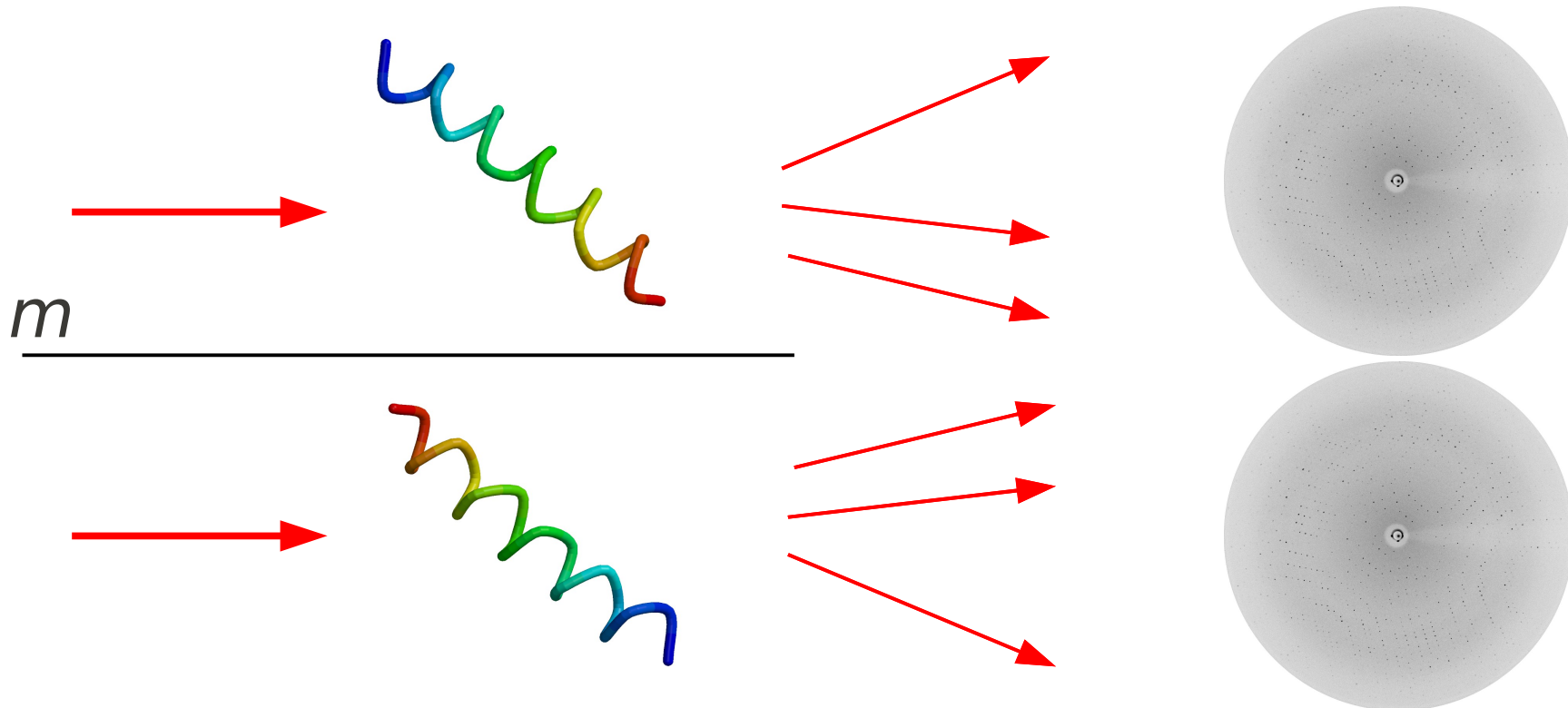


*Makromolekulių struktūros
&
Rentgeno kristalografija
(4)
Modelio konstravimas ir
interpretacija*

