Metalloenzymes as teaching tools for coordination chemistry From bonding to spectroscopy

Jordi Cirera
Departament de Química Inorgànica i Orgànica
Insitut de Química Teòrica i Computacional (IQTC-UB)

Bio-Chem-Comp School 2018, Bansko

Outline

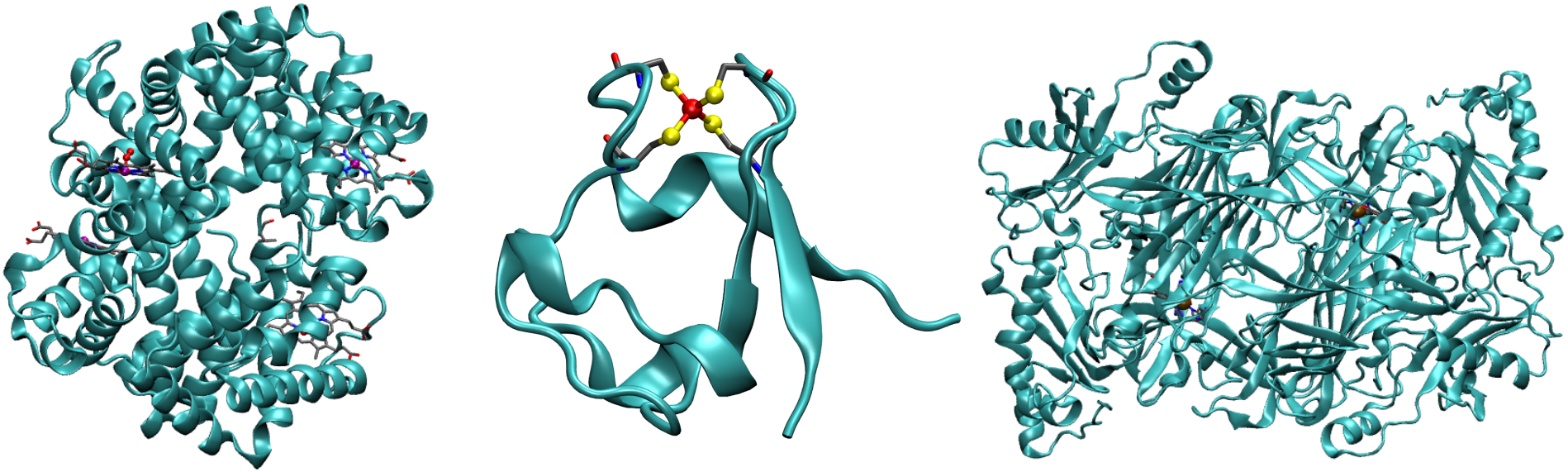
1. *The Protein Data Bank*
   - *1.1 What is the Protein Data Bank? (RCSD-PDB)*
   - *1.2 Information contained on the PDB*
   - *1.3 Working with the PDB*
   - *1.4 Ligand Explorer*
   - *1.5 The metal-PDB database*
   - *1.6 Additional Tools*

2. *Integrating the PDB as a teaching tool for coordination chemistry*
   - *2.1 Metalloenzymes as illustrative examples*
   - *2.2 Coordination motifs in active sites*
   - *2.3 Geometrical data/Coordination modes*
   - *2.4 Spin-State Information*
   - *2.5 Absorption Spectroscopy*
   - *2.6 Electron Paramagnetic Resonance Spectroscopy*
   - *2.7 Practical Cases*

3. *Summary and Outlook*

Metalloenzymes in coordination chemistry?

Metalloenzymes are usually constrained to bioniroganic courses



They all contain coordination compounds!

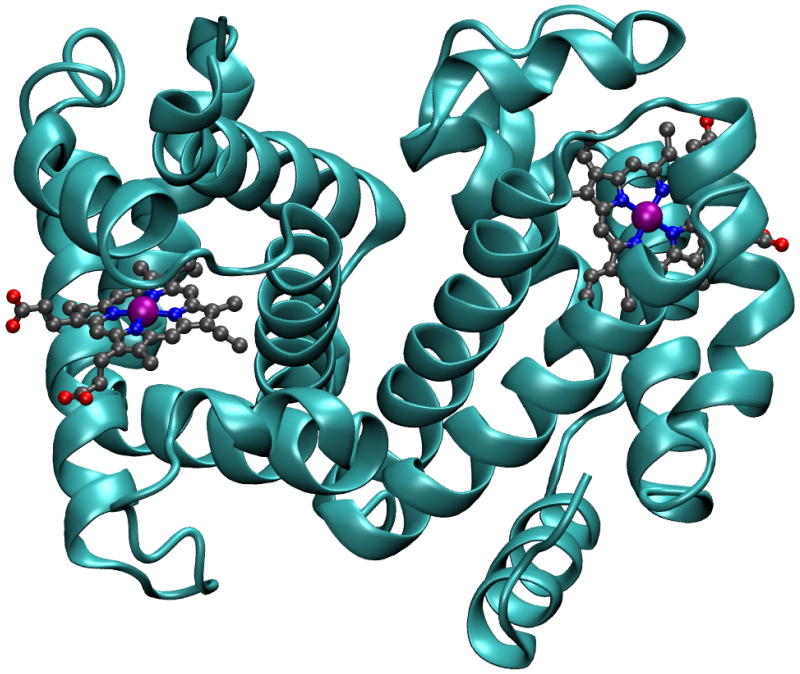
- We can use them to illustrate basic coordination chemistry concepts
- We can use them to illustrate advanced bonding concepts
- We can use them to illustrate spectroscopic techniques
- ….

The Protein Data Bank

1.1 What is the Protein Data Bank? (RCSD-PDB)

RCSB PDB
PROTEIN DATA BANK

The ***PDB*** is a database for the three-dimensional structural data of large biological molecules, such as proteins and nucleic acids

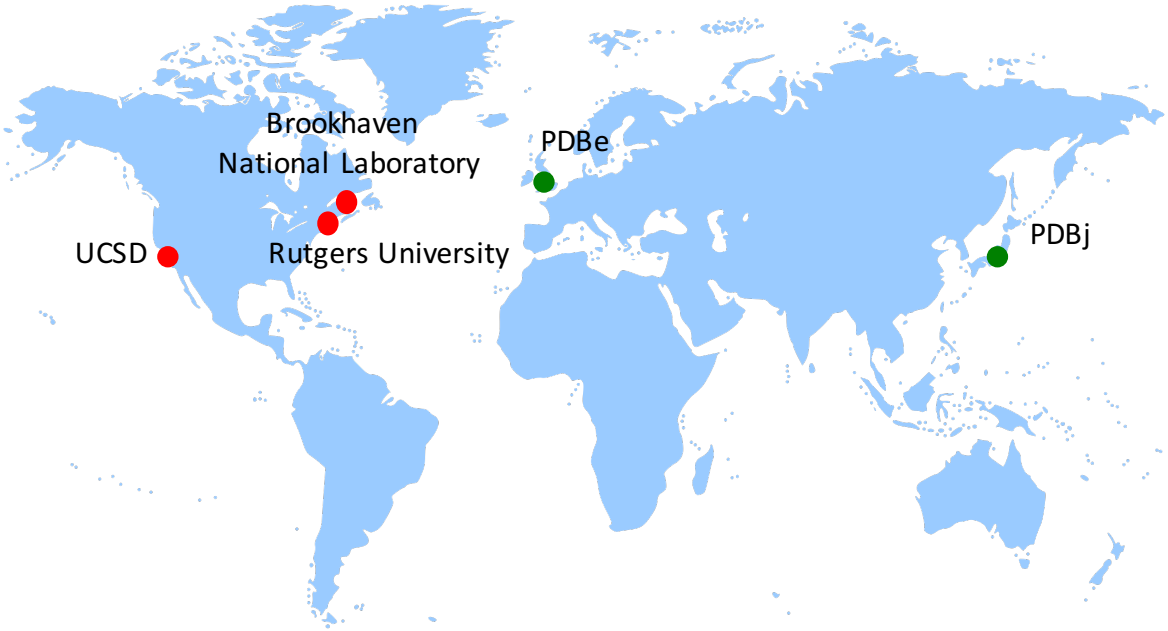


- Stuctural biology
- Genomics
- Bioniorganic Chemistry
- Reactivity of biological systems
- Drug Design

....

https://www.rcsb.org/

1.1 What is the Protein Data Bank? (RCSD-PDB)

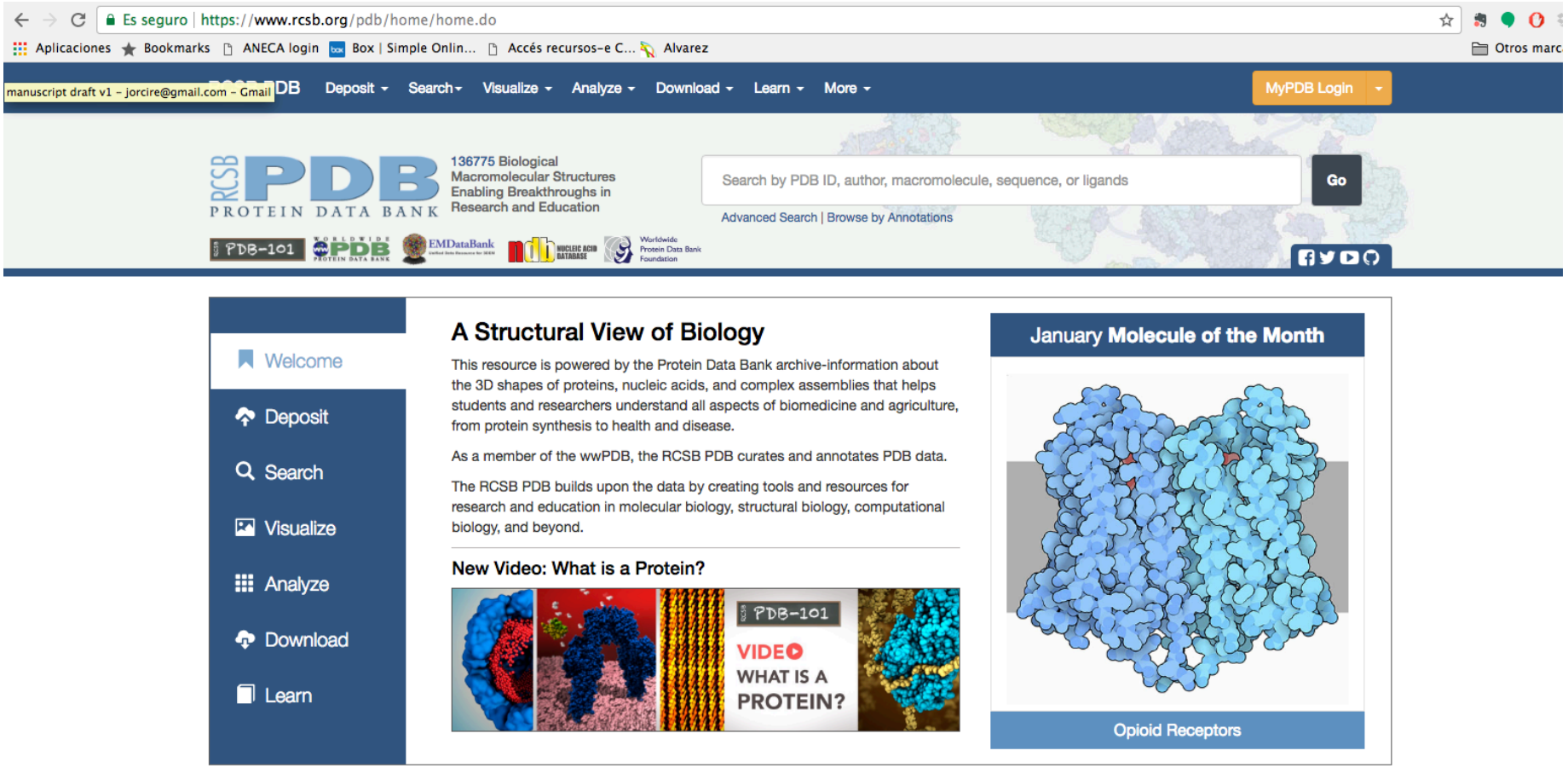


Gordon (SDSC)

https://www.rcsb.org/

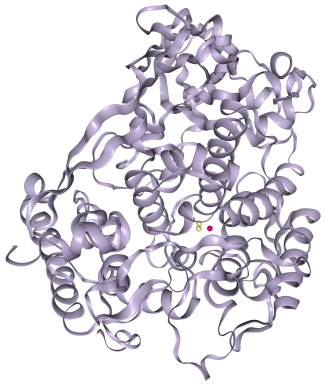
1.1 What is the Protein Data Bank? (RCSD-PDB)

- Open access database
- Web based interface (with its pros and cons…)

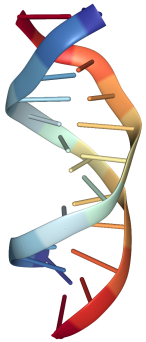


https://www.rcsb.org/

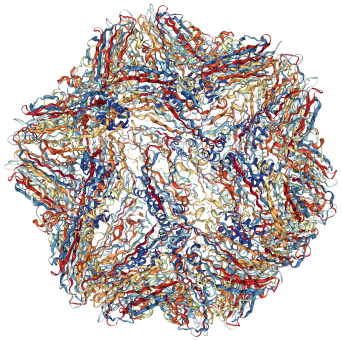
1.2 Information contained on the PDB



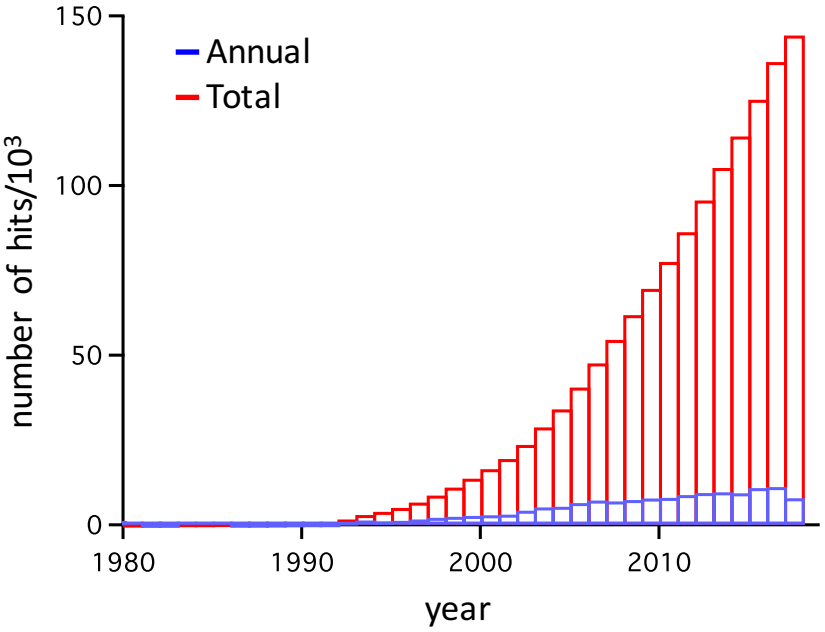
Proteins



DNA/RNA



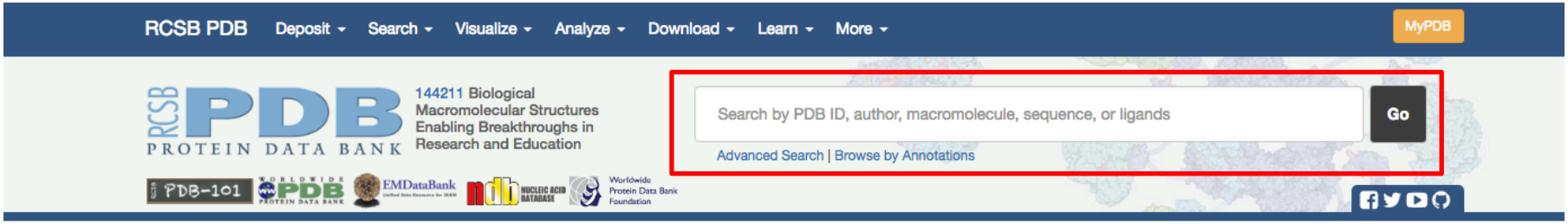
Virus

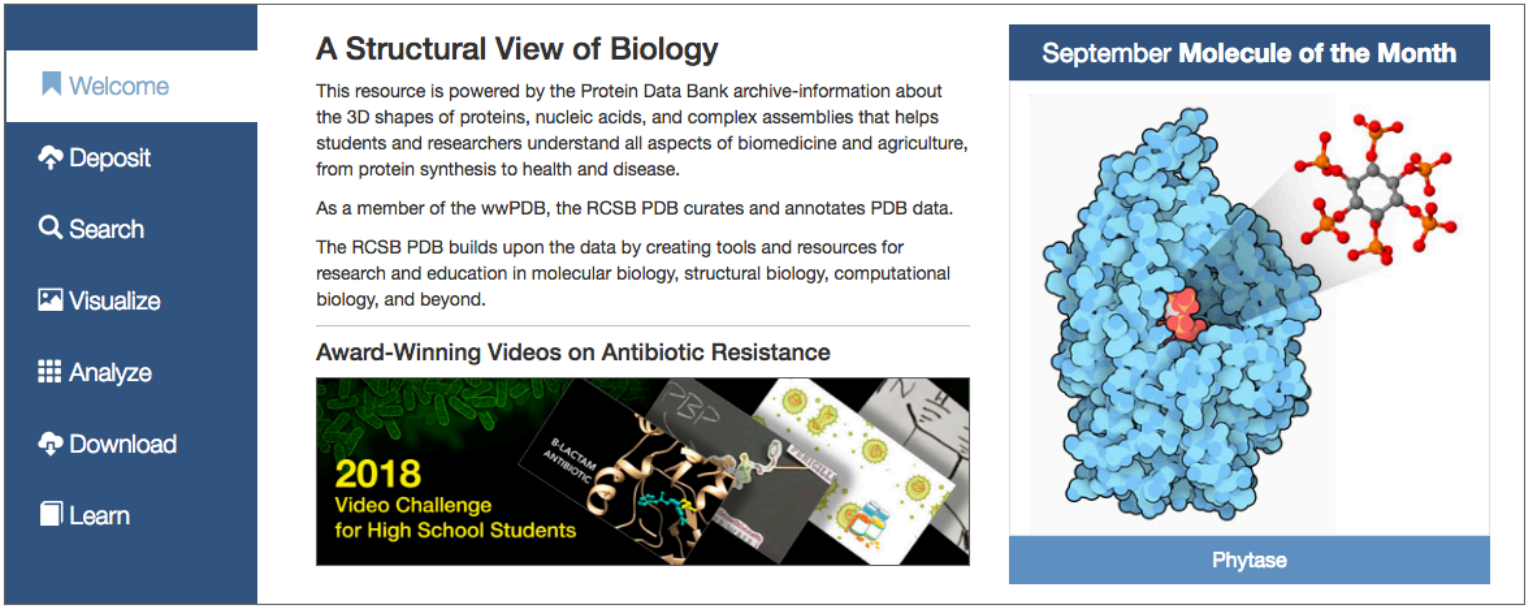


Exp. Method	*Proteins*	*Nucleic Acids*	*Protein/NA Complex*	*Other*	*Total*
X-Ray	120875	1958	6242	10	129085
NMR	10837	1254	250	8	12349
Electron Microscopy	1735	31	619	0	2385
Other	244	4	6	13	267
Multi Method	117	5	2	1	125
Total	133808	3252	7119	32	**144211**

1.3 Working with the PDB

- Web based search: we can search by any field!



