



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Drug-design - racionální návrh léčiv KFC/DD

03 – molekulární cíl

RNDr. Karel Berka, Ph.D.
ZS 2012/2013



Motto

Bez cíle se ani Robin Hood netrefí.

Nejmenovaný autor tohoto kurzu

Osnova

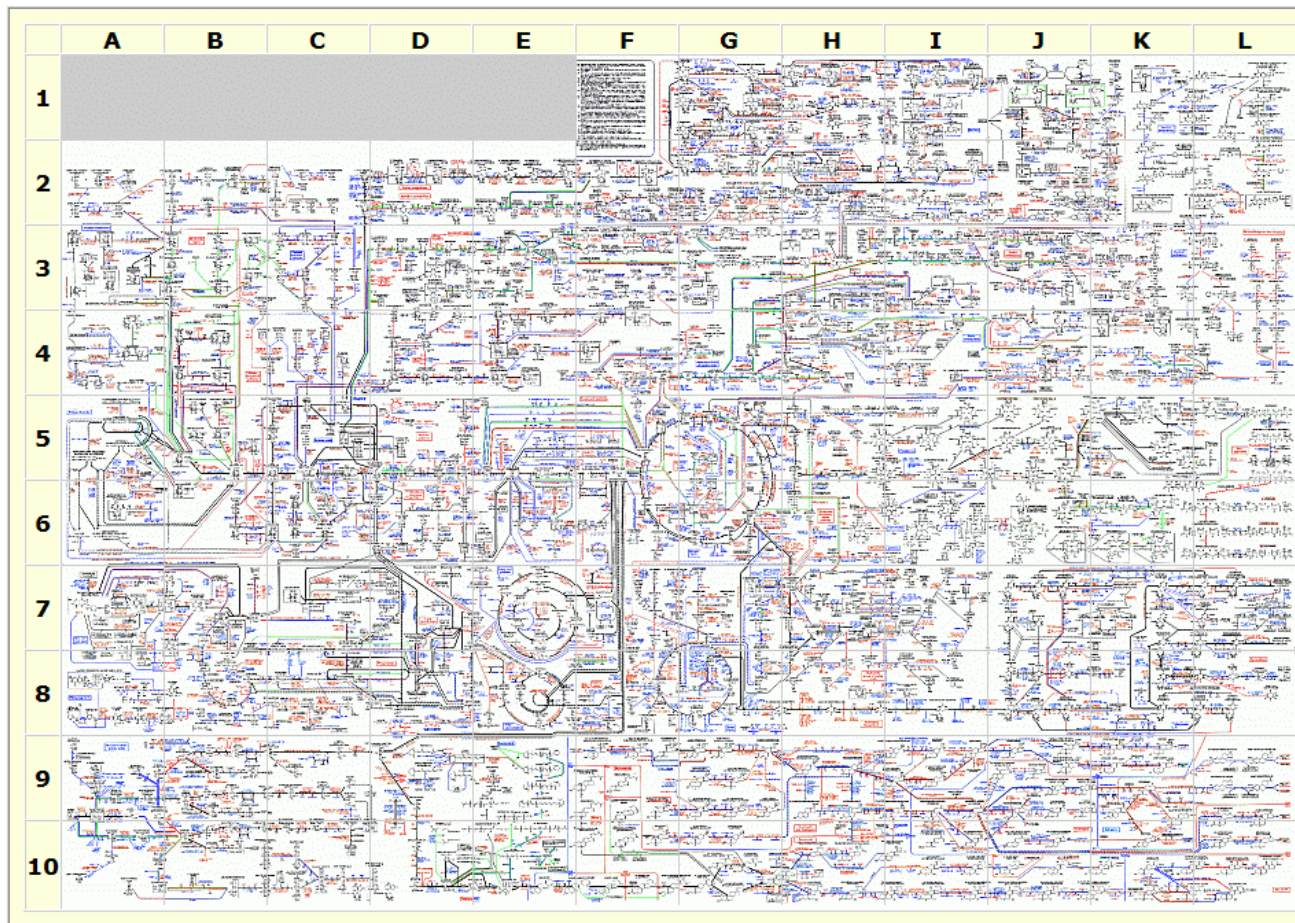
- Znalost biologického cíle
- Identifikace cíle
 - metabolické cesty
 - Metody určení struktury
- Charakterizace cíle
 - Mechanismus působení

Identifikace cíle

- Protein, který je v metabolické dráze spojené s chorobou
- Známe jeho primární sekvenci (dnes už skoro jisté, pokud se nejedná o nový patogen)
- Potřebujeme znát strukturu



Biochemical Pathways - Metabolic Pathways



For enquiries on how to obtain a paper copy of the wall chart, please contact Roche Applied Science directly, and do not email the ExpAsy staff.

http://web.expasy.org/cgi-bin/pathways/show_thumbnails.pl

Metabolické dráhy II

- Databáze

- KEGG - <http://www.genome.ad.jp/kegg>

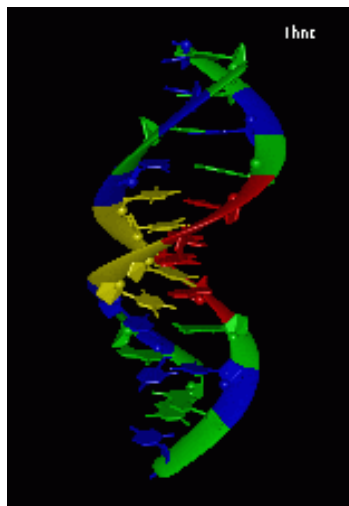
- MetaCyc - <http://metacyc.org>

- známé metabolické dráhy

- Hledáme úzké hrdlo – enzym, který bychom mohli zastavit

Jak získávat strukturu

- RTG
 - xyz souřadnice atomů
 - krystal, rozlišení,
 - interpretace mezimolekulové interakce
- EM
 - elektronový obal
 - nízké rozlišení
 - velké komplexy
 - Ne příliš vhodné pro drug design
- NMR
 - torzní uhly a vzdálenosti
 - dynamická informace
 - zpracování pomocí MD
- Predikce
 - Homologní modelování
 - Získání modelu ze znalosti sekvenční podobnosti k již existující struktuře
 - Protein threading
 - Poskládání z kousků
 - De novo predikce – protein folding