S. Gražulis
2012 m.

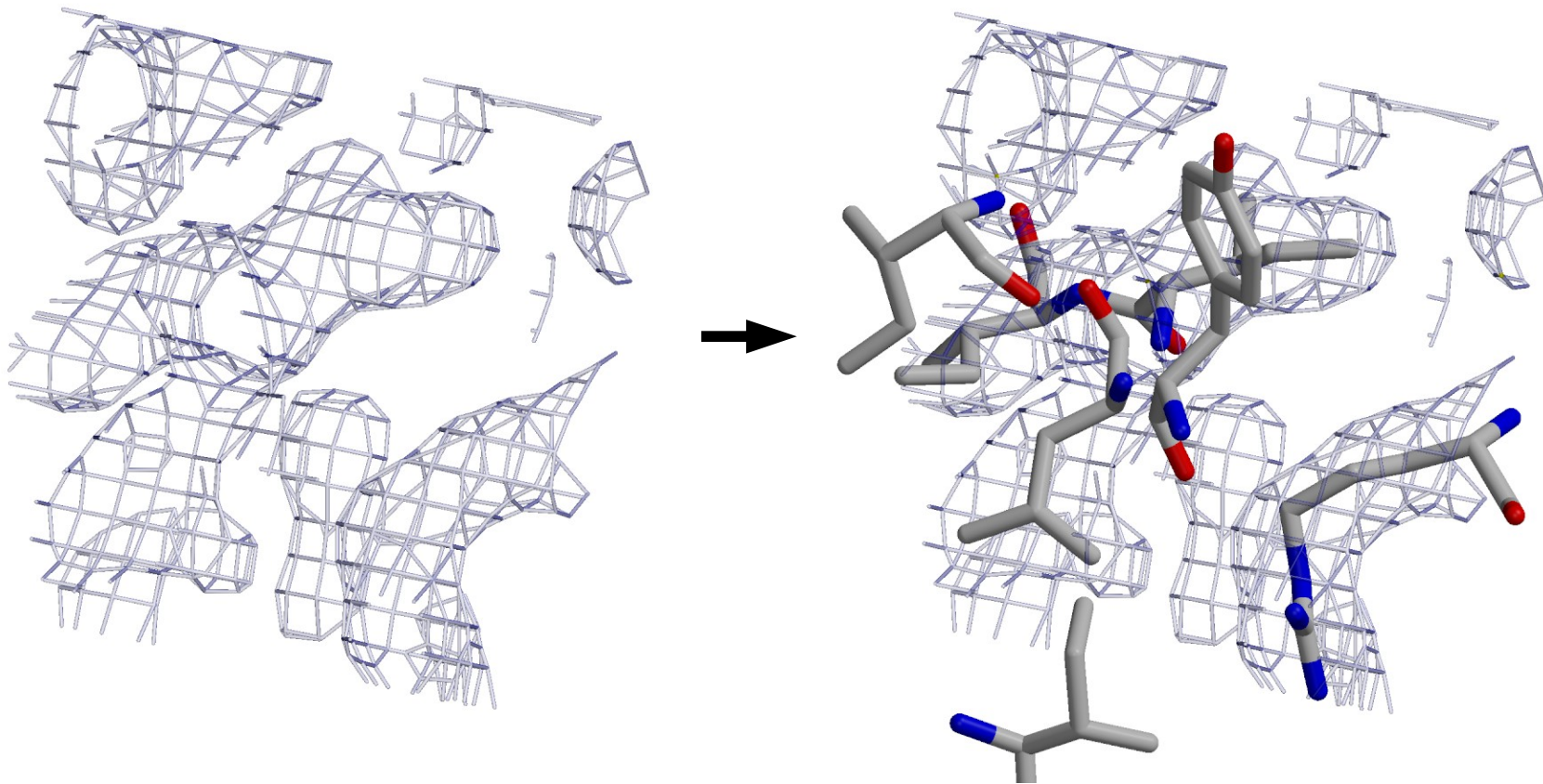
Difrakcijos uždavinio sprendinys – ne vienintelis!

Pavyzdys:



Atspindėta struktūra turi identišką difrakcijos vaizdą
(jei nėra anomalaus sklaidymo...)

Elektronų tankis – o kas toliau?



Rankinis modelio konstravimas

The screenshot displays a molecular modeling software interface. The main window shows a 3D wireframe model of a protein structure, rendered in blue and green. The interface includes several control panels and a command window.