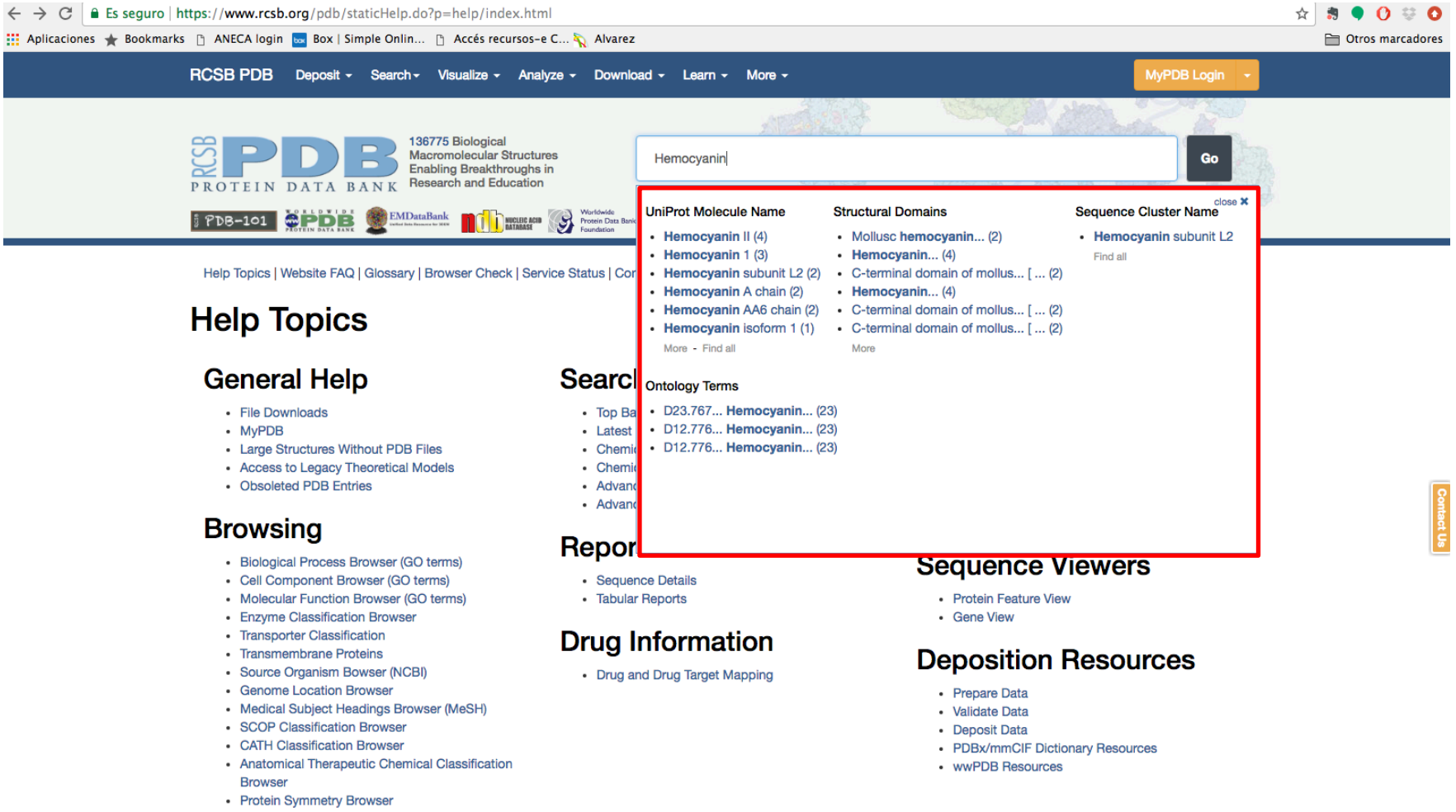


PDB ID, author, name, DOI, ligands, etc….

Very convenient, but lacks precision

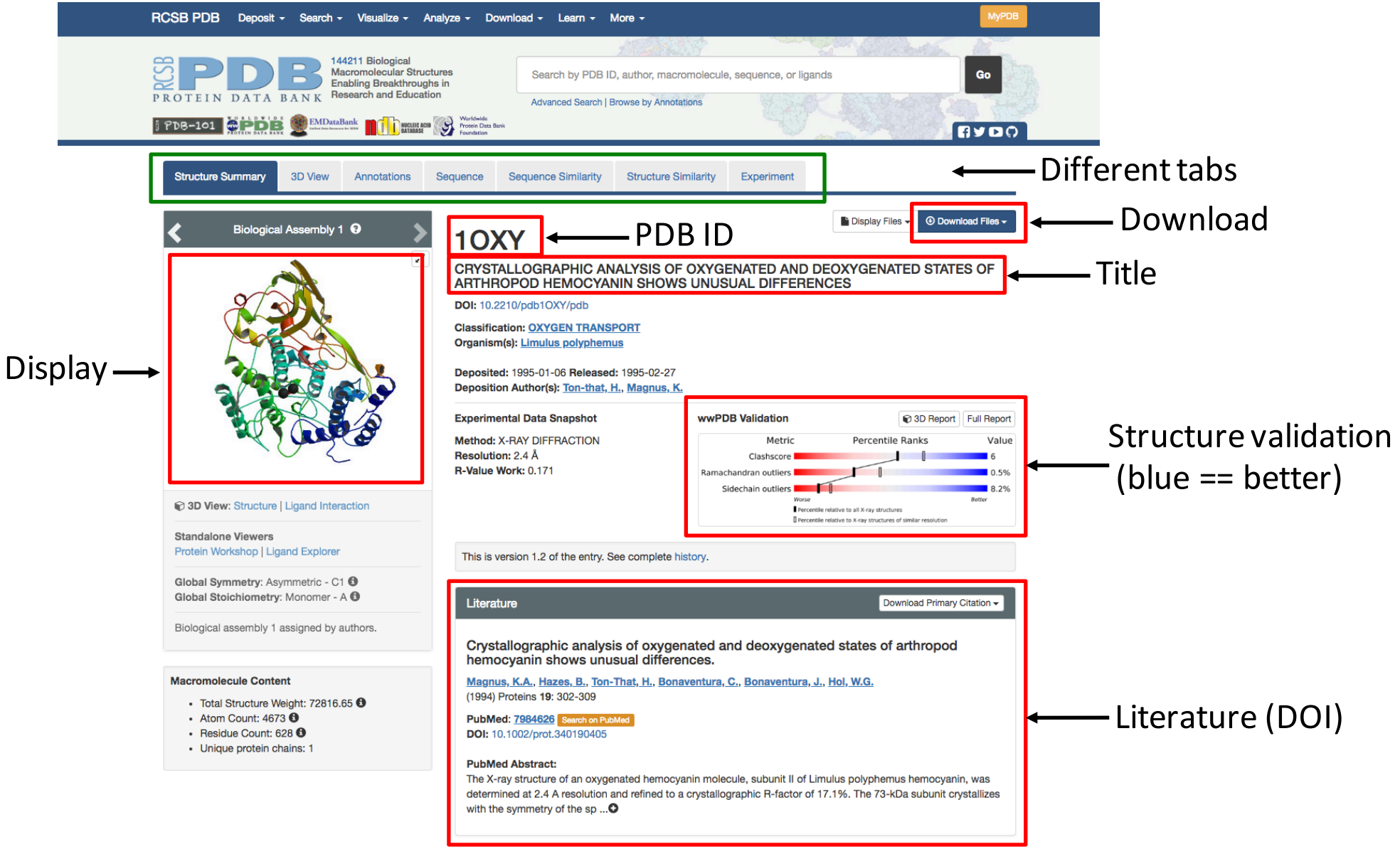
1.3 Working with the PDB

- Web based search: Hemocyanin (the web browser makes a guess….)



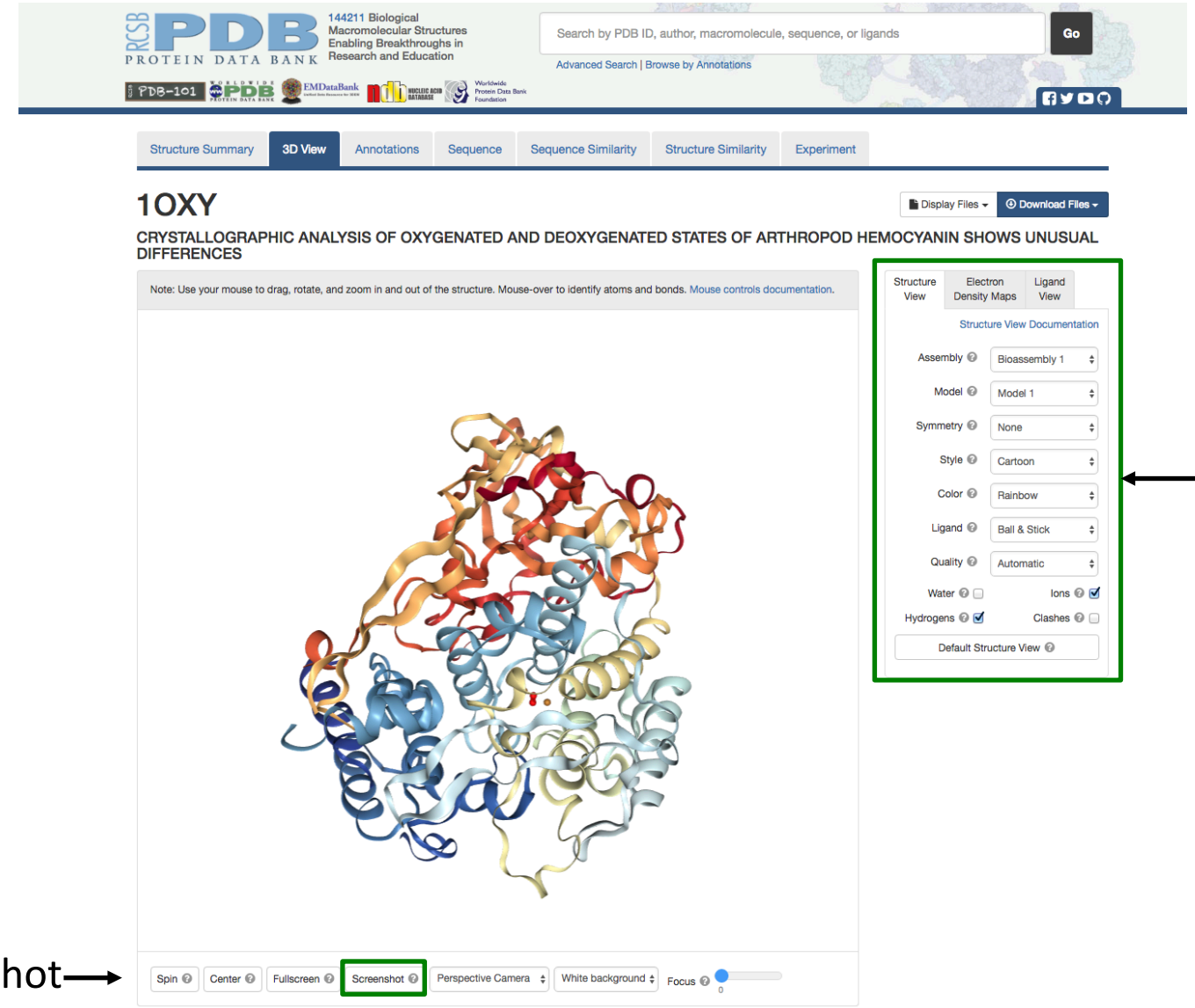
1.3 Working with the PDB

- Select a particular structure: **1OXY**



1.3 Working with the PDB

- Select a particular structure: **1OXY**
  Easy to get a very nice picture

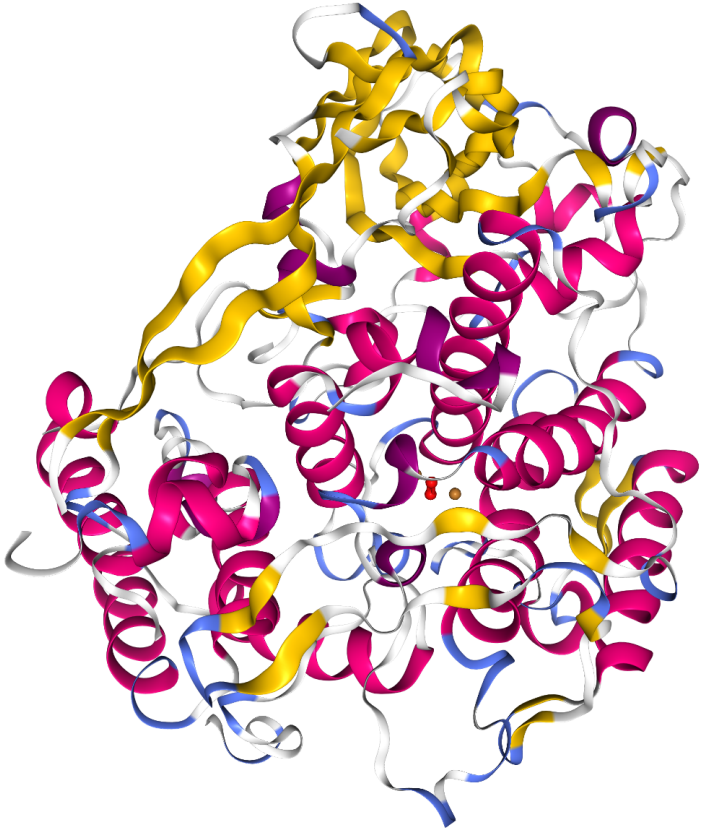


Display Options

Snapshot

1.3 Working with the PDB

- Select a particular structure: **1OXY**



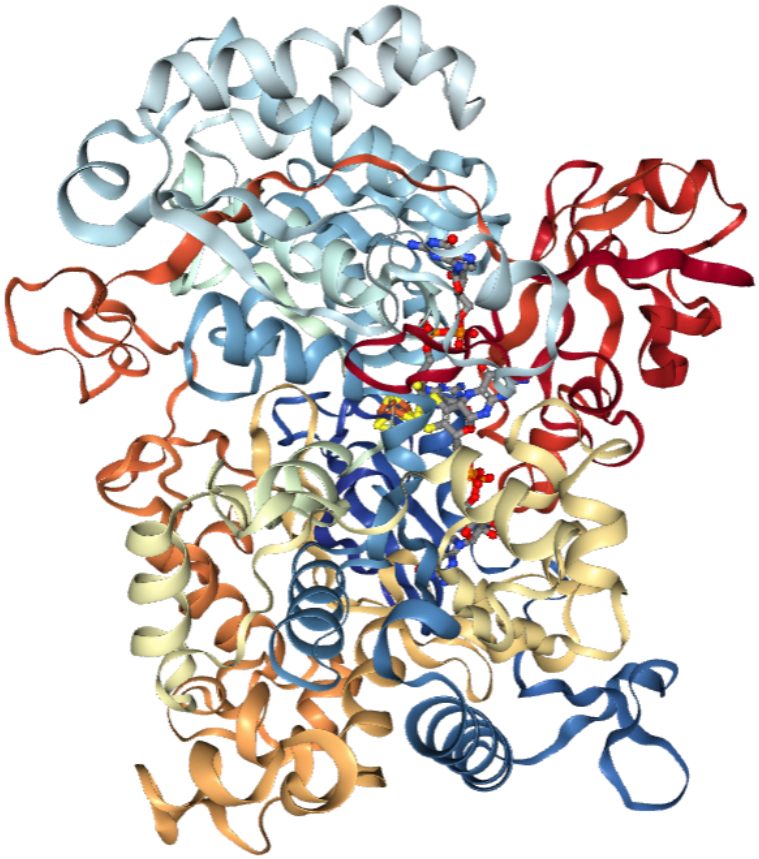
- Title
- Experimental data Snapshot
- Literature (DOI)
- Macromolecules
- Small molecules (Ligand Explorer!!)
- Structure validation (blue == better)
- Display
- Download files (open w/ Mercury)
- Download Citations (Mendeley/EndNote)

PDB Identifier: 4 positions alphanumeric string

1.3 Working with the PDB

- Search for Nitrate Reductase (escherichia coli) (14hits)

Crystal structure of the periplasmic nitrate reductase (NAP) from Escherichia coli (**2NYA**)



Literature (DOI): 10.1074/jbc.M607353200

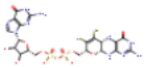
PDB considers everything is not protein is a ligand (ions, metals, drugs, etc…)

1.3 Working with the PDB

- **2NYA: small molecules menu**

Small Molecules

Ligands 3 Unique

ID	Chains	Name / Formula / InChI Key	2D Diagram & Interactions	3D Interactions
SF4 **Query on SF4** Download SDF File Download CCD File	A, F	**IRON/SULFUR CLUSTER** $Fe_4\ S_4$ LJBDFODJNLIPKO-VKOJMFJBAC		Ligand Interaction
6MO **Query on 6MO** Download SDF File Download CCD File	A, F	**MOLYBDENUM(VI) ION** Mo HCNGUXXTNNIKCQ-UHFFFAOYSA-N	Mo^{6+}	Ligand Interaction
MGD **Query on MGD** Download SDF File Download CCD File	A, F	**2-AMINO-5,6-DIMERCAPTO-7-METHYL-3,7,8A,9-TETRAHYDRO-8-OXA-1,3,9,10-TETRAAZA-ANTHRACEN-4-ONE GUANOSINE DINUCLEOTIDE** *MOLYBDOPTERIN GUANOSINE DINUCLEOTIDE* $C_{20}\ H_{26}\ N_{10}\ O_{13}\ P_2\ S_2$ VQAGYJCYOLHZDH-ILXWUORBSA-N		Ligand Interaction

