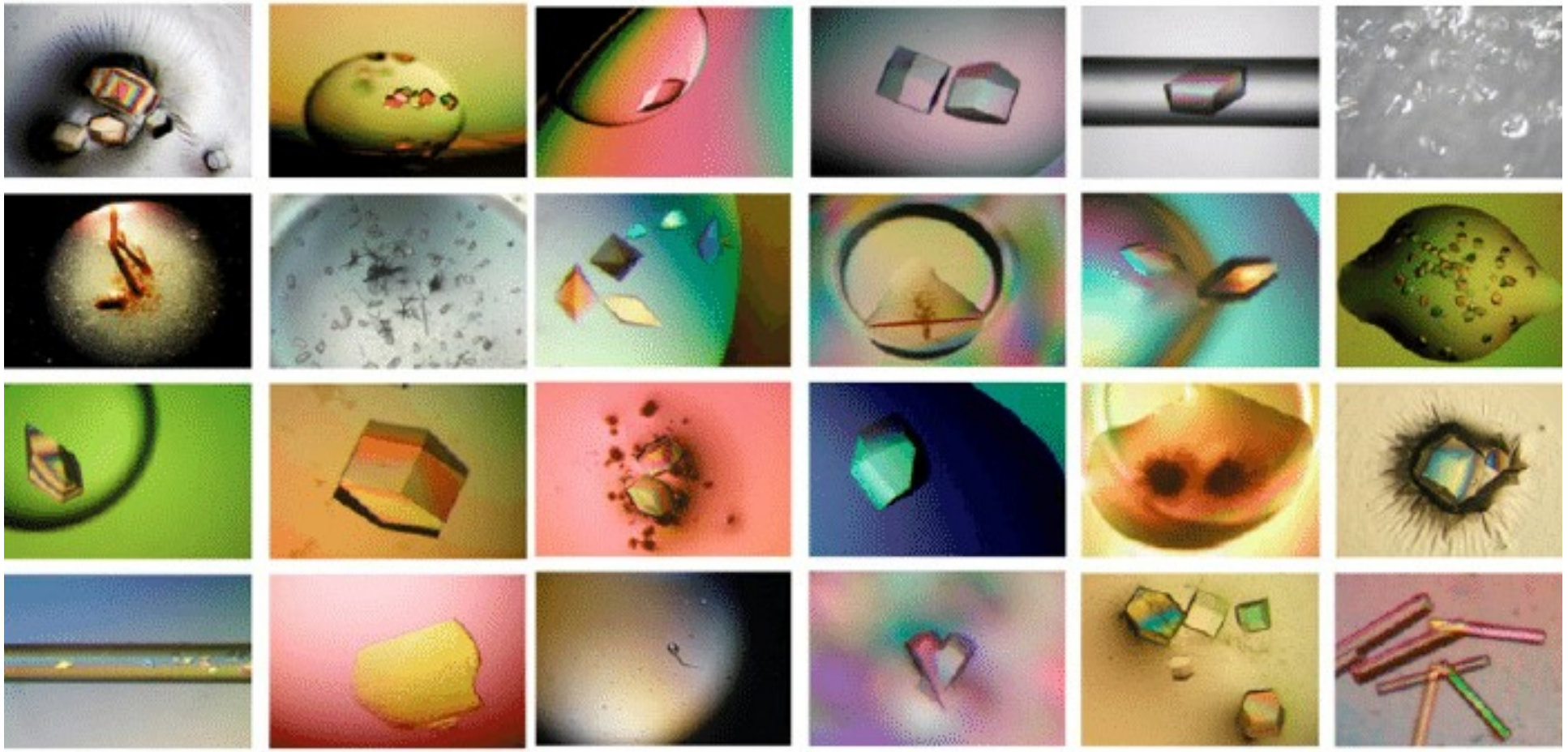
Crystal geometry

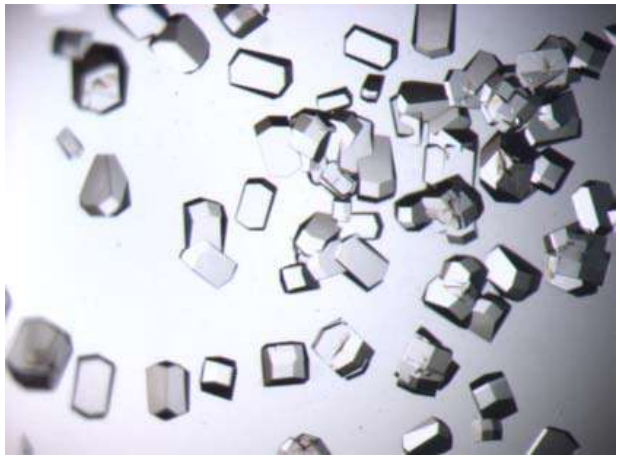


Corso di Biocristallografia e Microscopia Elettronica rdezorzi@units.it

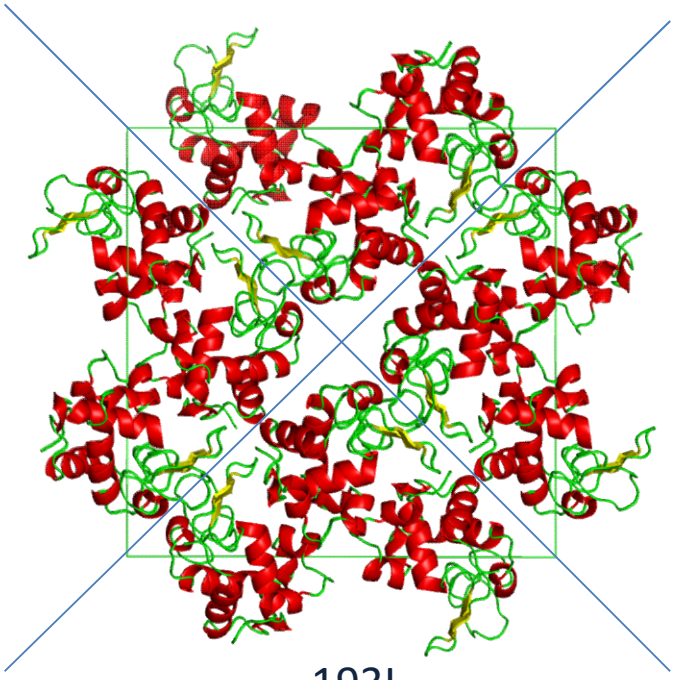
Sharp edges and plane faces

Regular crystal habit depends on internal long-range order of crystal structure.

Diffraction properties of crystals depend on their internal order: X-ray interaction with the ordered molecules forming the crystal.



Combination of the requirements for:
long-range order
and
symmetries between molecules in the crystal
=
230 **space groups** for crystals,
But only **65** allowed for protein crystals



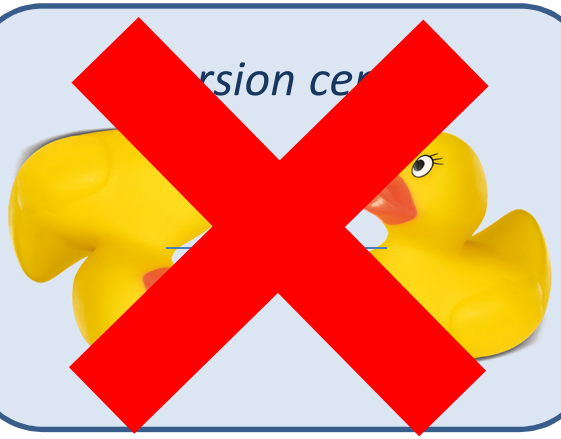
193L

To understand diffraction phenomena and to analyze crystallographic data:
analyze crystal order.

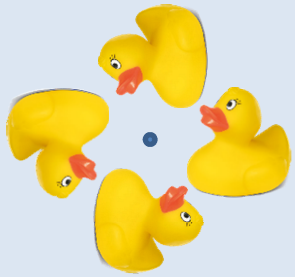
Point group symmetry elements

Point group symmetry elements are invariant points of transformations able to relate a point (x,y,z) to its symmetry related (x',y',z').

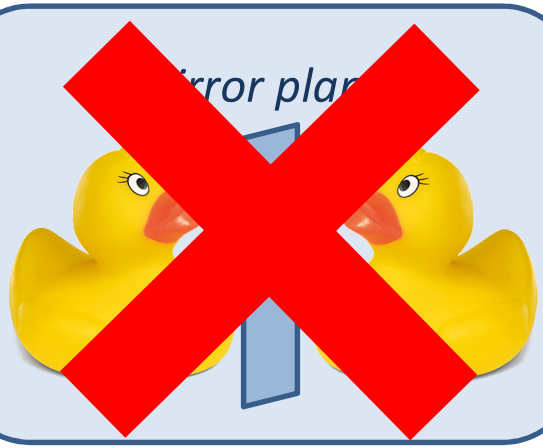
rsion cen



Rotation axis:



rror plan



During each of this transformations, a generic point of coordinates (x,y,z) is transformed in a symmetric point, with coordinates (x',y',z'), by a matrix operator:

Inversion center, i:

$$\boldsymbol{x'} = i\,\boldsymbol{x}$$

$$i = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Rotation axis, e.g. 4 along z:

$$\boldsymbol{x'} = A\,\boldsymbol{x}$$

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Mirror plane, e.g. in yz:

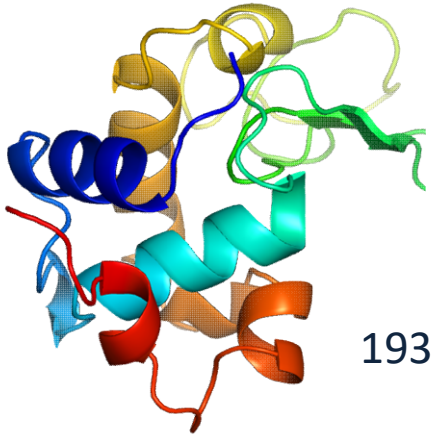
$$\boldsymbol{x'} = m\,\boldsymbol{x}$$

$$m = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Rotation axes

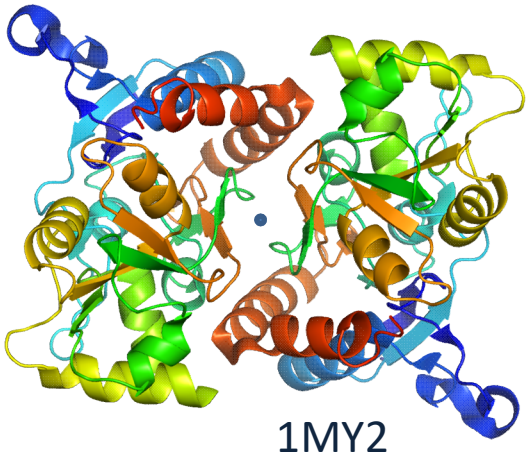
A rotation operation *n* (or of order *n*) describes a rotation of each point of *360°/n* around the rotation axis:

1 Rotation of 360° (identity!)



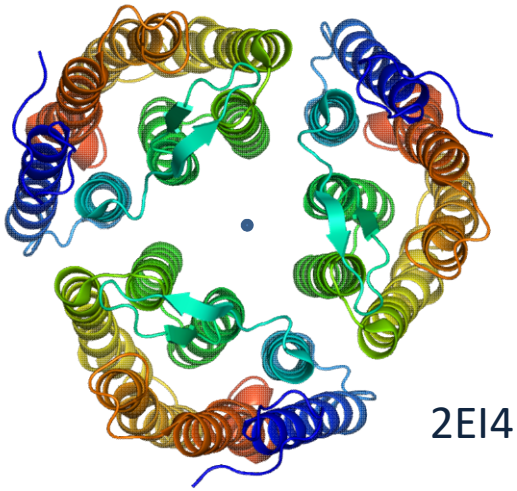
193L

2 Rotation of 180°



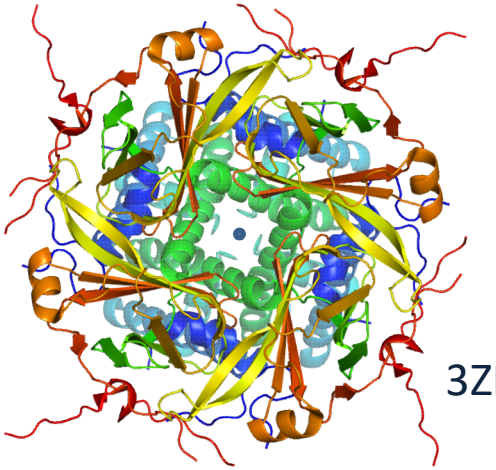
1MY2

3 Rotation of 120°



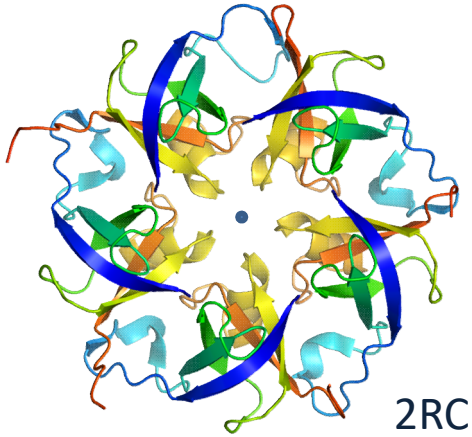
2EI4

4 Rotation of 90°



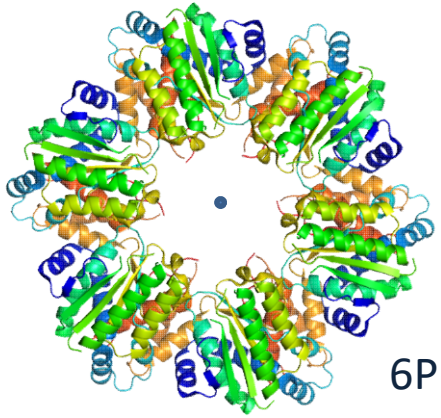
3ZRS

5 Rotation of 72°



2RCF

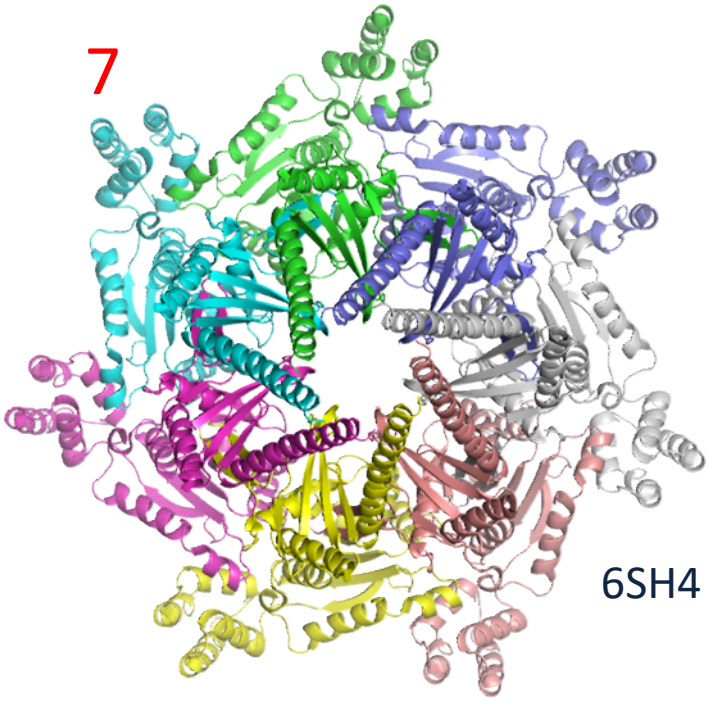
6 Rotation of 60°



6PB3

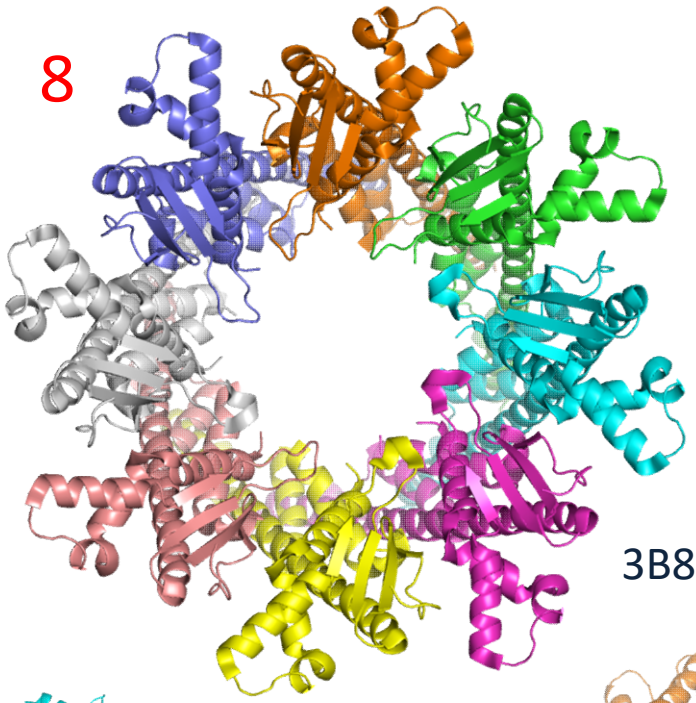
And more:

7



6SH4

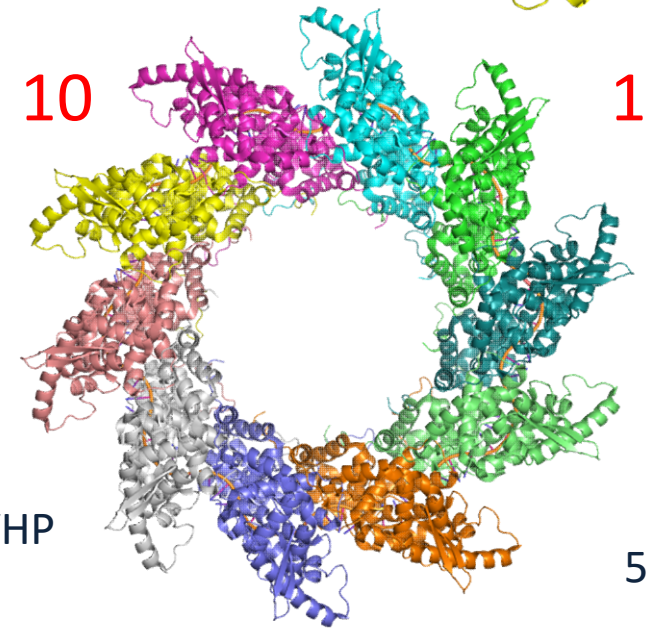
8



3B8O

When symmetry elements are present, they involve more than one protein chain: proteins are ASYMMETRIC!

10



2YHP

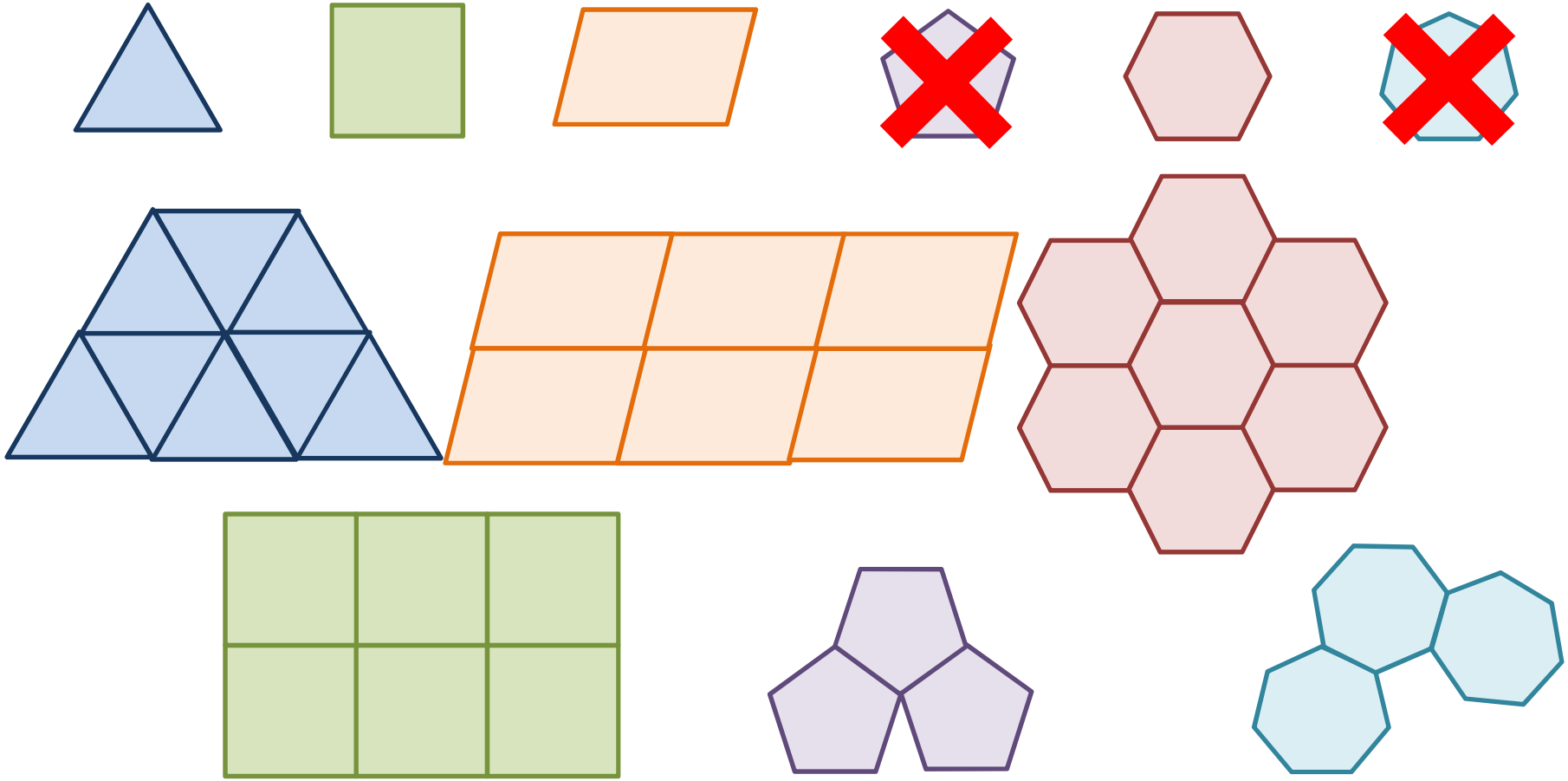
12

5NBZ

Translational periodicity: lattice and unit cell

Crystals are formed by repeated units along the three directions, but to simplify the problem we can start to describe lattice in the 2D case

How to cover a floor with tiles??



Translational periodicity: lattice and unit cell

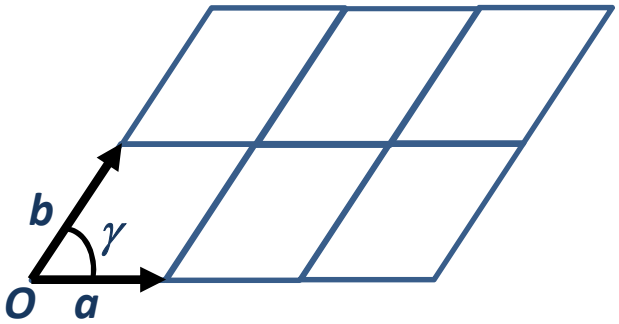
Crystals are formed by repeated units along the three directions, but to simplify the problem we can start to describe lattice in the 2D case

How to cover a floor with tiles??