Top Panel: Controls, Display, Graph, Rebuild, Bones, Density, Menus

Object_menu (Left Panel):

- BUILD on
- MYS1 off
- MYS2 off
- MINI off
- REFMAC off
- WAT on
- NEG on
- HG on
- NA off
- NB off
- AONB off
- BONA off
- UNREF off
- OLD off
- GLYC on
- TRIS on
- MES on
- SRT on
- SST on
- RRT on
- DELFWT on
- FWT on

User Menu (Right Panel):

- Centre_ID
- Clear_ID
- Clear_flags
- @remap
- @delpos
- @delneg
-
- @mir
- @mod
- @aomit
- @omit
- @make
- @remake
- @regenerate
- @next-ca
- @prev-ca
- @next-O
- @prev-O
- Move_zone
- Move_atom
- Dial_next
- Dial_preiou
- Yes
- Refi_zone
- Lego_Side_Ch
- RSR_rigid
- RSR_rotamer
- @store
- @ci
- @rcsdiff
- Dist_define
-
- @on_startup
- @ldmask
- @ldomac
- @reload
- @ldbuild
- @ldmaps

Command Window (Bottom Left):

```
Mol> Created
Sam> File ty
Sam> Databa
Sam> Space f
Sam> Space f
Sam> Molecul
Sam>
Mol> Nothin
Mol> Created
Sam> File ty
Sam> Databa
Sam> Space f
Sam> Space f
Sam> Molecul
Sam>
Mol> Nothin
Mol> Created
Sam> File ty
Sam> Databa
Sam> Space f
Sam> Space f
Sam> Molecul
Sam>
Mol> Nothin
Mol> Created
Mol> Macro i
As2> Macro i
As2> Macro i
Map> New map
Map>
Map> New map
Map>
centre-atom
As3> Macro in database.
Map> New map format :- )
Map> red
Map> New map format :- )
Map> blue
```

Bottom Panel: Zoom, Slab, Rot Z, Trans Z, Rot Y, Trans Y, Rot X, Trans X

Papildoma informacija modelio konstravimui

- Molekulių skaičius asimetriniame vienetė
- Baltymo dydis ir seka
- Individualių a.r. struktūros ;-)
- Kristalinimo tirpalų sudėtis (jonai, mažos molekulės) ir sudedamųjų dalių struktūros/savybės

Struktūrų patikslinimas

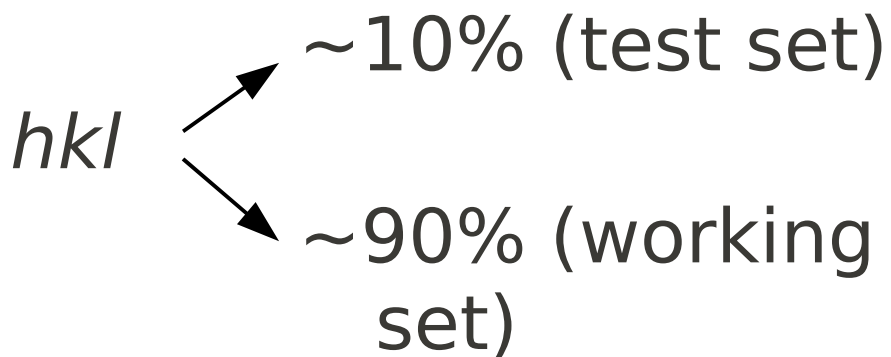
- Mažiausių kvadratų metodas
- Didžiausio tikėtinumo (maximum likelihood) metodai
- Baijeso statistika (Bayesian statistics)

Patikslinimo kriterijai

Kristalografinis R-faktorius:

$$R_{cryst} = \sum_{hkl} |F_{hkl}^{obs} - F_{hkl}^{calc}| / \sum_{hkl} |F_{hkl}^{obs}|$$

R-free:



$$R_{free} = \sum_{hkl \in Test\ set} \dots$$

$$R_{cryst} = \sum_{hkl \in Working\ set} \dots$$

Patikslinimo parametrai

- Atomų koordinatės
- Temperatūriniai (B) faktoriai
- Atomų pozicijų užimtumai
- Svarbu:
Parametrų/matavimų skaičiaus santykis