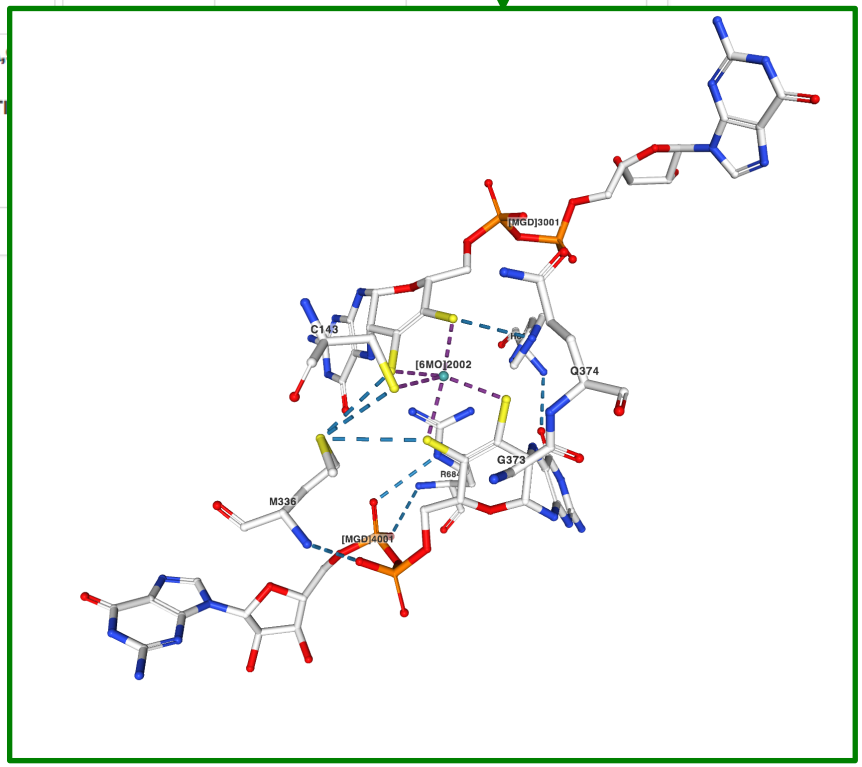
1.3 Working with the PDB

- **2NYA: small molecules menu**

Small Molecules

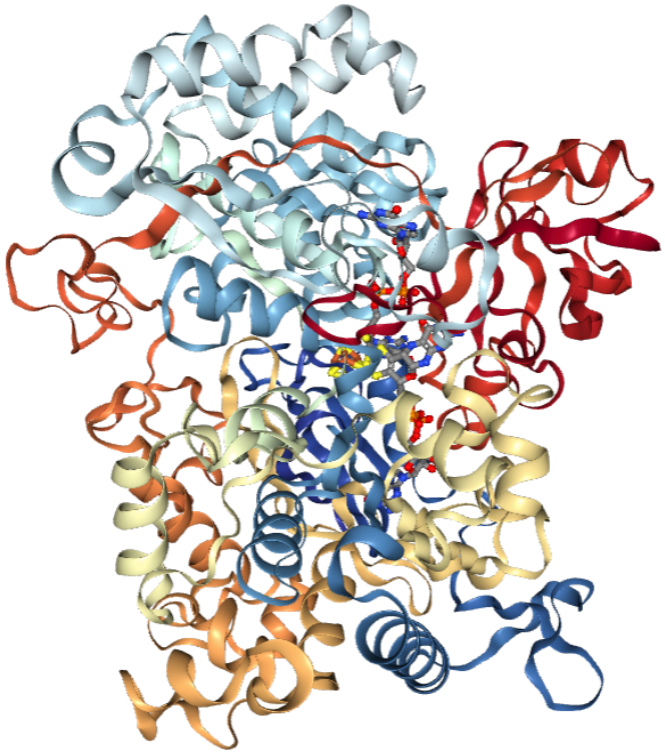
Ligands 3 Unique

ID	Chains	Name / Formula / InChI Key	2D Diagram & Interactions	3D Interactions
SF4 Query on SF4 Download SDF File Download CCD File	A, F	**IRON/SULFUR CLUSTER** $Fe_4 S_4$ LJBDFODJNLIPKO-VKOJMFJBAC		Ligand Interaction
6MO Query on 6MO Download SDF File Download CCD File	A, F	**MOLYBDENUM(VI) ION** Mo HCNGUXXTNNIKCQ-UHFFFAOYSA-N	Mo^{6+}	Ligand Interaction
MGD Query on MGD Download SDF File Download CCD File	A, F	**2-AMINO-5,6-DIMERCAPTO-7-METHYL-3,7,8A,** **TETRAHYDRO-8-OXA-1,3,9,10-TETRAAZA-** **ANTHRACEN-4-ONE GUANOSINE DINUCLEOTI** *MOLYBDOPTERIN GUANOSINE DINUCLEOTIDE* $C_{20} H_{26} N_{10} O_{13} P_2 S_2$ VQAGYJCYOLHZDH-ILXWUORBSA-N		

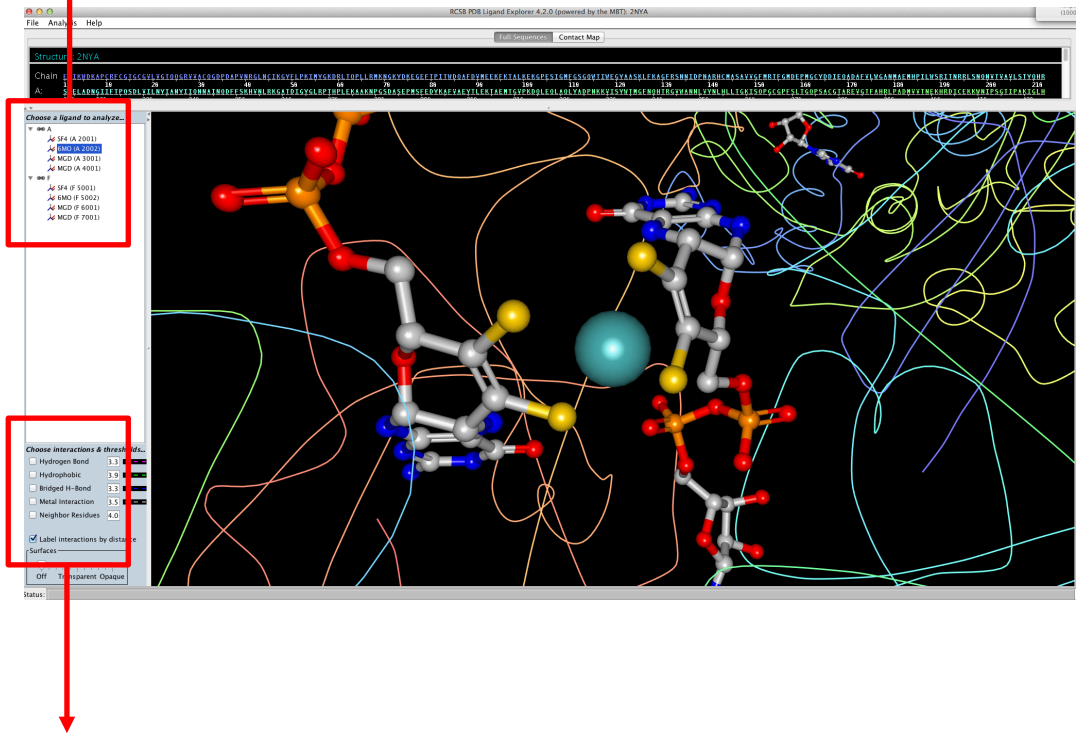


1.4 Ligand Explorer

- 2NYA: Ligand Explorer: Focuses on small molecules



Small molecules in 2NYA

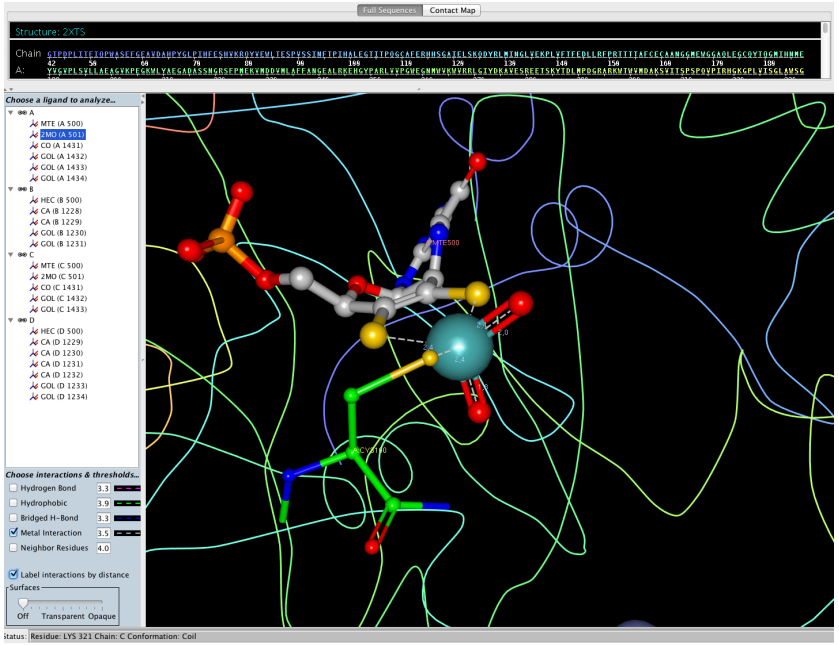


Visualization options: - Metal interactions
- Hidrogen bonds
- Neighbour residues

1.4 Ligand Explorer

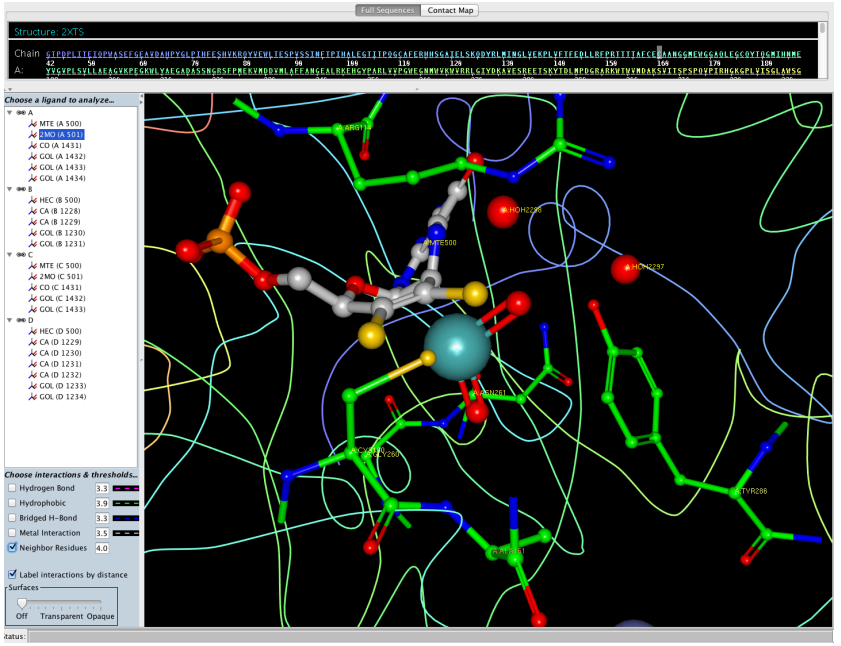
- Visualizing the MoO_2 group (ID: 2MO) on the 2XTS strcuture

Metal Interactions



3S (2.4 Å)
2O (2.0 and 1.8Å)

Neighbor residues

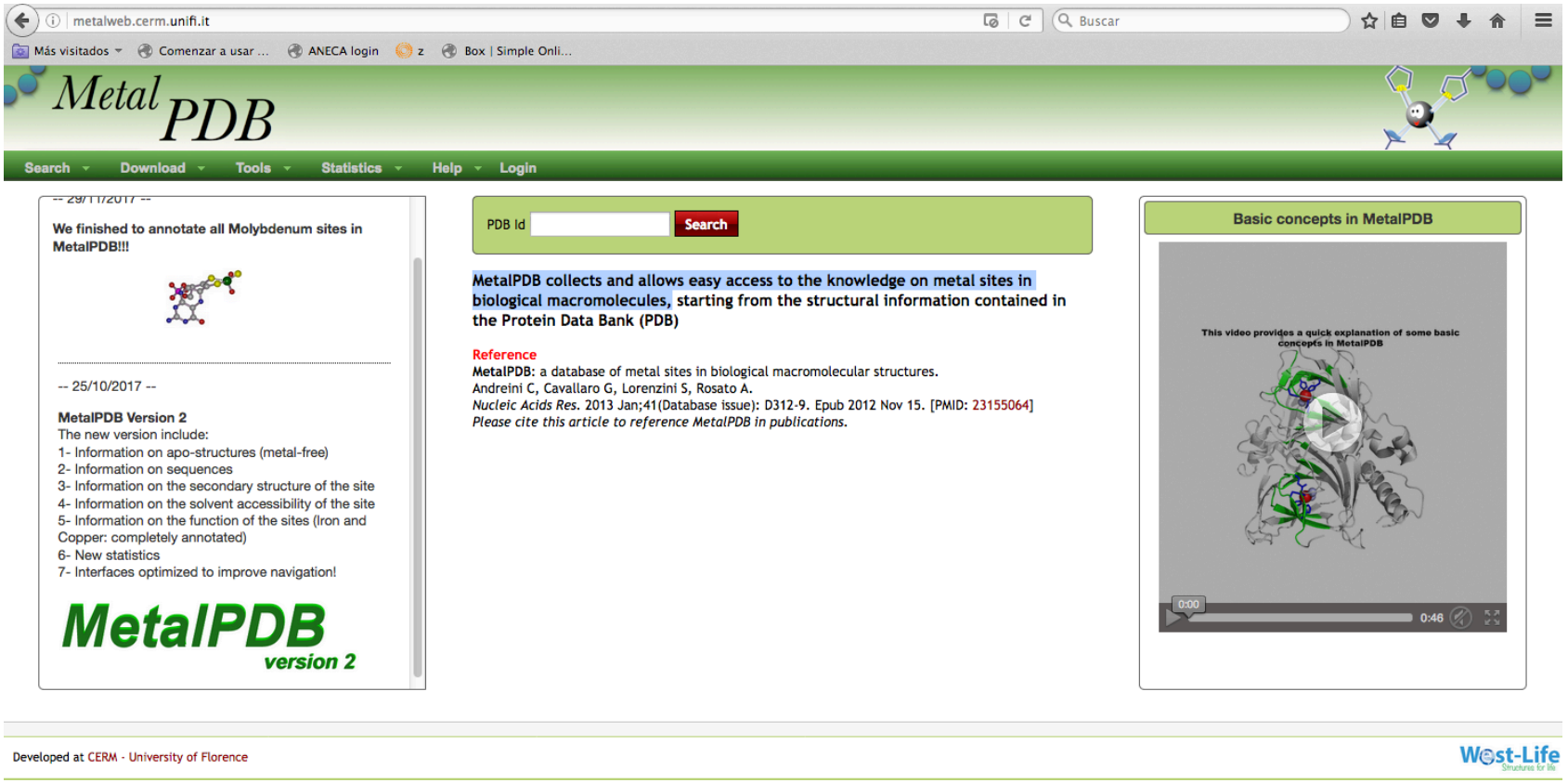


Backbone: ARG114, MTE500, ASN261, CYS160, GLY260, ALA161, TYR282

Two water molecules: HOH2298 HOH2297

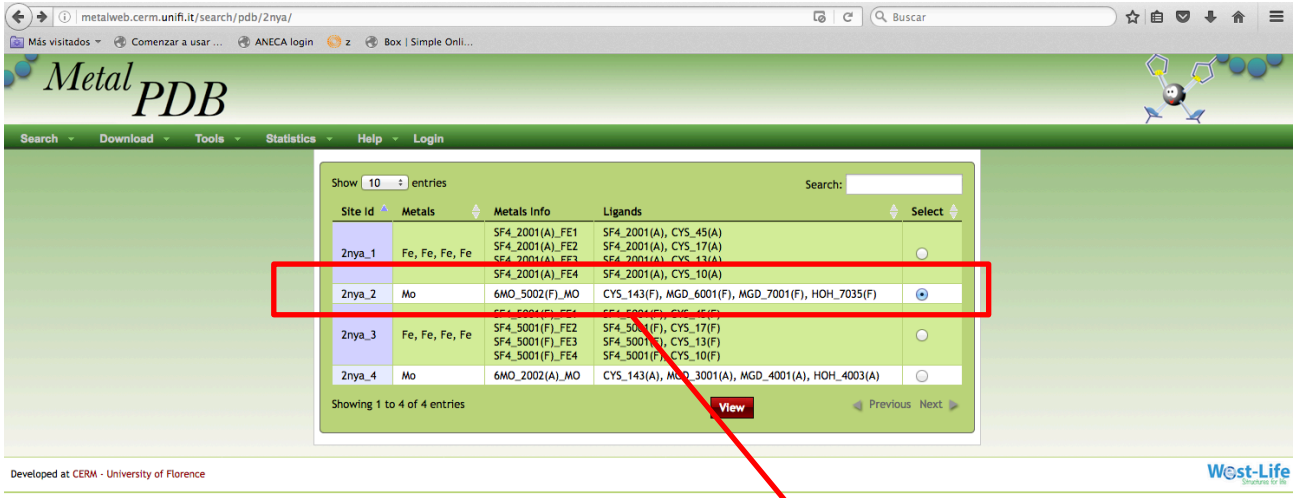
1.5 Metal Protein Data Bank (Metal-PDB)

Uses PDB information, but centered in providing with quick access to metal centers in biological systems



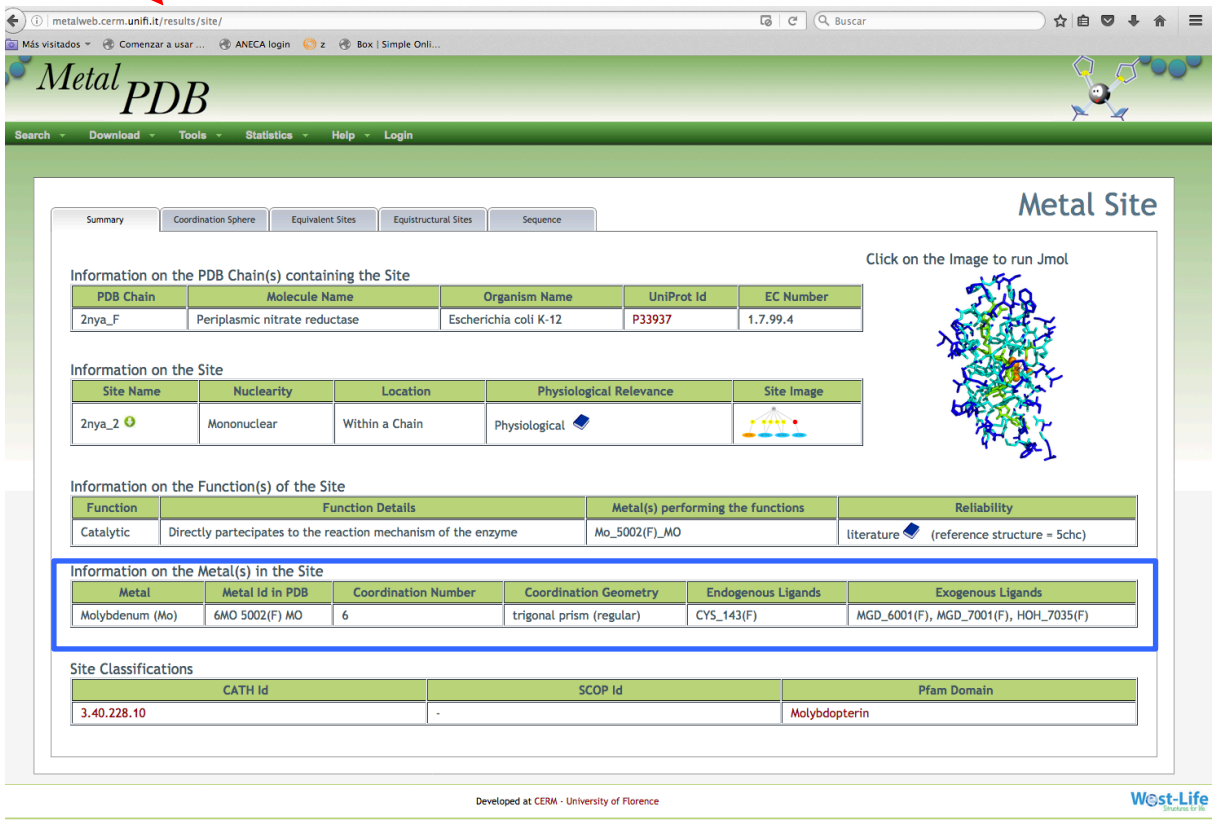
metalweb.cerm.unifi.it

1.5 Metal Protein Data Bank (Metal-PDB)



2NYA(nitrate reductase)

Metal site coordination chemistry



1.5 Metal Protein Data Bank (Metal-PDB)

- How can we search for exogenous groups?

