```
1EK5        MAEKVLVTGGAGYIGSHTVLELLEAGYLPVVIDNFHNAFRGGGSLPESLRRVQELTGRSV
1HDR        EARRVLVYGGRGALGSRCVQAFRARNWWVASVDVVENEEASAS-------IIVK---MTD
1T2A        R-NVALITGITGQDGSYLAEFLLEKGYEVHGIVRRSSSFNTG--------RIEHLYNMKL
              . .*: *  *  **  .  :   .:    :    .    .         : .    .

1EK5        EFEEM-DILDQGALQRLFKKYSFMAVIHFAGLKAVGESVQKPLDYYRVNLTGTIQLLEIM
1HDR        SFTEQAD-QVTAEVGKLLGEEKVDAILCVAGGWAGGNAKSKSL----------------F
1T2A        HYG---DLTDSTCLVKIINEVKPTEIYNLGAQSHVKISFDLAEYTADVDGVGTLRLLDAV
             :    *      : ::: : .   :  ...      : . .                 .

1EK5        KAHGVKNLV---FSSSATVYGNPQYLPLDEAHPTGGCTNPYGKSKFFIEEMIRDLCQADK
1HDR        KNCDLMWK----------------------------------------------------
1T2A        KTCGLINSVKFYQASTSELYGKVQEIPQKETTPFYP-RSPYGAAKLYAYWIVVNFREAYN
            *  .:

1EK5        TW--NAVLLRYFNPT-GAHASGCIGEDPQGIPNNLMPYVSQVAIGRREALNVFGNDYDTE
1HDR        -----------------Q----------SIWTSTISSHLATKHLKEGGLLTL-AGAKAAL
1T2A        LFAVNGILFNHESPRRGA----------NFVTRKISRSVAKIYLGQLECFSL-GNLDAKR
                                        .  . .:   ::   : .   :.: ..

1EK5        DGTGVRDYIHVVDLAKGHIAALRKLKEQCGCRIYNLGTGTGYSVLQMVQAMEKASGKKIP
1HDR        DG-------------------------TPGMIGYGMAKGAVHQLCQSL------------
1T2A        DWGHAKDYVEA----------MWLMLQNDEPEDFVIATGEVHSVREFVEKSFLHIGKTIV
            *                                : :..*  :.: : :

1EK5        YKVVARREGDVAAC-------------YANPS----------LAQEE-LGWTAALGLDRM
1HDR        -A--GKNSGMPPGAAAIAVLPVTLDTPMNRKS----------MPEADFSSWTPLEFLVET
1T2A        WE--GKNENEVGRCKETGKVHVTVDLKYYRPTEVDFLQGDCTKAKQK-LNWKPRVAFDEL
                .:...    .               . :           .: .  .*..   : .

1EK5        CEDLWRWQKQNPS-----------------GFGT
1HDR        FHDW-ITGKNRPSSGSLIQVVTTEGRTELTPAYF
1T2A        VREMVHADVEL-------------------MRTN
              .      .
```



UDP-galactose-4-epimerase (GALE)

A nucleotide-binding protein

Nterminal (Met1-Thr189) →NAD binding

7-stranded parallel β-sheet
5-α-helix
} Rossmann fold

Cterminal (Gly190-Ala348) → UDPsugar binding

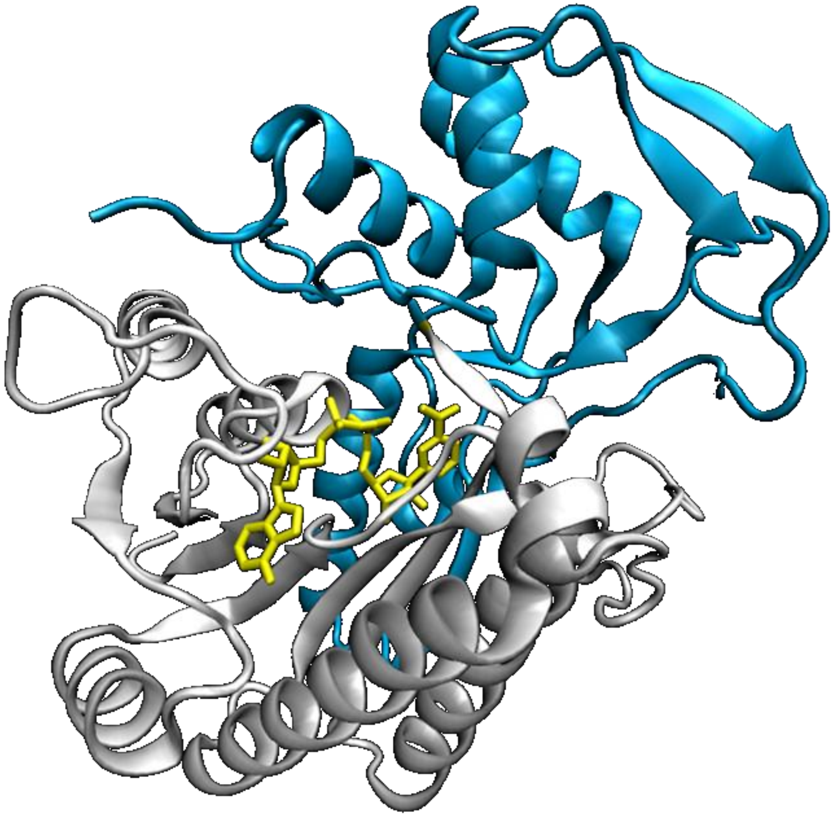


Fig.5 Single Nterminal and Cterminal GALE domains

Conserved features among these proteins:

- GXGXXG motif

-Side chains interactions with the dinucleotide

UDP-GALACTOSE **GALE** → UDP-GLUCOSE

Thoden BJ, et al. 2000.; Dym O, et al. 2001.

NAD-binding proteins

Nicotinamide adenine dinucleotide (NAD)

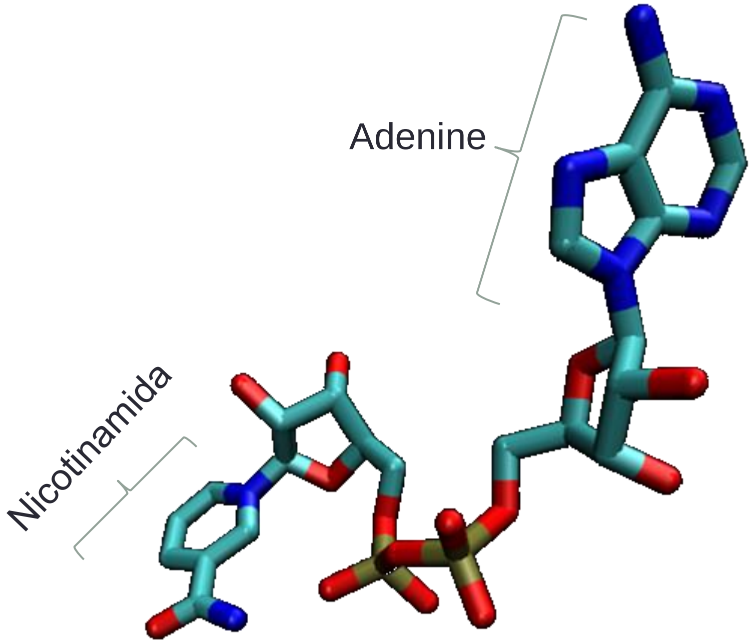


Fig.2 Nicotinamide adenine dinucleotide

Binding domain for NAD consists of **two paired Rossmann fold**.

The main role of NAD+ cofactor is the **redox reactions.**

Reactions of this type are catalyzed by a large group of enzymes called **oxidoreductases**.

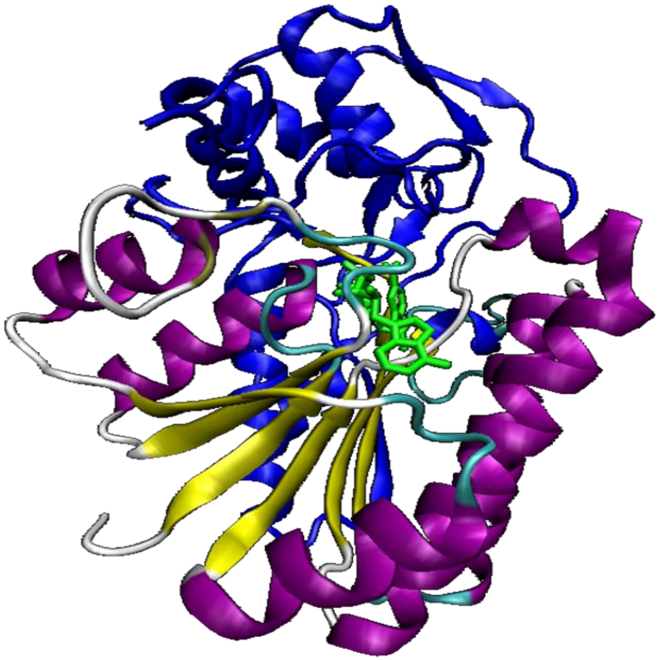


Fig.6 Single Rossman fold: A β-α-β-α-β unit.

Thoden BJ, et al. 2000.; Dym O, et al. 2001.