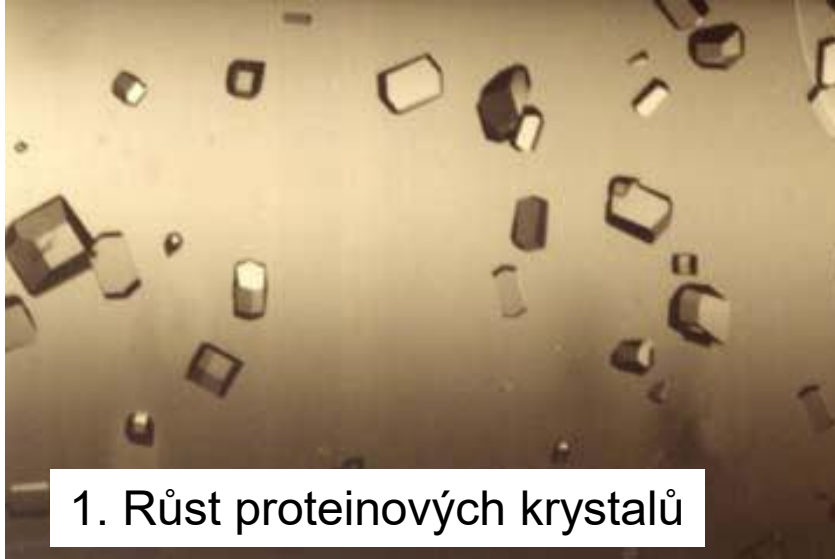


X-ray

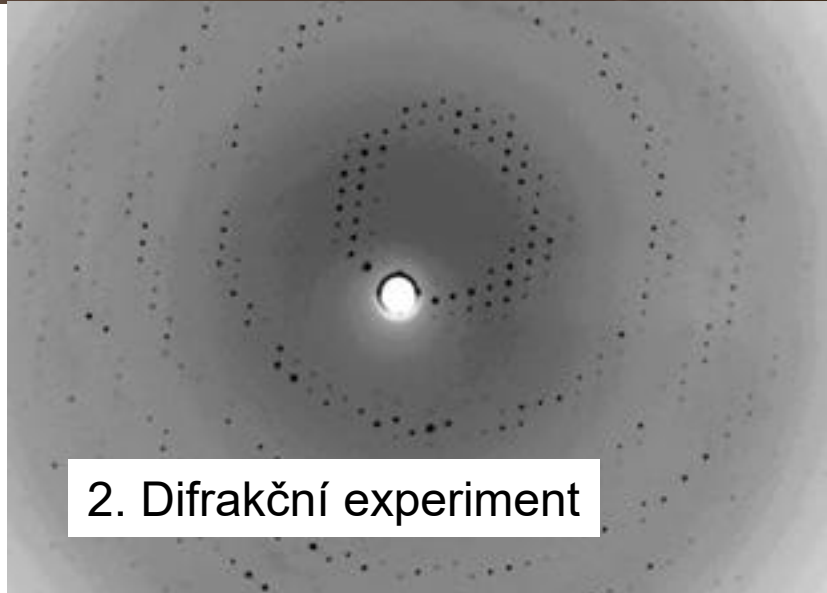
aneb

rentgenová krystalografie

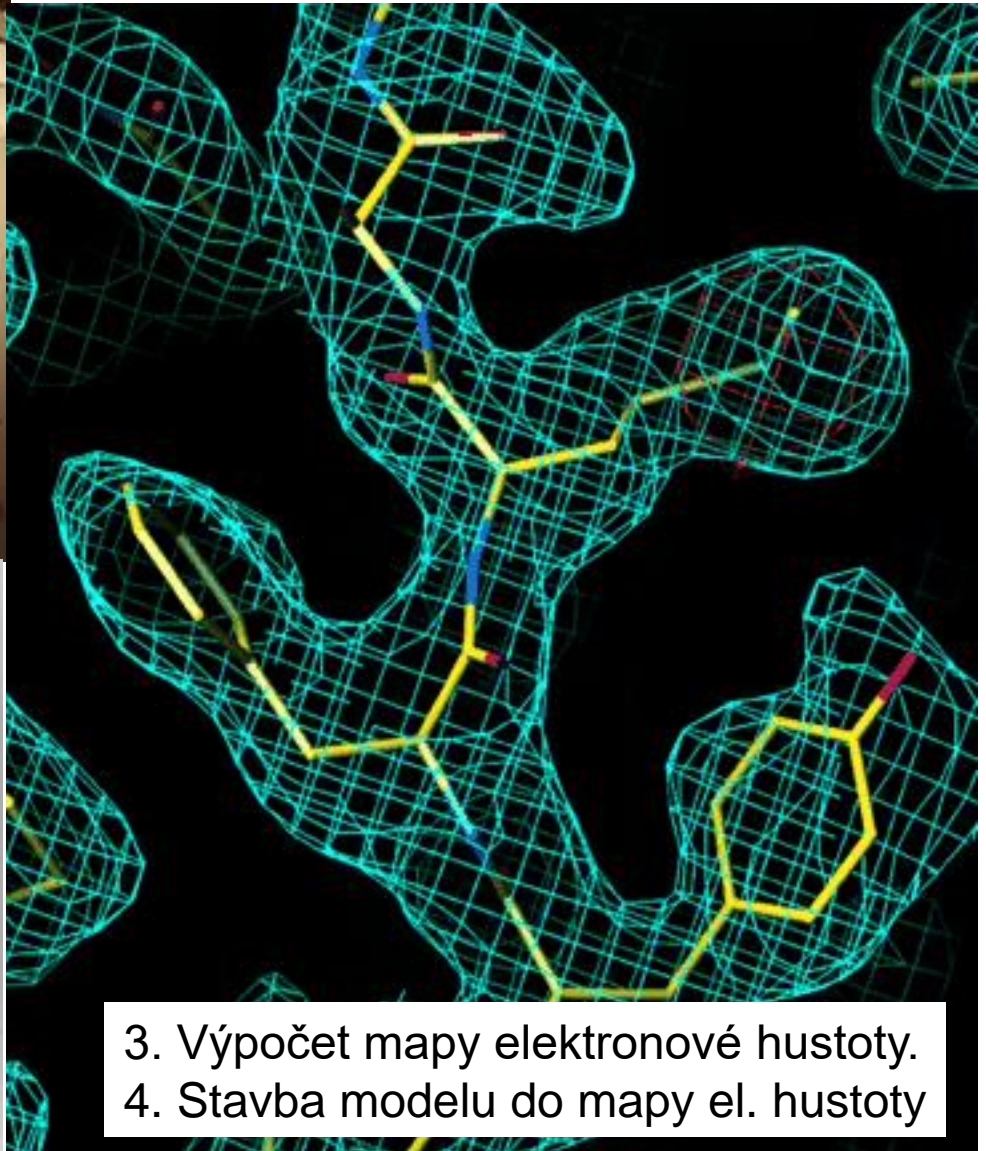
Rentgenová krystalografie



1. Růst proteinových krystalů



2. Difrakční experiment



3. Výpočet mapy elektronové hustoty.
4. Stavba modelu do mapy el. hustoty

Proč X-Ray

Elektromagnetické záření interaguje s objekty jejichž velikost je srovnatelná s vlnovou délkou (λ).

Např. viditelné světlo má vlnovou délku přibližně od 400 do 700 nm.

Vzdálenosti mezi atomy:

$$\text{C-C} = 1,54 \text{ \AA}$$

$$\text{C=C} = 1,23 \text{ \AA}$$

$$1 \text{ \AA} (\text{\AA}ngstrom) = 0.1 \text{ nm}$$

$$\text{C-N} = 1,45 \text{ \AA}$$

$$1 \text{ \AA} = 10^{-10} \text{ m}$$

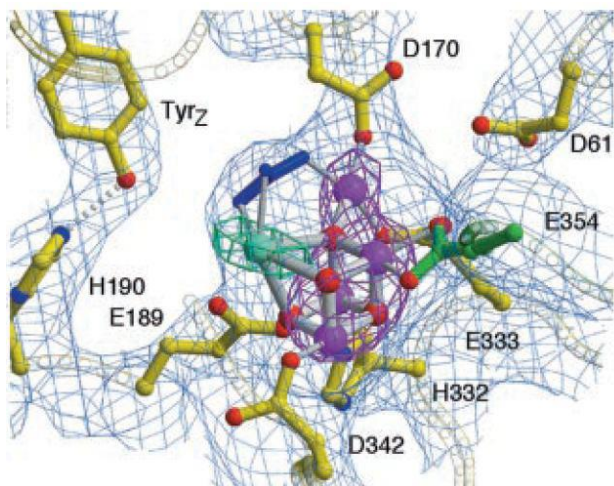
$$\text{N-(H).....O} = 2,8 \text{ \AA}$$

V laboratoři se běžně používá $\text{CuK}\alpha$ - $\lambda = 1,54 \text{ \AA}$.

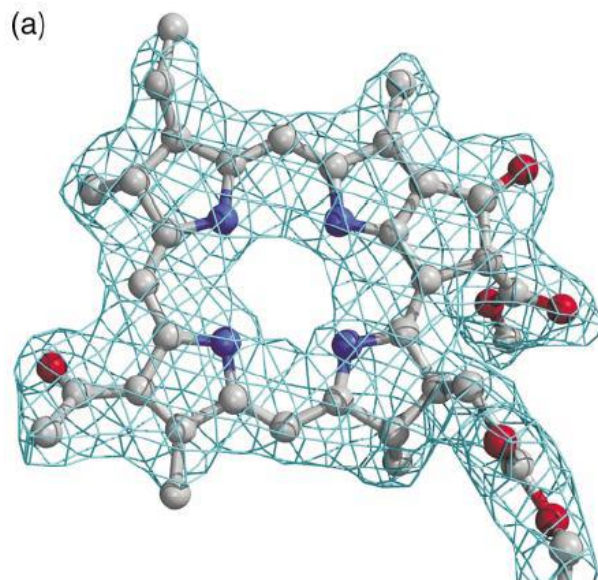
Synchrotron – $\lambda = 0,5 \text{ \AA} - 2,5 \text{ \AA}$.

Rozlišení (R)

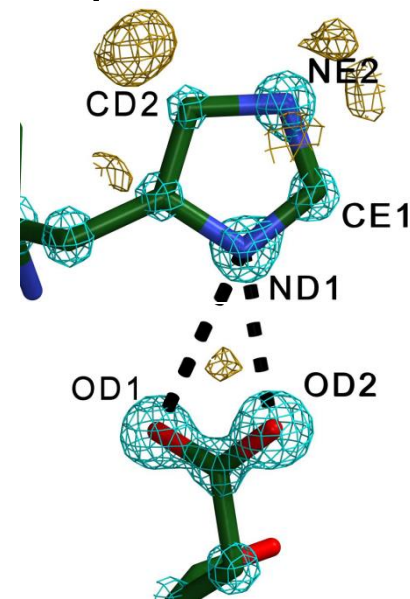
- v jednotkách Å, (více reflexí-lepší zesílení signál-šum)
- schopnost rozlišit detaily na vzdálenost. Teoreticky rozlišitelné detaily separované nejméně 0.7x rozlišení.
- Čím lepší rozlišení, tím lepší mapu dostaneme - snadnější stavba modelu!



3.5 Å mapa fotosystému II



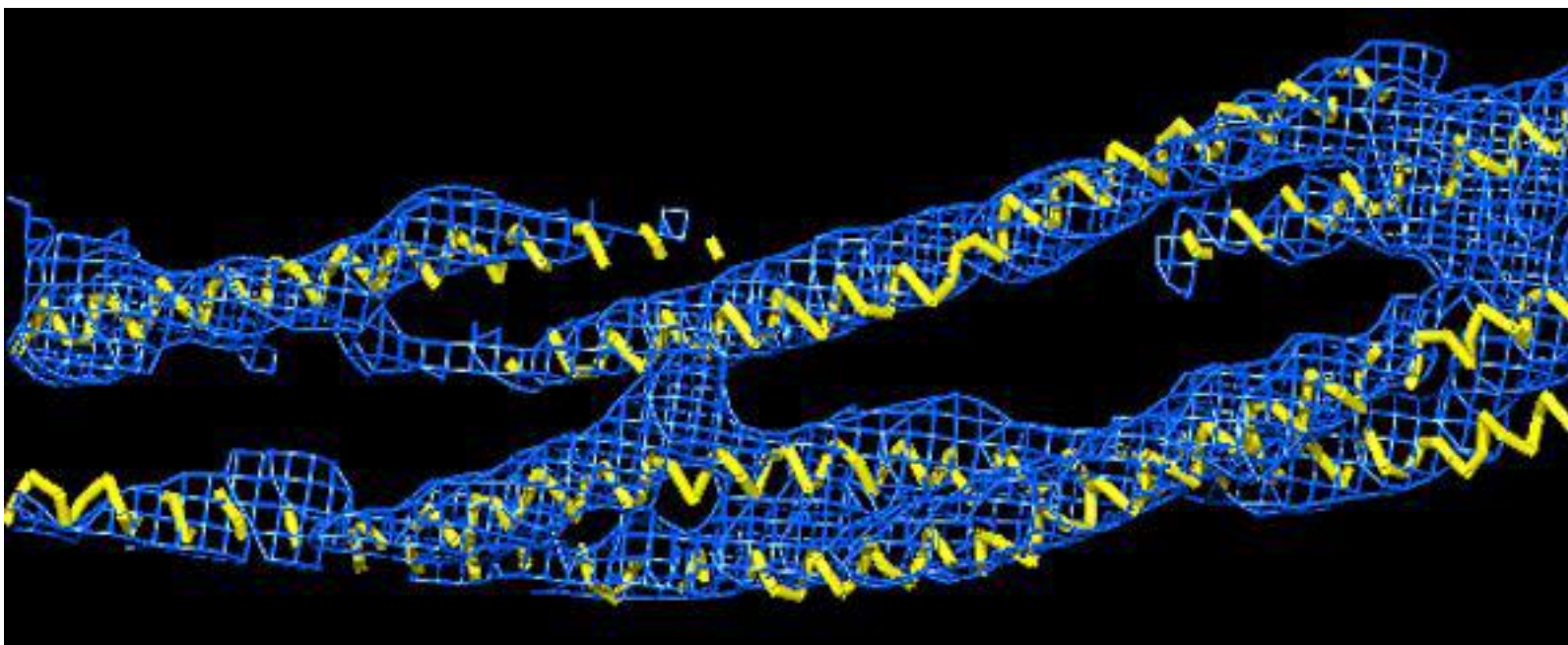
2.3 Å mapa fotosyntetického reakčního centra



0.95 Å mapa elastasy

Nízké rozlišení

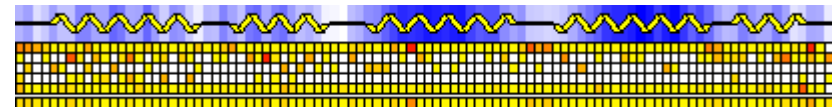
- Mapy s nízkým rozlišením ukazují pouze obecné vlastnosti jako je např. tvar molekuly a umístění elementů sekundární struktury.