  Chemical Component Dictionary (htpps://www.wwpdb.org/data/ccd)

  - PDB Dictionary: Residues, small molecules, ….
  - Contains accurate chemical descriptors
  - Organizes via 3 alphanumeric characters
  - Can be accessed via Ligand Explorer

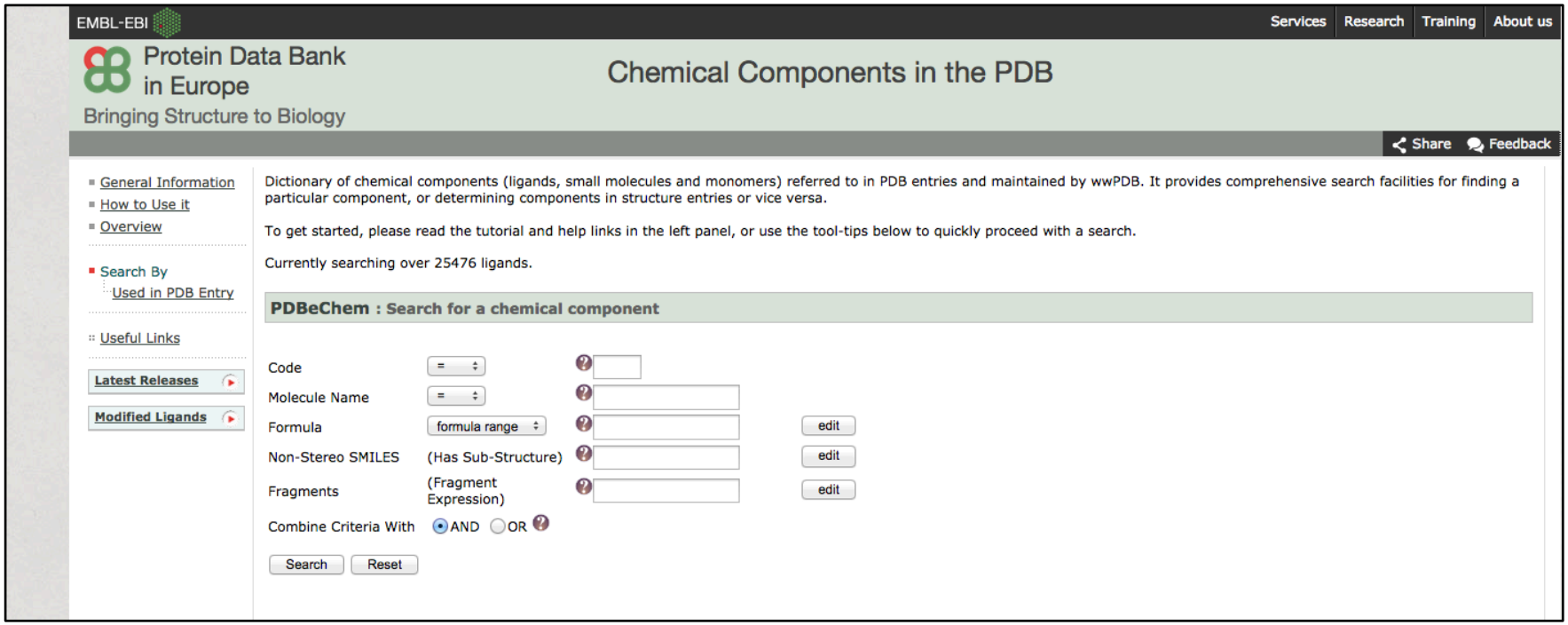
1.5 Metal Protein Data Bank (Metal-PDB)

- How can we search for exogenous groups?

  Chemical Component Dictionary (htpps://www.wwpdb.org/data/ccd)

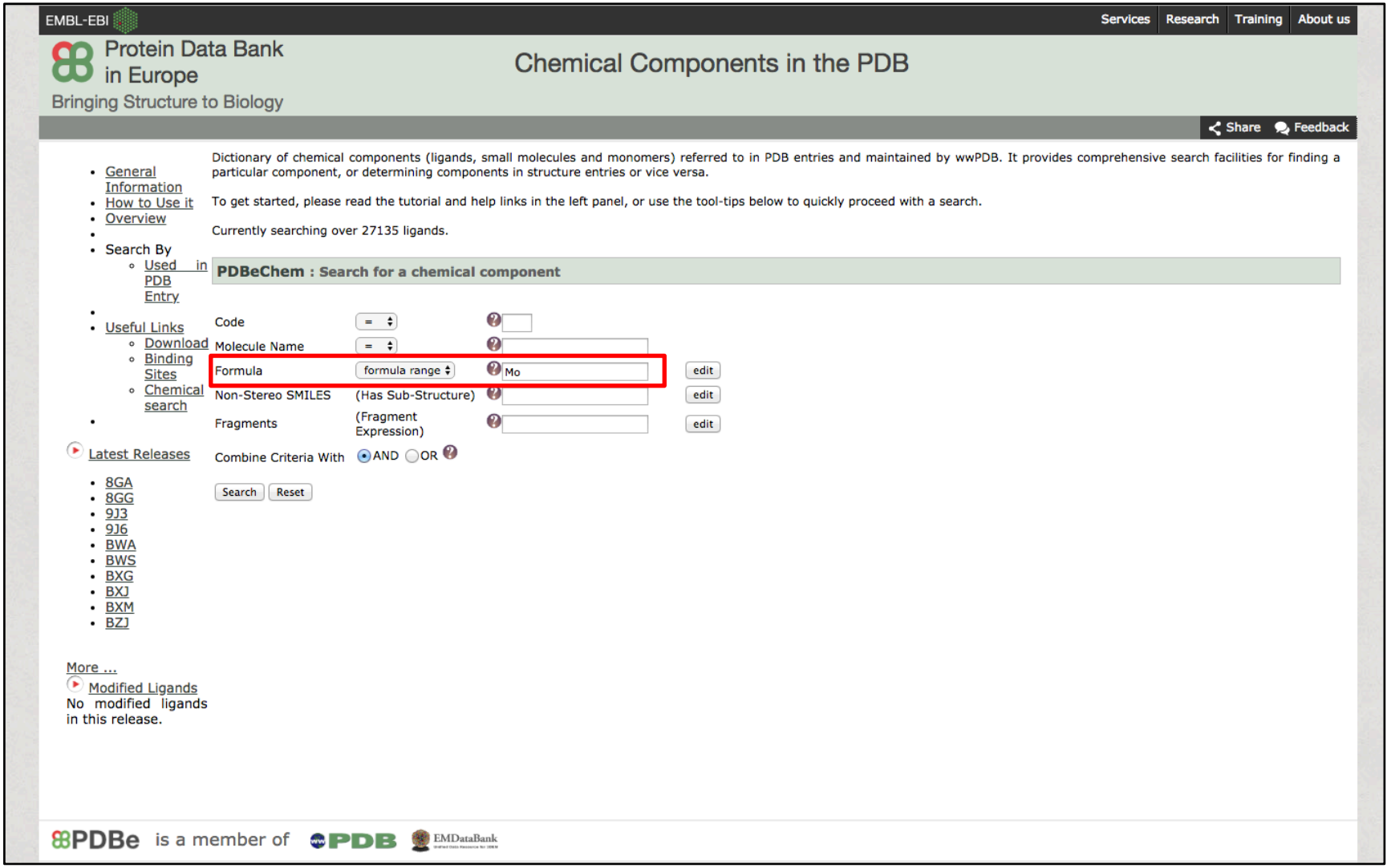
  - PDB Dictionary: Residues, small molecules, ….
  - Contains accurate chemical descriptors
  - Organizes via 3 alphanumeric characters
  - Can be accessed via Ligand Explorer

- PDBeChem (*http://www.ebi.ac.uk/pdbe-srv/pdbechem/)*



1.5 Metal Protein Data Bank (Metal-PDB)

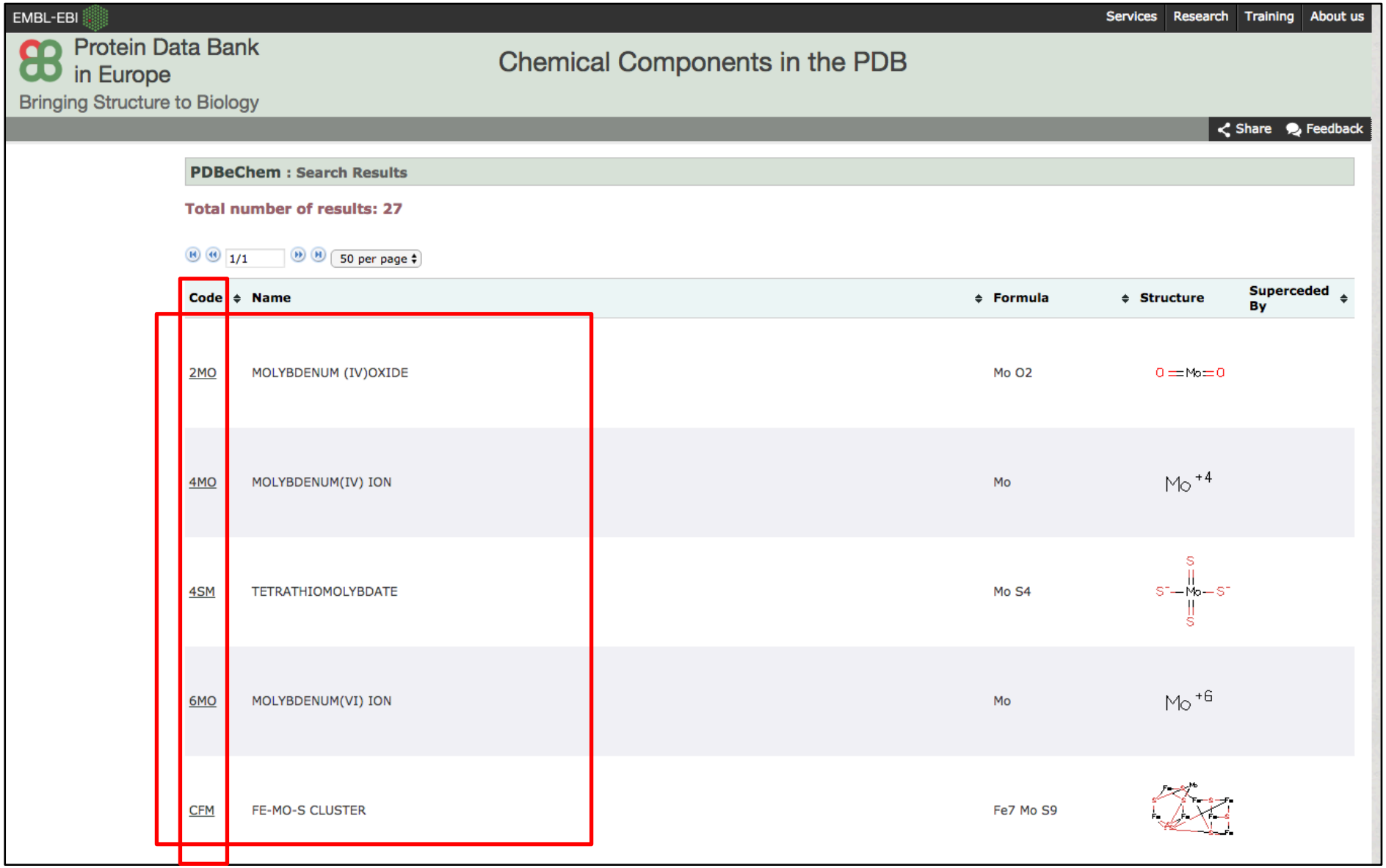
- ***Formula***: Allows us to search for metal/metals and recover the ligands ID



PDBeChem (*http://www.ebi.ac.uk/pdbe-srv/pdbechem/*)

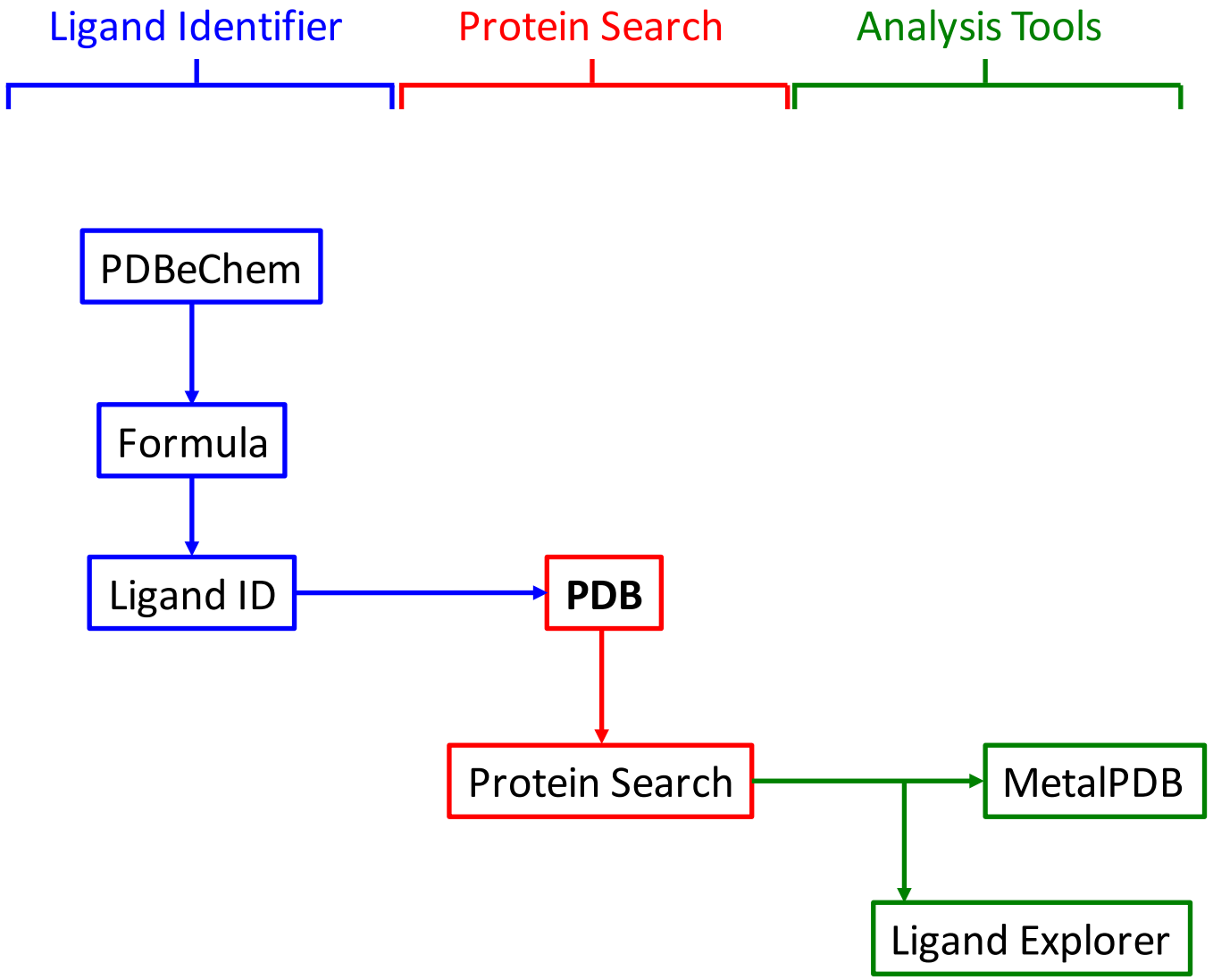
1.5 Metal Protein Data Bank (Metal-PDB)

- The ID codes allows to search in the PDB for proteins with a given cofactor



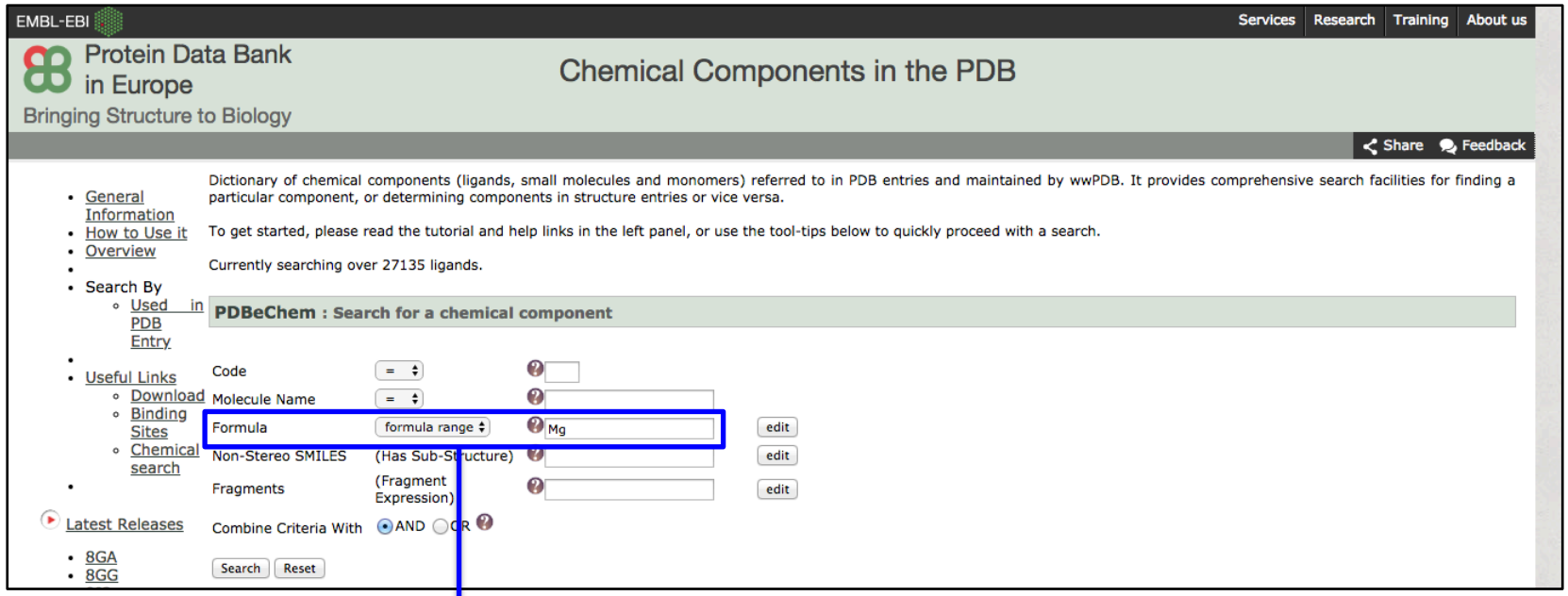
PDBeChem (*http://www.ebi.ac.uk/pdbe-srv/pdbechem/*)