Due to the requirement for **translational periodicity**, lattices can have only the following rotational symmetries:

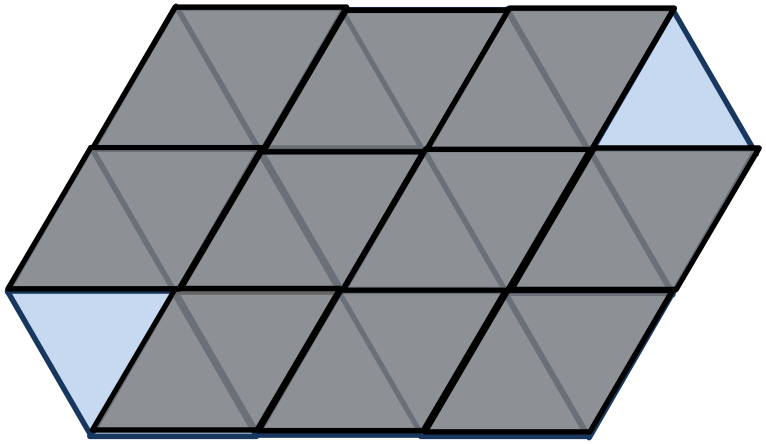
1, **2**, **3**, **4** and **6**

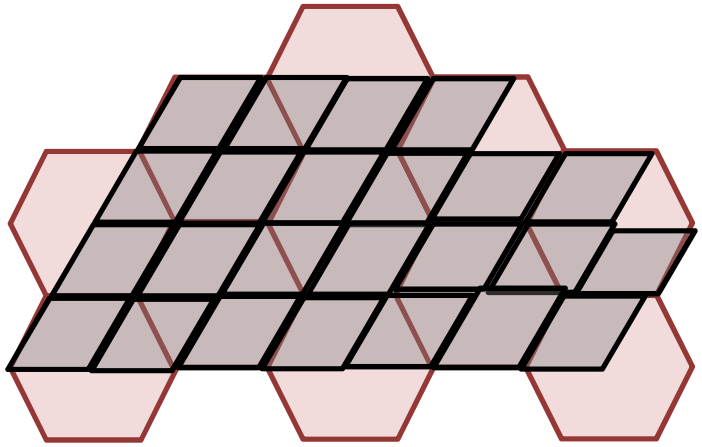
In 2D, the lattice is formed by a repeted unit, translated in 2 directions:



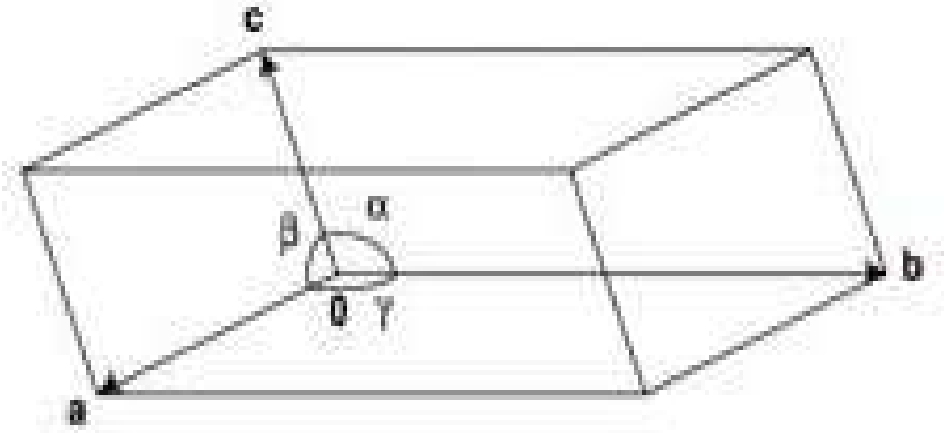
The repetitive unit is called **unit cell**. The lattice is defined by an **origin** of the translations and the **unit cell parameters** (or dimensions): (in 2D) a, b and γ

Even when symmetry elements include a 3-fold axis or a 6-fold axis, unit cell can be described as a parallelogram:





In 3D:



Unit cell parameters:
a, *b*, *c* and α, β, γ

Crystal families and crystal systems

Crystal family	Crystal system	Unit cell dimentions	Minimal symmetry elements
Triclinic	Triclinic	***a***, ***b***, ***c***, α, β, γ	none
Monoclinic	Monoclinic	***a***, ***b***, ***c***, β (α, γ=90°)	2
Orthorhombic	Orthorhombic	***a***, ***b***, ***c*** (α=β=γ=90°)	three perpendicular 2 axes
Tetragonal	Tetragonal	***a***, ***c*** (*b*=*a*, α=β=γ=90°)	4
Hexagonal	Trigonal	***a***, ***c*** (*b*=*a*, α=β=90°, γ=120°)	3
	Hexagonal		6
Cubic	Cubic	***a*** (*b*=*a*, *c*=*a*, α=β=γ=90°)	four 3 axes along diagonal directions

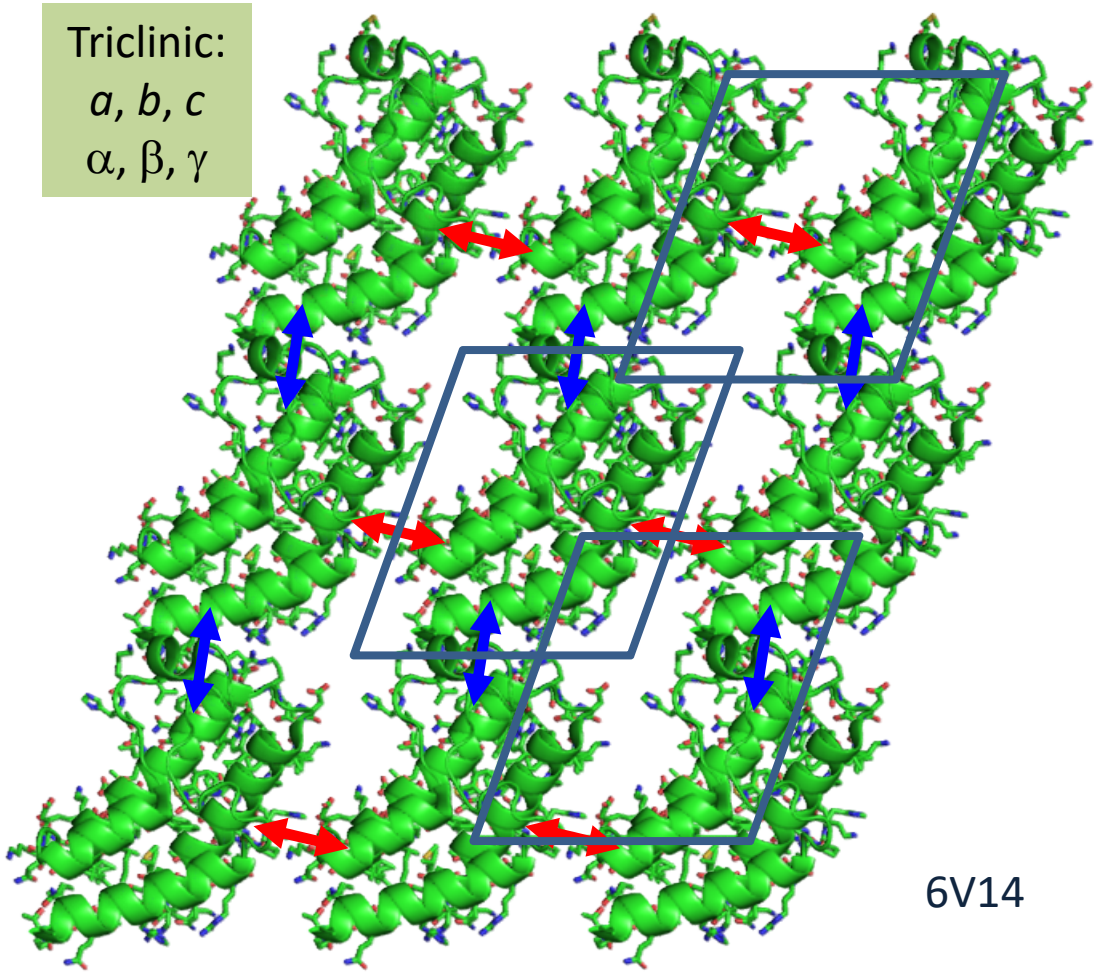
Intermolecular contacts and crystal packing

Crystals are held together by intermolecular contacts, that determine their packing:

Crystal packing must be evaluated:

1) to evaluate the effect of crystal contacts on protein conformation

2) to obtain biologically active unit (for symmetry related oligomers)



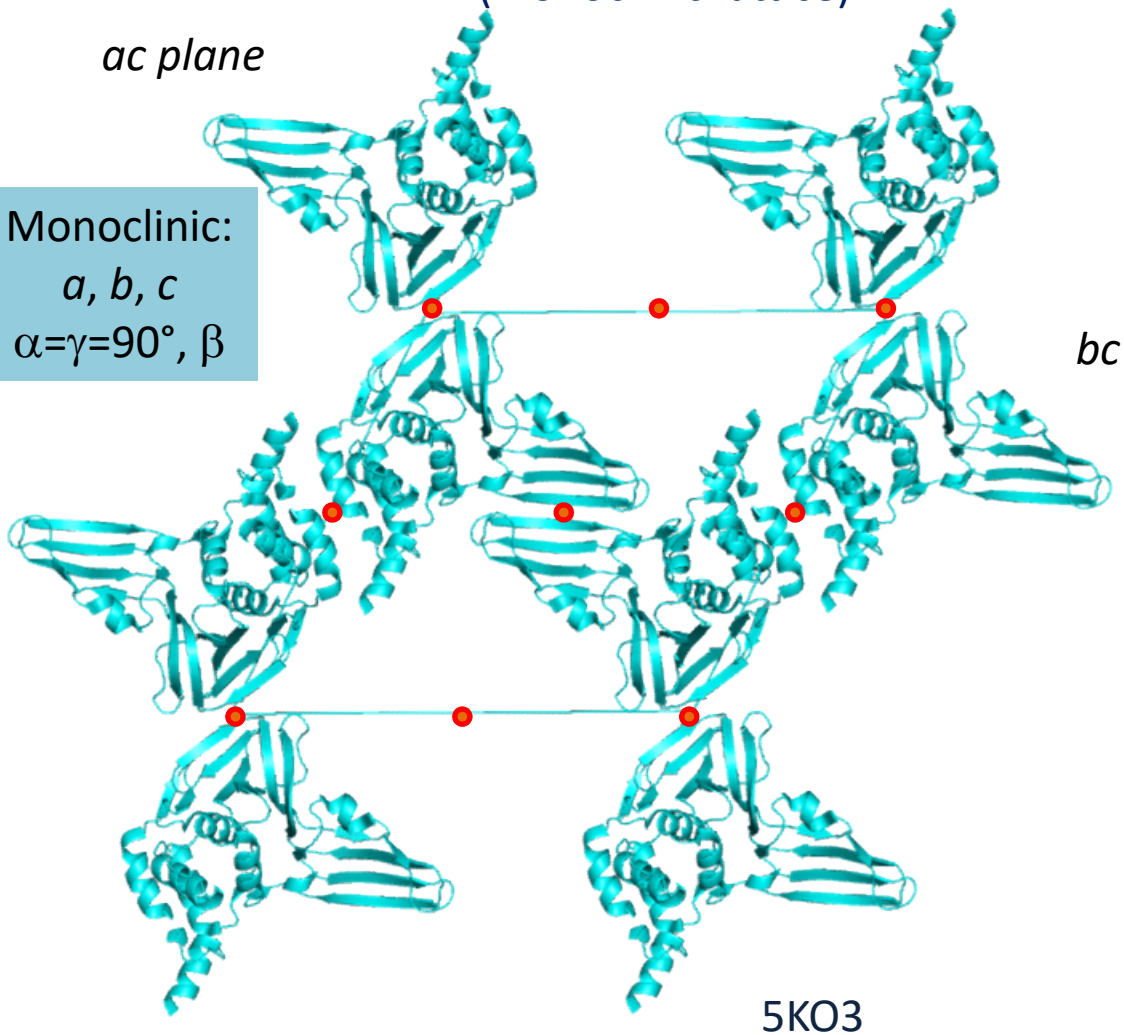
Different **origin** choices are possible (while unit cell dimensions are the same!) for a lattice with no additional symmetry.

Choice of the origin

The choice of the origin depends on the symmetry of the crystal lattice.