$R = 7 \text{ \AA}$, tropomyosin.

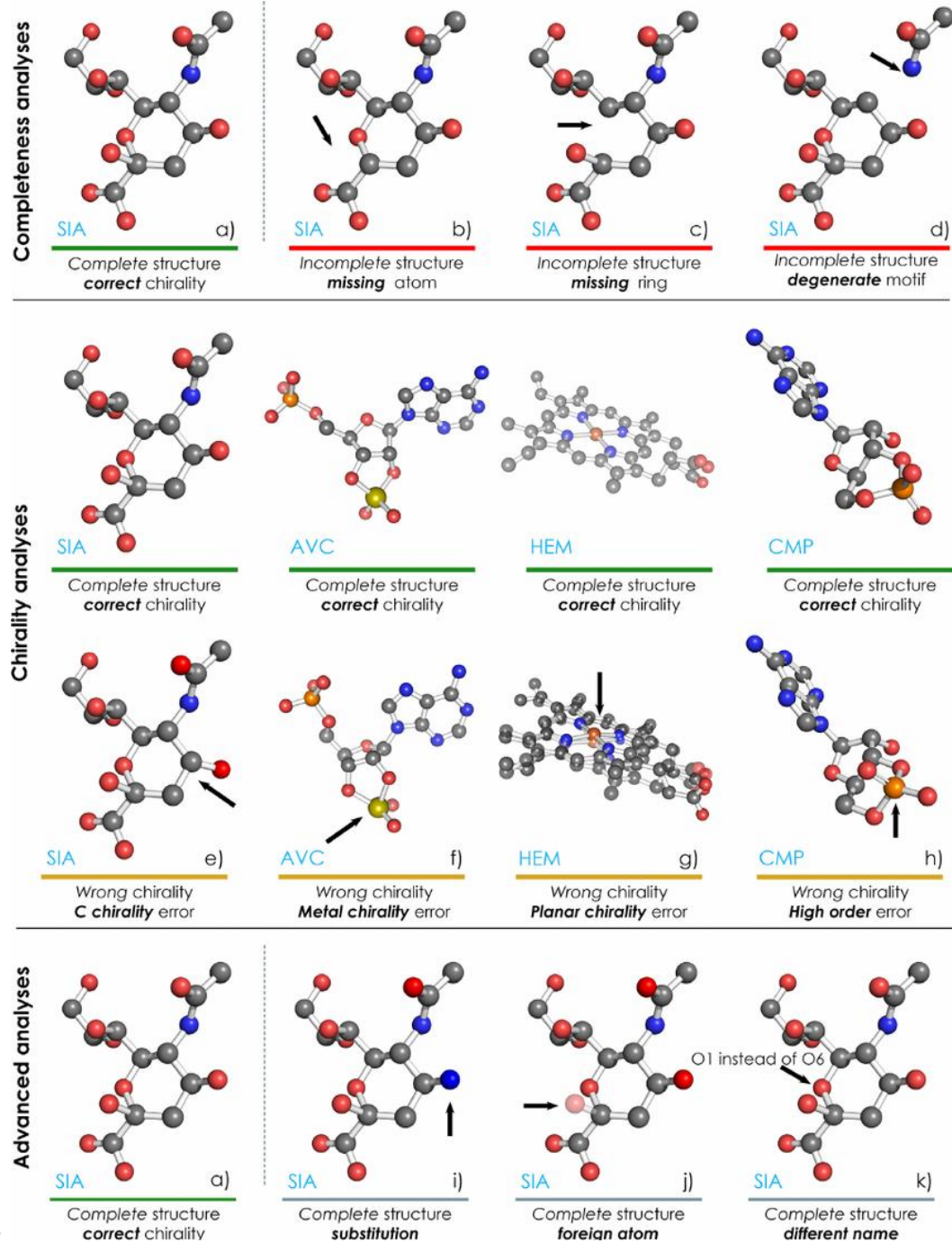
kontrola kvality

- Model quality is assessed with Knowledge-Based (KB) and Model-vs-Data (MvD) metrics
 - KB metrics compare structural features **with the database of protein structures**
 - MvD metrics **compare models against experimental data**
- Assessment requires annotation of well-defined regions the protein structure model
- Stereochemická – PROCHECK 3.5
 - [1. Ramachandran plot](#)
 - [2. Ramachandran plots by residue type](#)
 - [3. Chi1-Chi2 plots](#)
 - [4. Main-chain parameters](#)
 - [5. Side-chain parameters](#)
 - [6. Residue properties](#)
 - [7. Main-chain bond length distributions](#)
 - [8. Main-chain bond angle distributions](#)
 - [9. RMS distances from planarity](#)
 - [10. Distorted geometry plots](#)
- R_{free} – cross-validace - 5-10 % dat vynecháno.



ValidatorDB

- Ligandy
- kontrola oproti DB
 - kompletnost
 - kontrola chirality
 - substituce



Vařeková, R.S., Jaiswal, D., Sehnal, D., Ionescu, C.-M., Geidl, S., Pravda, L., Horský, V., Wimmerová, M. and Koča, J. (2014) [MotiveValidator: interactive web-based validation of ligand and residue structure in biomolecular complexes](#). *Nucleic Acids Research*, 42(Web Server issue), W227–33.

NMR

aneb

Nukleární magnetická rezonance

Nukleární magnetická rezonance

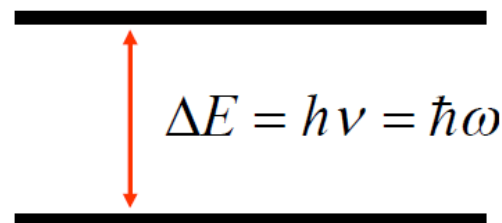
NMR

- NMR spektroskopie využívá magnetických vlastností jader atomů.
- Absorpční(emisní) spektroskopie, podobně jako IČ nebo UV. Detekuje absorpci radiofrekvenčního záření jádry atomů v molekule.
- Radiofrekvenční energie (ΔE přechodů jaderného spinu):

$$\lambda = 1011 \text{ až } 3 \times 10^7 \text{ nm}$$

$$\nu = 10^6 \text{ až } 10^{10} \text{ Hz}$$

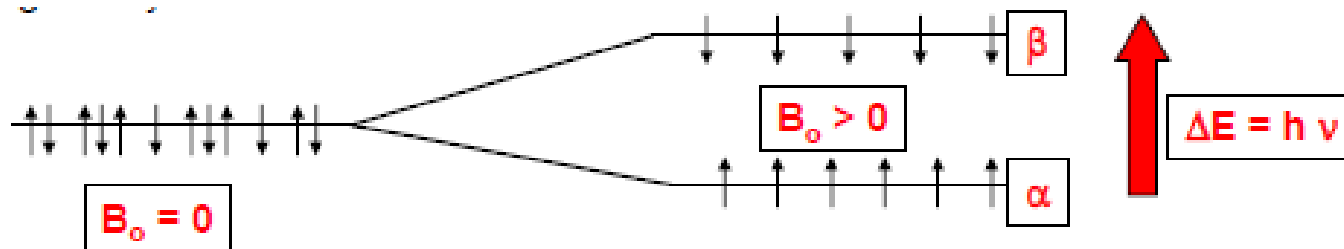
dovolené spinové stavy



Nastavením frekvence elektromagnetického záření (ν , nebo ω) na rezonanční podmínku dojde k indukci přechodů mezi hladinami jaderného spinu (*tzn.* můžeme měřit NMR spektrum!).

NMR

- Externí magnetické pole => energetický rozdíl mezi spinovými stavy jaderných magnetických momentů (m)



- Rozdíl v populaci stavů je dán rozdílem energií

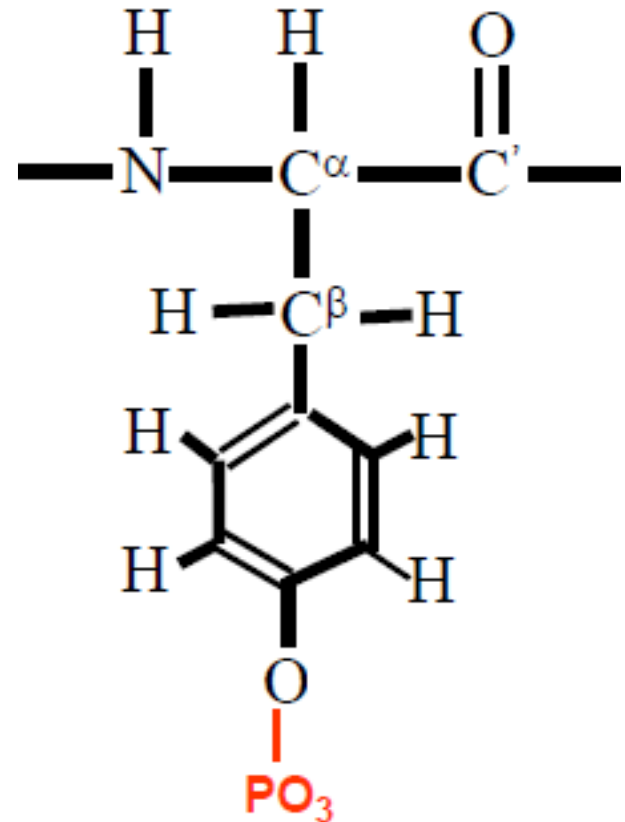
$$N_{\alpha} / N_{\beta} = e^{\Delta E / kT}$$