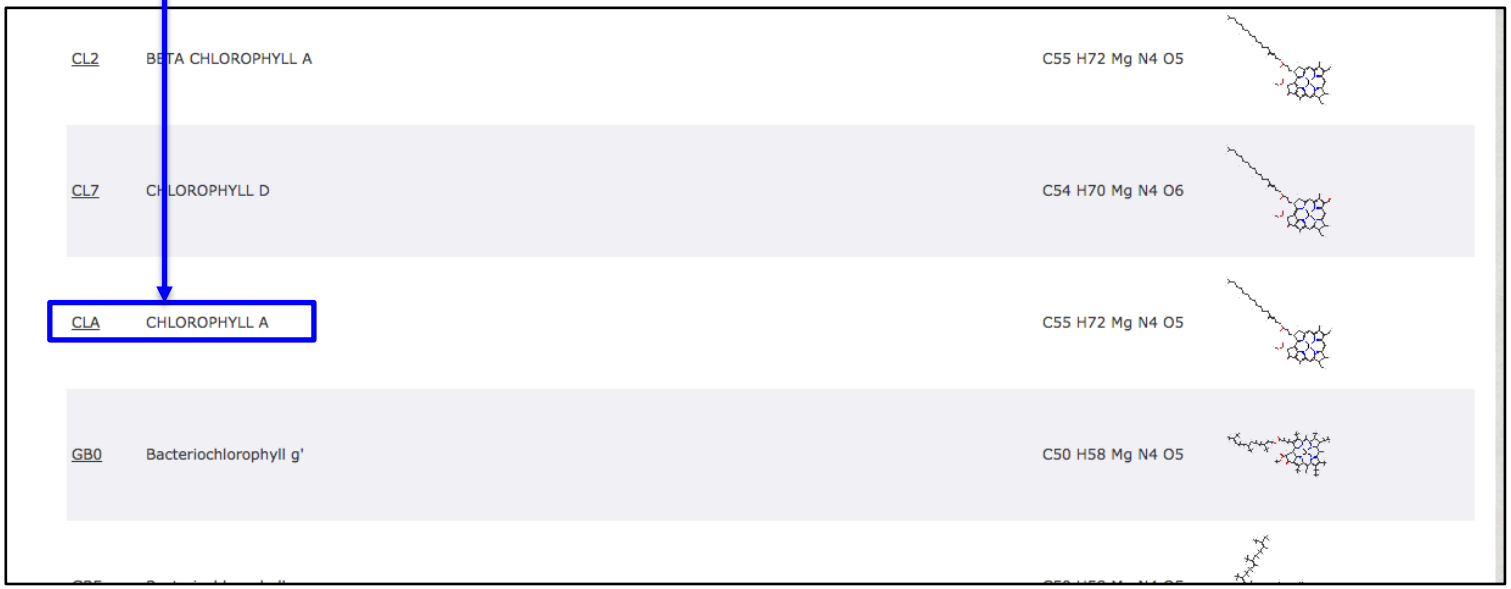
1.5 Metal Protein Data Bank (Metal-PDB)



1.5 Metal Protein Data Bank (Metal-PDB)

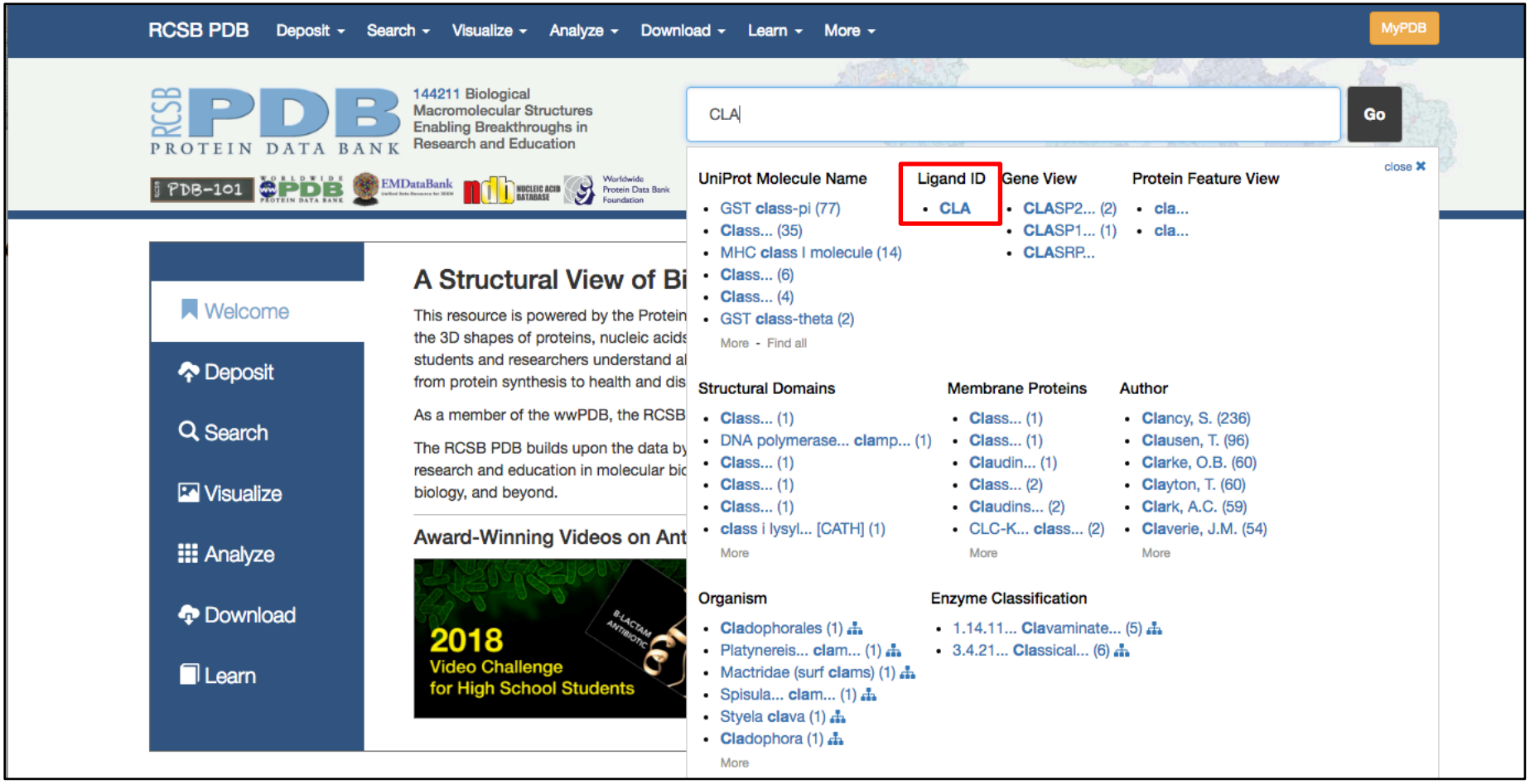
- ***Practical case 1:*** *Search for a structure with chlorophyll*





1.5 Metal Protein Data Bank (Metal-PDB)

- ***Practical case 1:*** *Search for a structure with chlorophyll*



1.5 Metal Protein Data Bank (Metal-PDB)

- ***Practical case 1:*** *Search for a structure with chlorophyll*

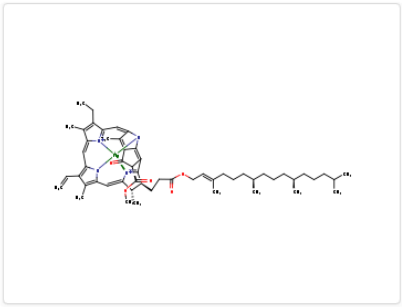
RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB

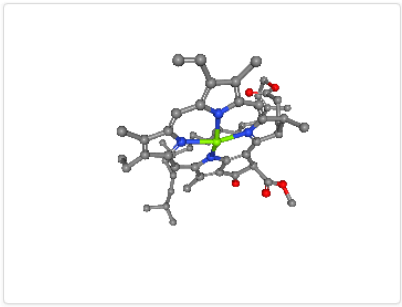
RCSB PDB PROTEIN DATA BANK 144211 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Search by PDB ID, author, macromolecule, sequence, or ligands Go

Advanced Search | Browse by Annotations

PDB-101 EMDataBank Nucleic Acid Database Worldwide Protein Data Bank Foundation





Rotate Hydrogens Labels

View / Download Files

CLA

CHLOROPHYLL A

CLA as a free ligand exists in 94 entries. Examples include: 2E75 2WSE 5XNO

Find related ligands: Stereoisomers Similar ligands Chemical Structure Search

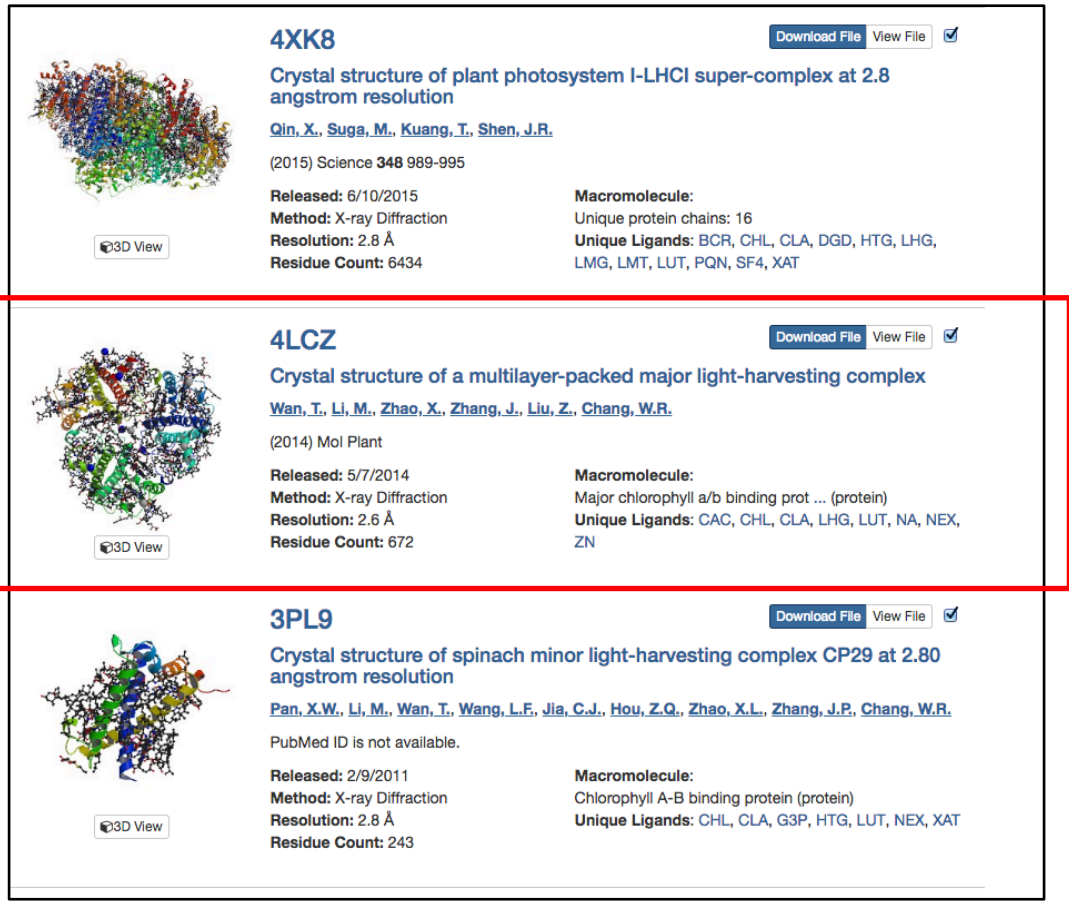
View summary at Ligand Expo

Chemical Component Summary	
Name	CHLOROPHYLL A
Identifiers	n/a
Formula	C_{55} H_{72} Mg N_4 O_5
Molecular Weight	893.49
Type	NON-POLYMER
Isomeric SMILES	CCC1=C(C)C2=Cc3c(C=C)c(C)c4C=C5[C@@H](C)[C@H](CCC(=O)OC\C=C(/C)CCC[C@H](C)CCC[C@H](C)CCCC(C)C)C6=[N+]5[Mg@]5(n34)n3c(=CC1=[N+]25)c(C)c1C(=O)[C@H](C(=O)OC)C6=c31
InChI	InChI=1S/C55H73N4O5.Mg/c1-13-39-35(8)42-28-44-37(10)41(24-25-48(60)64-27-26-34(7)23-17-22-33(6)21-16-20-32(5)19-15-18-31(3)4)52(58-44)50-51(55(62)63-12)54(61)49-38(11)45(59-53(49)50)30-47-40(14-2)36(9)43(57-47)29-46(39)56-42;/h13,26,28-33,37,41,51H,1,14-25,27H2,2-12H3,(H-,56,57,58,59,61);/q-1;+4/p-1/b34-26+;/t32-,33-,37+,41+,51-;/m1./s1
InChIKey	VIQFHHZSLDFWDU-AENOIHSZSA-M

Chemical Details	
Formal Charge	2
Atom Count	137
Chiral Atom Count	8
Chiral Atoms	C13, C2A, C3A, C8, CBD, NA, NC, ND
Bond Count	145
Aromatic Bond Count	5
Leaving Atoms	n/a

1.5 Metal Protein Data Bank (Metal-PDB)

- ***Practical case 1:*** *Search for a structure with chlorophyll*



4XK8 Download File | View File

Crystal structure of plant photosystem I-LHCI super-complex at 2.8 angstrom resolution

Qin, X., Suga, M., Kuang, T., Shen, J.R.

(2015) Science **348** 989-995

Released: 6/10/2015
Method: X-ray Diffraction
Resolution: 2.8 Å
Residue Count: 6434

Macromolecule:
Unique protein chains: 16
Unique Ligands: BCR, CHL, CLA, DGD, HTG, LHG, LMG, LMT, LUT, PQN, SF4, XAT

3D View

4LCZ Download File | View File

Crystal structure of a multilayer-packed major light-harvesting complex

Wan, T., Li, M., Zhao, X., Zhang, J., Liu, Z., Chang, W.R.

(2014) Mol Plant

Released: 5/7/2014
Method: X-ray Diffraction
Resolution: 2.6 Å
Residue Count: 672

Macromolecule:
Major chlorophyll a/b binding prot ... (protein)
Unique Ligands: CAC, CHL, CLA, LHG, LUT, NA, NEX, ZN

3D View

3PL9 Download File | View File

Crystal structure of spinach minor light-harvesting complex CP29 at 2.80 angstrom resolution

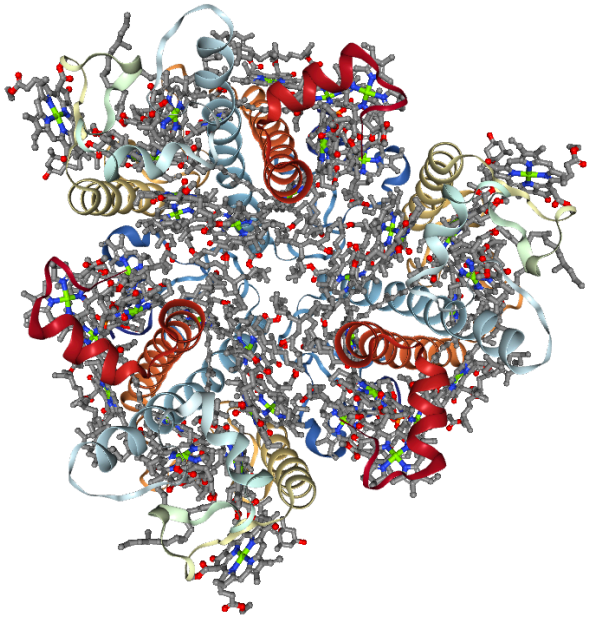
Pan, X.W., Li, M., Wan, T., Wang, L.F., Jia, C.J., Hou, Z.Q., Zhao, X.L., Zhang, J.P., Chang, W.R.

PubMed ID is not available.