NAD-binding proteins

Aminoacid residue interactions

SIDE CHAIN	RESIDUES
Polar	Tyr13, Tyr37, Asn157
Apolar	Ile14
+ charge (basic)	Lys161
- charge (acidic)	Asp33, Asp66

These amino acid residues are not conserved, but so do their properties.

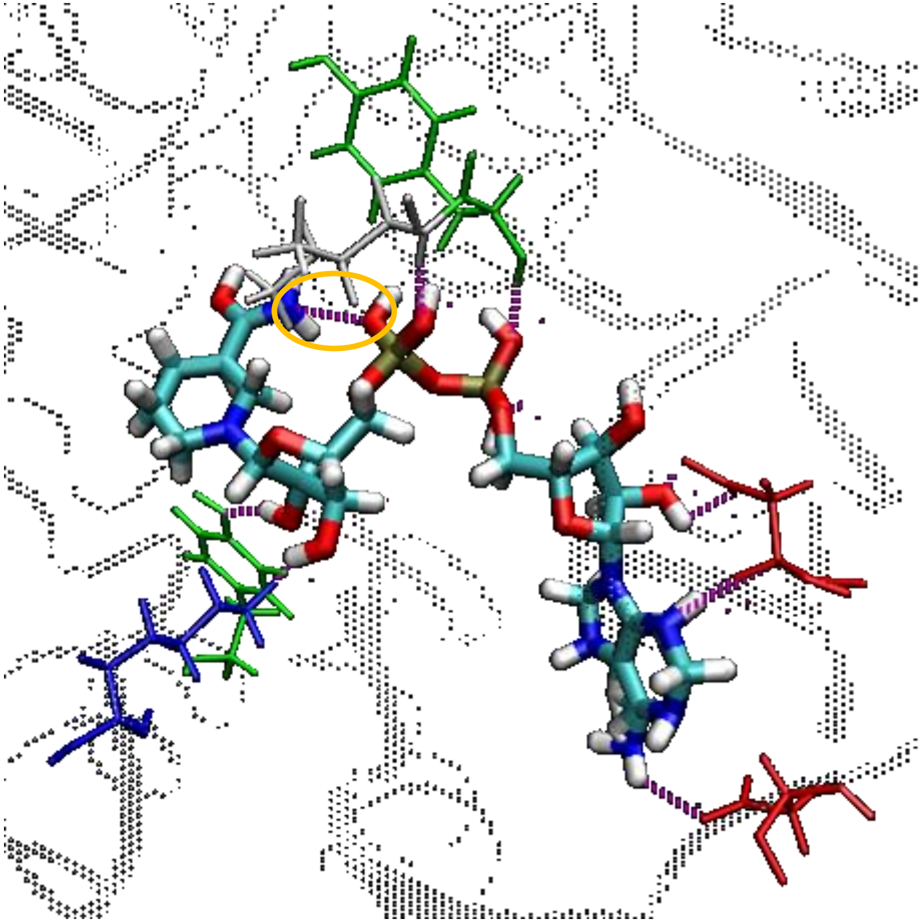


Fig.2 Main amino acid interactions in the NAD+ binding pocket.

Thoden BJ, et al. 2000.

NAD-binding proteins

The role of water in the binding pocket

Water molecules mediate about **30% of hydrogen bonds**.

Concentrated **around the pyrophosphate group**.

Important parameter for the **dinucleotide recognition**.

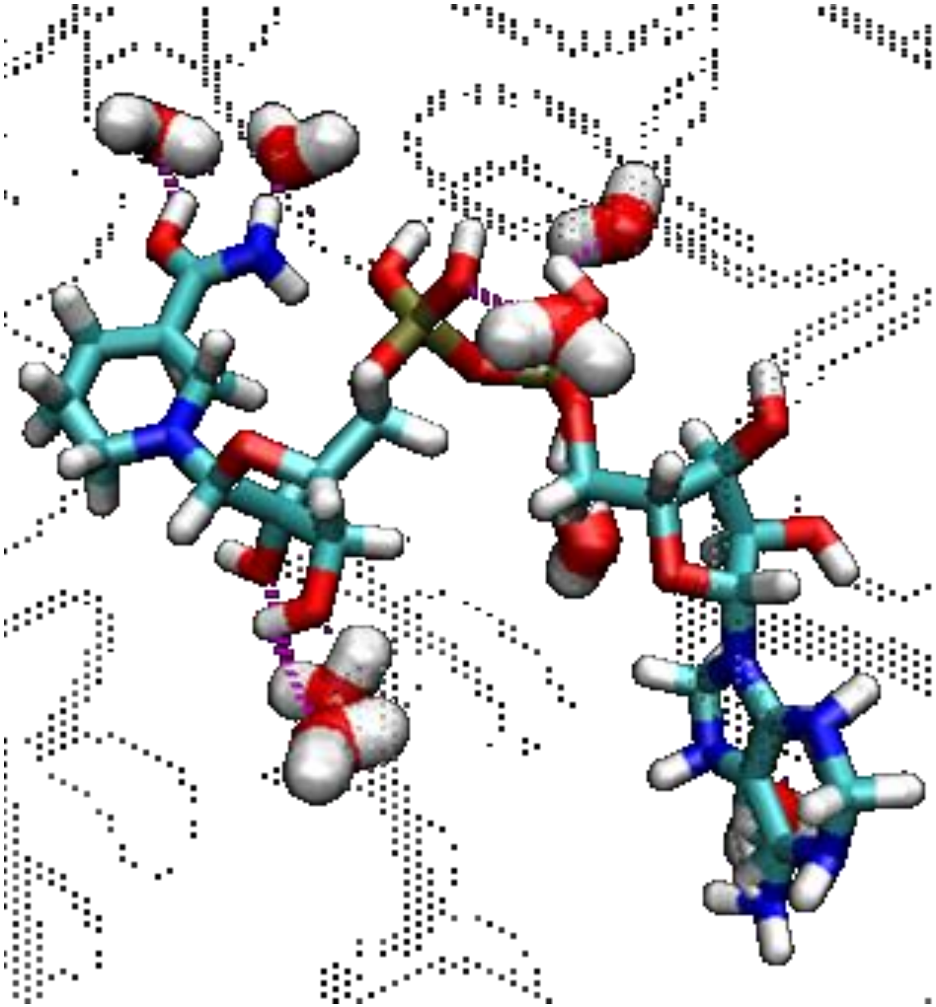


Fig.2 Water molecules interacting with NAD+.

Thoden BJ, et al. 2000.; Bottoms CA,, et al. 2002.

NAD-binding proteins

How does NAD cofactor stabilize itself into the protein?

GLYCINES → no side chains

A-helix → overall positive dipole moment

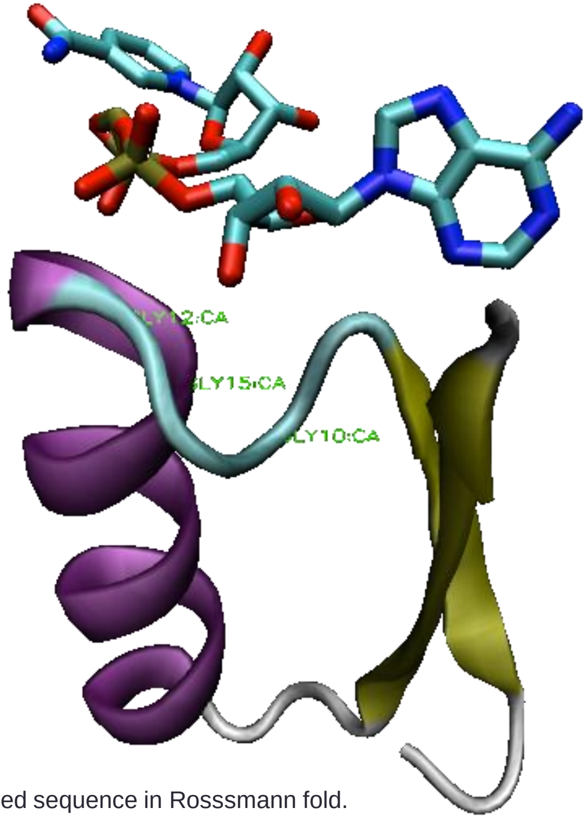


Fig.2 Gly conserved sequence in Rosssmann fold.