Case in point: lattice with 2-fold axis along *b*
(monoclinic lattice)

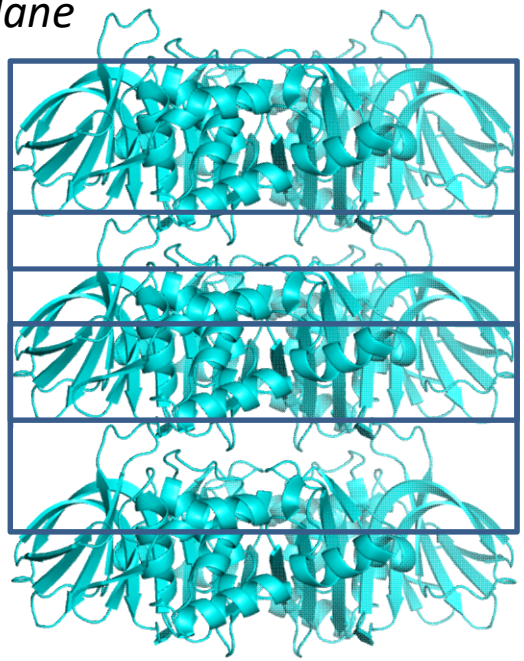
ac plane



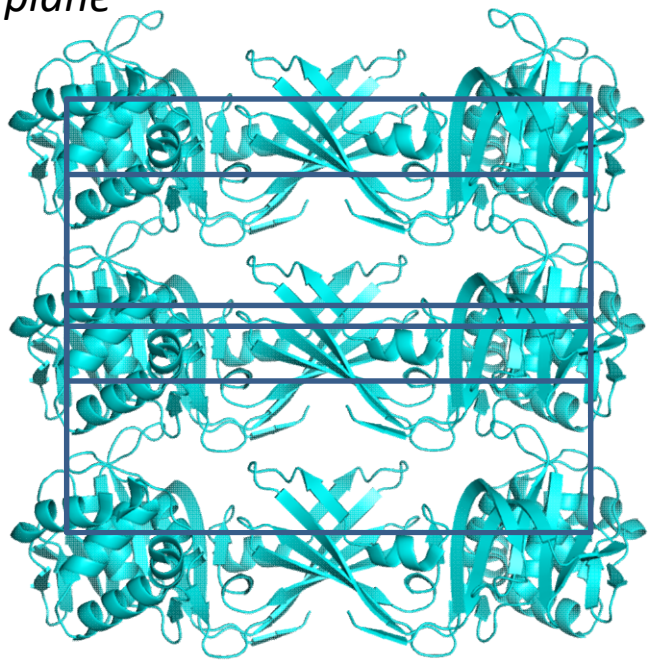
Monoclinic:
a, *b*, *c*
$\alpha=\gamma=90°$, β

5KO3

ab plane



bc plane

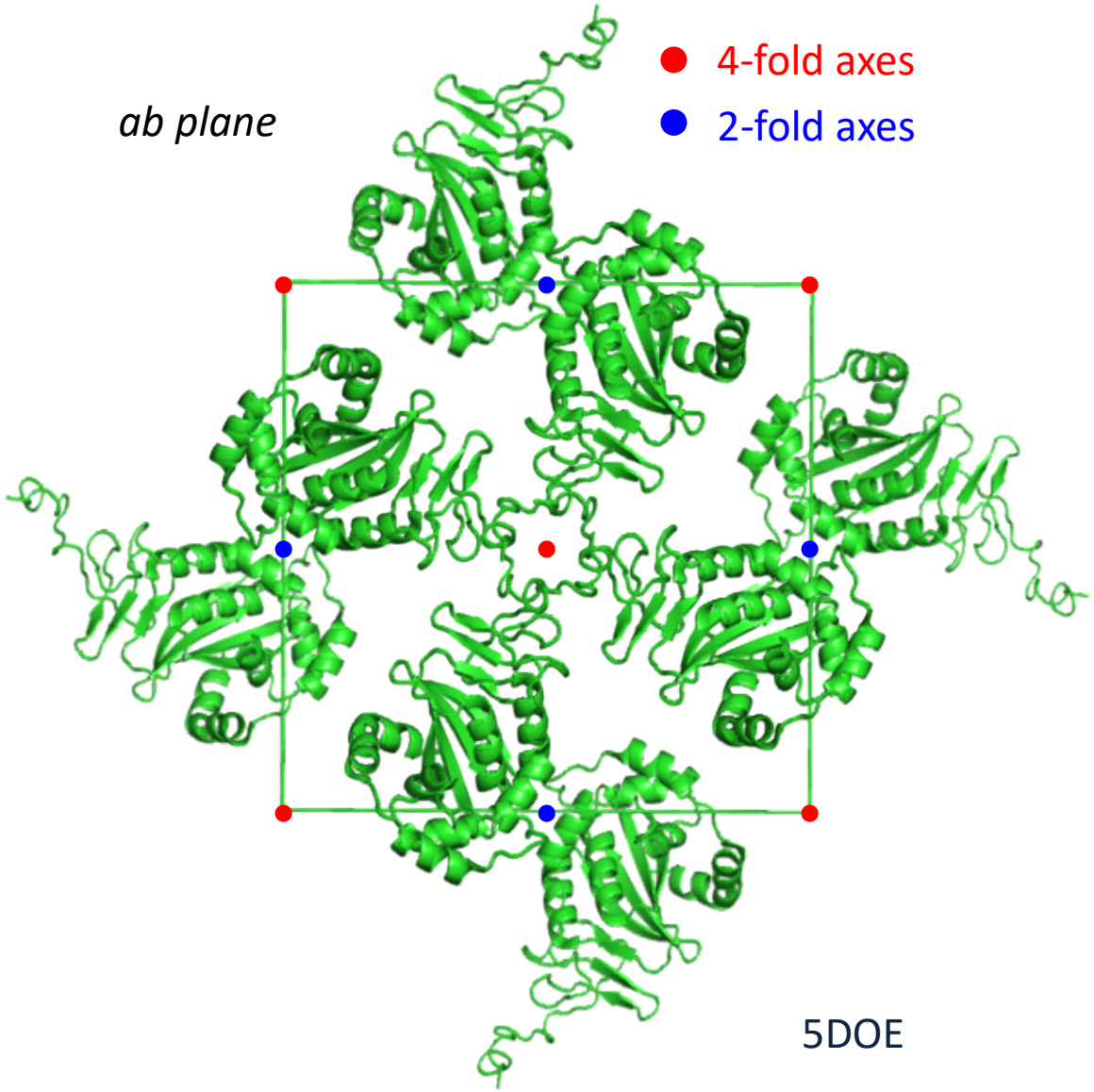


Choice of the origin

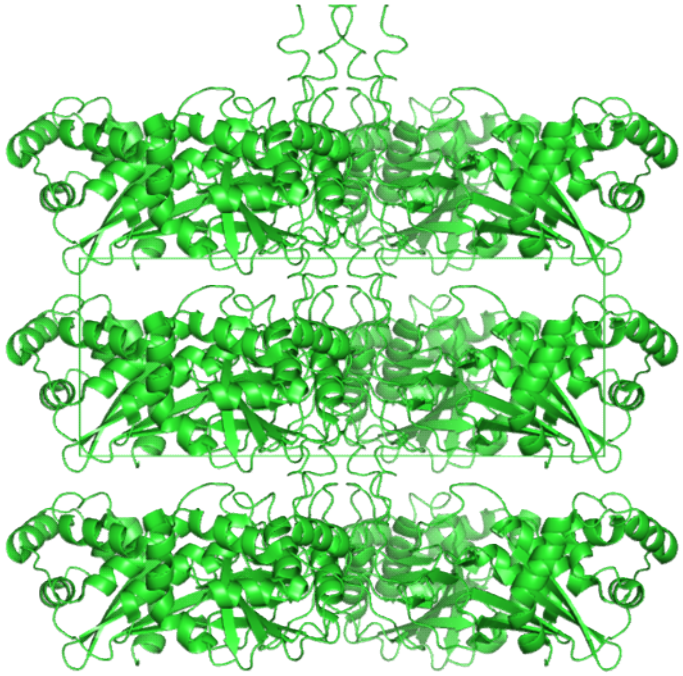
Tetragonal:
a=*b*, *c*
α=β=γ=90°

Case in point: lattice with 4-fold axis along *c* (tetragonal lattice)

ab plane



ac and bc planes

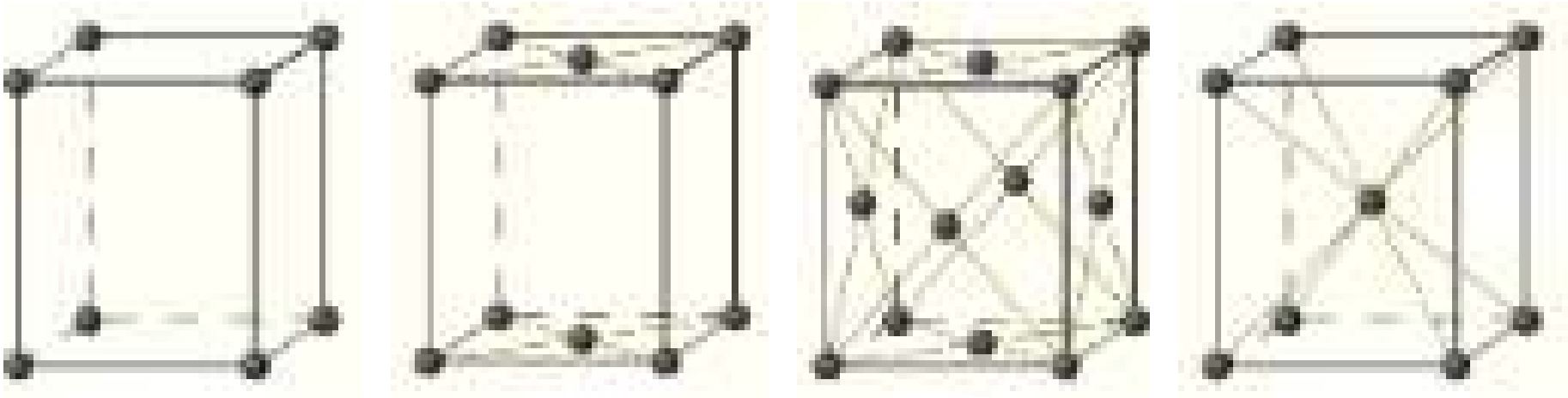


5DOE

Primitive and centered lattices

Primitive lattices have *nodal points* (*nodes*) at the vertices of the unit cell.

But further translations are possible in the unit cell, forming **centered lattices**, with additional nodes:



Primitive lattice

Base centered

Face centered

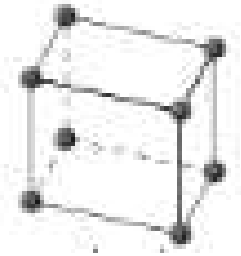
Face(s) centered lattices

Body centered lattice

A primitive lattice can **always** describe the structure, but in presence of specific elements of symmetry it can be useful to describe the lattice as centered.

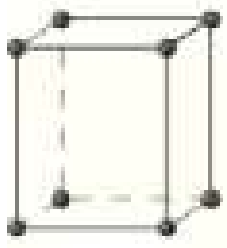
Crystal families + centering = 14 Bravais lattices