Papildomi apribijimai

- Jungčių ilgiai
- Kampai tarp jungčių
- Dvisieniai kampai
- B-faktorių apribojimai
- Anizotropinių B-faktorių koreliacijos
- Nekristalografinės simetrijos apribojimai

Patikslinimo savybės

- Rentgeno duomenų svoris
- Apribinomų griežtumas
- Patikslinamų parametrų skaičius (izotropiniai/anizotropiniai B-faktoriai, alternatyvios konformacijos, tirpiklio/ligandų molekulių skaičius)

Nevaltyminės molekulės baltymo struktūroje

- Tirpiklio (vandens) molekulės
- Jonai
- Ligandai, kofaktoriai
- Modifikuotos amino rūgštys ar
DNR bazės

PDB faile patalpinta informacija

- Pastabos (**REMARK**)
- Seka, antrinė struktūra
- Kristalografinė informacija (**CRYST** įrašas)
- Koordinatės – **ATOM** ir **HETAM** įrašai

PDB failų interpretavimas

- Techniniai momentai:
 - ortogonalizacija
- Modelio turinys:
 - Nesutvarkyti rajonai
 - Vanduo ar jonas?
 - Rotamerai
 - Alternatyvios konformacijos

Chaosu liudininkai

- Nuliniai užimtumai
- Milžiniški B-faktoriai
- Liekanos be šoninių grandinių

PDB failo turinys

- Asimetrinio vieneto turinys
- Išimtys – virusų struktūros (vieno kapsomero turinys)
- makromolekulių struktūros (macromolecular structure, .mmol) failai – biologiškai aktyvios molekulės rekonstrukcija

PDB failų formatas

- ASCII koduotės tekstiniai failai
- Fiksuotų kolonėlių formatas
- Įrašas – viena eilutė
- Kiekvienas įrašas prasideda raktiniu žodžiu

<http://www.wwpdb.org/docs.html>

<http://www.wwpdb.org/documentation/format23/v2.3.html>

<http://www.wwpdb.org/documentation/format2.3-0108-a4.pdf>

<http://www.wwpdb.org/documentation/format3.1-20070719.pdf>

PDB failo pavyzdys

```
HEADER      HYDROLASE                      15-SEP-05  2C1L
TITLE      STRUCTURE OF THE BFII RESTRICTION ENDONUCLEASE
..
JRNL       AUTH  S.GRAZULIS,E.MANAKOVA,M.ROESSLE,M.BOCHTLER,
JRNL       AUTH 2 G.TAMULAITIENE,R.HUBER,V.SIKSNYS
...
REMARK     2 RESOLUTION. 1.90 ANGSTROMS.
...
CRYST1    138.925  138.925  94.135  90.00  90.00  90.00 I 4          16
SCALE1     0.007198  0.000000  0.000000          0.000000
SCALE2     0.000000  0.007198  0.000000          0.000000
SCALE3     0.000000  0.000000  0.010623          0.000000
ATOM       1  N  AMET A    1      40.881  1.095  49.888  0.33 24.33      N
ATOM       2  N  BMET A    1      40.265  1.169  49.581  0.33 24.33      N
...
END
```