– např. $\Delta E = 3.8 \times 10^{-5}$ kcal/mol pro ^1H při 400 MHz ($B_0 = 9.5\text{T}$) $N_{\alpha}/N_{\beta} = 1.000064$

- Rozdíly populací velmi malé

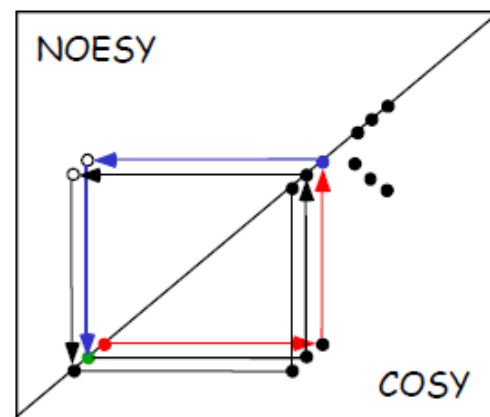
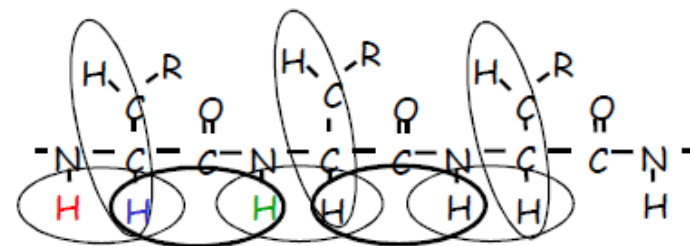
NMR-aktivní jádra v proteinech

- Přirozená:
 - ^1H , spin $\frac{1}{2}$
 - ^{31}P , spin $\frac{1}{2}$
- Obohacená díky bakteriální expresi:
 - ^2H , spin 1
 - ^{13}C , spin $\frac{1}{2}$
 - ^{15}N , spin $\frac{1}{2}$



Určení struktury pomocí NMR

- Přiřazení sekvence (určení vedlejších řetězců)
 - COSY
 - NOESY
 - 3D-NMR experimenty
- Sekundární struktura
 - chemické posuny (backbone)
 - dipolární spřáhnutí => prostor
 - J-spřáhnutí => torzní úhly
- Terciární struktura (do modelu)
 - NOE intenzity (neodpovídající COSY pro vedlejší řetězce)

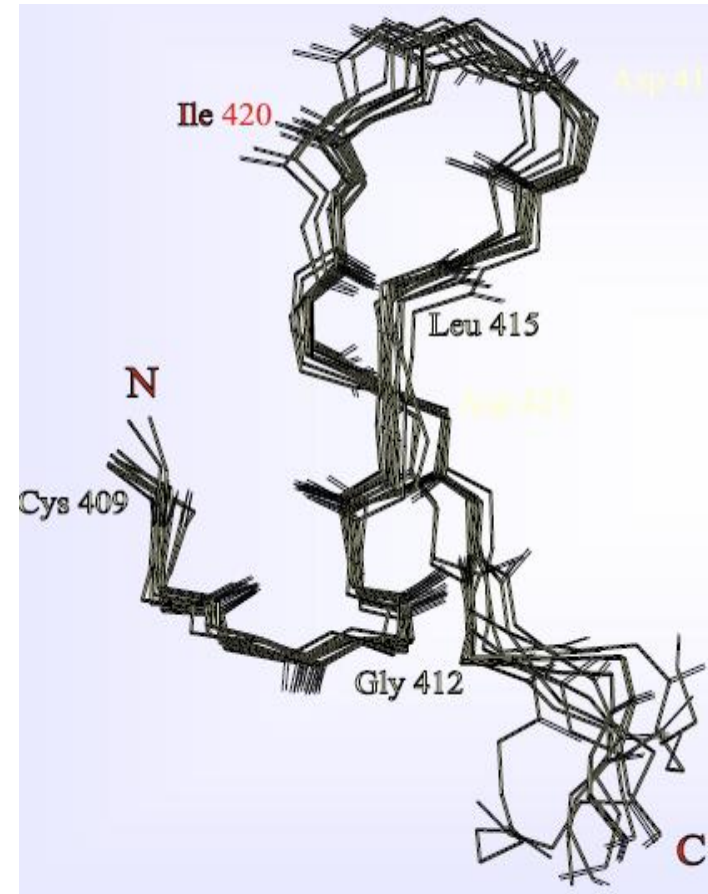


^1H
Chemický
posun

^1H Chemický posun

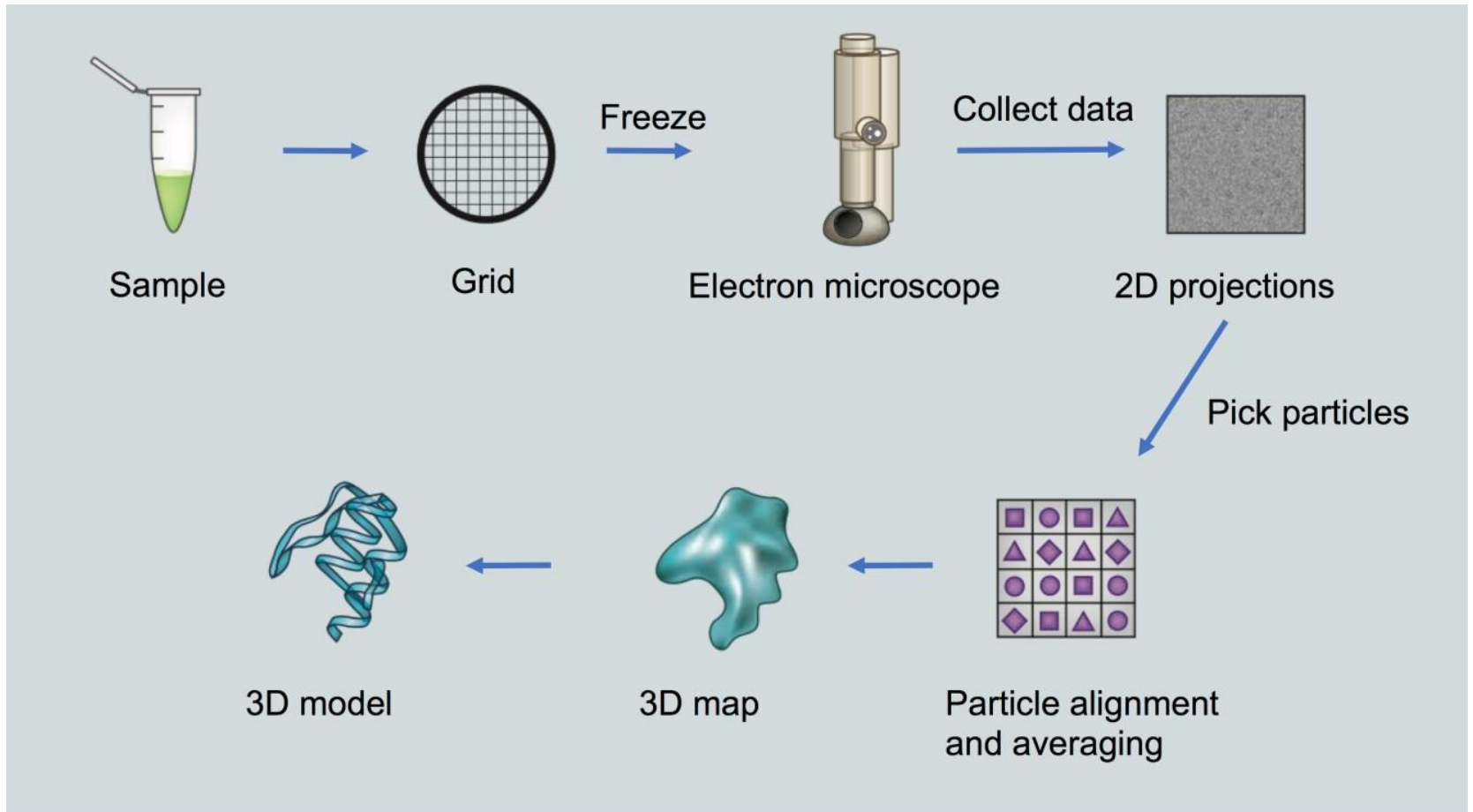
Kvalita struktury NMR

- Výsledkem NMR experimentu
= Získ celé sady struktur, které splňují podmínky
- Kvalita:
 - Stereochemie – Ramachandran
 - Ekvivalent R_{free} – vynechá se část dat při určování struktury a pak se to přes ně kontroluje
není standardizováno!



Cryo-EM

Cryo EM



2017 NOBEL PRIZE IN CHEMISTRY



The Nobel Prize in Chemistry 2017 was awarded to **Jacques Dubochet**, **Joachim Frank**, and **Richard Henderson** for the development of cryo-electron microscopy for determining biomolecule structures.

X-RAY CRYSTALLOGRAPHY



Structures of proteins that form crystals

NMR SPECTROSCOPY



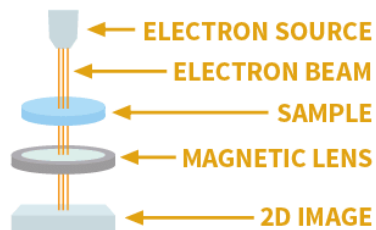
Structures of small proteins in solution

CRYO-EM

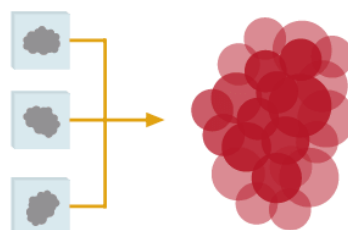


Structures of large, non-crystalline proteins

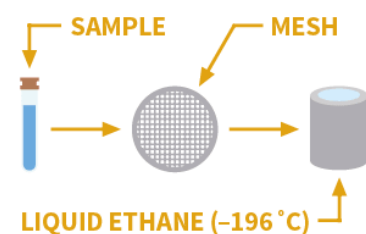
Cryo-electron microscopy (cryo-EM) is a technique that makes it possible to produce 3D images of biomolecules at atomic resolution. It can be used to capture images of biomolecules which could not be visualised with previously existing techniques.