Triclinic lattice

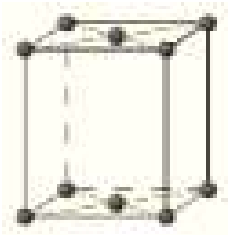


Primitive, P

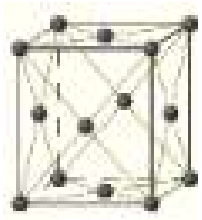
Orthorhombic lattice



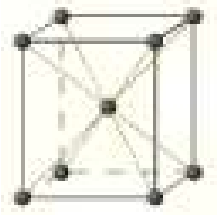
Primitive, P



Base centered, C

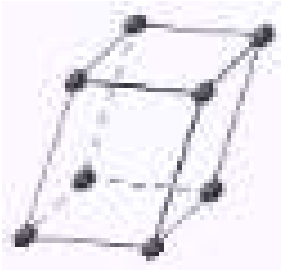


Face centered, F

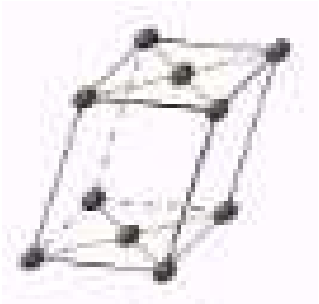


Body centered, I

Monoclinic lattice

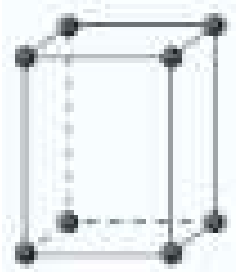


Primitive, P



Base centered, C

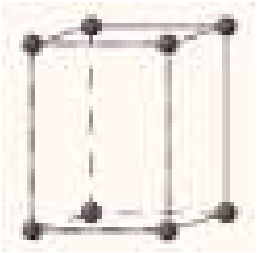
Tetragonal lattice



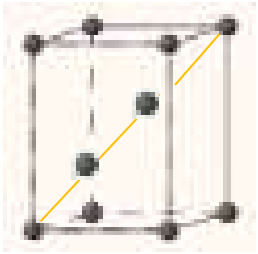
Primitive, P

Body centered, I

Hexagonal lattice

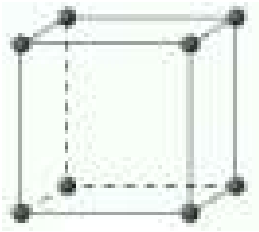


Primitive, P

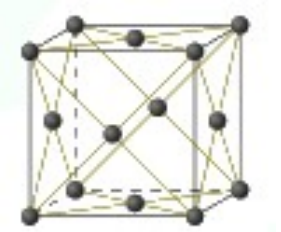


Rhombohedral, R

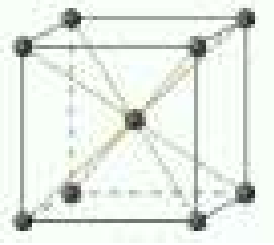
Cubic lattice



Primitive, P



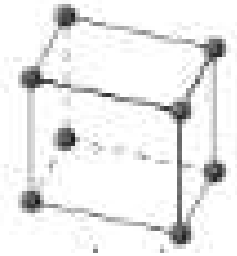
Face centered, F



Body centered, I

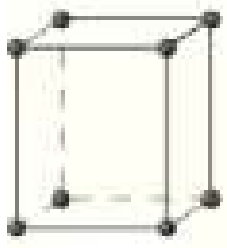
Crystal families + centering = 14 Bravais lattices

Triclinic lattice

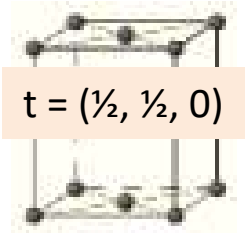


Primitive, P

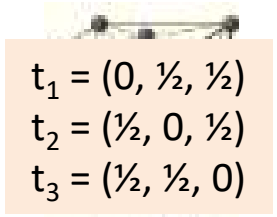
Orthorhombic lattice



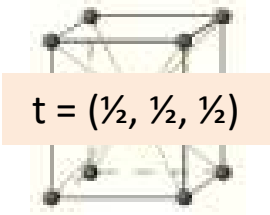
Primitive, P



Base centered, C

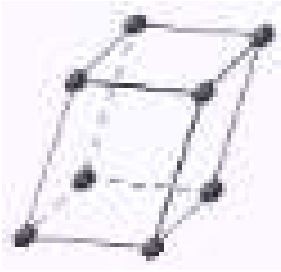


Face centered, F

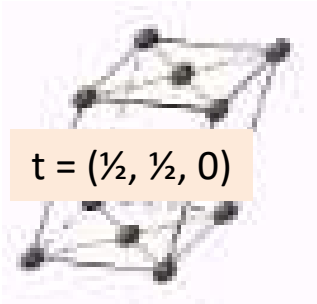


Body centered, I

Monoclinic lattice

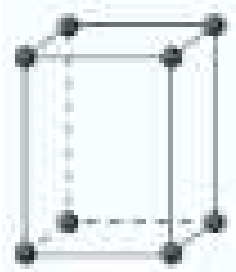


Primitive, P

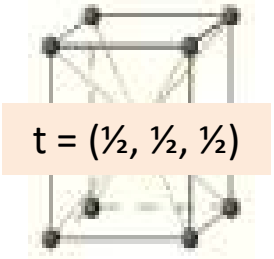


Base centered, C

Tetragonal lattice

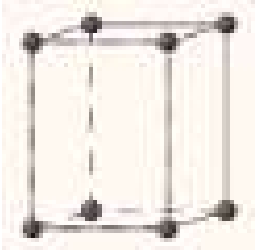


Primitive, P

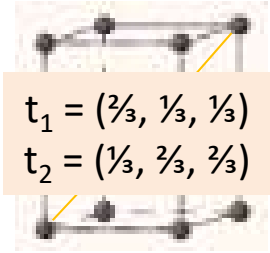


Body centered, I

Hexagonal lattice

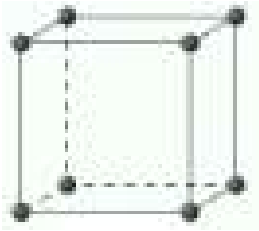


Primitive, P

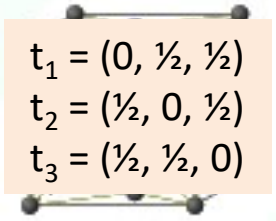


Rhombohedral, R

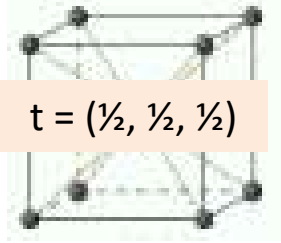
Cubic lattice



Primitive, P



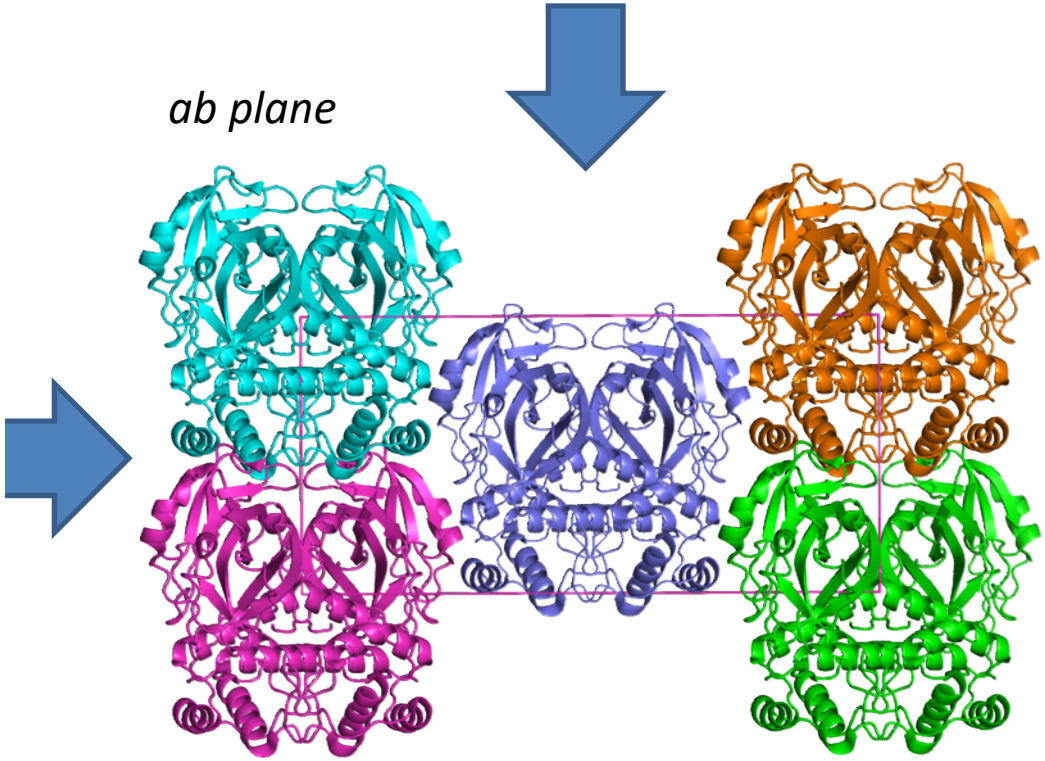
Face centered, F



Body centered, I

Case in point: monoclinic *C2* lattice

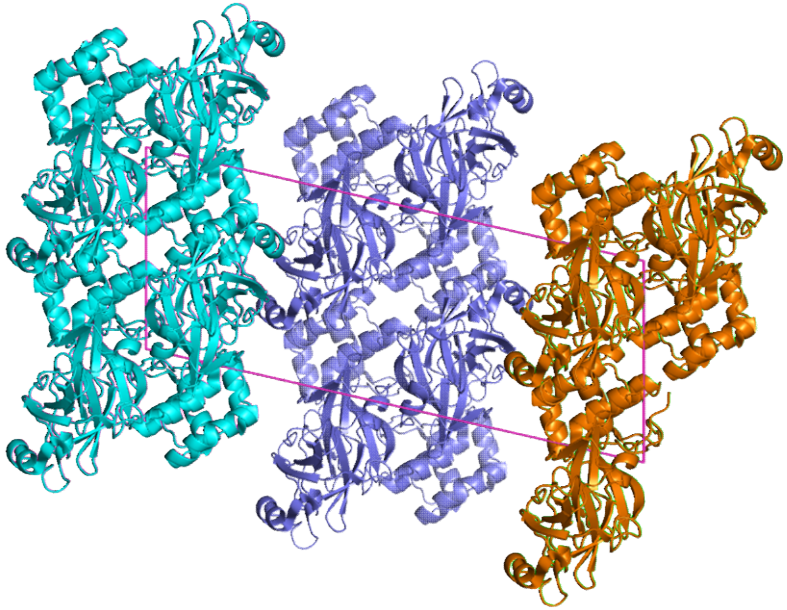
ab plane



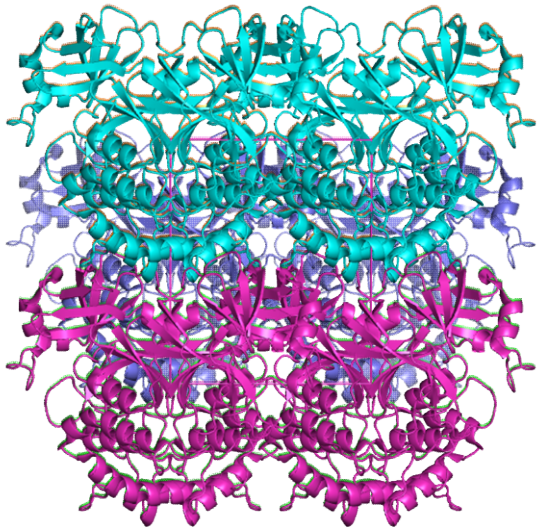
5RE9

Monoclinic:
a, *b*, *c*
$\alpha=\gamma=90°, \beta$

ac plane



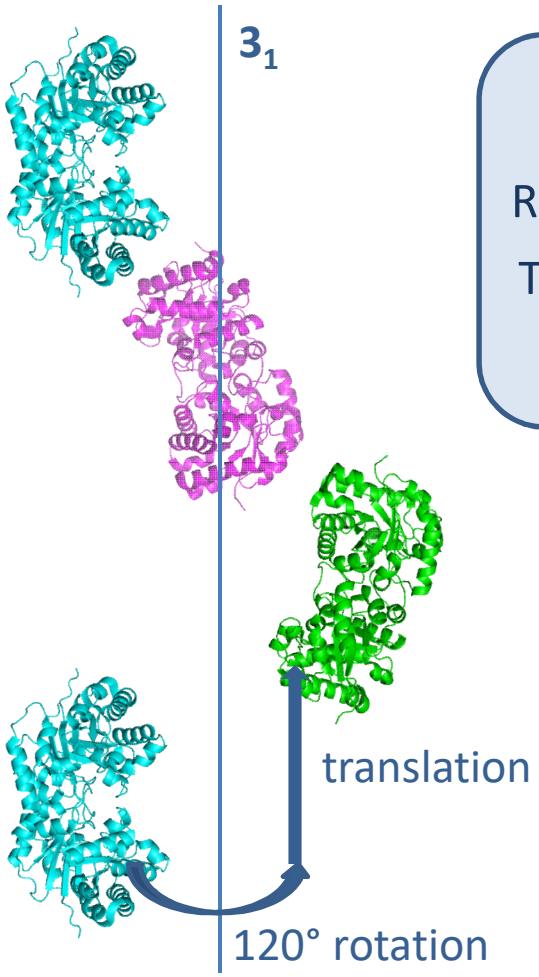
bc plane



Rotation + translation = screw axis

Besides the point group symmetry elements, additional symmetry elements are possible in crystal cells:

- ***screw axis*** *= rotation + translation*
- *glide planes = mirror + translation*