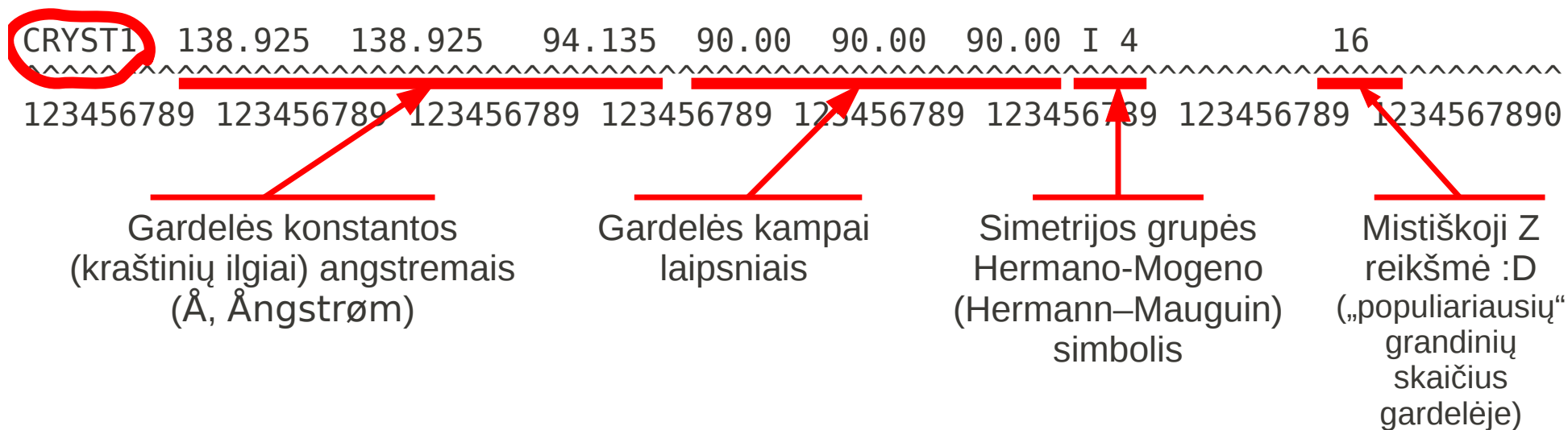
PDB formato ATOM įrašai

Iš originalių PDB 1KNV bei 2EZV įrašų:

ATOM	1	N	ASN	A	4	3.407	40.303	50.109	1.00	66.19	N
123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	1234567890
raktinis žodis	atomo tipas, numeris, liekana – unikalus atomo identifikatorius				ortogonalios koordinatės, Å			užimtumas ir B-faktorius		atomo cheminis simbolis	
ATOM	1	N	ASN	A	4	3.407	40.303	50.109	1.00	66.19	1KNV N
ATOM	1501	N	ACYS	A	186	48.353	52.281	47.983	0.61	20.47	A001 N
ATOM	1502	N	BCYS	A	186	48.355	52.281	47.983	0.39	22.86	A002 N
123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	1234567890
alternatyvios padėties žymė				segmento vardas							
ATOM	2	CA	MET	A	1	64.171	0.298	-93.738	1.00	21.86	C
HETATM	4853	CA	CA	201	77.279	-24.071	-72.974	1.00	36.59	CA	
HETATM	4778	CL	CL	3001	46.959	58.438	4.909	1.00	27.44	1KNVCL	-1
123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	123456789	1234567890
cheminis simbolis seniau buvo žymimas vado pozicijos				atomo cheminis simbolis				krūvis			

Kristalografinė informacija PDB faile – CRYST1 įrašas

Iš PDB 2C1L:



REMARK įrašai PDB failuose

- Iki 1992 – laisvas tekstas (tik žmogui)
- Nuo 1992 – griežtesnė struktūra (žmogui ir mašinai).

```
REMARK 2 RESOLUTION. 2.5 ANGSTROMS. 155CE 1
REMARK 3 155C 15
REMARK 3 REFINEMENT. THESE ATOMIC COORDINATES MUST BE CONSIDERED AS 155CE 2
REMARK 3 PRELIMINARY. THEY WERE OBTAINED BY RUNNING SEVERAL CYCLES 155C 17
REMARK 3 OF THE DIAMOND MODEL BUILDING ROUTINE ON GUIDE POINTS FOR 155C 18
REMARK 3 ATOMS MEASURED FROM THE WIRE KENDREW MODEL. ...
...

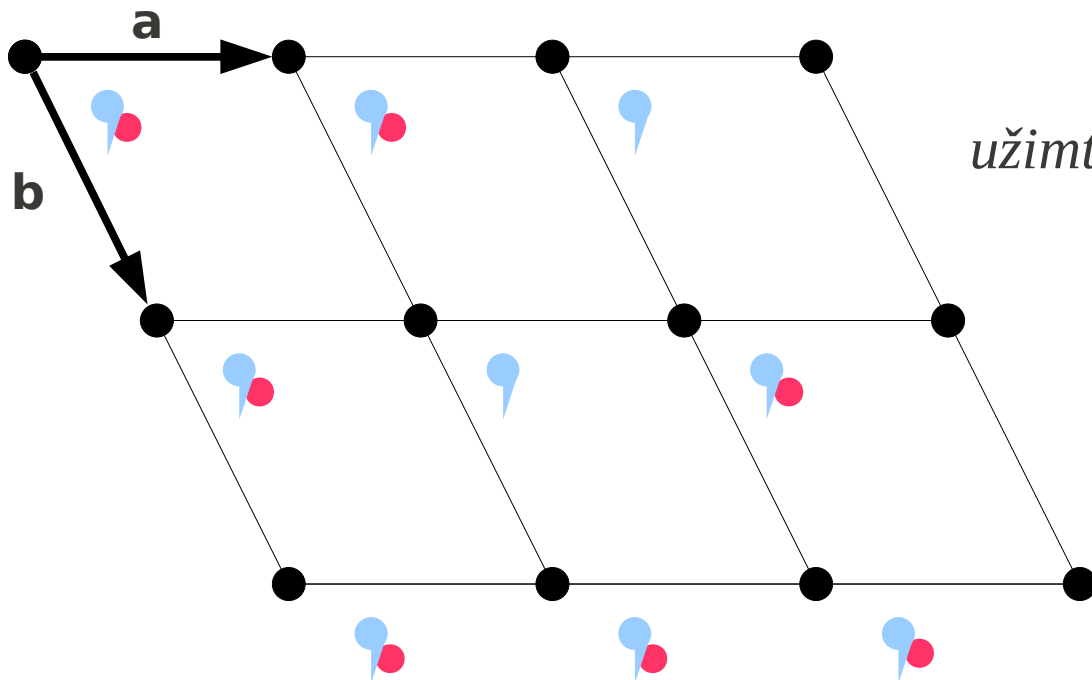
```

```
REMARK 2 RESOLUTION. 2.17 ANGSTROMS.
...
REMARK 3 RESOLUTION RANGE HIGH (ANGSTROMS) : 2.17
REMARK 3 RESOLUTION RANGE LOW (ANGSTROMS) : 24.61
REMARK 3 DATA CUTOFF (SIGMA(F)) : 2.000
REMARK 3 DATA CUTOFF HIGH (ABS(F)) : 2011306.160
REMARK 3 DATA CUTOFF LOW (ABS(F)) : 0.0000
REMARK 3 COMPLETENESS (WORKING+TEST) (%) : 93.6
REMARK 3 NUMBER OF REFLECTIONS : 42686

```

Užimtumas (*angl.* occupancy)

- Idealiame kristale visų gardelių turinys yra vienodas
- Realiame kristale kai kurie atomai gali būti ne visose gardelėse:

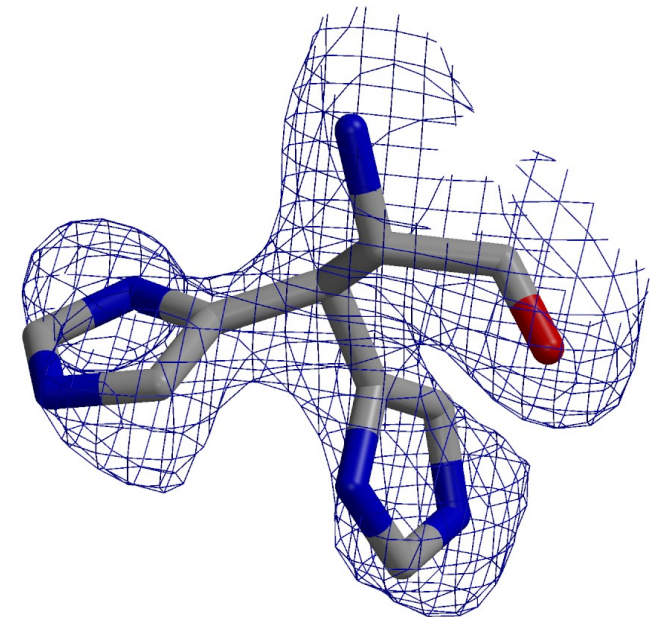
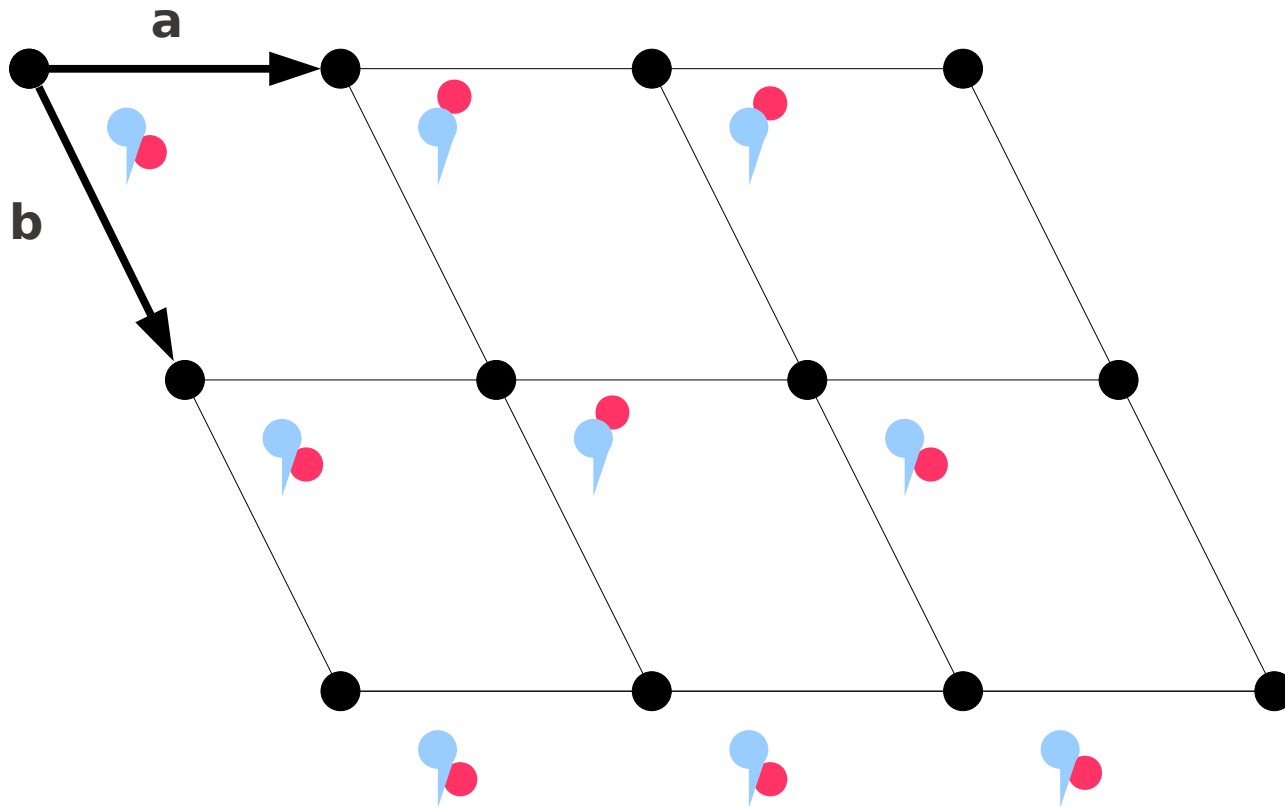


$$\text{užimtumas} = q = \frac{\text{užimtų vietų skaičius}}{\text{bendras vietų skaičius}}$$

$$q = \frac{7}{9} = 0.778$$

Alternatyvios padėtys (*angl.* alternative locations)