Henderson pioneered the use of electron microscopy (EM) to visualise proteins. Using it, he produced the first atomic resolution image of a protein, bacteriorhodopsin, in 1990.



Frank developed an image analysis method that allowed computers to assemble a high resolution 3D image from many 2D EM images, improving the quality of biomolecule images.



Biological samples dry out and are damaged when in vacuum during EM. **Dubochet** solved this by rapidly freezing samples in water at -196°C to form an icy glass instead of crystals.



WHY DOES THIS RESEARCH MATTER?

Cryo-EM allows scientists to reveal how proteins move and interact with other molecules, freezing and observing them mid-process. It could improve our understanding of drug targets and biological processes.

Nobel Prize in Chemistry Press release: https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2017/press.html

EMDB

EM Resources

- Home
- Statistics
- Validation
- EMDataBank
- EMPIAR
- Test data

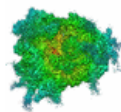
EMDB

- Latest maps
- Latest headers
- Latest updates
- Search
- Browse
- FTP archive
- Deposit EM map/model
- EMDb data model

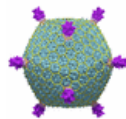
The Electron Microscopy Data Bank (EMDB) at PDBe

Quick access

Click on one of these categories:



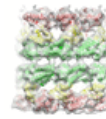
Ribosome



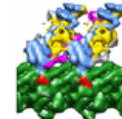
Virus



Phage



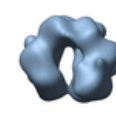
GroEL



Microtubule



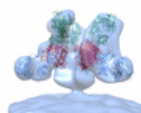
Polymerase



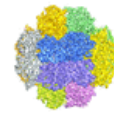
Helicase



Human



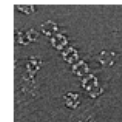
HIV



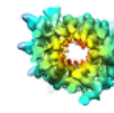
Entries with
fitted models



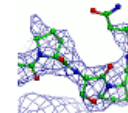
Single particle



Tomography



Helical
reconstruction



<5Å
resolution

or enter 4-digit EMDb entry number:

[Entry summary](#)

[Visual analysis of map](#)

[Volume viewer](#)

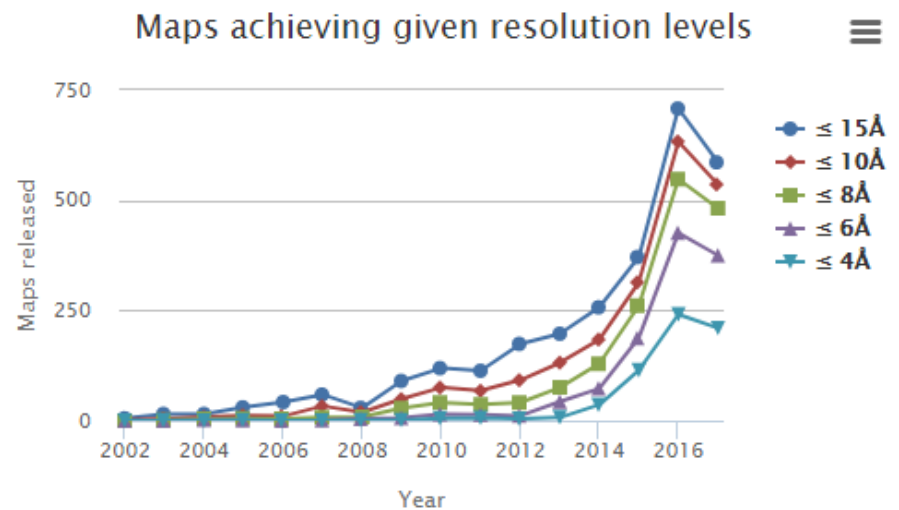
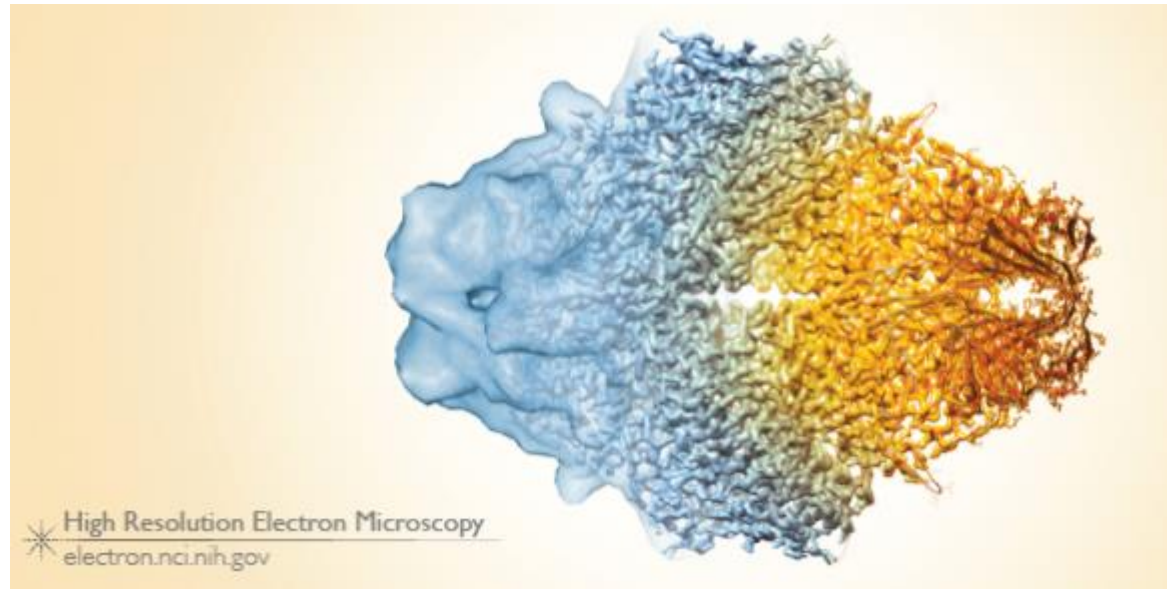
Introduction

The Electron Microscopy Data Bank (EMDB) is a public repository for electron microscopy density maps of macromolecules and subcellular structures. It covers a variety of techniques, including single-particle analysis, electron tomography (2D) crystallography.

The EMDb was founded at EBI in 2002, under the leadership of Kim Henrick. Since 2007 it has been operated jointly with the [Research Collaboratory for Structural Bioinformatics \(RCSB PDB\)](#) as a part of [EMDataBank](#) which is funded by a grant to PDBe, the RCSB and the [National Center for Macromolecular Imaging \(NCMI\)](#).

EMDB

- V poslední době výrazné vylepšení rozlišení
- teoretické limity jsou v řádu Å



Predikce struktury

Predikce struktury proteinů

- Predikce struktury
 - vycházející ze známých stabilních struktur
 - Homologní modelování
 - SwissMODEL, I-TASSER
 - threading
 - Modeller,
 - vycházející z fyzikálních modelů
 - de novo modelování (ab initio)
 - Quark, Robbeta, molekulární mechanika
 - skládání proteinů (protein folding)
 - Folding@Home, FoldIt

Homologní modelování

- také komparativní, nebo knowledge-based modelování
- strukturu neznámého proteinu sestaví na základě znalosti struktury homologního proteinu
- Swiss-MODEL
- http://spdbv.vital-it.ch/modeling_tut.html



Obecný protokol

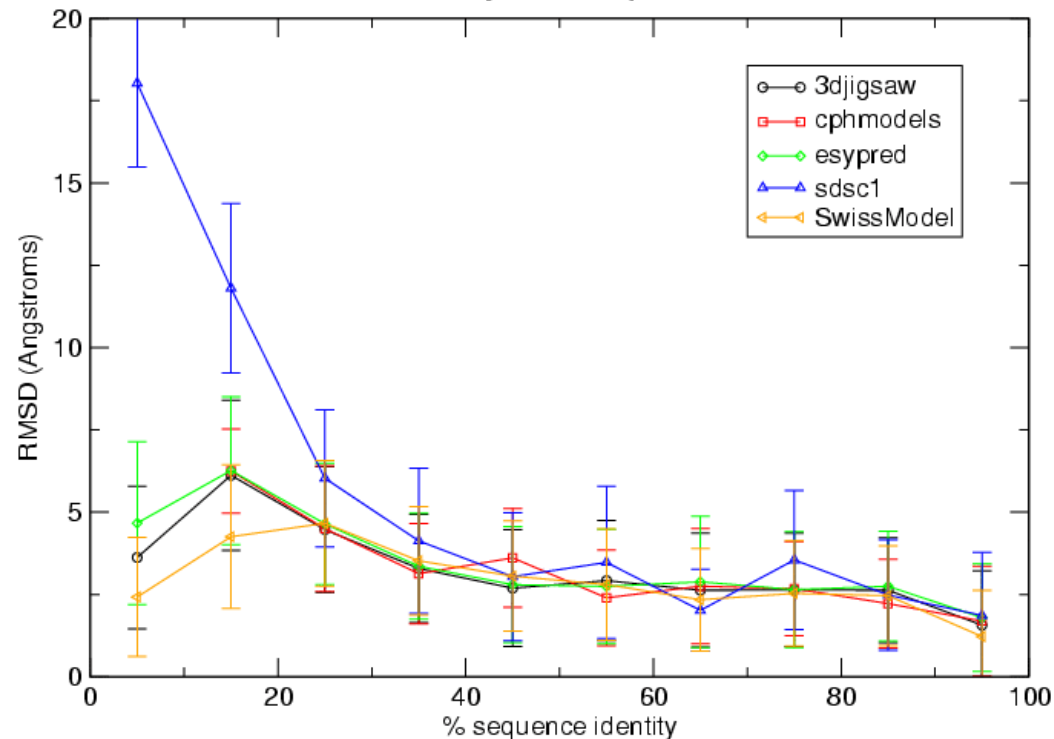
- vybrat protein k modelování
- hledání homologů
 - ne – ab initio, ev. threading
 - ano – Clustal W – alignment
- modelování centrálního regionu
 - (analýza očima – odpovídá to experimentu?)
- if ano – modelování loopů (hledání fragmentů)
- modelování vedlejších řetězců (rotamery)
- minimalizace (molekulové modelování, dynamika)
- stereochemická kontrola modelu (PROCHECK, Ramachandran plot)

Kdy použít Homologní modelování

- pokud je podobnost mezi sekvencí templátu a cílové sekvence dostatečně vysoká

- alespoň
30% IDENTITY je
nutnou podmínkou
tvorby modelu

RMSD vs % sequence identity (target/best template)
[bins of 10%]



Protein threading

- Sestavení databáze strukturních templátů
- Sestavení skórovací funkce
 - Měří jak dobře odpovídá sekvenci struktura
 - Obsahuje možnost popsat mutace, okolí, párové interakce mezi blízkými aminokyselinami, sekundární strukturu a rozdíly v délkách
 - Kvalita skórovací funkce blízce souvisí s přesností predikce!
- Threading alignment
 - Iterativní porovnání (alignment) cílové sekvence s templáty struktury pomocí optimalizace skórovací funkce
- Threading predikce
 - Nejpravděpodobnější alignment => threading predikce
 - Posléze se vytvoří strukturní model (ev. víc modelů) umístěním atomů páteře na místa templátu.