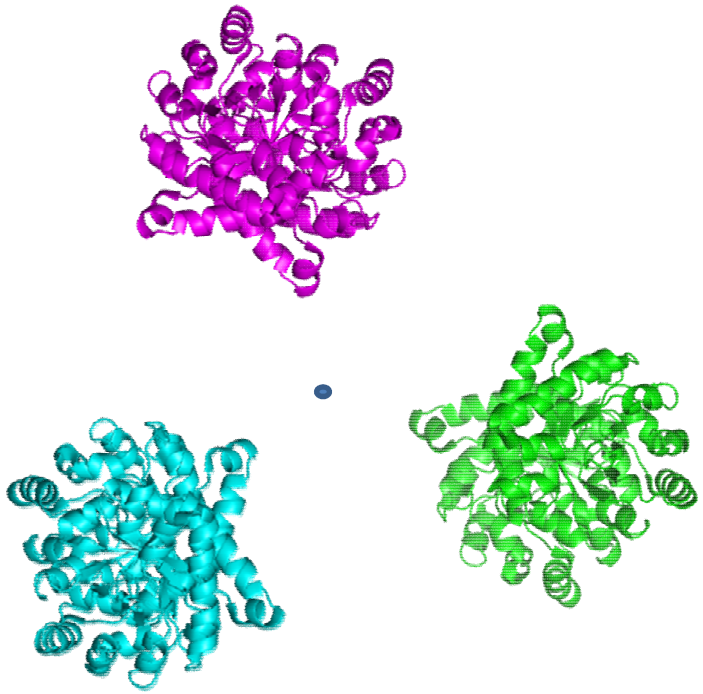


6R8H

Screw axis N_m

Rotation of order *N*

Translation of *m/N* of the unit cell



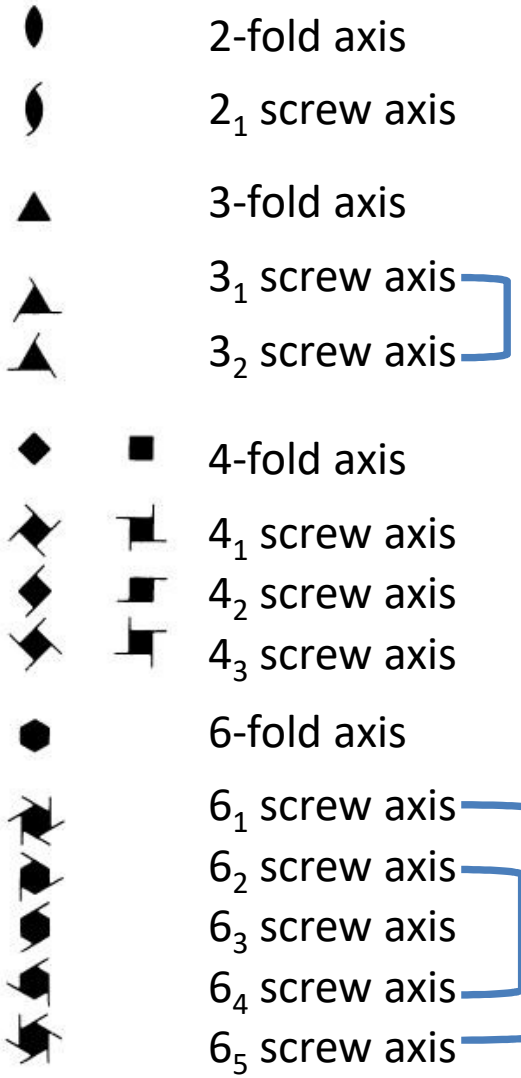
Screw axis 3_1

Rotation of order *3*

Translation of *1/3* of the unit cell

Symmetry operations

Symmetry operations are represented with specific figures when perpendicular to the plane, or arrows when parallel to the plane of the figure:



2-fold axis

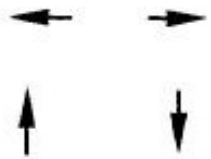
2_1 screw axis

3-fold axis

3_1 screw axis

3_2 screw axis

Rototranslations in opposite directions!



Rotation axes

4-fold axis

4_1 screw axis

4_2 screw axis

4_3 screw axis

6-fold axis

6_1 screw axis

6_2 screw axis

6_3 screw axis

6_4 screw axis

6_5 screw axis

Rototranslations in opposite directions!

Rototranslation axes

Fractional coordinates

Each symmetry operation can be represented by:

(1) a square matrix **S** and (2) a translation vector **t**

that transform a generic point ***x*** with coordinates (*x*,*y*,*z*) in its symmetric ***x'*** with coordinates (x',y',z'): $\boldsymbol{x'} = \boldsymbol{S} \cdot \boldsymbol{x} + \boldsymbol{t}$

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

The coordinate system of this equation, however, is not the orthogonal system in which we describe the atomic structure with distances in Å (i.e. in the pdb file).

The matrices **S** and vectors **t** are defined in the coordinate system of each specific lattice, i.e. having the *x*,*y*,*z* directions along the translational vectors that define the lattice. In addition, this coordinate system has **fractional coordinates**, that are calculated as fractions of the unit cell parameters in each direction.

Example: the coordinate system of a monoclinic lattice has 2 directions (x and z) that are not perpendicular, but form an angle equal to β.

Transformation: 4-fold axis along *c* direction of a tetragonal lattice

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

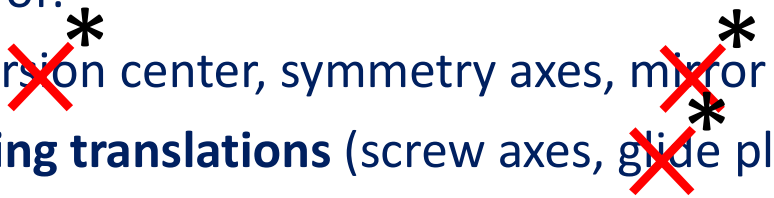
Transformation: 2_1 screw axis along the *b* direction

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ {}^{1}/_{2} \\ 0 \end{pmatrix}$$

Space groups

Space groups are a combination of:

- **point group operations** (inversion center, symmetry axes, mirror planes),
- **symmetry operations involving translations** (screw axes, glide planes),
- **centering translations** (base-, face-, body-centering),
- **lattice translations**.



A space group:

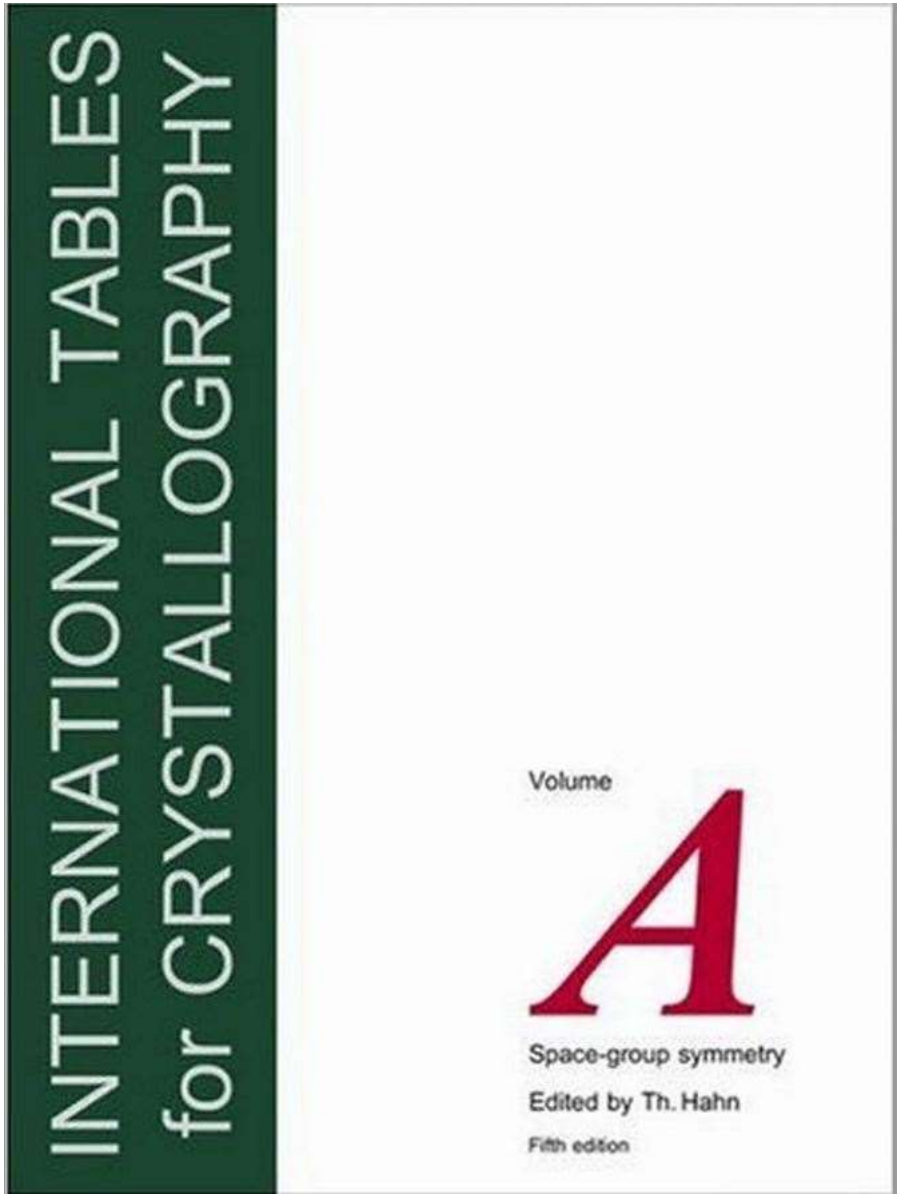
(1) is a **closed** group (combinations of its elements yield the identity transformation),

(2) contains the **identity** transformation,

(3) contains inverse transformations of its elements (**inversion** property),

(4) contains combinations of its elements (**associativity** property).

In 3D, 230 space groups,
but only 65 for protein structures*



International Tables of Crystallography Volume A

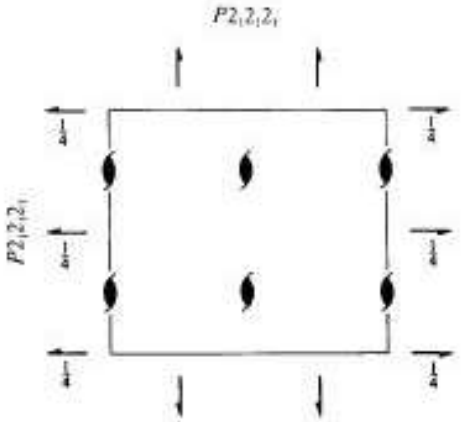
Volume containing information for each of the 230 space groups:

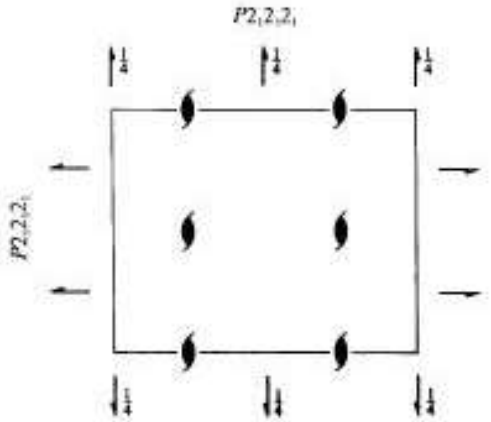
- Symmetry elements and their position
- Origin choice
- General positions and their multiplicity
- Asymmetric unit
- Systematic absences
- ...