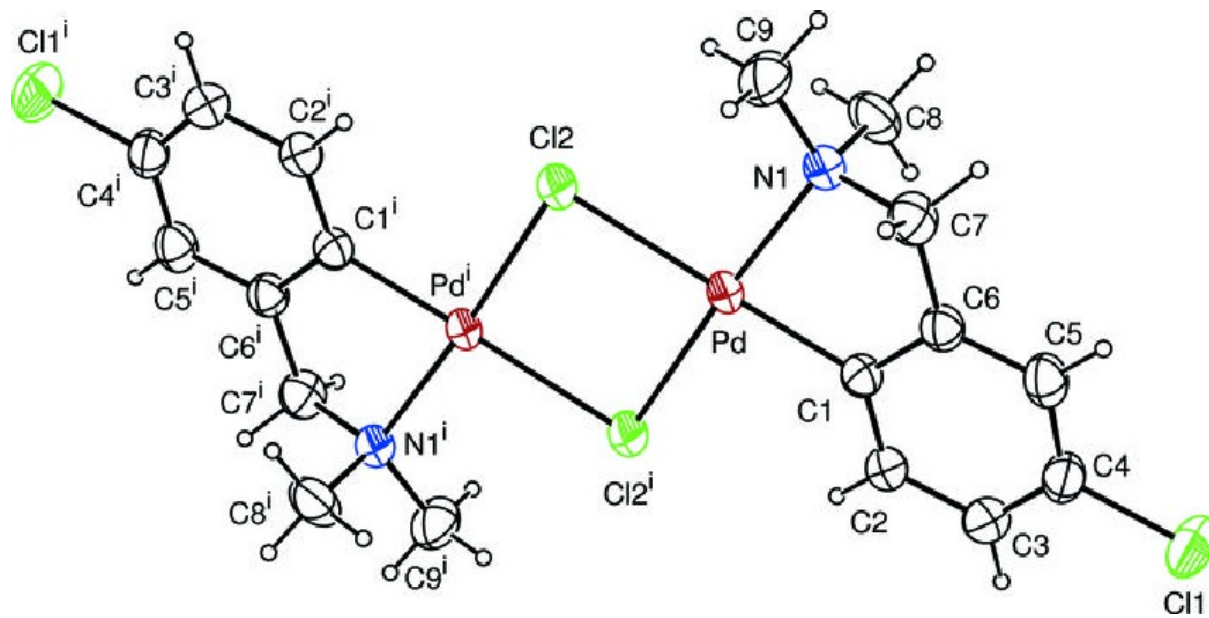
- Kai kurie atomai skirtingose gardelėse gali būti išsidėstę skirtingai:



PDB ID 1KNV, His B169
Grazulis *et al.*

„Temperatūriniai faktoriai“ (*angl.* “temperature factors”, B-factors; IUCr: Debye-Waller factor)

$$B = 8\pi^2 \langle u^2 \rangle$$



Sang et al. Acta Cryst. (2010). E66, m252

<http://journals.iucr.org/e/issues/2010/03/00/bq2191/index.html>

B-faktorių paaiškinimas:

[http://spdbv.vital-it.ch/TheMolecularLevel/ModQual/#Temperature%20factor%20\(crystallogr](http://spdbv.vital-it.ch/TheMolecularLevel/ModQual/#Temperature%20factor%20(crystallogr)

<http://pldserver1.biochem.queensu.ca/~rlc/work/teaching/definitions.shtml>

B – temperatūrinis faktorius
ATOM įrašė
u – atomo pozicijos nuokrypis
nuo vidutinės pozicijos.

⟨ ⟩ – vidurkis laike

$$B = 79 \text{Å}^2 \Leftrightarrow \sqrt{\langle u^2 \rangle} = 1.0 \text{Å}$$

$$\langle u^2 \rangle = \lim_{T \rightarrow +\infty} \int_{-T/2}^{T/2} u^2(t) dt$$

$$p(u) = \frac{1}{\sqrt{2\pi \langle u^2 \rangle}} e^{-\frac{u^2}{2\langle u^2 \rangle}}$$

Literatūra tolesniam skaitymui

- Crystallography Made Crystal Clear by Gale Rhodes (Author)
- Protein Crystallography by T. Blundell (Author), Lewis Johnson (Author)
- Practical Protein Crystallography, Second Edition by Duncan McRee (Author) (September 1999)
- Principles of Protein X-Ray Crystallography (Springer Advanced Texts in Chemistry) by Jan Drenth (February 1999)
- X-Ray Analysis and the Structure of Organic Molecules by Jack D. Dunitz (Author)
- Fundamentals of Crystallography by Carmelo Giacovazzo (Editor), et al.

Resursai tinkle

Protein Data Bank (PDB) <http://www.rcsb.org/pdb/>

Macromolecular Structure Database <http://www.ebi.ac.uk/msd/>

International Union of Crystallography <http://www.iucr.org/>