Modeller

Program for Comparative Protein
Structure Modelling by Satisfaction
of Spatial Restraints



```

A I L V G S M P R R D G M E R K D L L K A N V K I F K C Q G A
Y E V C P V D C F Y E G P N F L V I H P D E C I D C A L C E P
L A C K P E C P V N I I Q G S - - Y A I D A D S C I D C G S
G - - I A C G A C K P E C P V N I I Q G S - - I Y A I D A D S
  
```

- homologní modelování s constraints (např. NMR, EM, apod)
 - používá tzv. i-Sites (krátké kousky, pro které zná strukturu)

<http://salilab.org/modeller/>

[Comparative protein modelling by satisfaction of spatial restraints.](#) Šali A, Blundell TL. *J Mol Biol.* 1993 Dec 5;234(3):779-815.

Kdy použít threading?

- Pokud nemám dost sekvenční identity k jednomu templátu
- nejlépe po jednotlivých doménách
- nejlépe zkusit několik programů a porovnat, který fold je nejčastější – konsensus
- použít další znalosti o proteinu (funkce) – opět to může napomoci vybrat správný fold.

Ab initio modeling

- ab initio = bez předchozích znalostí (templátu)
- masivní hledání správné konformace a k tomu fyzikální (pseudo-fyzikální) energetická funkce na popis volné energie

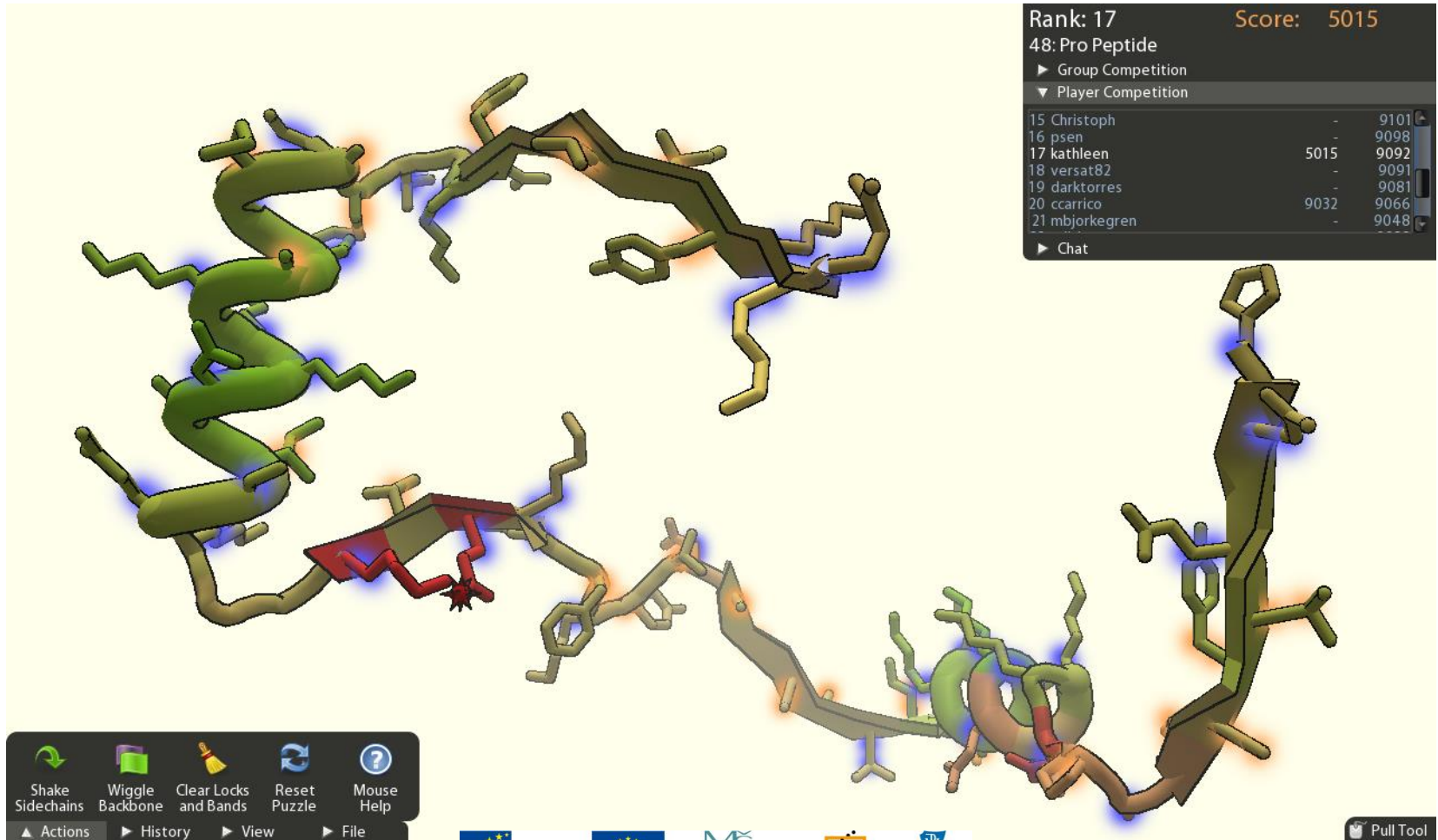
www.bakerlab.org



- <http://robetta.bakerlab.org/>
- *ab initio* and comparative models of protein domains
- Nejméně přesné ale jediné použitelné, pokud neznám templát

FoldIt

- skládání proteinů jako hra



The screenshot displays the FoldIt game interface. The main area shows a protein puzzle with several chains in different colors (green, yellow, red, grey) and glowing blue and orange spots. The top right corner shows the player's rank and score: Rank: 17, Score: 5015. Below this, there is a 'Player Competition' table with the following data:

Rank	Player	Score	Target
15	Christoph	-	9101
16	psen	-	9098
17	kathleen	5015	9092
18	versat82	-	9091
19	darktorres	-	9081
20	ccarrico	9032	9066
21	mbjorkegren	-	9048

At the bottom left, there is a toolbar with icons for 'Shake Sidechains', 'Wiggle Backbone', 'Clear Locks and Bands', 'Reset Puzzle', and 'Mouse Help'. At the bottom right, there is a 'Pull Tool' icon.

- <http://fold.it/portal/>

Která metoda je nejlepší?

1. X-Ray (s dobrým rozlišením)
2. NMR
3. Modely
 1. nejlepší bývá kombinace (I-TASSER)
 2. Threading (Modeller – ideální pokud je dobrý templát)
 3. Homologní modelování (Swiss-Model)
 4. Ab initio (Robetta)

Nobelovské porovnání jednotlivých metod

X-Ray	Others
1901: Wilhelm C. Röntgen (Physics) – X-ray	1943: Otto Stern (Physics) magnetic moment of the proton (NMR)
1914: Max von Laue (Physics) diffraction of X-rays by crystals	1944: Isidor I. Rabi (Physics) resonance method for recording the magnetic properties of atomic nuclei (NMR)
1915: William H. Bragg and William L. Bragg (Physics) – Bragg's equation	1952: Felix Bloch, Edward M. Purcell (Physics) nuclear magnetic precision measurements (NMR)
1964: Dorothy C. Hodgkin (Chemistry) structures of penicillin and vitamin B-12.	1982: Aaron Klug (Chemistry) crystallographic electron microscopy (EM)
1985: Herbert A. Hauptman and Jerome Karle (Chemistry) phase problem	1986: Ernst Ruska, Gerd Binnig, Heinrich Rohrer (Physics) TEM, STM
1954: Linus Pauling (Chemistry) – chemical bond, peptide bond, and the structures of the alpha helix and beta strand	1991: Richard R. Ernst (Chemistry) high resolution nuclear magnetic resonance (NMR) spectroscopy
1962 Francis H.C. Crick, James D. Watson, Maurice H.F. Wilkins (Physiology or Medicine) – DNA	1994: Bertram N. Brockhouse and Clifford G. Shull (Physics) neutron scattering
1962: Max F. Perutz and John C. Kendrew (Chemistry) globular proteins – myoglobin, hemoglobin	1998: Walter Kohn a John A. Pople (Chemistry) – computational methods in quantum chemistry (in silico)
1988: Johann Deisenhofer, Robert Huber, and Hartmut Michel (Chemistry) photosynthetic reaction centre (1PRC).	2002: John B. Fenn, Koichi Tanaka (Chemistry) soft ionization mass spectrometry (MS)
1996: Paul D. Boyer, John E. Walker, and Jens C. Skou (Chemistry) F1-ATPase (1bmf , 1cow)	2002: Kurt Wüthrich (Chemistry) nuclear magnetic resonance (NMR)
2003: Peter Agre and Roderick MacKinnon (Chemistry) membrane channels (1bl8 , 2f2b , 2evu)	2003: Paul C. Lauterbur, Peter Mansfield (Physiology or Medicine) magnetic resonance imaging (MRI)
2006: Roger Kornberg (Chemistry) molecular basis of eukaryotic transcription (1i3q , 1i50 , 1i6h)	2013: Martin Karplus, Michael Levitt and Arieh Warshel (Chemistry) - multiscale models for complex chemical systems (in silico)
2009 Venkatraman Ramakrishnan, Thomas A. Steitz, and Ada E. Yonath (Chemistry) ribosome (1ffk , 1fjg , 1fka , 1gix , 1giy)	2017: Jack Dubochet, Joachim Frank, Richard Henderson (Chemistry) cryo-electron microscopy for macromolecules (CryoEM)
2012 Robert J. Lefkowitz, Brian K. Kobilka (Chemistry) GPCR (3sn6 , 3uon , 4daj , 4dkl)	