Released: 2/9/2011
Method: X-ray Diffraction
Resolution: 2.8 Å
Residue Count: 243

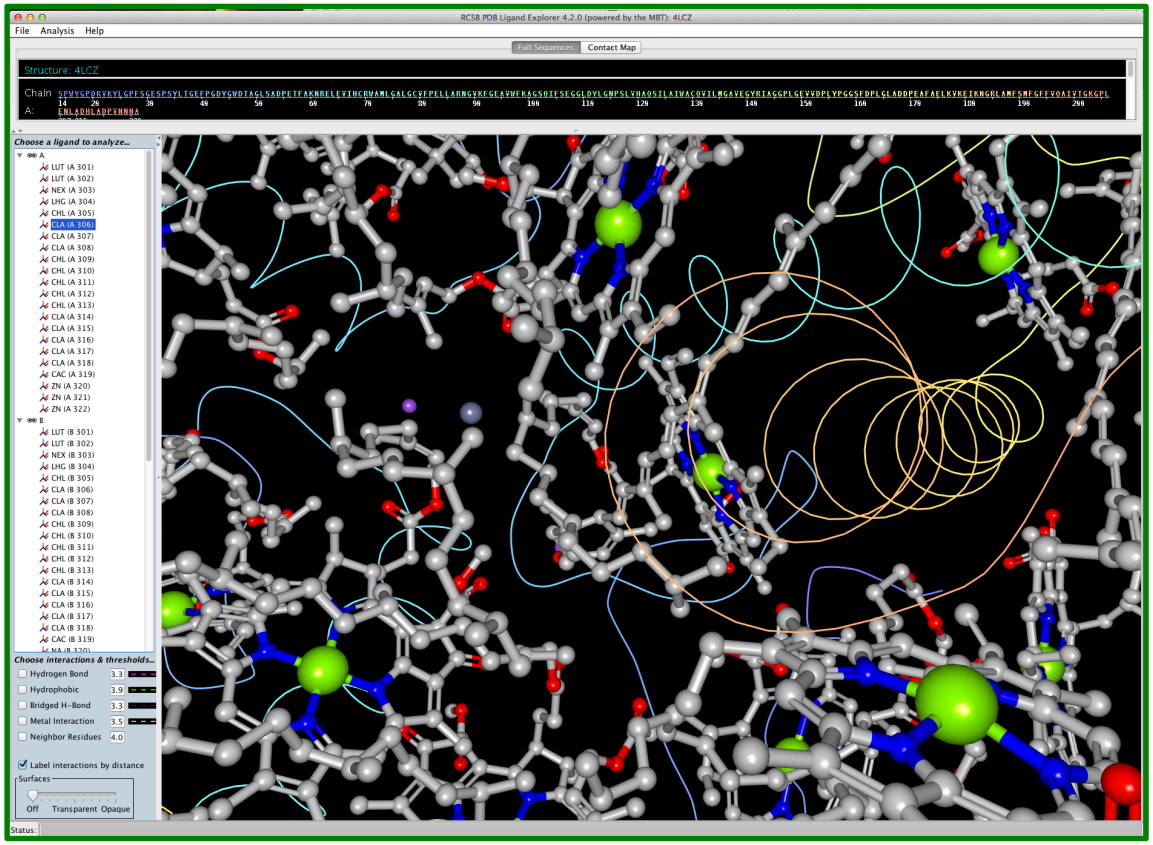
Macromolecule:
Chlorophyll A-B binding protein (protein)
Unique Ligands: CHL, CLA, G3P, HTG, LUT, NEX, XAT

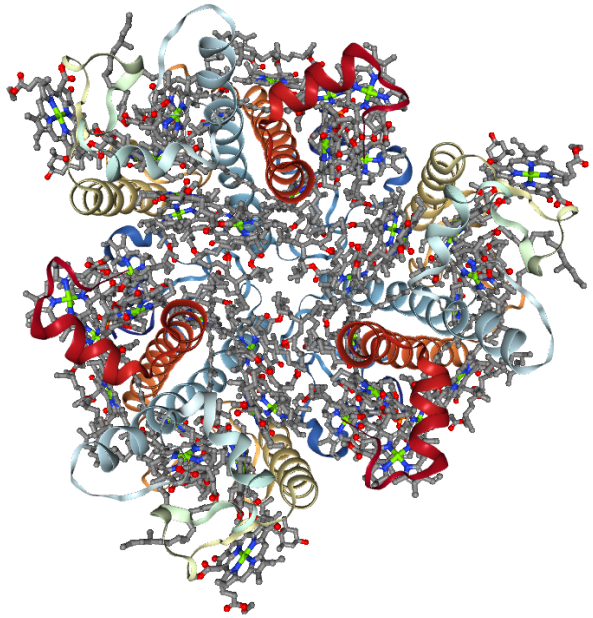
3D View



1.5 Metal Protein Data Bank (Metal-PDB)

- ***Practical case 1:*** *Search for a structure with chlorophyll*





1.6 Additional Tools

VMD (Visual Molecular Dynamics)
https://www.ks.uiuc.edu/Research/vmd/

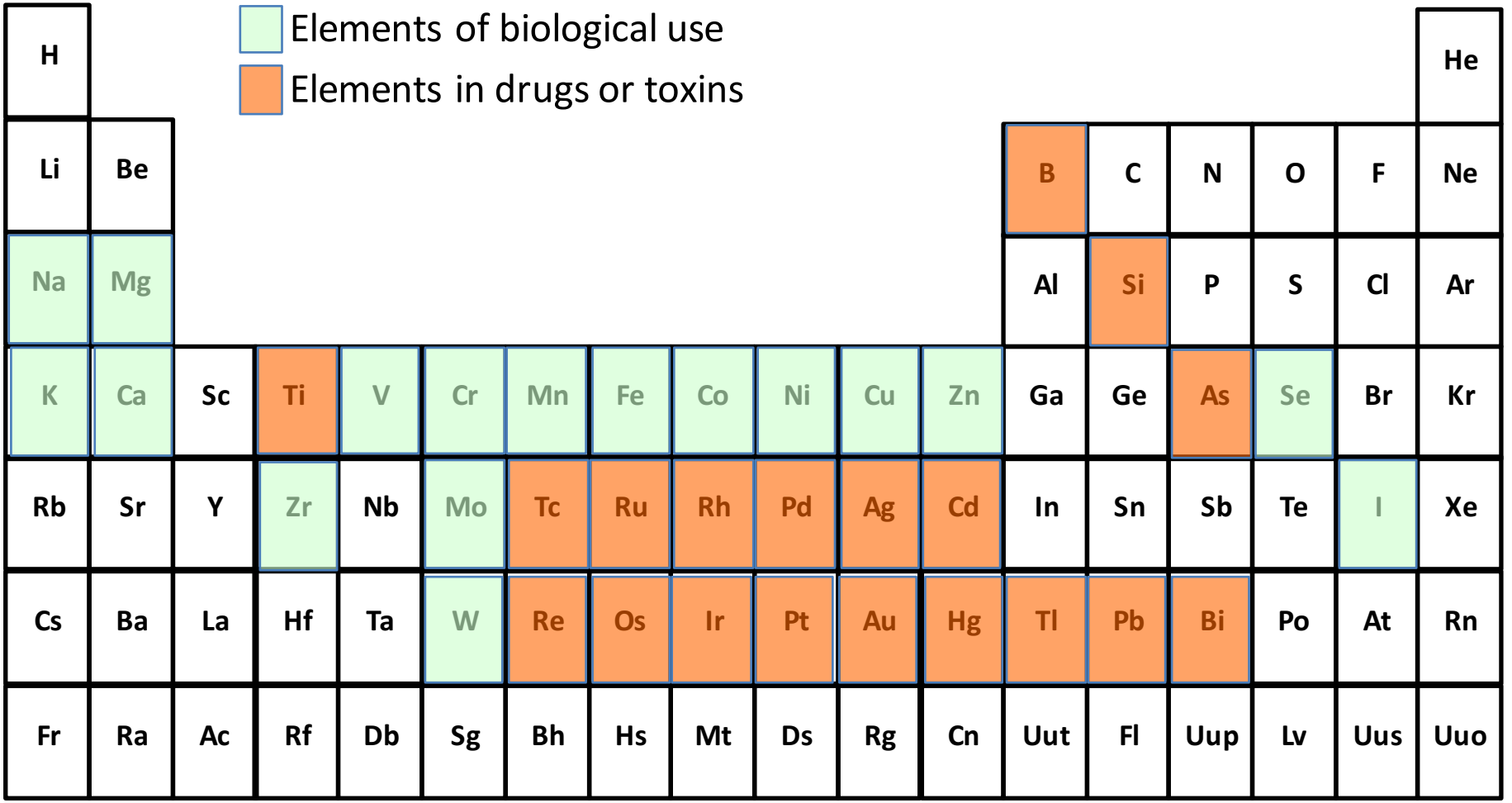
Mercury
https://www.ccdc.cam.ac.uk/solutions/csd-system/components/mercury/

Chimera
https://www.cgl.ucsf.edu/chimera/

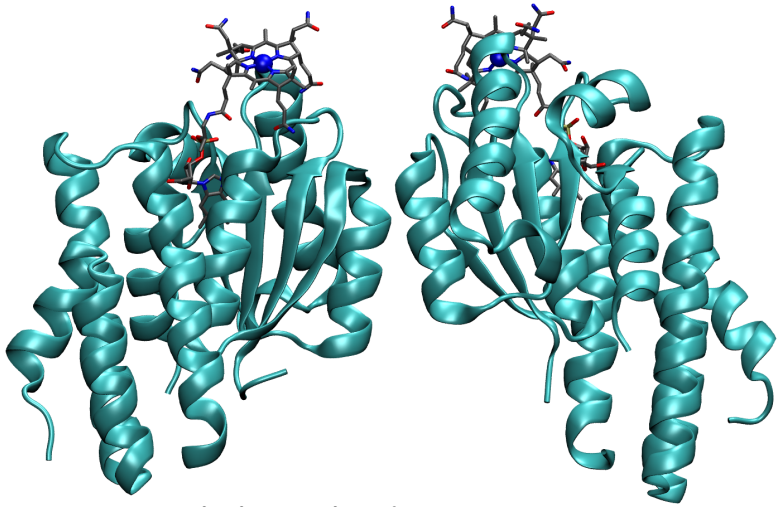
Integrating the PDB as a teaching tool for coordination Chemistry

2.1 Metalloenzymes as illustrative examples

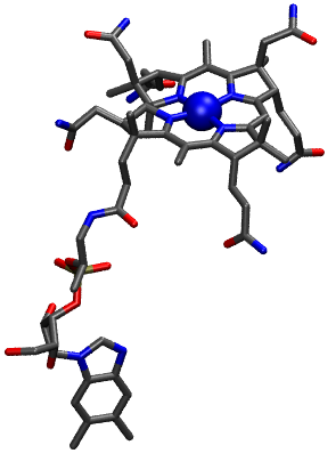
- Metalloenzymes contain metal coordination sites: They can be used as illustrative examples

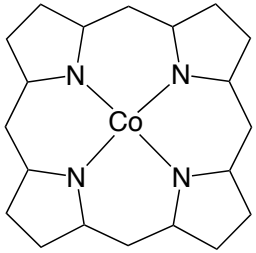


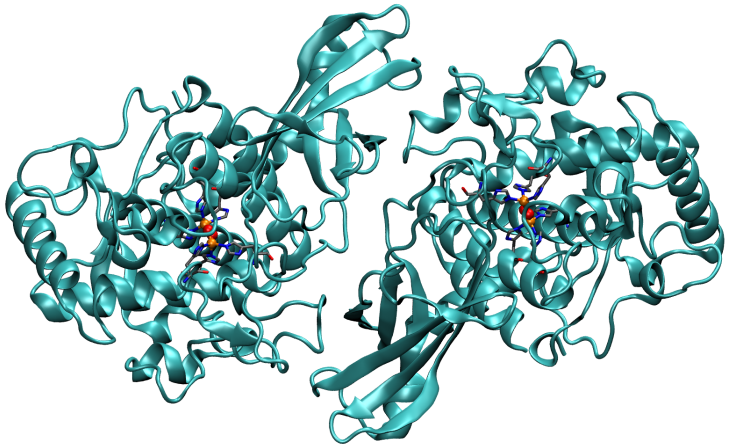
2.1 Metalloenzymes as illustrative examples



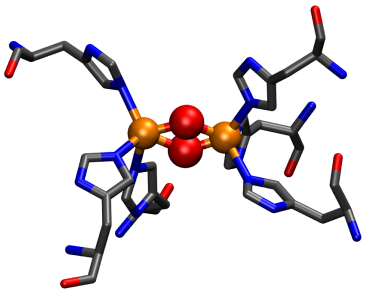
cobalamin-binding protein

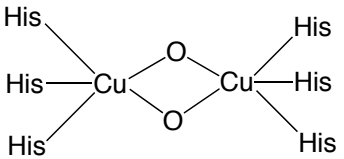






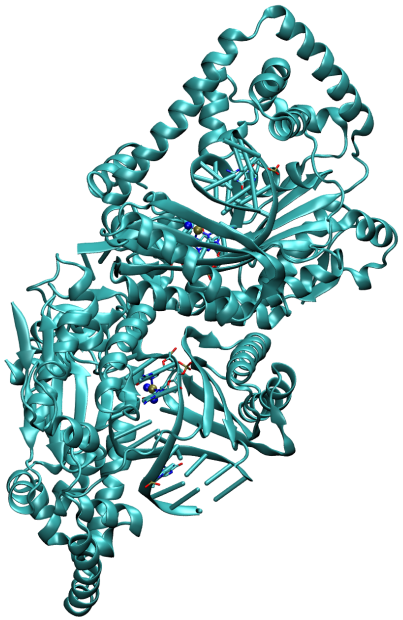
Hemocyanin

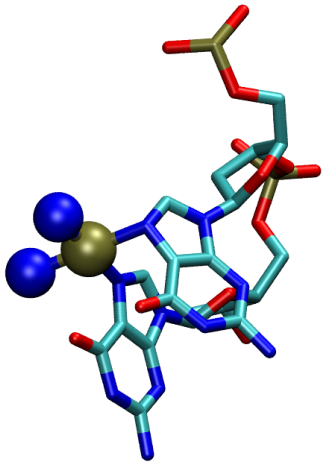


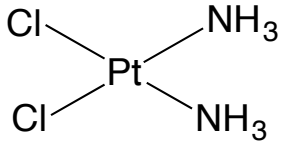


Biological systems allow for a contextualization of coordination chemistry
Visually appealing systems

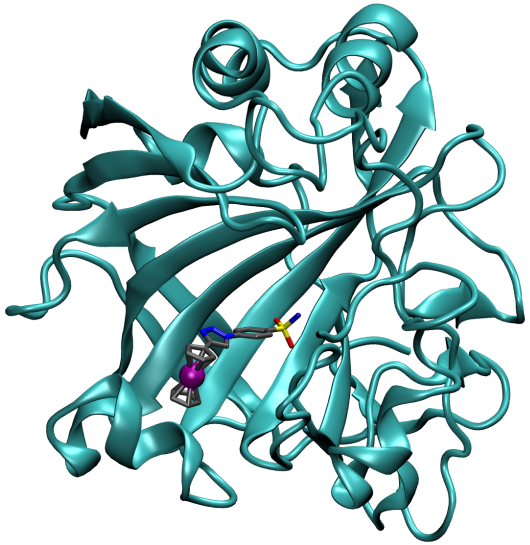
2.1 Metalloenzymes as illustrative examples

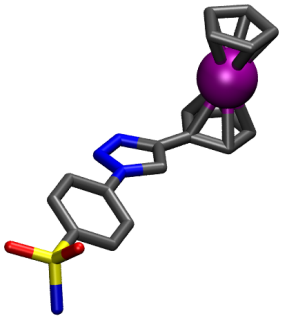


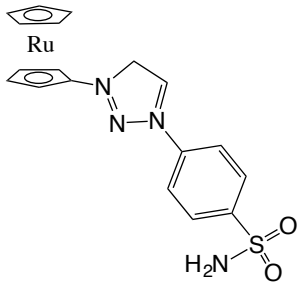




DNA polymerase with DNA containing
a major *cis*-platin lesion

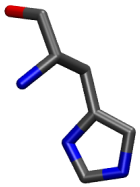




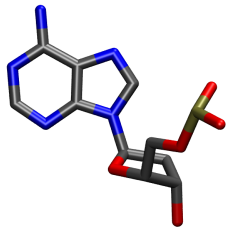


Carbonic anhydrase II in complex with
p-(4-ruthenocenyl-1H-1,2,3-triazol-1yl)benzenesulfonamide