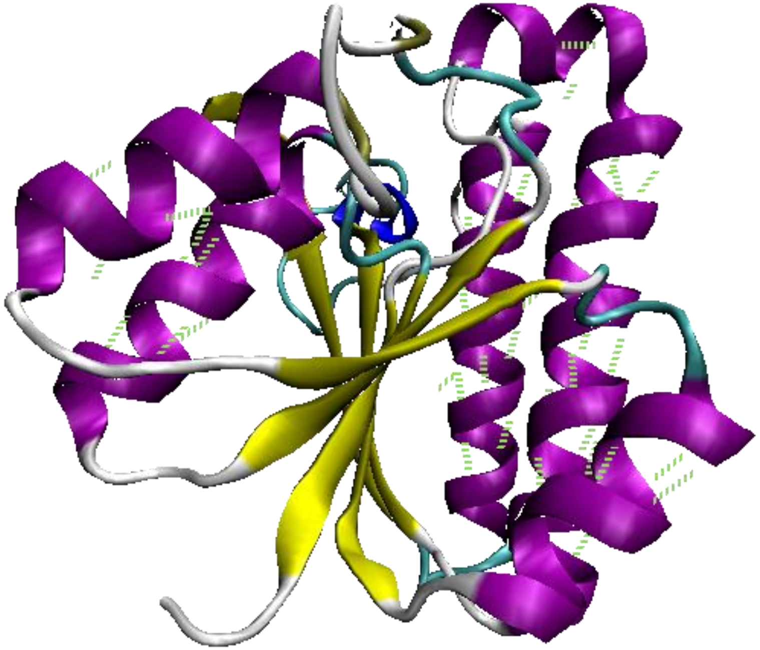
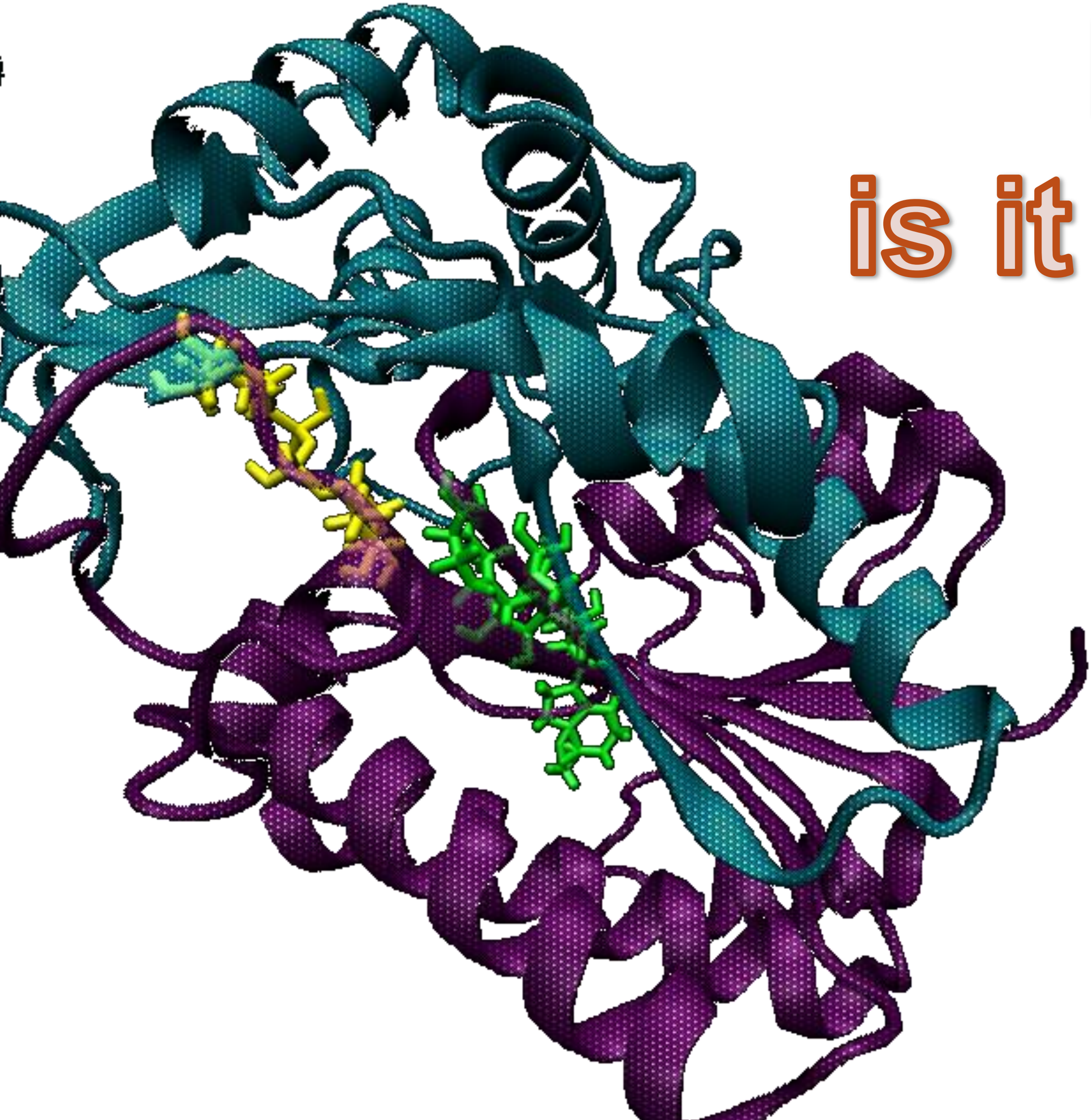


Fig.2 α-helix peptide dipole are parallel to the helix axis.

Bottoms CA,, et al. 2002.



but…

is it enough?

NO

NAD-binding proteins

The role of water is crucial

- **Structurally conserved water molecule**
- **Conserved hydrogen-bonding pattern**

Invariant hydrogen bonds

- Last conserved Gly (αA)
- Dinucleotide pyrophosphate

Variant hydrogen bonds

- 2nd/3rd conserved Gly
- Cterminal residue of 4β

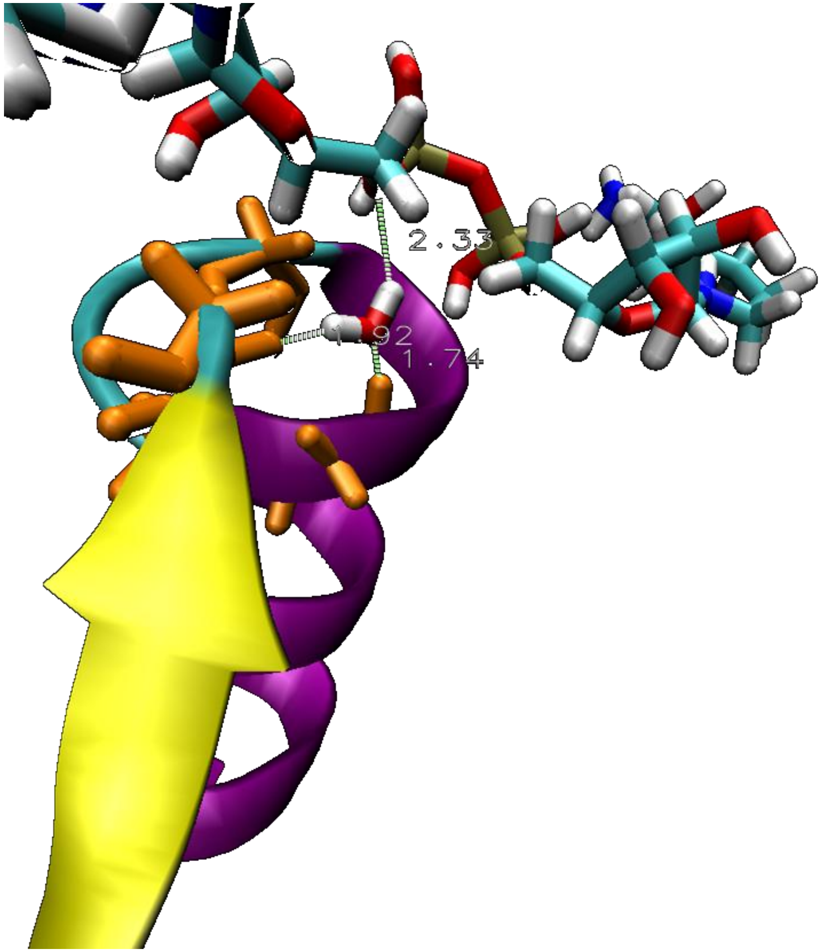
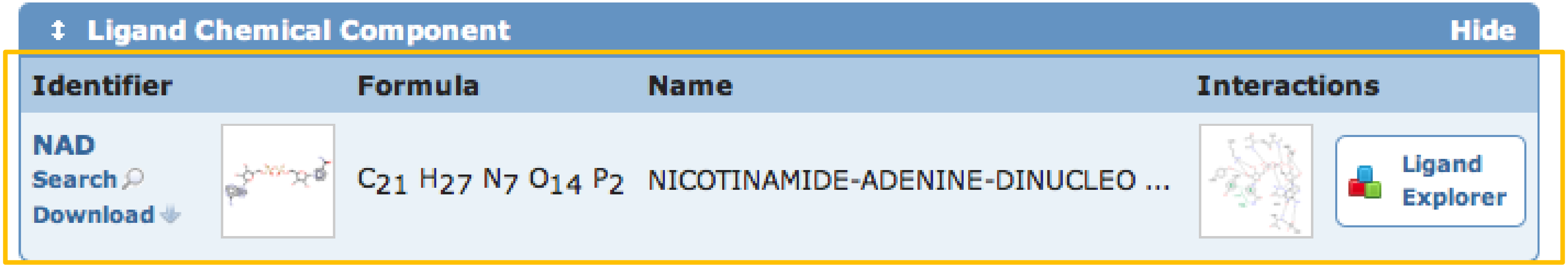


Fig.7 A single water conserved molecule.

Bottoms CA,, et al. 2002.