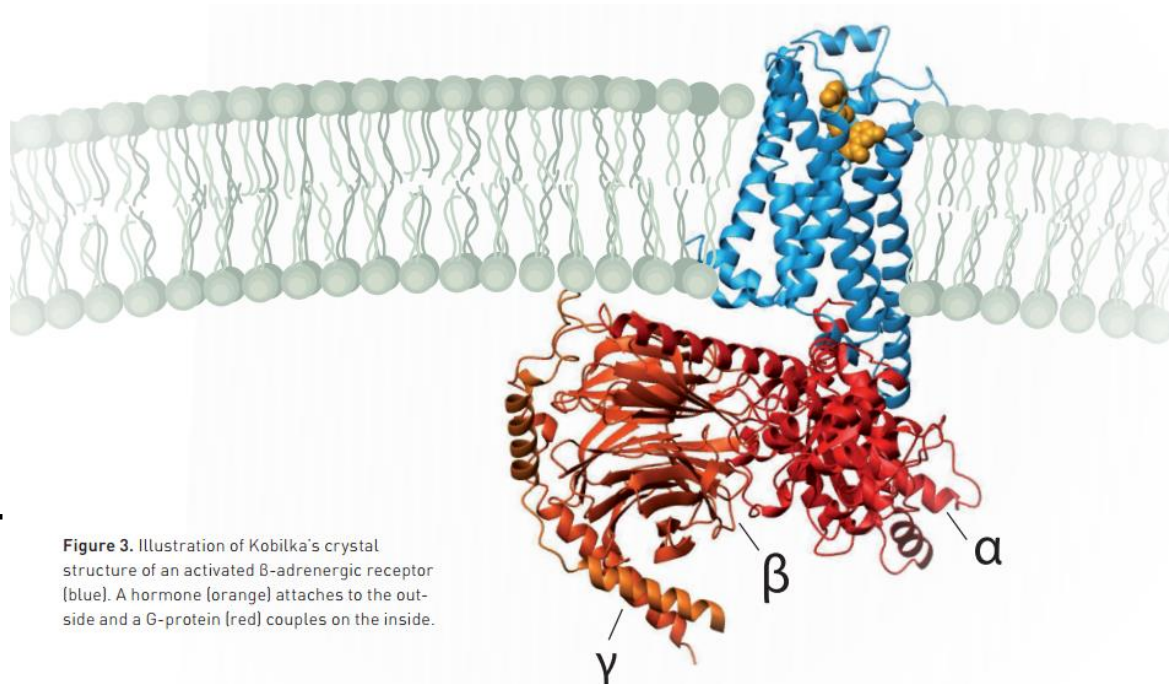


Figure 3. Illustration of Kobilka's crystal structure of an activated B-adrenergic receptor (blue). A hormone (orange) attaches to the outside and a G-protein (red) couples on the inside.

The image, published in *Nature*, reveals new details about GPCRs, for instance, what the activated receptor looks like when it opens up a void where the G-protein likes to bind (figure 4). Such knowledge will be very useful in the future for the development of new pharmaceutical drugs.

- Brian K. Kobilka
- Robert J. Lefkowitz



"for studies of G-protein-coupled receptors"

Příště:

- Jaké strategie zvolit v závislosti na tom, jaké informace známe?

Pravidla pro určení spinu izotopu

Nukleon.č.(A)	Proton.č.(Z)	I	Detekce
liché	sudé nebo liché	$1/2, 3/2, 5/2$	ano
sudé	sudé	0	ne
sudé	liché	1, 2, 3	ano

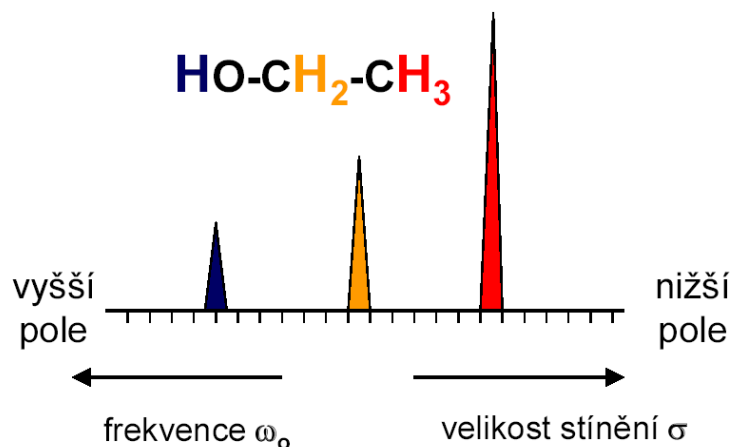
Možný počet spinových stavů = $2I + 1$:

Jádro	Spin.kvant.č.	Počet stavů	Mag.spin.č.
^1H	$I = 1/2$	$2(1/2) + 1 = 2$	$m = \pm 1/2$
^{14}N	$I = 1$	$2(1) + 1 = 3$	$m = -1, 0, 1$

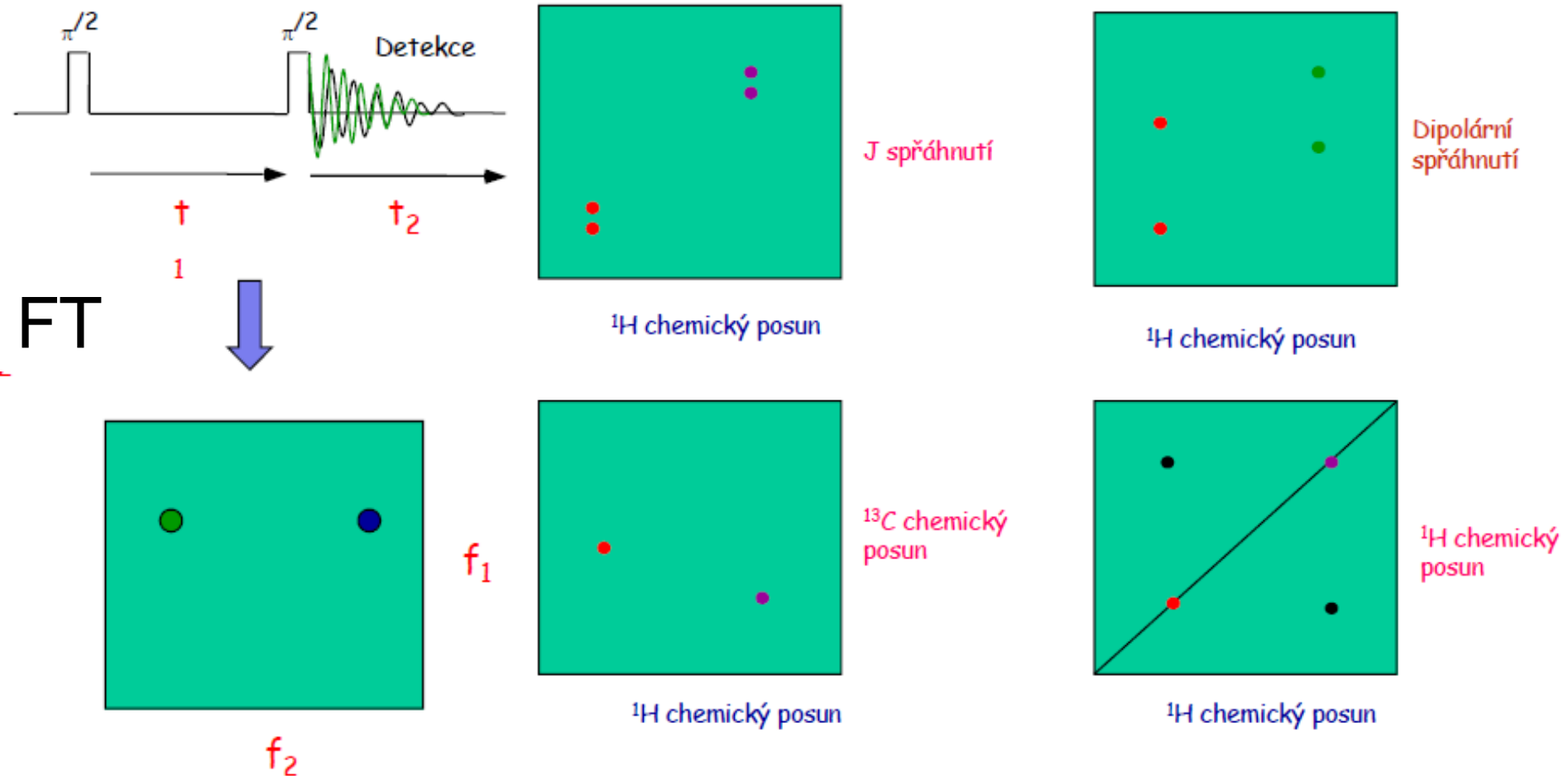
Chemický posun

- Různá jádra mají různé rezonanční frekvence
- Magnetické pole, ve kterém se jádro nachází není rovno vnějšímu magnetickému poli.
 - Elektrony v okolí jádra (chemické okolí) stíní vnější pole – výsledné efektivní magnetické pole \mathbf{B}_{eff} je tvořeno vnějším polem \mathbf{B}_0 a polem lokálním \mathbf{B}_{loc} .