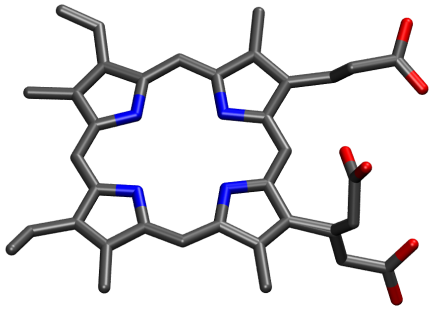
2.2 Coordination motifs in active sites



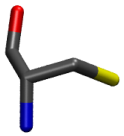
Histidine



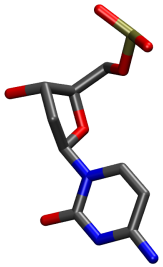
Adenine



chloro/heme group

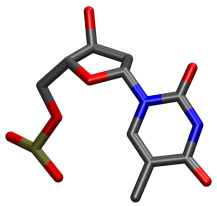


Cysteine

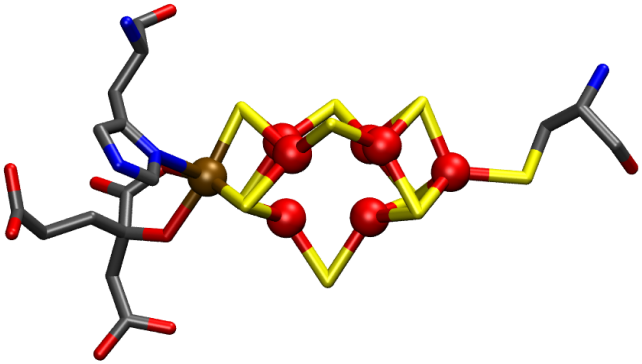


Cytosine

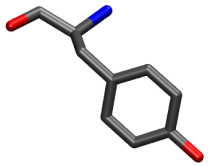
Methyonine



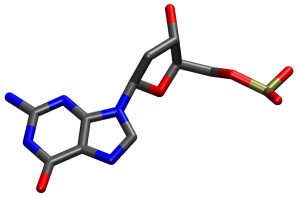
Thymine



Fe-Mo cofactor

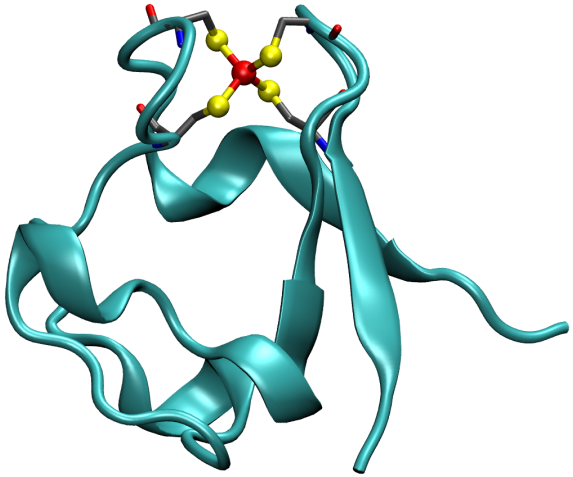


Tyrosine

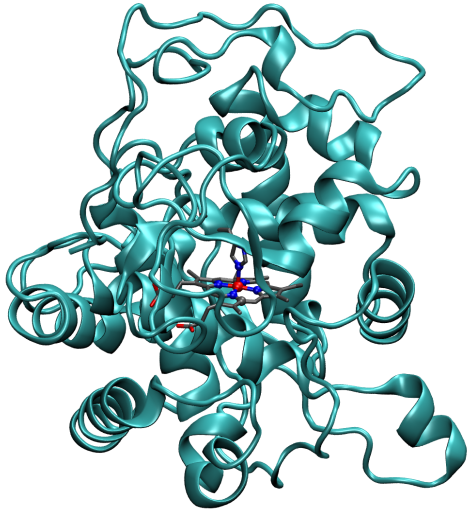


Guanine

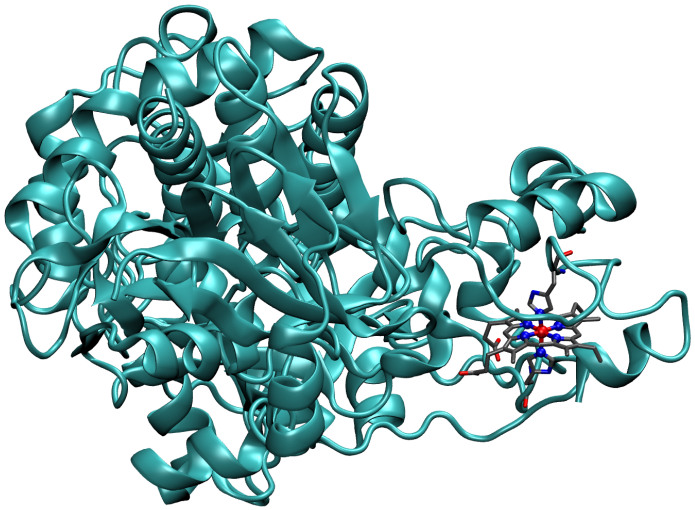
2.3 Geometrical Data/Coordination Modes



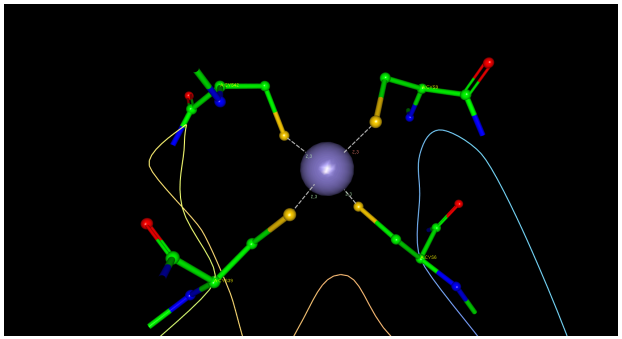
$[FeS_4]$-Td



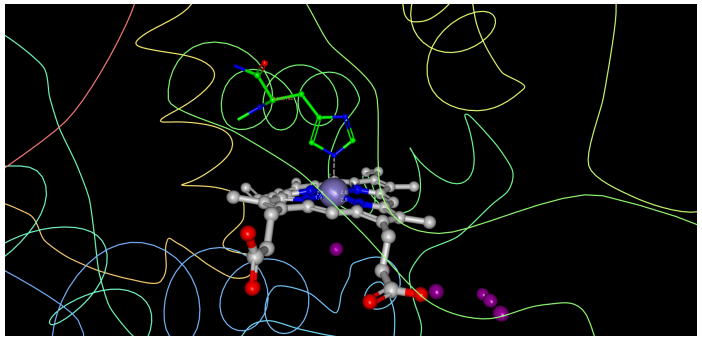
$[FeN_5]$-SPY



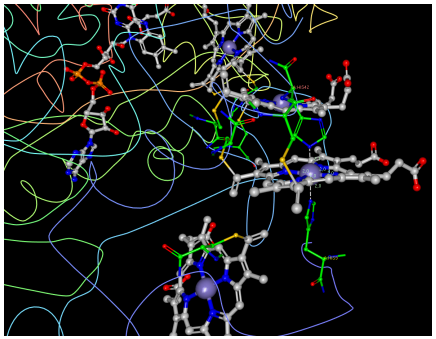
$[FeN_6]$-Oh



r(Fe-S) = 2.3 Å



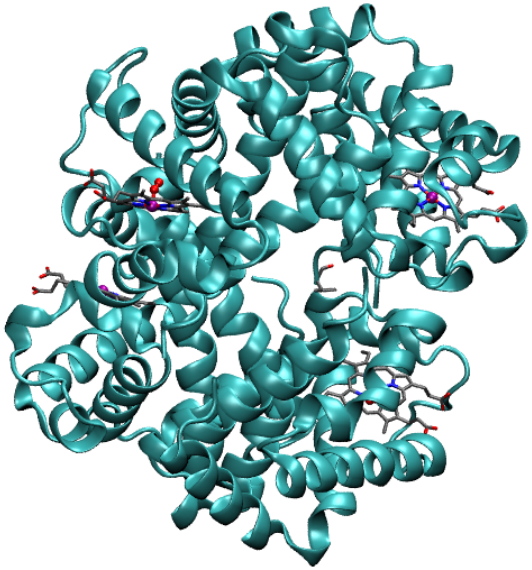
r(Fe-N) = 2.0-2.1 Å



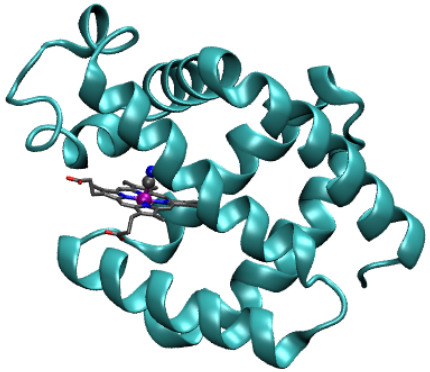
r(Fe-N) = 2.0-2.1 Å

Study of coordination numbers
Donor-acceptor, bonding concepts, metal-ligand bond-length

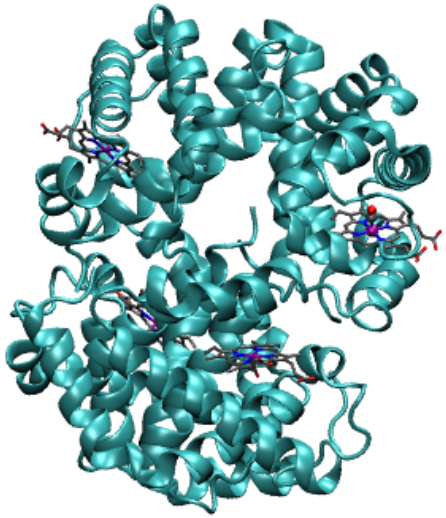
2.3 Geometrical Data/Coordination Modes



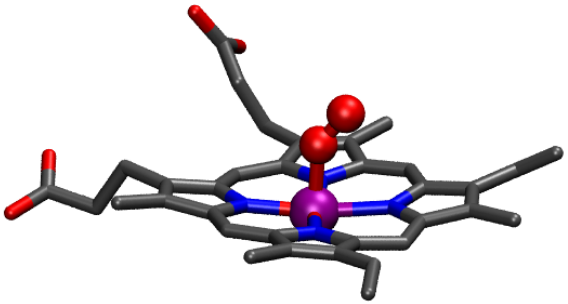
oxy-heme



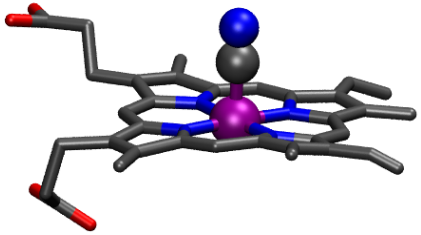
cyano-heme



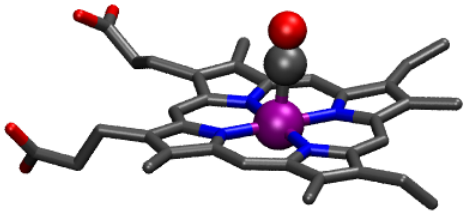
carbonyl-heme



r(Fe-CN) = 1.949 Å



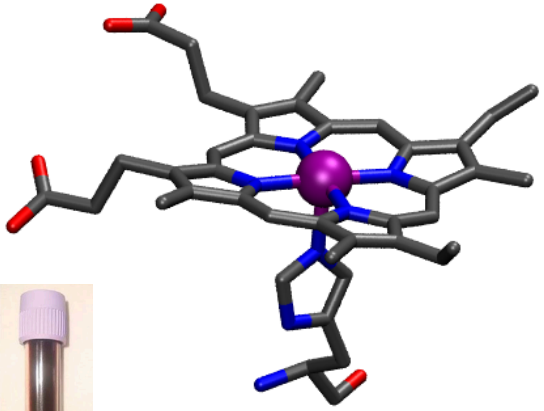
r(Fe-CN) = 1.949 Å



r(Fe-CO) = 2.018 Å

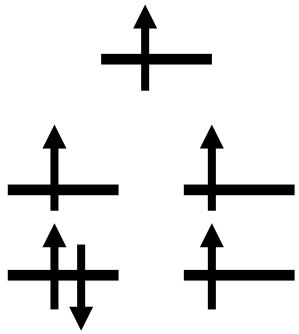
σ-bonding, π-bonding and backbonding

2.4 Spin-State Information

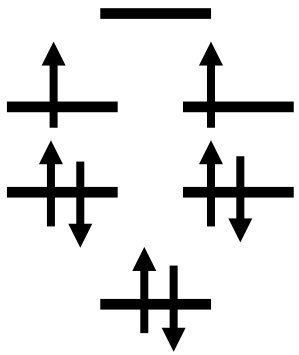


deoxy-heme

+ O_2



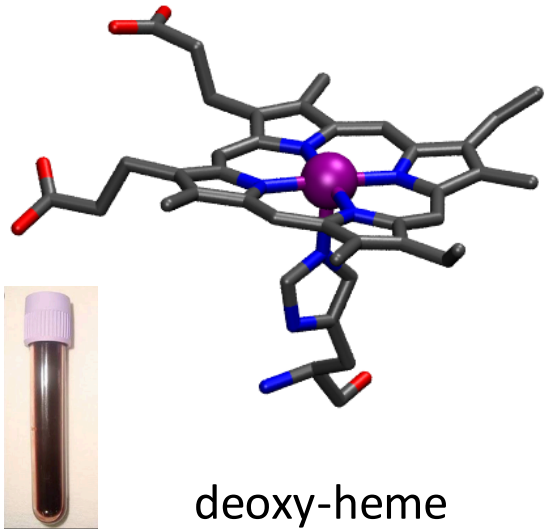
high-spin Fe^{II} (S=2)



O_2 (S=1)

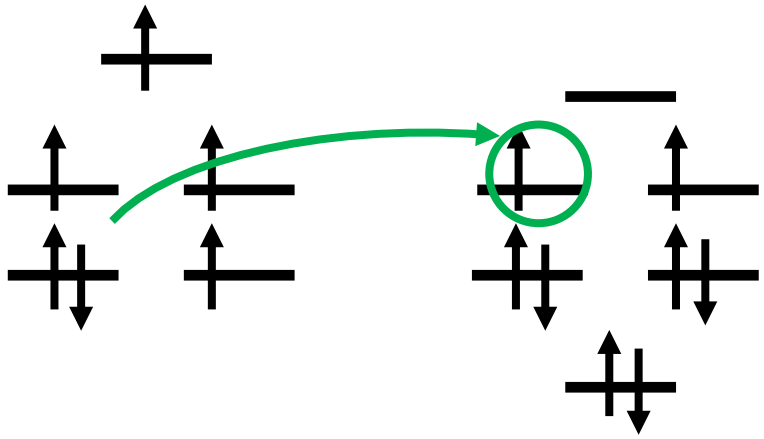
Electronic structure, spin-state, oxidation

2.4 Spin-State Information



deoxy-heme

+ O_2

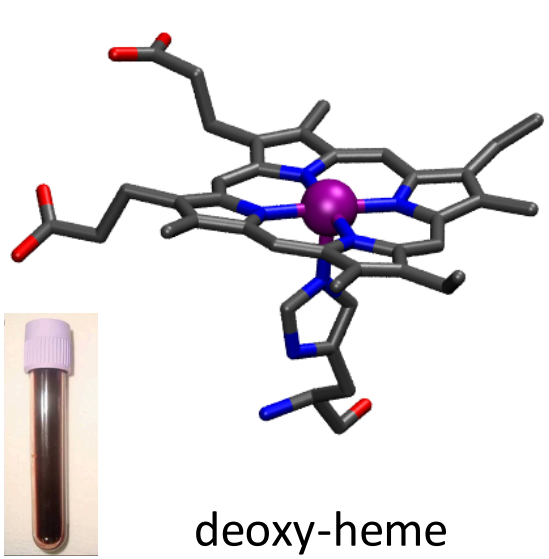


high-spin Fe^{II} (S=2)

O_2 (S=1)

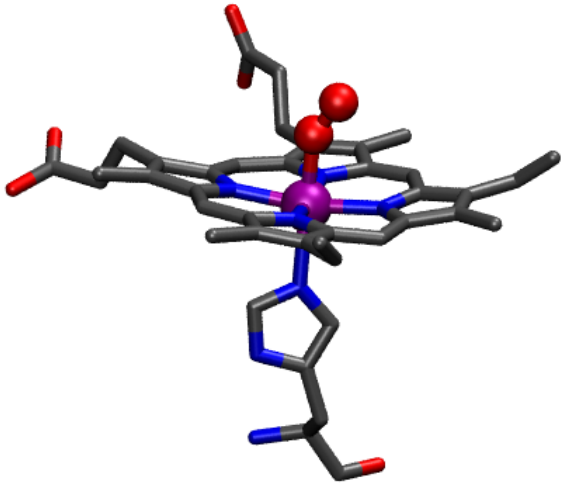
Electronic structure, spin-state, oxidation

2.4 Spin-State Information

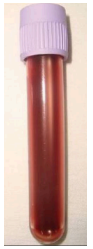


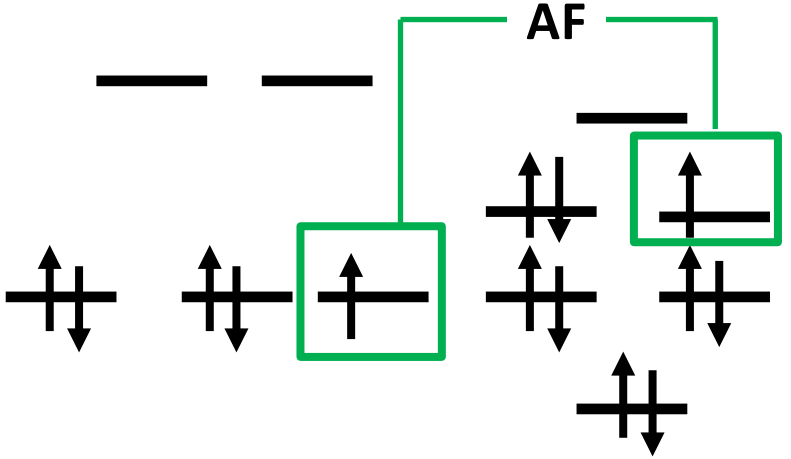
deoxy-heme

+ O_2



oxy-heme

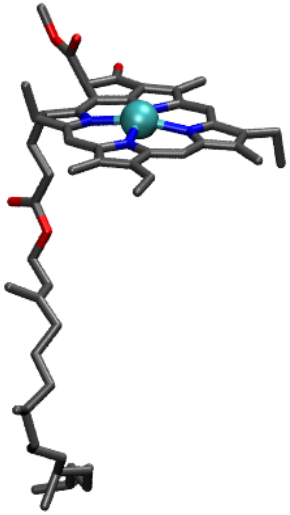




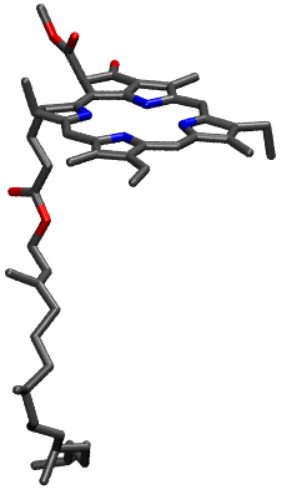
low-spin Fe^{III} (S=1/2) $O_2^{\bullet}$ (S=1/2)