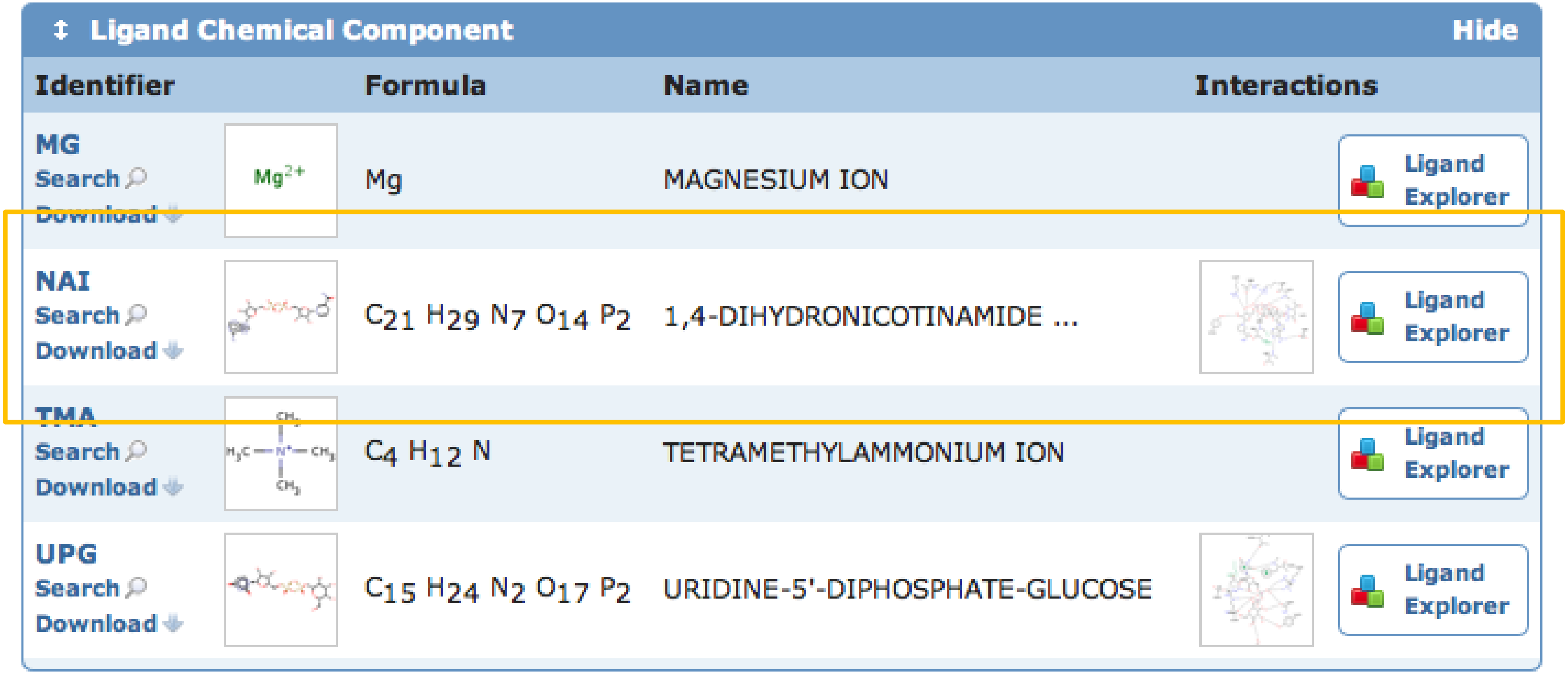
1EK5: STRUCTURE OF HUMAN UDP-GALACTOSE 4-EPIMERASE



Ligand Chemical Component Hide

Identifier	Formula	Name	Interactions
NAD Search Download	$C_{21} H_{27} N_7 O_{14} P_2$	NICOTINAMIDE-ADENINE-DINUCLEO ...	Ligand Explorer

1EK6: STRUCTURE OF HUMAN UDP-GALACTOSE 4-EPIMERASE



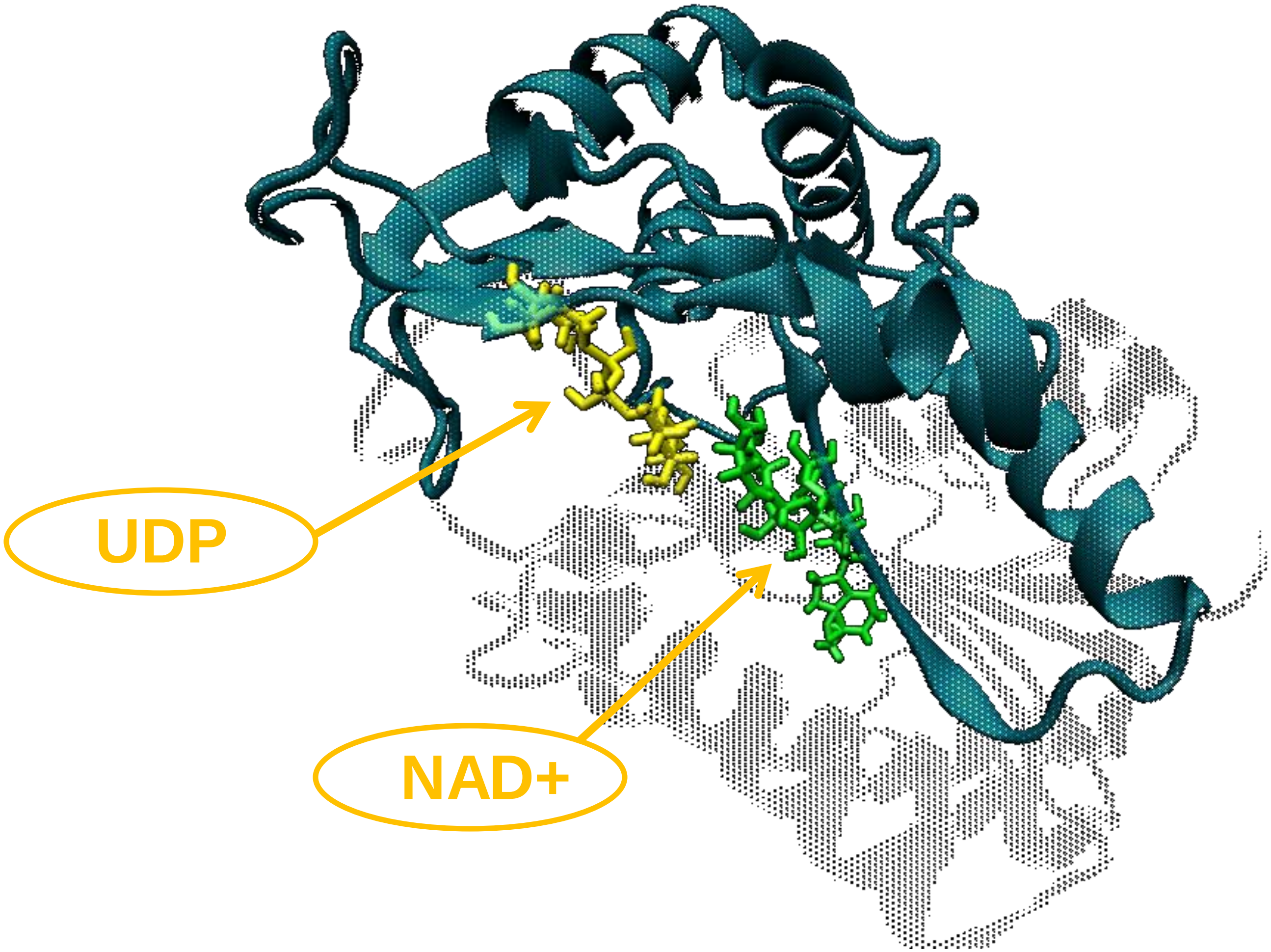
Ligand Chemical Component Hide

Identifier	Formula	Name	Interactions
MG Search Download	Mg	MAGNESIUM ION	Ligand Explorer
NAI Search Download	$C_{21} H_{29} N_7 O_{14} P_2$	1,4-DIHYDRONICOTINAMIDE ...	Ligand Explorer
TMA Search Download	$C_4 H_{12} N$	TETRAMETHYLAMMONIUM ION	Ligand Explorer
UPG Search Download	$C_{15} H_{24} N_2 O_{17} P_2$	URIDINE-5'-DIPHOSPHATE-GLUCOSE	Ligand Explorer