$\mathbf{B}_{\text{eff}} = \mathbf{B}_0 - \mathbf{B}_{\text{loc}} = \mathbf{B}_0(1 - \sigma)$, kde σ je konstanta magnetického stínění



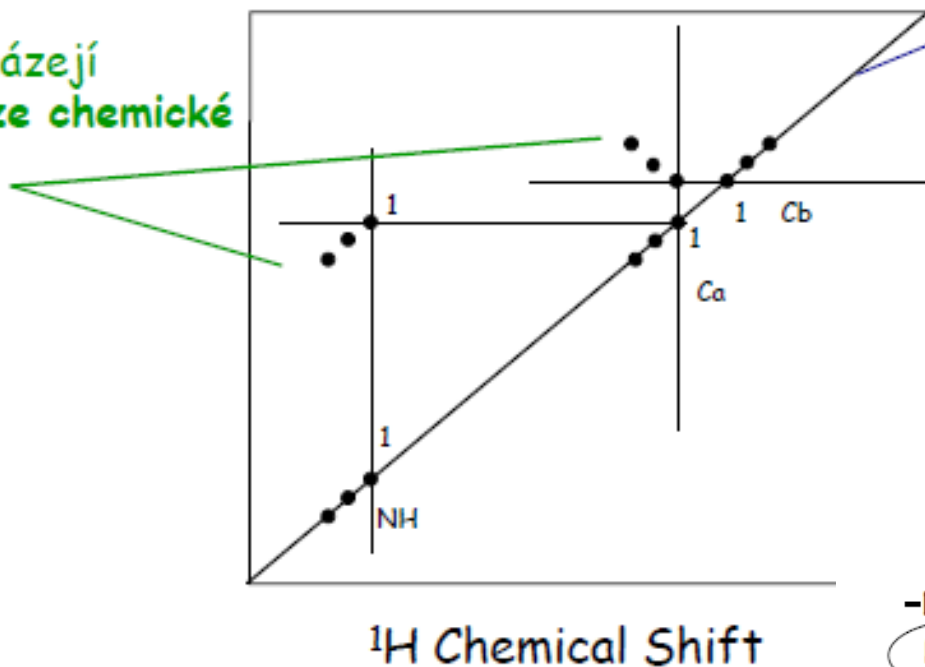
Obečný 2D experiment



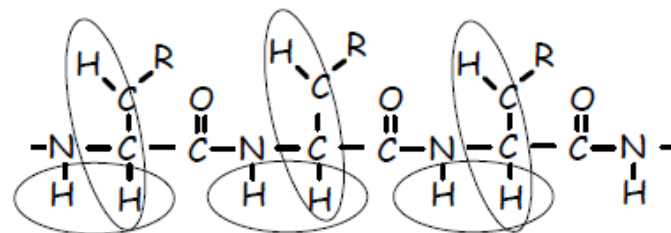
COSY (Correlated Spectroscopy)

1-dimensionální spektrum je ve směru diagonály

Cross píky pocházejí z korelace skrze chemické vazby



COSY Crosspeaks - ^3J Couplings



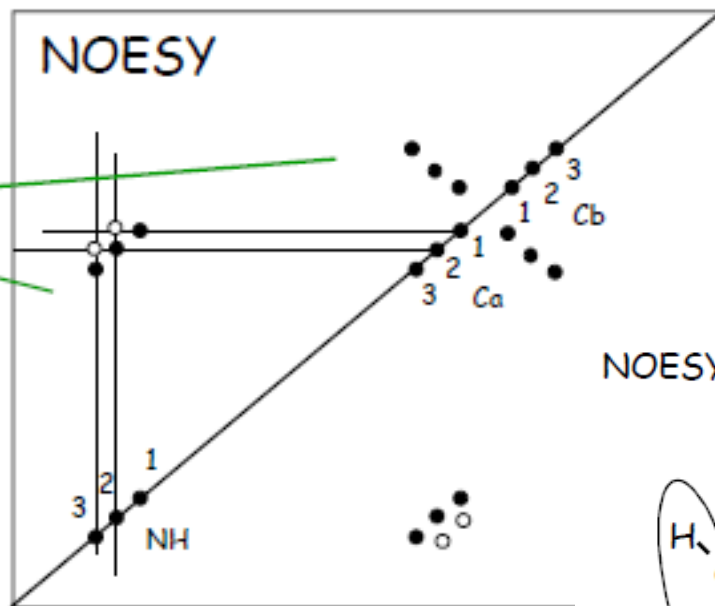
Cross píky jsou užitečné pro určení spinových systémů jednotlivých aminokyselin

NOESY

(Nuclear Overhauser Enhancement Spectroscopy)

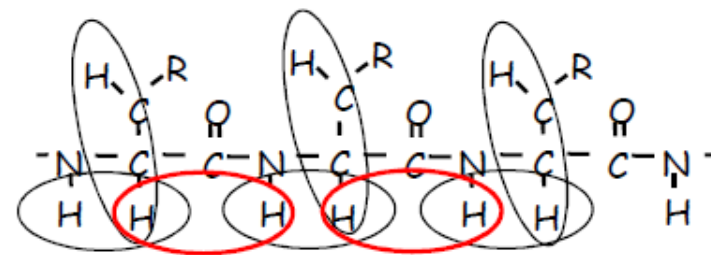
Cross píky vznikají díky korelaci skrze prostor

Cross píky v COSY spektru se nacházejí i v NOESY spektru



NOESY Crosspíky < 5 Å Skrze prostor

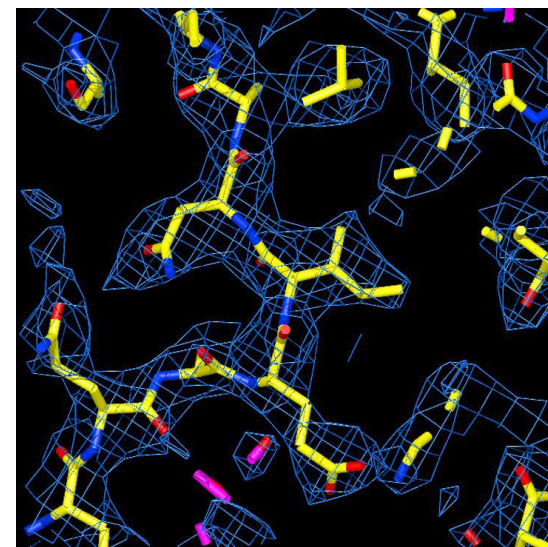
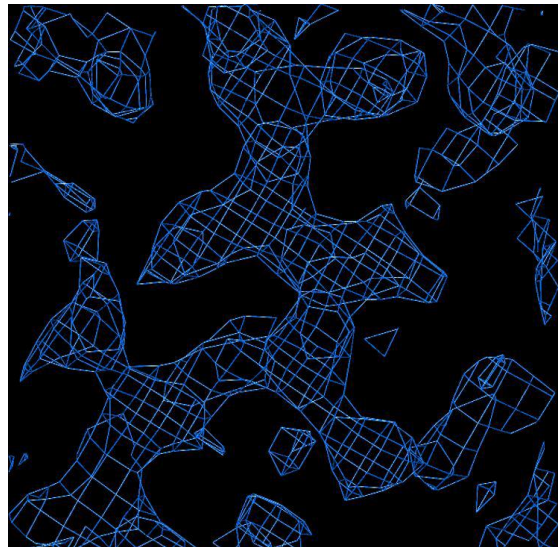
^1H Chemical Shift



Intenzita cross-píků odpovídá inter-jaderné vzdálenosti

Klasické rozlišení proteinů

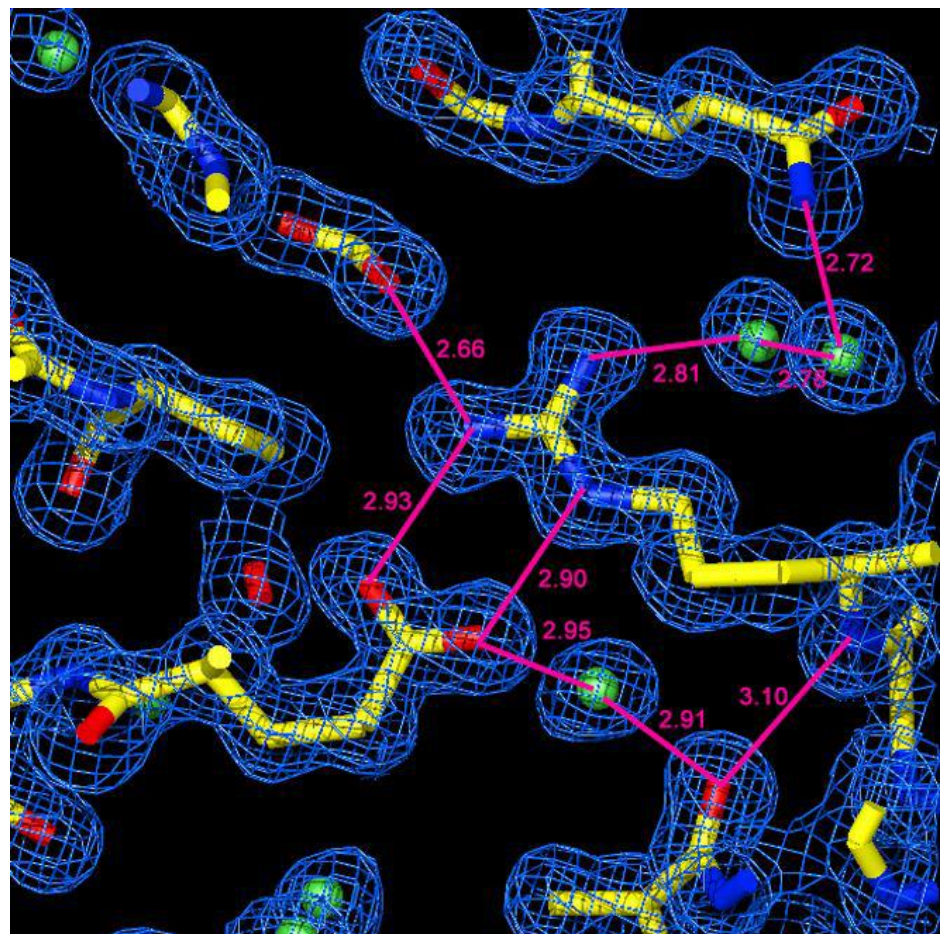
- běžné rozlišení u proteinových struktur pod 2.5 Å
- Při tomto rozlišení je snadné sledovat průběh hlavního řetězce a řada postraních řetězců má také dobře definovanou hustotu. U proteinů je limit pro publikaci struktury rozlišení 3.0 Å.
- Pod 2.5 Å většinou lze použít pro drug design (dokování)



R = 2.6 Å

Vysoké rozlišení

- Mapa elektronové hustoty s velmi vysokým rozlišením jasně ukazuje pozice jednotlivých atomů.
- Ideální pro drug design



$R = 1.2 \text{ \AA}$

H-vazby (fialová) mezi N a O atomy.