Electronic structure, spin-state, oxidation

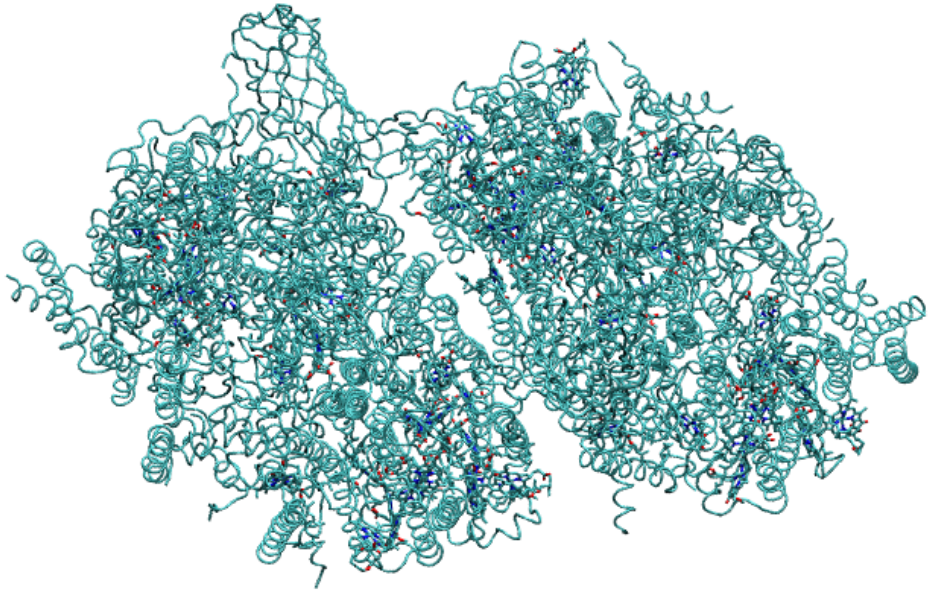
2.5 Absorption Spectroscopy



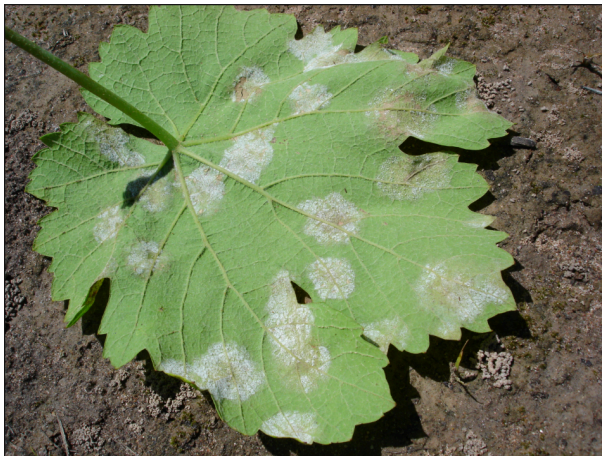
Clorophylla A



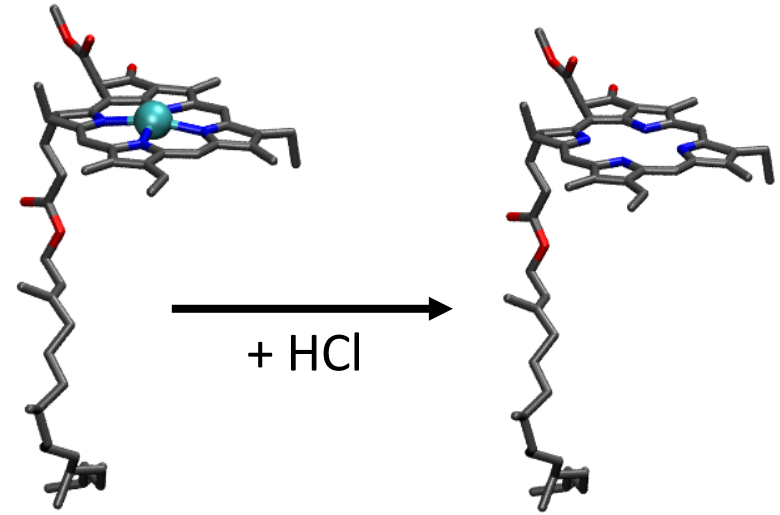
Pheophytin



Photosystem II

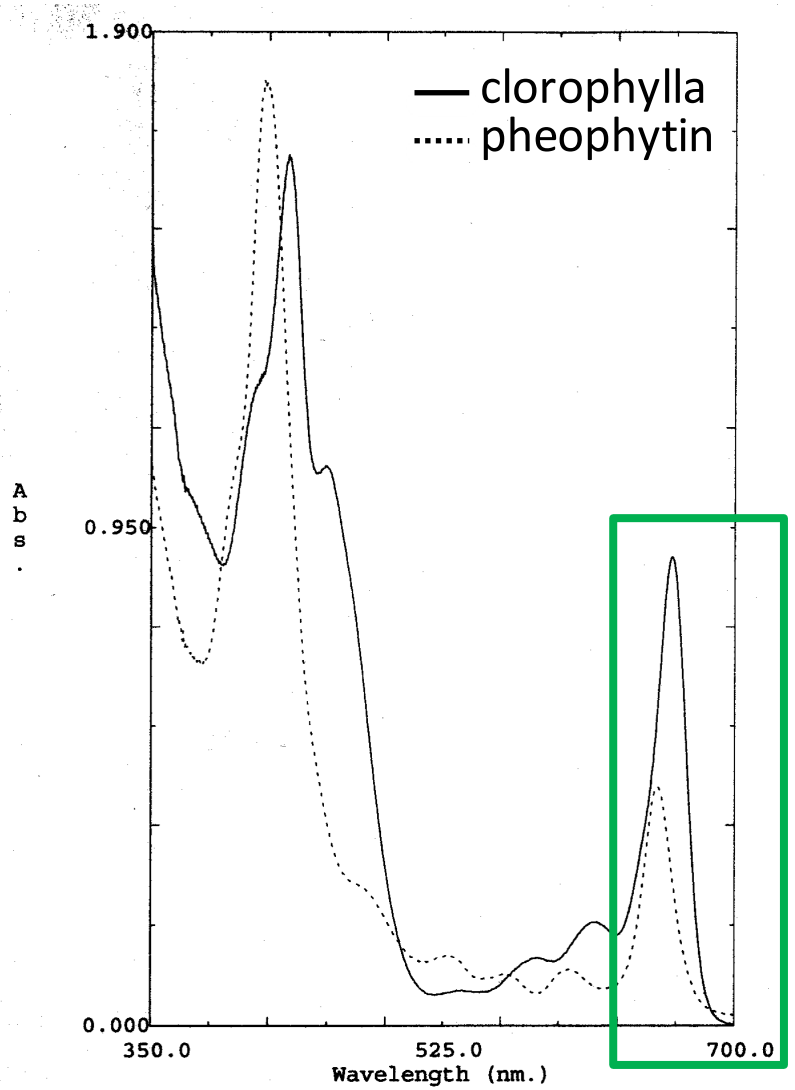


2.5 Absorption Spectroscopy

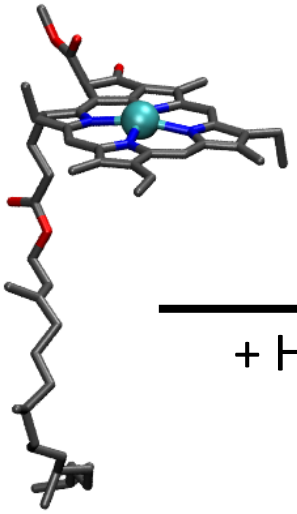


Clorophylla A

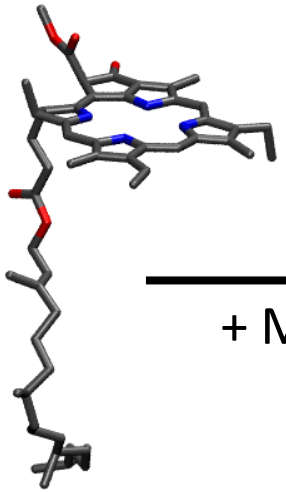
Pheophytin



2.5 Absorption Spectroscopy



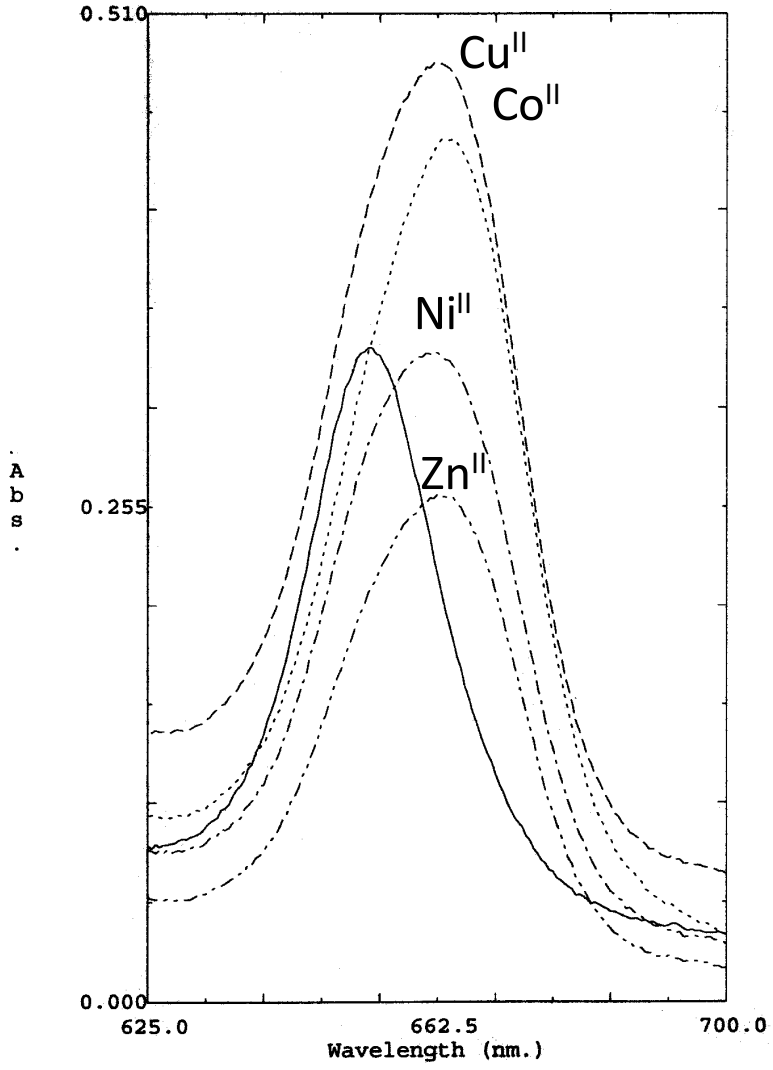
+ HCl



+ M^{n+}

Clorophylla A

Pheophytin



Color can be tuned via metal exchange

Changes in electronic spectra: Spectrochemical series of metals

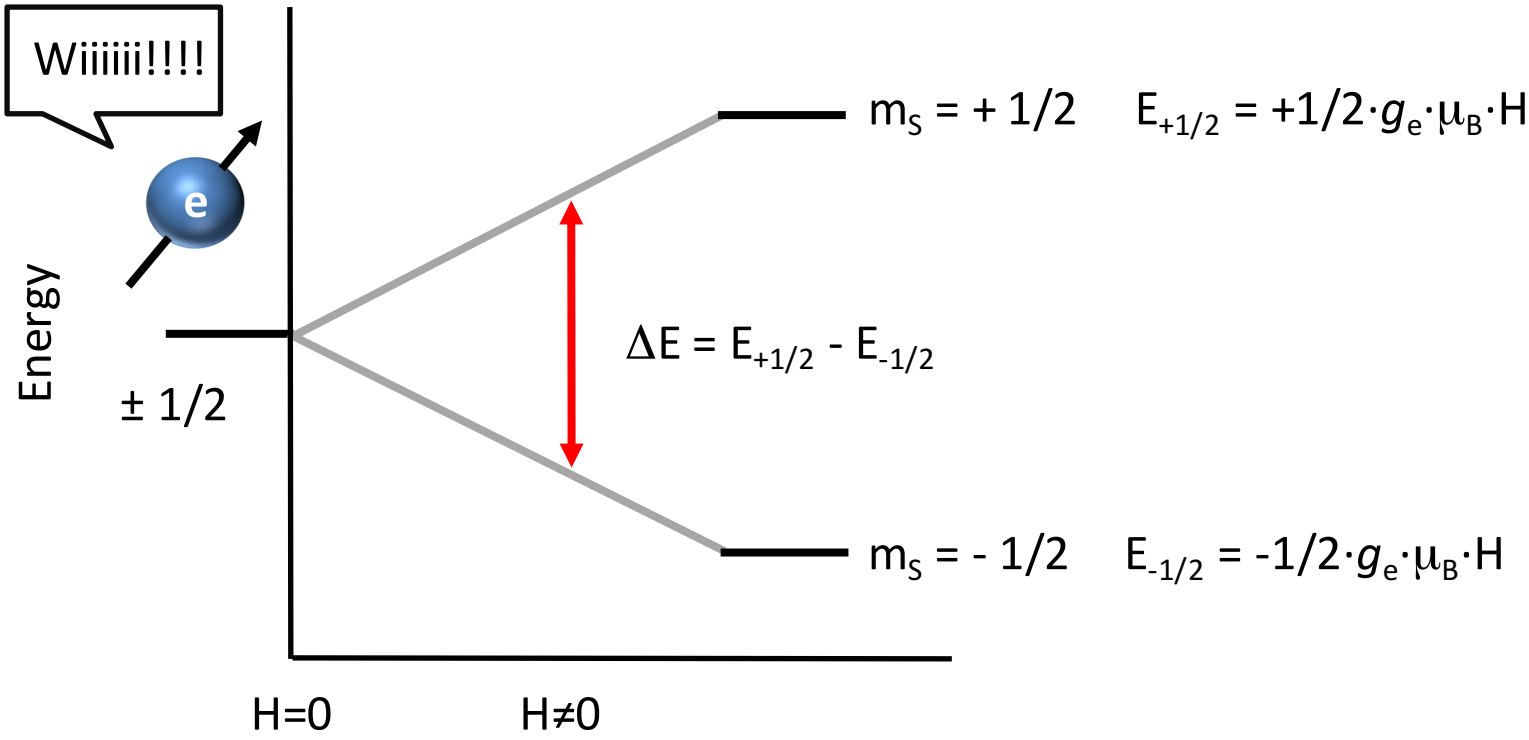
2.6 Electron Paramagnetic Resonance

Physical origin of the EPR signal

Zeeman effect: $\longrightarrow$ $\mathbf{H} = g_e \cdot \mu_B \cdot H \cdot \mathbf{S}$ $\longrightarrow$ $E = m_S \cdot g_e \cdot \mu_B \cdot H$

Selection Rule

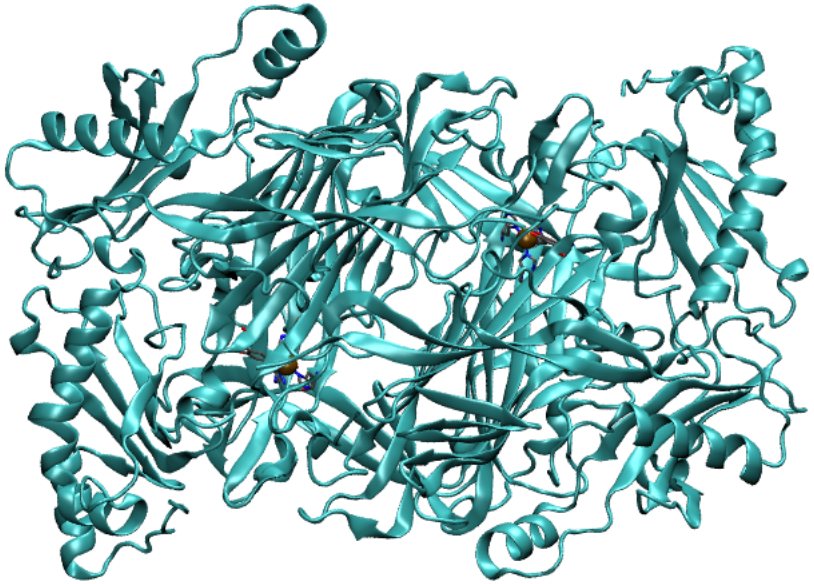
$\Delta M_S = \pm 1$

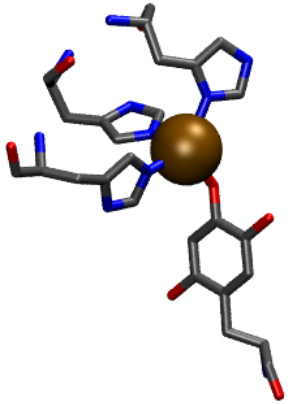


Fundamental equation of EPR spectroscopy

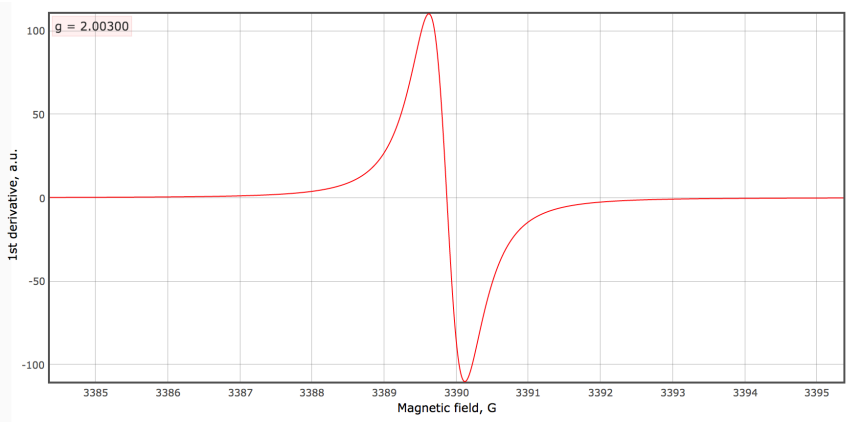
$$\left.\begin{aligned} \Delta E &= g_e \cdot \mu_B \cdot H \\ \Delta E &= h\nu \end{aligned}\right\} \boxed{h\nu = g_e \cdot \mu_B \cdot H}$$

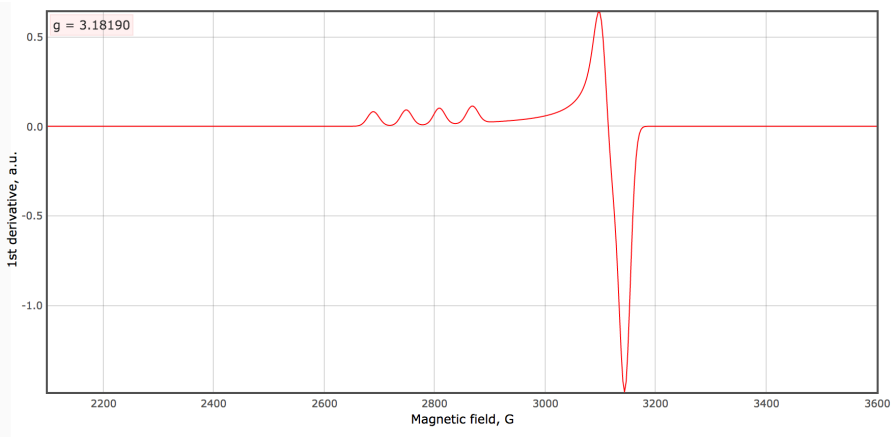
2.6 Electron Paramagnetic Resonance





Single Cu^{II} d^9 ion





Free electro EPR spectra

Cu^{II}spectra (g_x, g_y and g_z)

2.7 Practical Cases

- How many *homo sapiens structures contain the* $[WO_4]^{2-}$ motif?

- "Zinc fingers" are proteins involved in repairing and replication processes of DNA
  - Serach for the 2EPC structure
  - Visualize the Zn^{2+} cation using Ligand Explorer software
  - Write down the name of the Zn coordinating amino-acids
  - Write down the four Zn-L metal-ligand bond lengths

- Search for the *cis*-platin $[PtCl_2(NH_3)_2]$ identifier
  - Recover one structure in which *cis*-platin is interacting with the DNA (5N26)
  - Visualize the structure in Mercury
  - Visualize the structure using the Ligand Explorer software
  - Search for neighbour residues
  - Write down all Pt-L metal-ligand bond lengths

2.7 Practical Cases

- How many *homo sapiens structures contain the* $[WO_4]^{2-}$ motif?

  ID: WO4; homo sapiens 6

- "Zinc fingers" are proteins involved in repairing and replication processes of DNA
  - Serach for the 2EPC structure
  - Visualize the Zn^{2+} cation using Ligand Explorer software
  - Write down the name of the Zn coordinating amino-acids
  - Write down the four Zn-L metal-ligand bond lengths

  ID: Zn; CYS247, CYS250, CYS267 and HIS263: 2.3, 2.2, 2.3 y 2.1 Å

- Search for the *cis*-platin $[PtCl_2(NH_3)_2]$ identifier
  - Recover one structure in which *cis*-platin is interacting with the DNA (5N26)
  - Visualize the structure in Mercury
  - Visualize the structure using the Ligand Explorer software
  - Search for neighbour residues
  - Write down all Pt-L metal-ligand bond lengths

  ID: CPT; LYS68 and HIS136; 2.1, 2.1 Å (Pt-NH_3 2.1 and 2.0 Å)

2.7 Practical Cases

Written report 1:

1. The protein's name, PDB code, and its function
2. A picture of the binding site from the X-ray
3. ChemDraw representations of a number of inorganic complexes that display similar coordination chemistry and activity, giving Conquest refcode and IUPAC names
4. A brief account of the synthesis of the chosen analogue
5. Key literature references

Written report model 2:

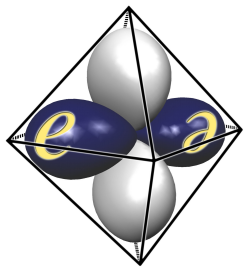
1. Protein function (e.g., electron transfer–ligand exchange, ligand modification) and its relation to the nature of the binding site: the coordination number and geometry, whether it is hard or soft, or has a vacant coordination site
2. The oxidation state of the metal
3. The suitability of the metal for the purpose for which it is used. Can other metals provide a similar function?
4. A critical evaluation of the model complexes available
5. A brief explanation of what can be added to our knowledge from wet chemistry, analysis, and spectroscopy

3. Summary and Outlook

- The **PDB** is an open access structural database with tons of information (allows the student work on his own!!)
- In combination with **Metal-PDB** and **PDBeChem** can be used to scan for metalloenzymes
- Metalloenzymes can be used to illustrate coordination chemistry basic ideas
- Metalloenzymes can be used to illustrate spectroscopic techniques
- Possibility of designing specific curricular problems for both coordination chemistry and bioinorganic courses

Acknowledgments

BG05M2OP001-2.009 and Sofia University crew



Institut de Química Teòrica i Computacional
UNIVERSITAT DE BARCELONA

UNIVERSITAT DE BARCELONA