

# Databáze v chemické a forenzní analýze

Tereza Uhlíková, tereza.uhlikova@vscht.cz

Tereza Uhlíková

Ústav analytické chemie

skupina teoretické spektroskopie

místnost B4335

<https://web.vscht.cz/~uhlikovt/>

[tereza.uhlikova@vscht.cz](mailto:tereza.uhlikova@vscht.cz)

# O čem to bude?

		Téma přednášky	Téma cvičení
5	17.3.	Základní principy AI	otázky - odpovědi - učení
6	24.3.	Použití ML	Algoritmy ML
7	31.3.	Chemické databáze	Analýza spekter pomocí databází
8	7.4.	Chemické databáze	Konkrétní použití ML - MS databáze
9	14.4.	Konkrétní použití ML, Forezní databáze - zkušková témata	
10	21.4.	dodělavky - opakování, ERA, normalizace, krystalografie	
11	28.4.	Ivan Raich	
12	5.5.	Ivan Raich	

## X-RAY DIFFRACTION ANALYSIS IN THE FORENSIC SCIENCE: THE LAST RESORT IN MANY CRIMINAL CASES

International Centre for Diffraction Data 2003, Advances in X-ray Analysis, Volume 46

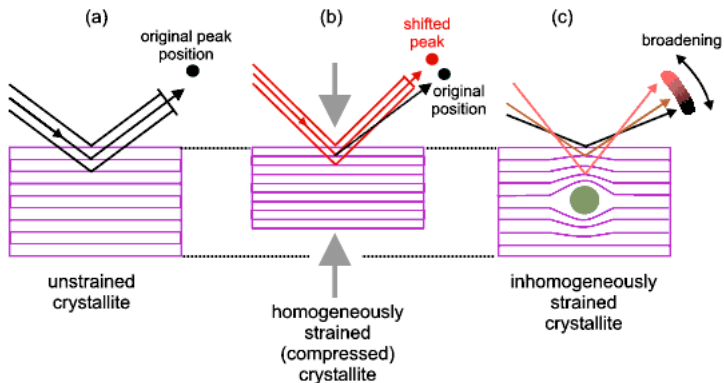
dle rozdílných krystalografických struktur (polymorfů) lze určit původ, padělky, změny vlivem působení různých chemicko-fyzikálních podmínek, biologickou aktivitu

existuje 230 krystalografických grup

<https://crystalsymmetry.wordpress.com/2014/08/15/the-space-group-list-project-as-a-poster/>

## Anorganická – minerály (1913) a ideální krystal

- Struktura krystalu - Konkrétní rozmístění stavebních částic (atomů, iontů) krystalických látek v prostoru
- Základní vzor
- Krystalová mříž (mřížka)
- Základní buňka (např. plošně centrovaná buňka)
- Krystalové (krystalografické) soustavy
- Symetrie vnějšího tvaru a struktury krystalů
- Bodové grupy symetrie krystalů
- Prostorové grupy symetrie krystalů
- Významné krystalografické směry



Je těžké vytvořit monokrystal – **Prášková difrakce**

Spektrum poskytuje – úhlovou polohu difrakčního maxima, intenzitu, tvar a šířku píků

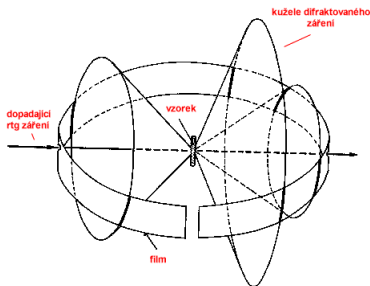
Aplikace a její výsledky – stupeň krystalinity, fázová identifikace, indexování základních buněk vyřešení krystalové struktury studium jedné látky v různých fázích (polymer)

Princip metody <http://www.xray.cz/kfes/vyuka/lp/pd.htm>

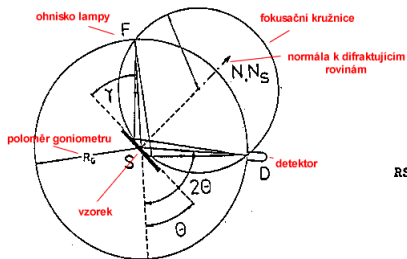
# Prášková difrakce

Podmínkou pro difrakci rentgenového záření o vlnové délce  $\lambda$  na osnově rovin s difrakčními indexy  $hkl$  a mezirovinnou vzdáleností  $d_{hkl}$  je, aby záření dopadalo na tyto roviny pod úhlem  $\theta_{hkl}$ , který splňuje Braggovu rovnici