NAD-binding proteins

Stereospecific transference of the H

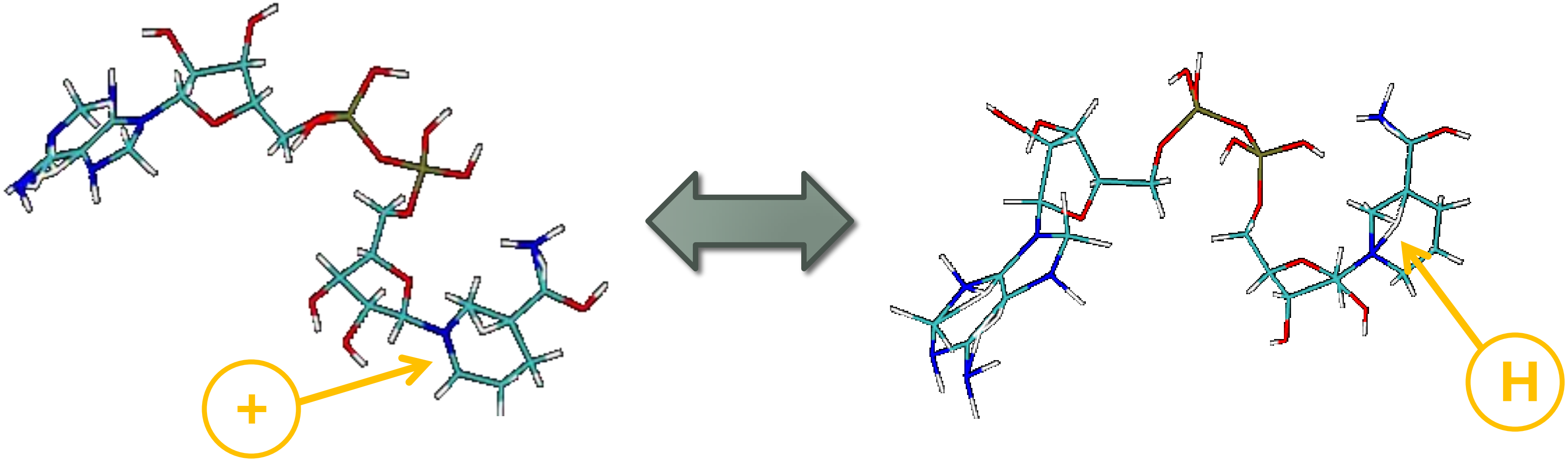
UDP-galactose-4-epimerase



Thoden BJ, et al. 2000.

NAD-binding proteins

Stereospecific transference of the H



NAD^+ OXIDIZING AGENT

accepts electrons from other molecules and becomes reduced

NADH REDUCING AGENT

donates electrons to other molecules and becomes oxidized

Thoden BJ, et al. 2000.

P-loop

Phosphate-binding motif

Fig.9 Phosphate-binding loop (P-loop)

The phosphate binding loop (**P-loop**) is the common motif in mononucleotide binding proteins.

The three-dimensional structure of the P-loop, **precedeed by a β-sheet and followed by an α-helix**, is similar in different protein families.

Fig.10 Superimposition of different protrein families with the P-loop.

C. Ramakrishnan, et al. 2001

P-loop

Walker A motif

The basic structural requirement of a P-loop in most ATP-binding enzymes is its consensus sequence **GXXXXGK T/S**.

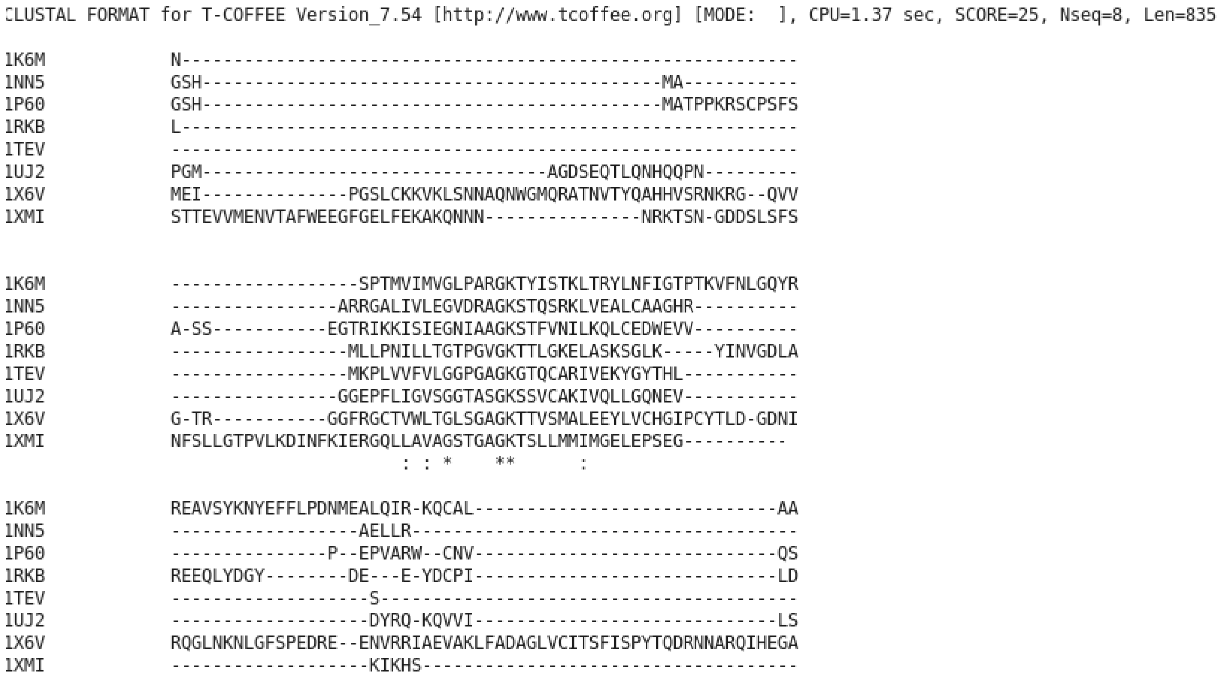
This pattern is mostly present as a **structurally and energetically favorable loop** that appears to provide more room to surround and manipulate the nucleotide phosphate.

Apha and beta proteins:

Fold : P-loop containing nucleoside triphosophate hydrolases

Superfamily: P-loop containing nucleoside triphosohate hydrolases

6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase, kinase
Tymidylate kinase
Deoxycytidine kinase
Adenylate kinase
UMP/CMP kinase
Uridine-cytidine kinase 2
Adenosine-5'phosphosulfate kinase
CFTR



```
CLUSTAL FORMAT for T-COFFEE Version_7.54 [http://www.tcoffee.org] [MODE:  ], CPU=1.37 sec, SCORE=25, Nseq=8, Len=835

1K6M            N-----------------------------------------------------------
1NN5            GSH--------------------------------------------MA-----------
1P60            GSH--------------------------------------------MATPPKRSCPSFS
1RKB            L-----------------------------------------------------------
1TEV            ------------------------------------------------------------
1UJ2            PGM---------------------------------AGDSEQTLQNHQQPN---------
1X6V            MEI--------------PGSLCKKVKLSNNAQNWGMQRATNVTYQAHHVSRNKRG--QVV
1XMI            STTEVVMENVTAFWEEGFGELFEKAKQNNN---------------NRKTSN-GDDSLSFS


1K6M            ------------------SPTMVIMVGLPARGKTYISTKLTRYLNFIGTPTKVFNLGQYR
1NN5            ----------------ARRGALIVLEGVDRAGKSTQSRKLVEALCAAGHR----------
1P60            A-SS-----------EGTRIKKISIEGNIAAGKSTFVNILKQLCEDWEVV----------
1RKB            -----------------MLLPNILLTGTPGVGKTTLGKELASKSGLK-----YINVGDLA
1TEV            -----------------MKPLVVFVLGGPGAGKGTQCARIVEKYGYTHL-----------
1UJ2            ----------------GGEPFLIGVSGGTASGKSSVCAKIVQLLGQNEV-----------
1X6V            G-TR-----------GGFRGCTVWLTGLSGAGKTTVSMALEEYLVCHGIPCYTLD-GDNI
1XMI            NFSLLGTPVLKDINFKIERGQLLAVAGSTGAGKTSLLMMIMGELEPSEG----------
                                      : : *    **      :

1K6M            REAVSYKNYEFFLPDNMEALQIR-KQCAL-----------------------------AA
1NN5            ------------------AELLR-------------------------------------
1P60            ---------------P--EPVARW--CNV-----------------------------QS
1RKB            REEQLYDGY--------DE---E-YDCPI-----------------------------LD
1TEV            -------------------S----------------------------------------
1UJ2            -------------------DYRQ-KQVVI-----------------------------LS
1X6V            RQGLNKNLGFSPEDRE--ENVRRIAEVAKLFADAGLVCITSFISPYTQDRNNARQIHEGA
1XMI            -------------------KIKHS------------------------------------
```

Fig.11 T-coffee alignment among different P-loop containing nucleoside triphosphate hydrolases.