$P2_12_12_1$ D_2^4 222 Orthorhombic

No. 19 $P2_12_12_1$ Patterson symmetry $Pmmm$

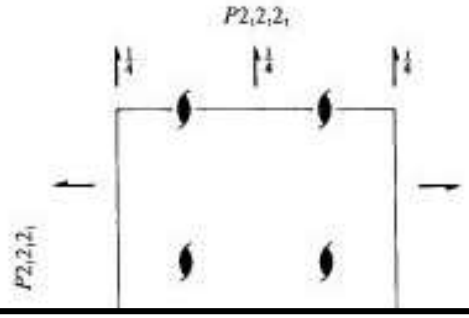
Name: type of lattice (*P*) and symmetry elements along the lattice main axes

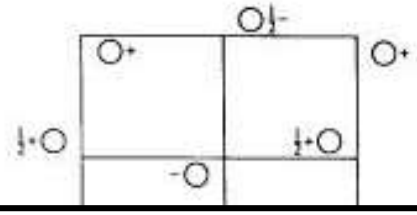




Crystal system and Patterson symmetry

Projections of the unit cell with symmetry elements





Origin at midpoint of three non-intersecting pairs of parallel 2_1 axes

Asymmetric unit $0 \le x \le \frac{1}{2}$; $0 \le y \le \frac{1}{2}$; $0 \le z \le 1$

Symmetry operations

(1) 1 (2) $2(0,0,\frac{1}{2})$ $\frac{1}{4},0,z$ (3) $2(0,\frac{1}{2},0)$ $0,y,\frac{1}{4}$

Origin position, asymmetric unit limits in fractional coordinates, symmetry operations

Generators selected (1); $t(1,0,0)$; $t(0,1,0)$; $t(0,0,1)$; (2); (3)

Positions

Multiplicity, Wyckoff letter, Site symmetry		Coordinates			
4	a 1	(1) x,y,z	(2) $x+\frac{1}{2},\bar{y},z+\frac{1}{2}$	(3) $\bar{x},y+\frac{1}{2},\bar{z}+\frac{1}{2}$	(4) $\bar{x}+\frac{1}{2},\bar{y}+\frac{1}{2},\bar{z}$

Symmetry equivalents for a generic point in fractional coordinates

Reflection conditions

General:

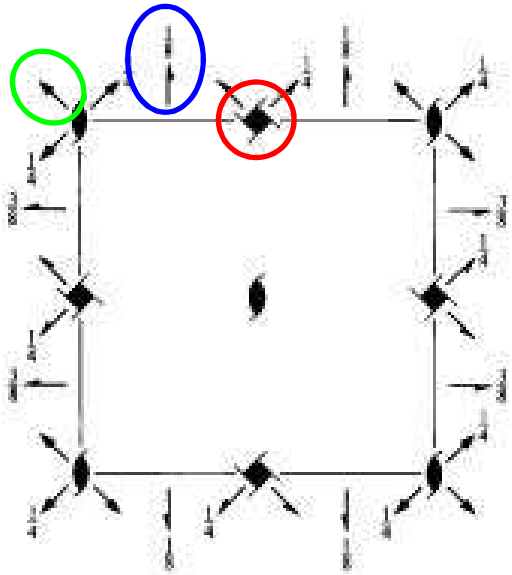
$h00$: $h = 2n$
$0k0$: $k = 2n$
$00l$: $l = 2n$

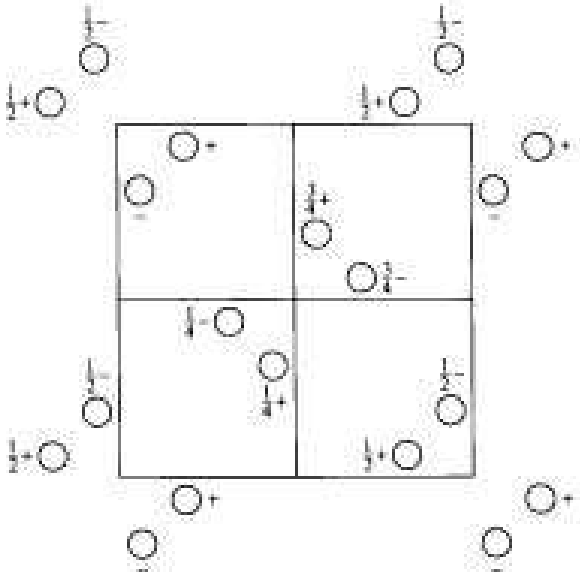
Reflections allowed by systematic absences

International Tables for Crystallography (2006). Vol. A, Space group 96, pp. 376–377.

$P4_32_12$ D_4^8 422 Tetragonal

No. 96 $P4_32_12$ Patterson symmetry $P4/mmm$





Origin on $2\,[1\,1\,0]$ at $2_1\,1\,(1,2)$

Asymmetric unit $0 \le x \le 1;\quad 0 \le y \le 1;\quad 0 \le z \le \frac{1}{8}$

Symmetry operations

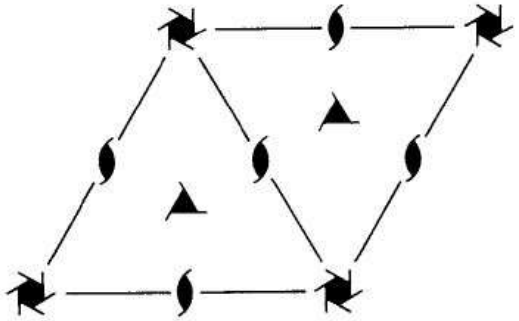
(1) 1
(2) $2(0,0,\frac{1}{2})$ $0,0,z$
(3) $4^+(0,0,\frac{1}{4})$ $0,\frac{1}{2},z$
(4) $4^-(0,0,\frac{1}{4})$ $\frac{1}{2},0,z$
(5) $2(0,\frac{1}{2},0)$ $\frac{1}{4},y,\frac{1}{8}$
(6) $2(\frac{1}{2},0,0)$ $x,\frac{1}{4},\frac{1}{8}$
(7) 2 $x,x,0$
(8) 2 $x,x,\frac{1}{4}$

Enantiomorphic space groups

Enantiomorphic space groups, with screw axis in opposite directions:

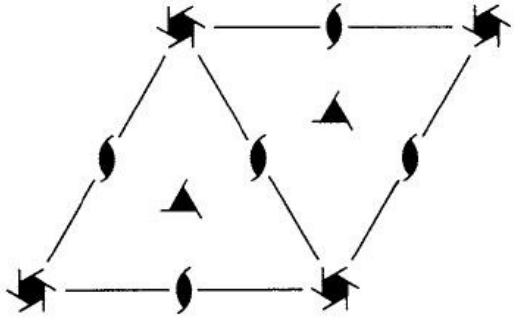
$P6_1$ C_6^2

No. 169 $P6_1$

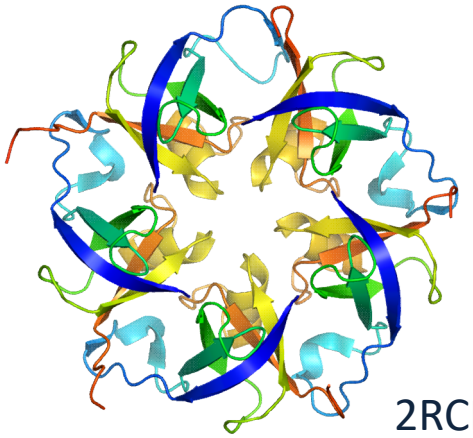


$P6_5$

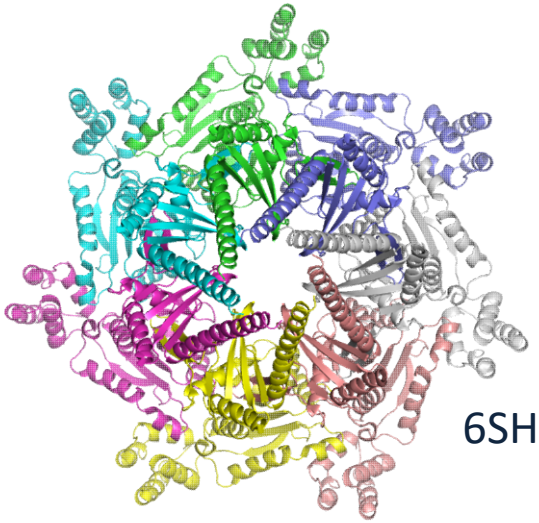
No. 170



Enantiomorphic pairs:
(tetragonal) *P* 4_1 and *P* 4_3, *P* 4_1 *2 2* and *P* 4_3 *2 2*, *P* 4_1 2_1 *2* and *P* 4_3 2_1 *2*, *I* 4_3 *2 2*, ...
(hexagonal) *P* 3_1 and *P* 3_2 , *P* 3_1 *2 1* and *P* 3_2 *2 1* , *P* 3_1 *1 2* and *P* 3_2 *1 2*, *P* 6_1 and *P* 6_5 , *P* 6_1 *2 2* and *P* 6_5 *2 2* , *P* 6_2 and *P* 6_4 , *P* 6_2 *2 2* and *P* 6_4 *2 2*
(cubic) *P* 4_1 *3 2* and *P* 4_3 *3 2.*



2RCF



6SH4

Non-crystallographic symmetry

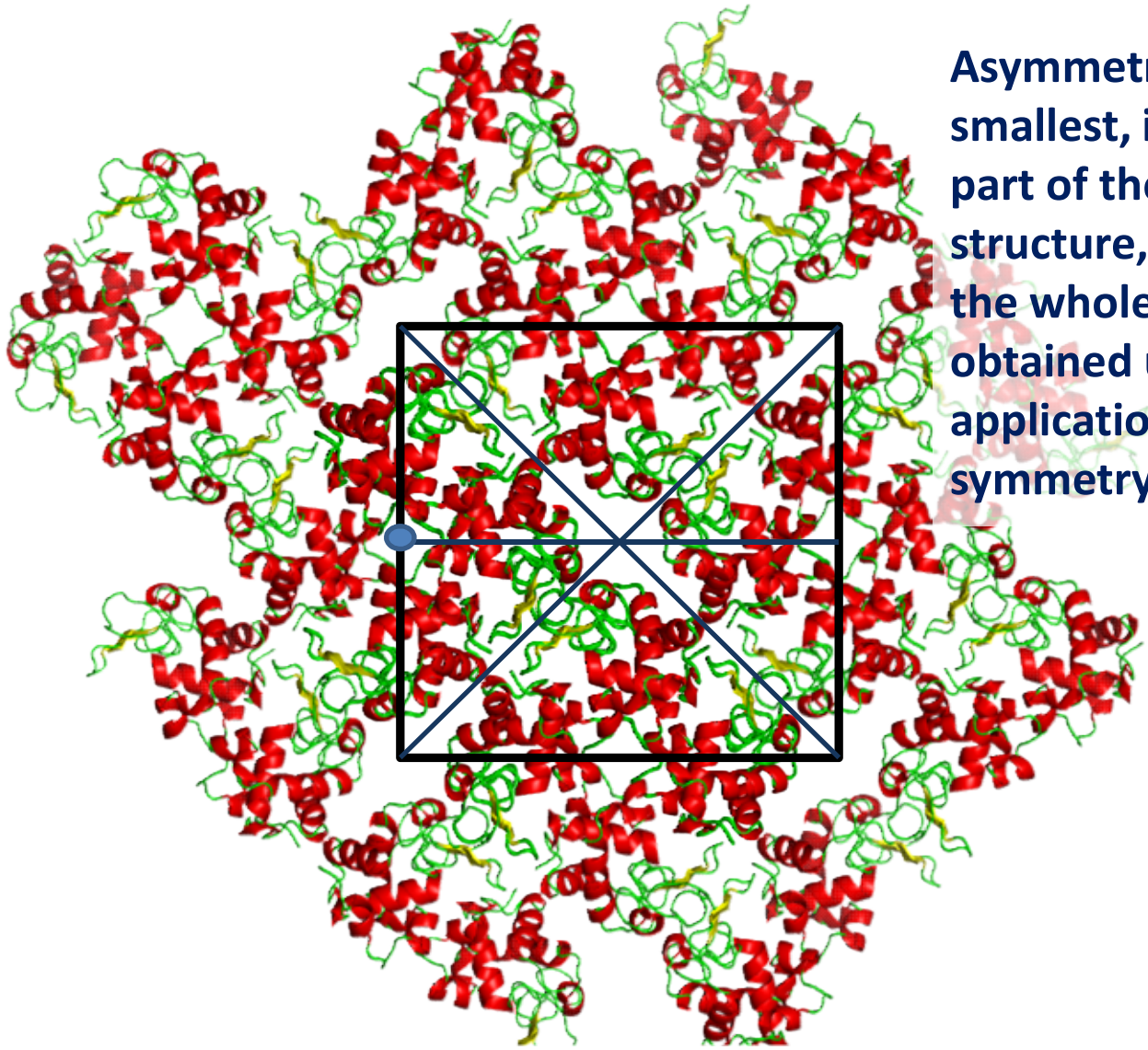
Besides the symmetry elements of the space group, other symmetry elements may occur in a crystal structure, including symmetry elements not allowed by the lattice periodicity (e.g. 5-fold axes, 7-fold axes, ...).