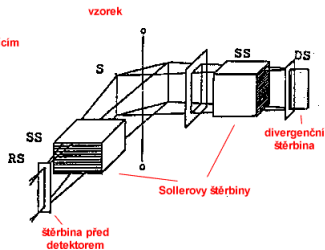
$$2d_{hkl} \sin \theta_{hkl} = \lambda$$



# Prášková difrakce



Schema Braggovy-Brentanovy geometrie



Systém štěrbin v BB geometrii

PDF - Powder Diffraction Files (není portable document od Adobe)  
standardně se dodává jako součást software pro rentgenové difraktometry  
obsahuje úhly reflexi ( $d_{hkl}$ ) a jejich intenzity a slouží pro identifikaci  
jednotlivých krystalických fází (fázovou analýzu)

David Sedmidubský

The International Centre for Diffraction Data \$5775.00 na rok

<http://www.icdd.com/>

[http://www2.fiz-karlsruhe.de/icsd\\_home.html](http://www2.fiz-karlsruhe.de/icsd_home.html)

Obsahuje kompletní strukturní data (grupa symetrie, mřížové parametry, polohy jednotlivých atomů a jejich obsazovací faktory, teplotní faktory)  
199000 prověřených dat včetně jejich atomových souřadnic od roku 1913  
2019 krystalových struktur prvků  
33809 data pro dvojčlenné směsi  
65126 trojčetné  
60669 čtyřčetné a pětičetné  
okolo 138500 položek (79.8%) má přiřazený strukturní typ a dalších 8230  
má aspoň prototyp  
.. Cena a demoverze

<http://www.nist.gov/srd/nist84.cfm>

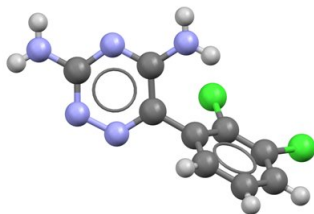
The Cambridge Crystallographic Data Centre

<http://www.ccdc.cam.ac.uk/pages/Home.aspx>

Cambridge Structural Database

<http://www.ccdc.cam.ac.uk/Solutions/CSDSystem/Pages/CSD.aspx>

Založena v 1965, CSD je světová úschovna krystalických struktur malých organických a organo-kovových. Nyní obsahuje přes 1 110 837 struktur.



Dělí se na několik částí:

- CSD – samotná databáze
- ConQuest – vyhledávání v CSD
- IsoStar – knihovna mezimolekulových interakcí
- WebCSD – online přístup k nově publikovaným strukturám
- Mercury – zobrazování a analýza CSD dat
- PreQuest – prostředí pro vytvoření vaší vlastní databáze
- Mogul – ověřování molekulové geometrie

http:

[//www.ccdc.cam.ac.uk/Solutions/CSDSystem/Pages/WebCSD.aspx](http://www.ccdc.cam.ac.uk/Solutions/CSDSystem/Pages/WebCSD.aspx)

ConQuest

- vyhledávání podle jména, vzorce, prvkového zastoupení, reference v literatuře, experimentálních detailů
- chemické podmínky jako náboj, hybridizace a počet cyklů
- mezimolekulové a nevazebné interakce

CCDC ConQuest (1) : search2 [Search]

File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists View Results

Refcode: ABUSAR CSD version 5.33 (November 2011)

ABUSAR

Analyse Hitlist

- ✓ ABADOX
- ✓ ABADUD
- ✓ ABEBIS
- ✓ ABEBOY
- ✓ ABEMOJ
- ✓ ABEMUP
- ✓ ABUKEN
- ✓ ABUPOC
- ✓ ABURUK
- ✓ **ABUSAR**
- ✓ ABUSOF
- ✓ ABZNP S
- ✓ ABZSLM
- ✓ ABZTCX
- ✓ ACANEY
- ✓ ACANIC
- ✓ ACANOI
- ✓ ACAPEAW
- ✓ ACAPEA
- ✓ ACENEB
- ✓ ACIFAT
- ✓ ACIROT