C. Ramakrishnan, et al. 2001

P-loop

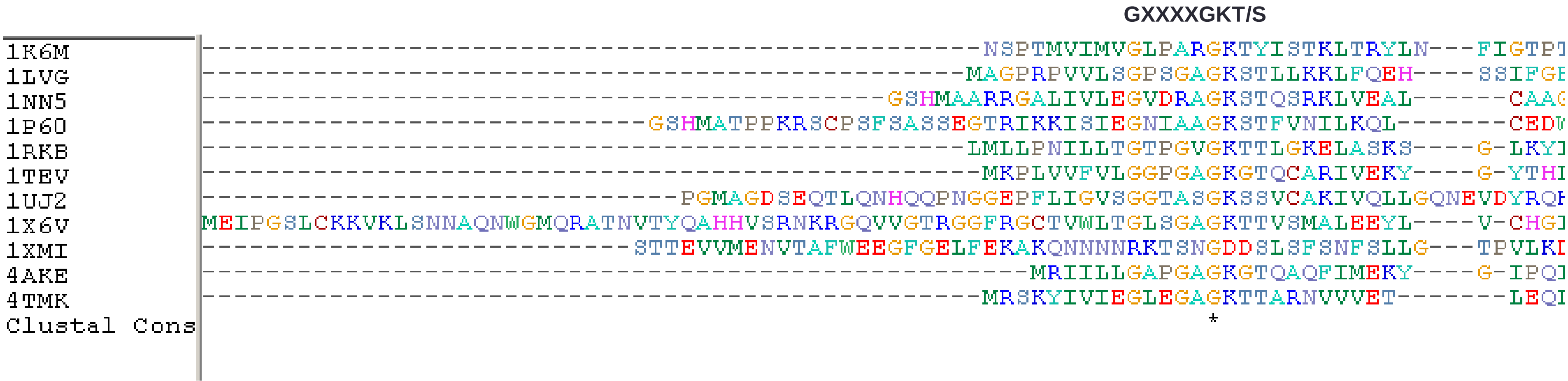
Walker A motif

What is important in the Walker A motif?

Variable quartet (XXXX), however G (12.3 %), A (11.9 %), S (9.8 %), V (8.4 %) and T (5.9 %) occur more commonly than other amino-acids.

A glycine-rich loop.

Lysine (K) and Threonine (T) residue.



```
1K6M         -------------------------------------------------NSPTMVIMVGLPARGKTYISTKLTRYLN---FIGTP
1LVG         ------------------------------------------------MAGPRPVVLSGPSGAGKSTLLKKLFQEH----SSIFG
1NN5         -------------------------------------------GSHMAARRGALIVLEGVDRAGKSTQSRKLVEAL------CAA
1P60         ----------------------------GSHMATPPKRSCPSFSASSEGTRIKKISIEGNIAAGKSTFVNILKQL-------CED
1RKB         ------------------------------------------------LMLLPNILLTGTPGVGKTTLGKELASKS----G-LKY
1TEV         -------------------------------------------------MKPLVVFVLGGPGAGKGTQCARIVEKY----G-YTH
1UJ2         ------------------------------PGMAGDSEQTLQNHQQPNGGEPFLIGVSGGTASGKSSVCAKIVQLLGQNEVDYRQ
1X6V         MEIPGSLCKKVKLSNNAQNWGMQRATNVTYQAHHVSRNKRGQVVGTRGGFRGCTVWLTGLSGAGKTTVSMALEEYL----V-CHG
1XMI         --------------------------STTEVVMENVTAFWEEGFGELFEKAKQNNNNRKTSNGDDSLSFSNFSLLG---TPVLK
4AKE         ----------------------------------------------------MRIILLGAPGAGKGTQAQFIMEKY----G-IPQ
4TMK         -------------------------------------------------MRSKYIVIEGLEGAGKTTARNVVVET-------LEQ
Clustal Cons                                                                 *
```

Fig.12 T-coffee alignment among different P-loop containing nucleoside triphosphate hydrolases.

C. Ramakrishnan, et al. 2001

P-loop

Walker A motif

The **glycine-rich residues** in the loop clearly play an important conformational role in maintaining the structure of the loop.

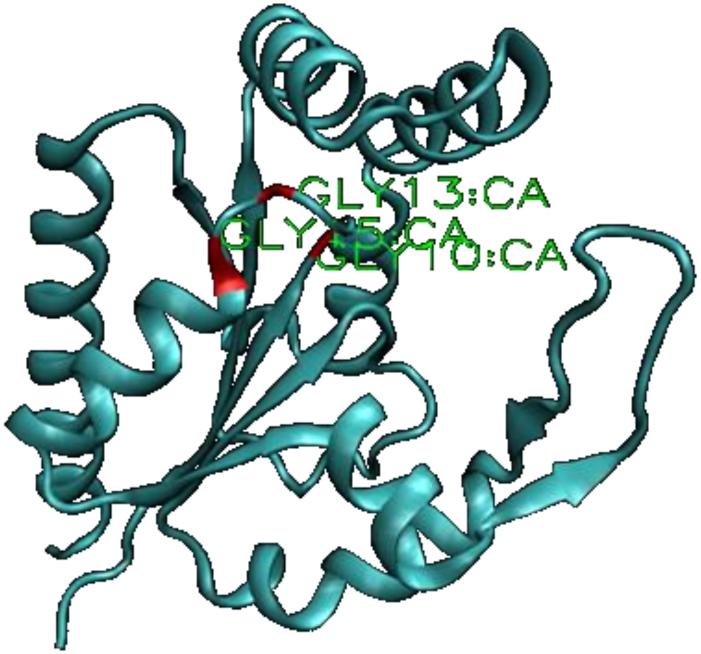


Fig.13 Gycine rich-loop

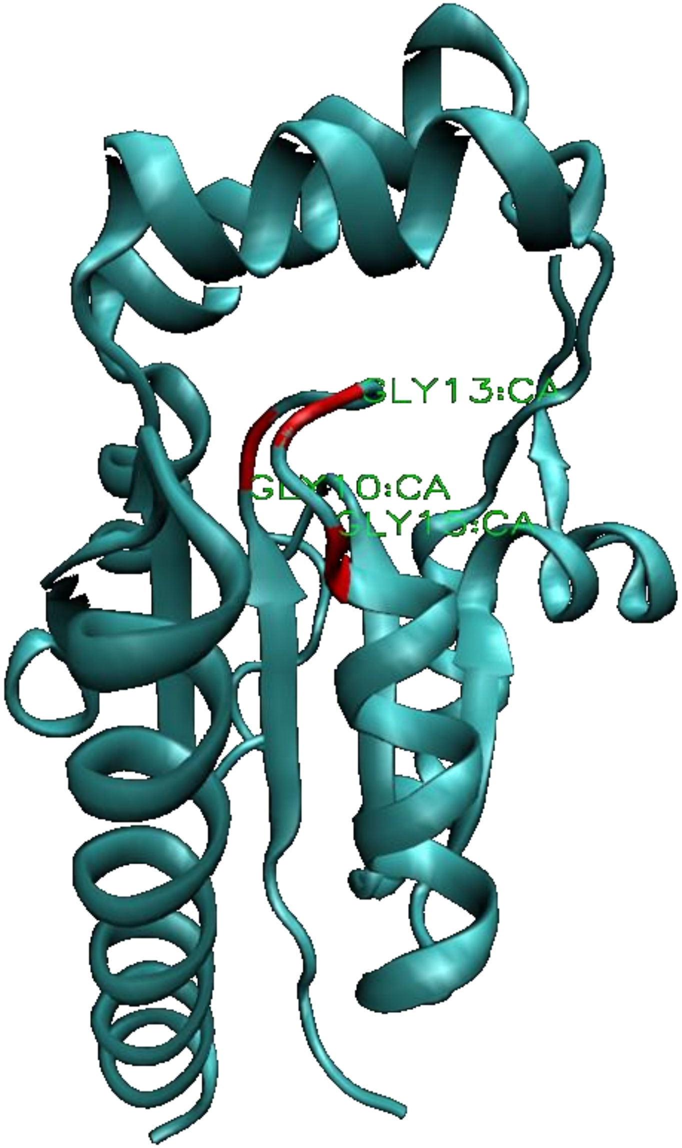


Fig.14 Glycine-rich loop.

Priva et al, 2011.

P-loop

Walker A motif

The **lysine residue** in the consensus sequence GXXXXGKT/S is crucial for the direct interaction with the phosphates of ATP.

The eight residue is usually a hydroxyl-containing residue: a **threonine** or **serine**.

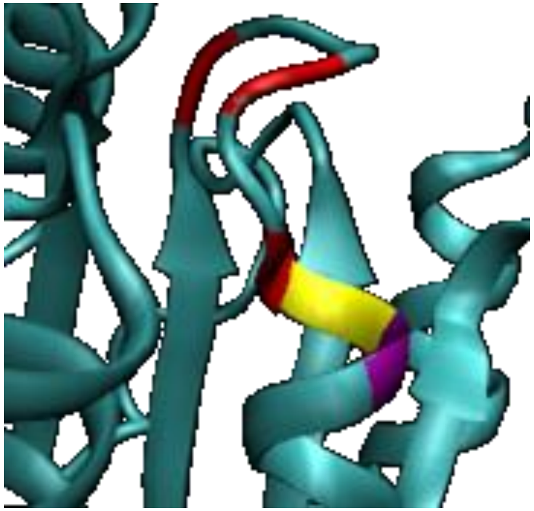


Fig.2 Walker A motif..