They are called Non-Crystallographic Symmetry (NCS).

Two protein chains related by NCS are crystallographically independent and both belong to the asymmetric unit. Usually, NCS is not perfect.

Asymmetric unit

Lysozyme (pdb: 193L):
tetragonal lattice, space group $P\,4_3\,2_1\,2$, unit cell a = 78.54 Å, c = 37.77 Å



Asymmetric units: smallest, independent part of the crystals structure, from which the whole structure is obtained upon application of the symmetry operations

pdb file

The pdb file contains:

- atomic Cartesian coordinates only of the asymmetric unit, in Å and in an orthogonal system (**ATOM** cards),
- lattice informations, i.e. unit cell dimensions (*a*, *b*, *c*, α, β, γ) in Å and °, and space group symbol and number (**CRYST** card)
- transformation matrix from Cartesian (orthogonal) coordinates in Å to fractional coordinates in the specific crystallographic system (**SCALE** card)

```
SITE     1 AC3  5 GLY A 142  TRP A 143  ASN A 144  ASN A 145
SITE     2 AC3  5 GLN A 148
CRYST1   76.465   98.869   79.516  90.00  93.49  90.00 C 1 2 1       4
ORIGX1      1.000000  0.000000  0.000000        0.00000
ORIGX2      0.000000  1.000000  0.000000        0.00000
ORIGX3      0.000000  0.000000  1.000000        0.00000
SCALE1      0.013078  0.000000  0.000798        0.00000
SCALE2      0.000000  0.010114  0.000000        0.00000
SCALE3      0.000000  0.000000  0.012599        0.00000
ATOM      1  N   SER A   6     -28.148   4.590  -7.800  1.00113.33           N
ATOM      2  CA  SER A   6     -26.785   4.701  -7.207  1.00115.91           C
ATOM      3  C   SER A   6     -26.419   6.159  -6.970  1.00114.54           C
ATOM      4  O   SER A   6     -26.103   6.886  -7.915  1.00110.10           O
ATOM      5  CB  SER A   6     -25.744   4.037  -8.113  1.00116.34           C
```

How to apply crystal symmetry with Pymol?

When opening a pdb file with the software Pymol, only the asymmetric unit appears. However, often the asymmetric unit is NOT the biologically active assembly. In addition, analysis of packing may be important to detect crystal artifacts.

The whole lattice can be reconstructed from (1) the contents of the asymmetric unit, (2) the unit cell parameters, and (3) the space group (i. e., information about the symmetry of the unit cell).

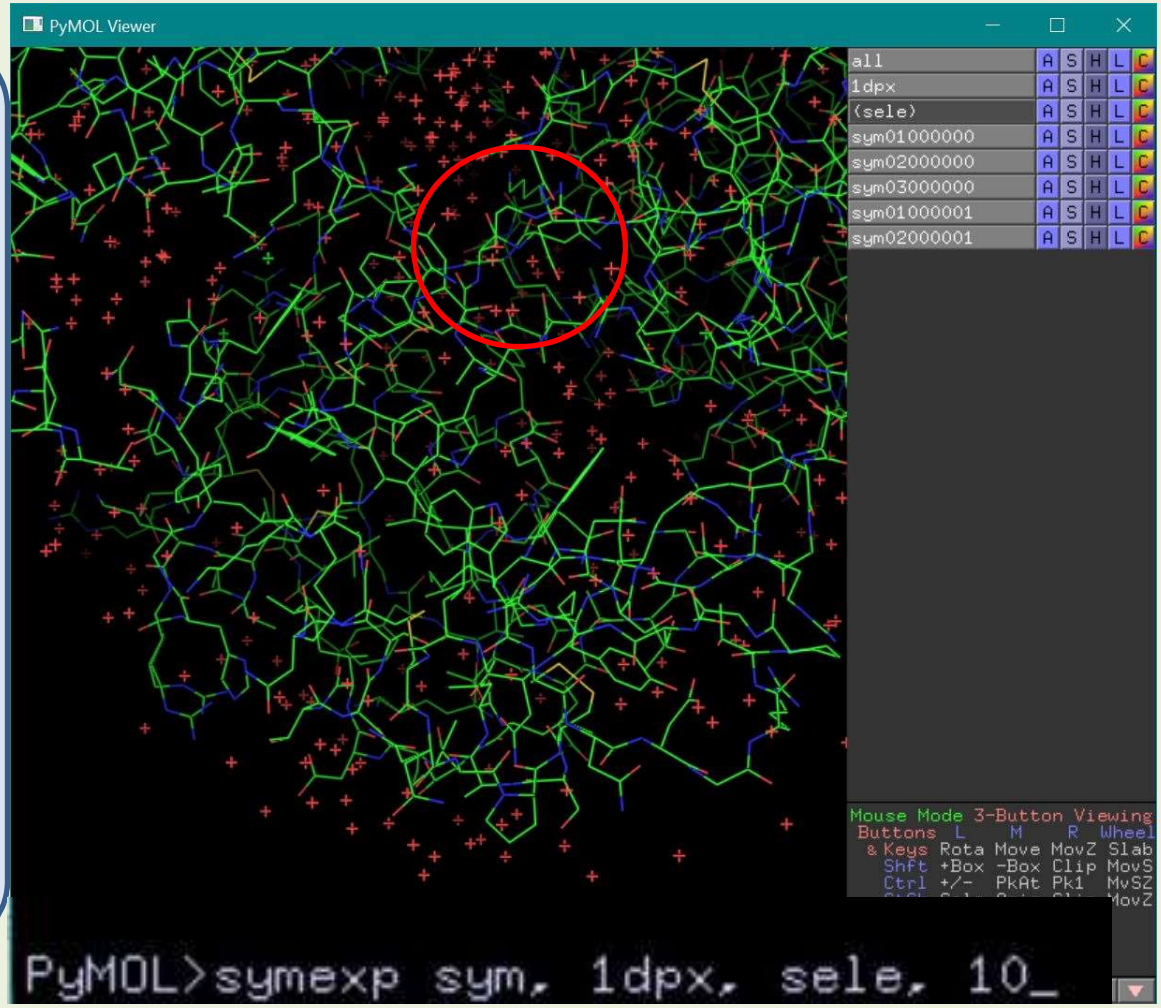
```
CRYST1   76.465   98.869   79.516  90.00  93.49  90.00 C 1 2 1       4
ORIGX1      1.000000  0.000000  0.000000        0.00000
ORIGX2      0.000000  1.000000  0.000000        0.00000
ORIGX3      0.000000  0.000000  1.000000        0.00000
SCALE1      0.013078  0.000000  0.000798        0.00000
SCALE2      0.000000  0.010114  0.000000        0.00000
SCALE3      0.000000  0.000000  0.012599        0.00000
ATOM      1  N   SER A   6     -28.148   4.590  -7.800  1.00113.33           N
ATOM      2  CA  SER A   6     -26.785   4.701  -7.207  1.00115.91           C
ATOM      3  C   SER A   6     -26.419   6.159  -6.970  1.00114.54           C
ATOM      4  O   SER A   6     -26.103   6.886  -7.915  1.00110.10           O
ATOM      5  CB  SER A   6     -25.744   4.037  -8.113  1.00116.34           C
ATOM      6  OG  SER A   6     -26.087   2.689  -8.379  1.00121.40           O
ATOM      7  N   SER A   7     -26.458   6.572  -5.703  1.00116.37           N
ATOM      8  CA  SER A   7     -26.153   7.952  -5.306  1.00112.95           C
ATOM      9  C   SER A   7     -24.758   8.391  -5.751  1.00109.79           C
ATOM     10  O   SER A   7     -24.497   9.586  -5.902  1.00104.20           O
ATOM     11  CB  SER A   7     -26.313   8.125  -3.794  1.00113.79           C
ATOM     12  OG  SER A   7     -25.474   7.228  -3.088  1.00117.64           O
```

How to apply crystal symmetry with Pymol?

The command *symexp* instructs the program to retrieve the information requires and show on the screen the symmetric molecules within a certain distance from a center.

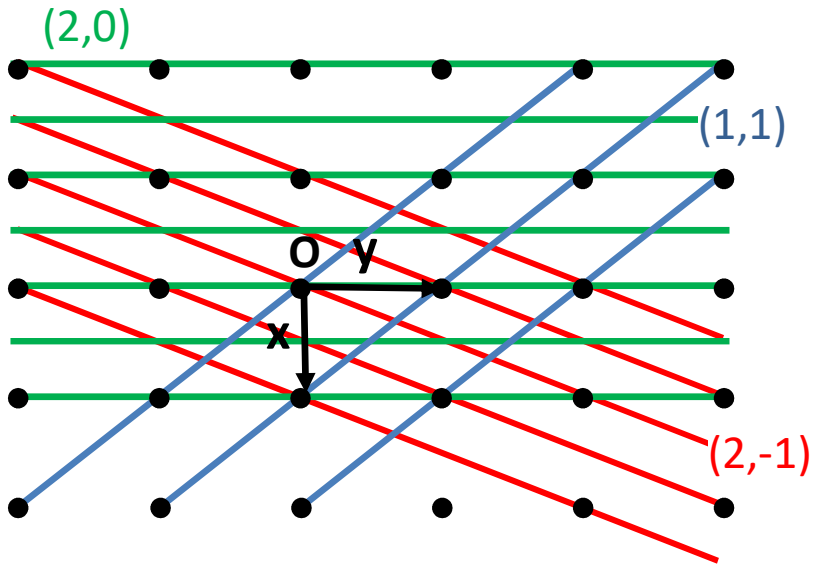
1) Open pdb file (1dpx.pdb, lysozyme)
2) Select and area of the structure for which crystal contacts should be analyzed (Arg residues on the surface)
3) Create a sphere with **radius n** (10) of the symmetric molecules around the **selected position** (center of sphere = sele) as new **objects** (name = sym), according to the symmetry elements present in the **original pdb file** (object name = 1dpx). Comand:

```
symexp sym, 1dpx, sele, 10
```



Miller indices

Crystal lattice: described in the real space (coordinates *x*,*y*,*z*)



In the real space of the crystal lattice, we can define **families of parallel planes**.

To identify this planes, we can use the **Miller indices (h,k)**:

starting from the origin of the lattice and moving in a lattice direction, we can count the number of planes until the next node.

For the blue set of planes: the first Miller index (in *x* direction) is 1, the second is 1.

For the red set of planes: the first Miller index is 2, the second is -1.

For the green set of planes: the first Miller index is 2, the second is 0.

Which of the families has the shorter distance d between planes?

The distance between planes depends on the Miller indices: planes with higher Miller indices have shorter distances.

From the real to the reciprocal lattice

The **reciprocal lattice** is a mathematical construction, but it is useful to describe diffraction phenomena.

Reciprocal lattice: described in the reciprocal space by the Miller indices (*h*,*k*,*l*).

The origin of the lattice is common to the real lattice.

Reciprocal lattice dimensions: $a^* = {}^1/_a\,, b^* = {}^1/_b\,, c^* = {}^1/_c$

Direction of the reciprocal lattice base vectors: $\boldsymbol{a}^* \perp \mathbf{b}$, $\boldsymbol{a}^* \perp \boldsymbol{c}$

$\boldsymbol{b}^* \perp \boldsymbol{a}$, $\boldsymbol{b}^* \perp \boldsymbol{c}$

$\boldsymbol{c}^* \perp \boldsymbol{a}$, $\boldsymbol{c}^* \perp \boldsymbol{b}$

A family of planes in real space can be described by a vector $\boldsymbol{d}^*$ in reciprocal space.

The vector $\boldsymbol{d}^*$ is obtained by the combination of the Miller indices and the base vectors of the lattice: $\boldsymbol{d}^* = h\boldsymbol{a}^* + k\boldsymbol{b}^* + l\boldsymbol{c}^*$.

The modulus of $\boldsymbol{d}^*$ is: $d^* = {}^1/_d$, with d distance between planes of the family in the real space.

(For an orthogonal system: $d_{hkl} = \sqrt{\frac{a^2}{h^2} + \frac{b^2}{k^2} + \frac{c^2}{l^2}}$)