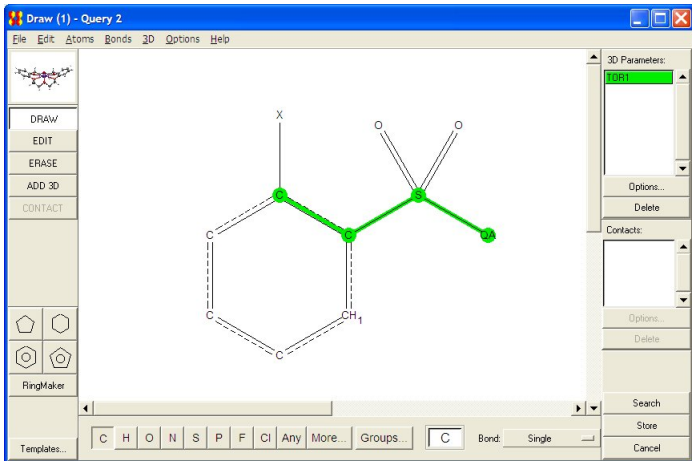
1910 hits

34%

Stop Search

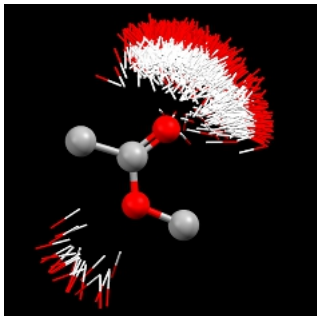
02 Param: All No Objects Detach

Show substructure matches

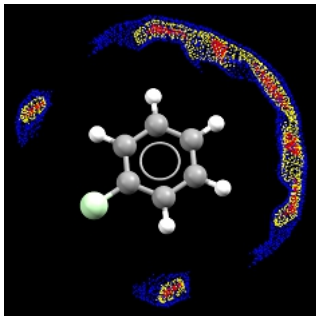


## Isostar – Knihovna mezimolekulových interakcí

O-H skupiny okolo alifatické esterové skupiny

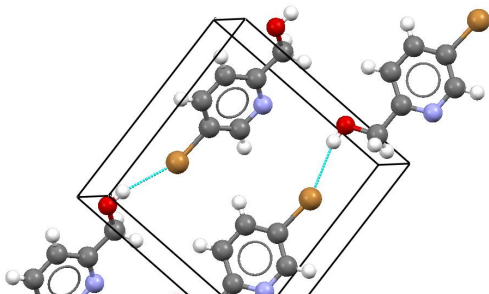


Elektronová hustota okolo fenylové skupiny



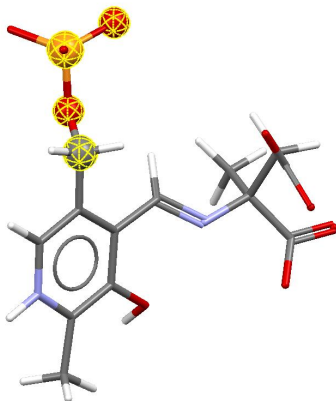
## Mercury – zobrazování a analýza CSD dat

- Zobrazuje elementy prostorových grup
- Vypočítá a zobrazí volná místa v krystalech, založená na kontaktech povrchů nebo s rozpouštědlem
- Ukazuje i vypočtenou geometrii jedné neinteragující molekuly pomocí MOPAC
- Počítá mezimolekulový potenciál a zobrazuje nejsilnější interakci v krystalu



Mogul – ověřování molekulové geometrie

Poskytuje přesné informace o dané molekulové geometrii prohledáváním chemicky klasifikovaných vazebných délek, valenčních úhlů, torzních úhlů a prstencových uspořádání odvozených z CSD



## WebCSD – vyhledávání přes web

**File Filter Help**

Find Entry

Entry

- BUFPQAE01
- BUVGII
- CAACTY
- CACWOS
- CADVEI
- CAFINE**
- CAFROR
- CAMHFA
- CAMXAP01
- CAQTET
- CARQOB
- CARTEN
- CARTEN02
- CATCOL13
- CBMZPN01
- CBMZPN02
- CBMZPN03
- CBMZPN10
- CBMZPN11
- CBMZPN12
- CCXAPT
- CEBGUL
- CECZEP
- CECZIT
- CEFXOA
- CEHTAK10

< >

500 Hits

100%

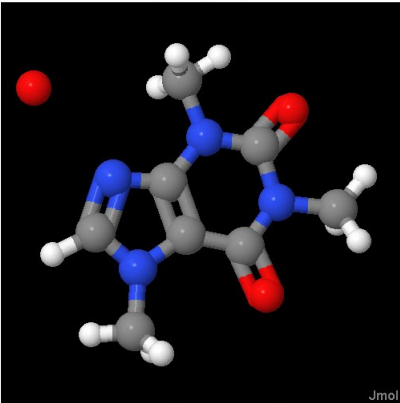
Stop Search

Entry loaded

**CAFINE** : 1,3,7-Trimethyl-purine-2,6-dione monohydrate

D.J.Sutor, *Acta Crystallogr.* (1958), **11**, 453, doi:10.1107/S0365110X58001286