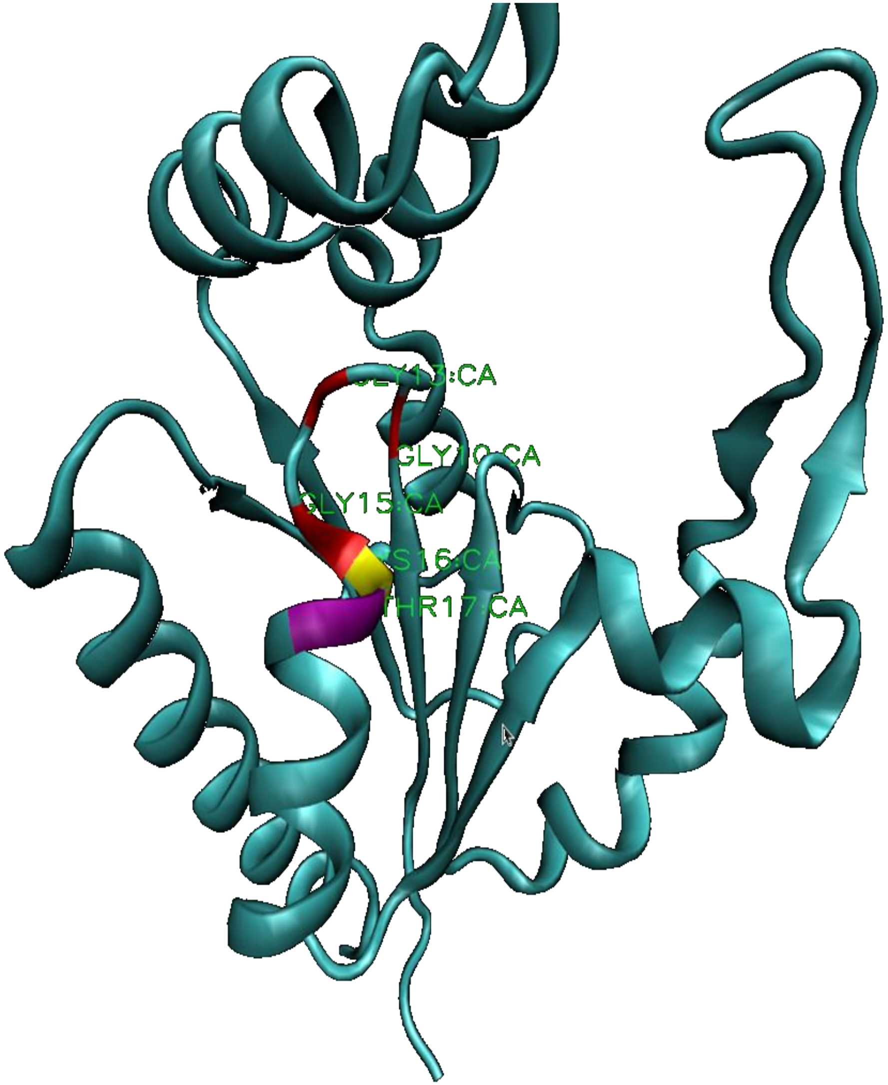


Fig.2 Walker A motif.

C. Ramakrishnan, et al. 2001

P-loop

P-loop and ATP

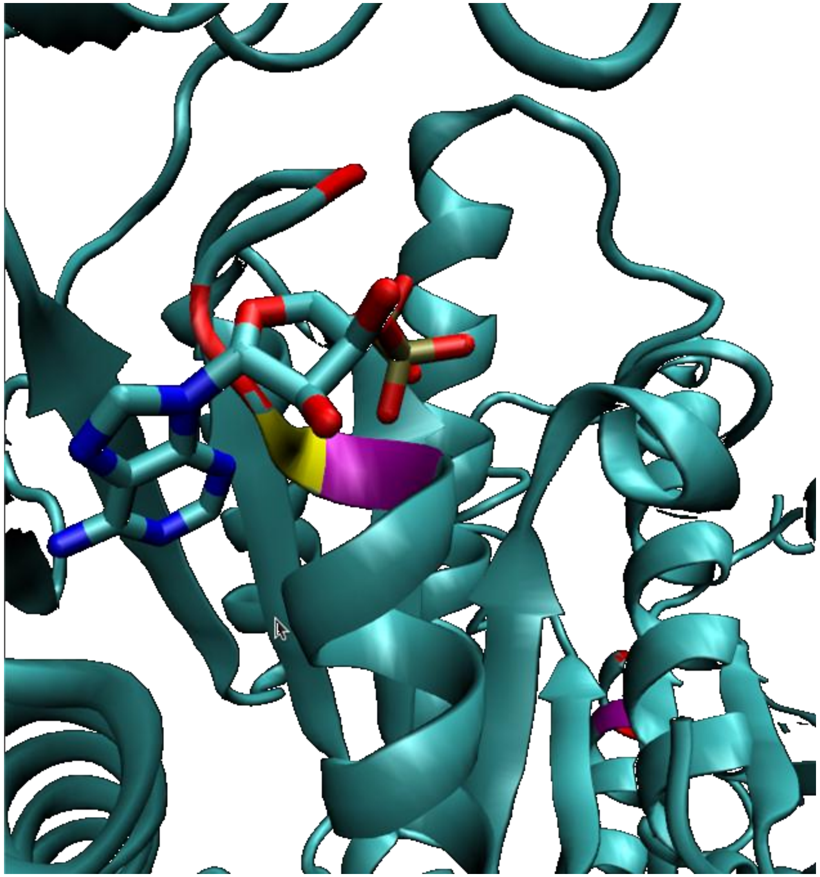


Fig.12 Interacion between ADP and an a.

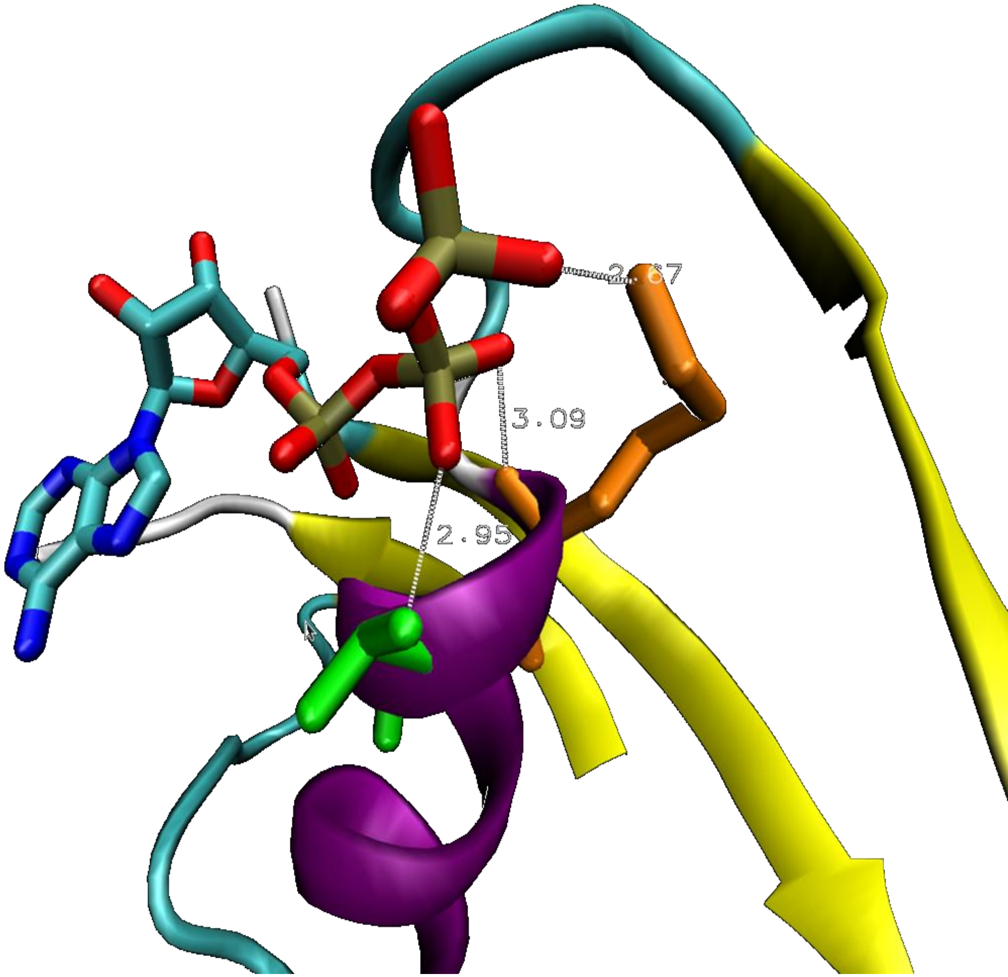


Fig.13 interaction between ATP and an adenylate kinase.

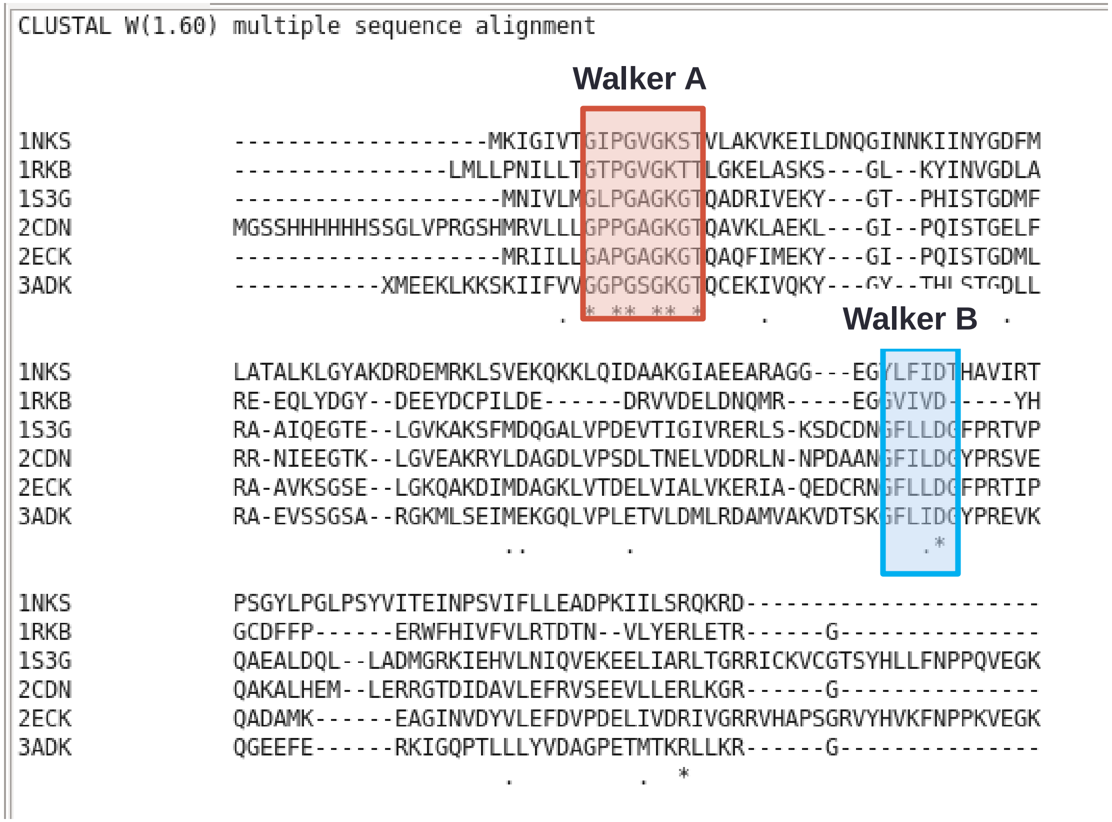
C. Ramakrishnan, et al. 2001

P-loop

Walker B motif

The Walker B motif is another integral part of the **ATP-binding site**.

The **aspartate residue** is required for ATP hydrolysis and is preceded by four hidrofobic residues: **hhhhD**



```
CLUSTAL W(1.60) multiple sequence alignment


1NKS            -------------------MKIGIVTGIPGVGKSTVLAKVKEILDNQGINNKIINYGDFM
1RKB            ----------------LMLLPNILLTGTPGVGKTTLGKELASKS---GL--KYINVGDLA
1S3G            --------------------MNIVLMGLPGAGKGTQADRIVEKY---GT--PHISTGDMF
2CDN            MGSSHHHHHHSSGLVPRGSHMRVLLLGPPGAGKGTQAVKLAEKL---GI--PQISTGELF
2ECK            --------------------MRIILLGAPGAGKGTQAQFIMEKY---GI--PQISTGDML
3ADK            -----------XMEEKLKKSKIIFVVGGPGSGKGTQCEKIVQKY---GY--THLSTGDLL
                                        . * ** ** *    .                 .

1NKS            LATALKLGYAKDRDEMRKLSVEKQKKLQIDAAKGIAEEARAGG---EGYLFIDTHAVIRT
1RKB            RE-EQLYDGY--DEEYDCPILDE------DRVVDELDNQMR-----EGGVIVD-----YH
1S3G            RA-AIQEGTE--LGVKAKSFMDQGALVPDEVTIGIVRERLS-KSDCDNGFLLDGFPRTVP
2CDN            RR-NIEEGTK--LGVEAKRYLDAGDLVPSDLTNELVDDRLN-NPDAANGFILDGYPRSVE
2ECK            RA-AVKSGSE--LGKQAKDIMDAGKLVTDELVIALVKERIA-QEDCRNGFLLDGFPRTIP
3ADK            RA-EVSSGSA--RGKMLSEIMEKGQLVPLETVLDMLRDAMVAKVDTSKGFLIDGYPREVK
                                    ..       .                     .*

1NKS            PSGYLPGLPSYVITEINPSVIFLLEADPKIILSRQKRD----------------------
1RKB            GCDFFP------ERWFHIVFVLRTDTN--VLYERLETR------G---------------
1S3G            QAEALDQL--LADMGRKIEHVLNIQVEKEELIARLTGRRICKVCGTSYHLLFNPPQVEGK
2CDN            QAKALHEM--LERRGTDIDAVLEFRVSEEVLLERLKGR------G---------------
2ECK            QADAMK------EAGINVDYVLEFDVPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGK
3ADK            QGEEFE------RKIGQPTLLLYVDAGPETMTKRLLKR------G---------------
                                    .         .  *
```

Fig.15 Clustal alignment of Adeniylate kinases of different species.

C. Ramakrishnan, et al. 2001

P-loop

Walker B motif

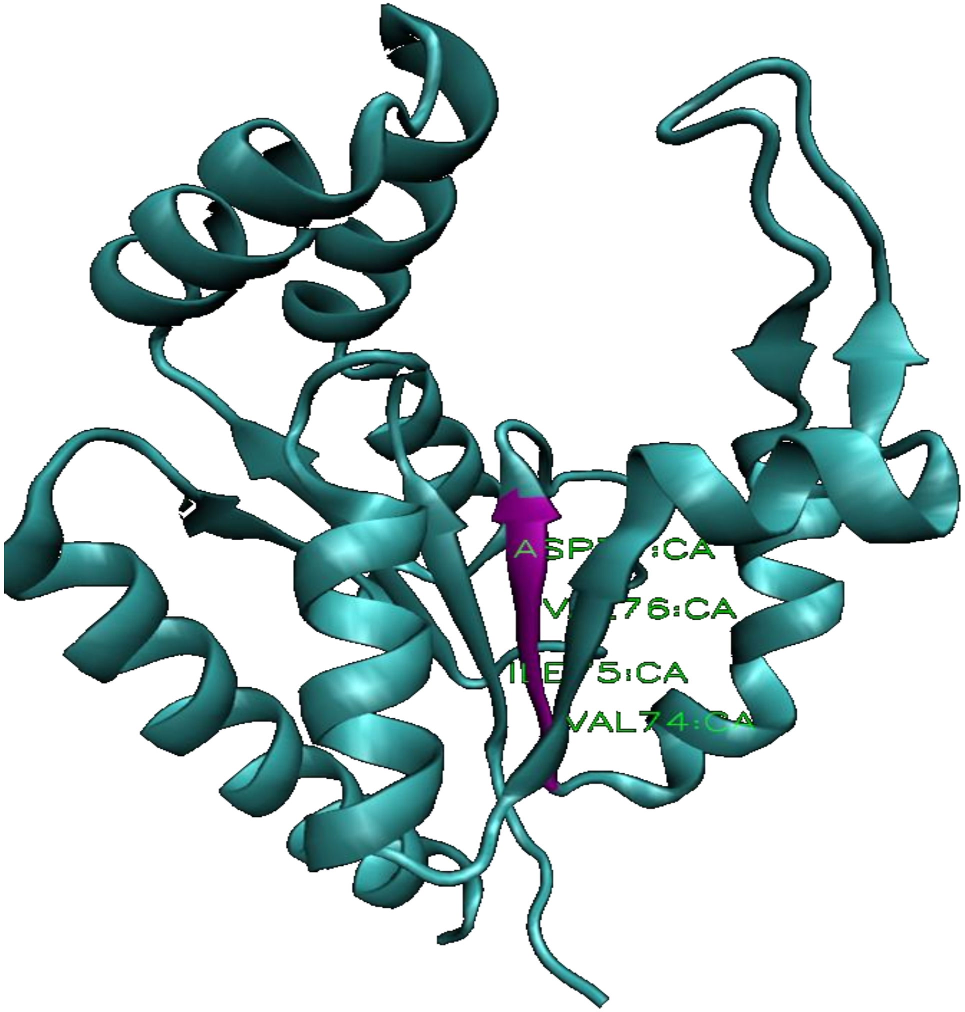


Fig.16 Walker B motif (1RKB)

C. Ramakrishnan, et al. 2001

Thank you for your attention!

References

Thoden BJ, Wohlers TM, Fridovich-Keil JL, Holden HM. Crystalographic Evidence for Tyr157 Functioning as the Active Site Base in Human UDP-Galactose 4-Epimerase. Am. Chem. Soc. 2000; 39: 5691-5701

Dym O, Eisenberg D. Sequence-structure analysis of FAD-containing proteins. Protein Sci. 2001; 10: 1712-1728

Bottoms CA, Smith PE, Tanner JJ. A structurally conserved water molecule in Rossmann dinucleotide-binding domains. Protein Sci. 2002; 11: 2125-213

Ramakdrishnan, Dani VS, Ramasarma T.A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins. Life Sciences Volume 15. 783-798.

Wierenga RK, De Mayer CH, Hol WG

Ragunathan Priya, Anil Kumar, Malathy Sony Subramanian Manimekalai and Gerhard Grüber. Conserved Glycine Residues in the P-Loop of ATP Synthases Form a Doorframe for Nucleotide Entrance. J. Mol. Biol. (2011) 413, 657–666