Hide Viewer

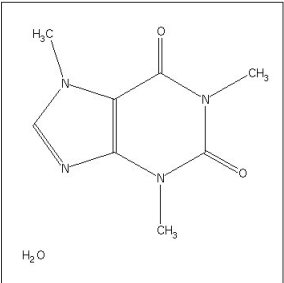


Jmol

Ball and Stick No Labels

Hydrogens  Bond types  Disorder

Diagram Details Viewer Export Options Help



$C_8H_{10}N_4O_2 \cdot H_2O$

**Space Group:** P 2<sub>1</sub>/a

**a** 14.8(1) **b** 16.7(1) **c** 3.97(3)

**$\alpha$**  90  **$\beta$**  97.0(5)  **$\gamma$**  90

**R-Factor:** 14.6%

**Temperature (K):** Room Temp.(283-303)

# Krystalografické - proteinové

1958 – první krystal myoglobin

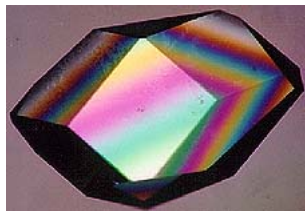
1962 - Nobelova cena v chemii Kendrew a Perutz



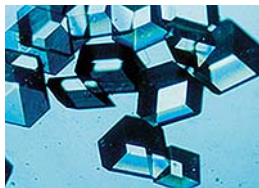
Krystalizace proteinů je obtížná – velká citlivost na teplotě, pH, iontové síle. . .

Proteinové krystaly jsou molekulární krystaly s velkým obsahem rozpouštědla

Lysozym

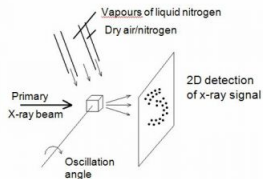
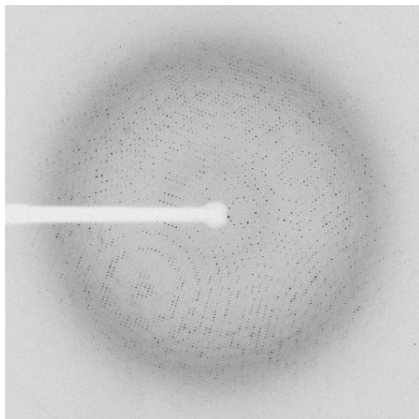


Inzulin

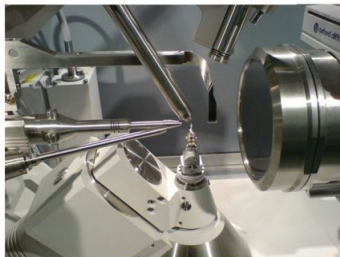


# Krystalografické - proteinové

<https://old.fzu.cz/oddeleni/oddeleni-strukturni-analyzy/research-subjects/krystalografie-proteinu>  
difrakce



Experimental set-up



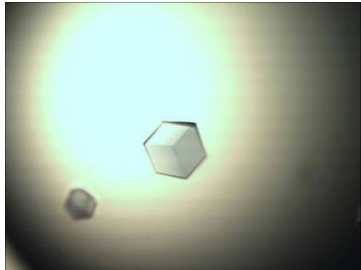
# Krystalografické - proteinové

<https://old.fzu.cz/oddeleni/oddeleni-strukturni-analyzy/research-subjects/krystalografie-proteinu>

Fázový problém u proteinů

Upřesňování proteinových struktur

Monokrystaly malé lakázy pěstované v malé kapce. Velikost krystalu uprostřed je asi 100 mikronů.



Laktáza z bakterie *Streptomyces coelicolor*, oranžové kuličky ionty mědi v aktivních místech enzymu, 12000 atomů

Protein Data Bank

<http://www.rcsb.org/pdb/home/home.do>

Obsahuje 151579 biologických makromolekulových struktur

Podle čeho se vyhledává

Příklad Hemoglobin

Ovo → ovotransferin

<http://www.rcsb.org/pdb/explore/explore.do?structureId=2D3I>

Najít příklady třeba viru Herpes

The Biologically Interesting Molecule Reference Dictionary (BIRD)

<http://www.wwpdb.org/data/bird>

# Krystalografické - proteinové

symetrie obalu lidského rhinoviru typu 14 složená z 240 jednotek (PDB 1k5m)

Vítěz soutěže: Věda je krásná

