Kristina Djinovic-Carugo

Structural Biology Laboratory

Sincrotrone Trieste

e-mail: kristina.djinovic@elettra.trieste.it

tel: +39 040 375 8059

Literature

- **Crystallographic methods and protocols, Methods in Molecular Biology vol 56, edited by C. Jones, B. Mulloy, M. R. Sandersson (1996)**
- Marcomolecular crystallography, Methods in Enzymology, vol 276-277, edited by C.W. Carter and R.M. Sweet (1997)

Literature

- X-ray structure determination (A practical guide), G.H Stout and L.H. Jensen) (1989)
- The principles of Protein X-ray crystallograph, J. Drenth (1999)
- Crystal structure analysis - A primer, J.P. Glusker, K.N. Trueblood (1985)

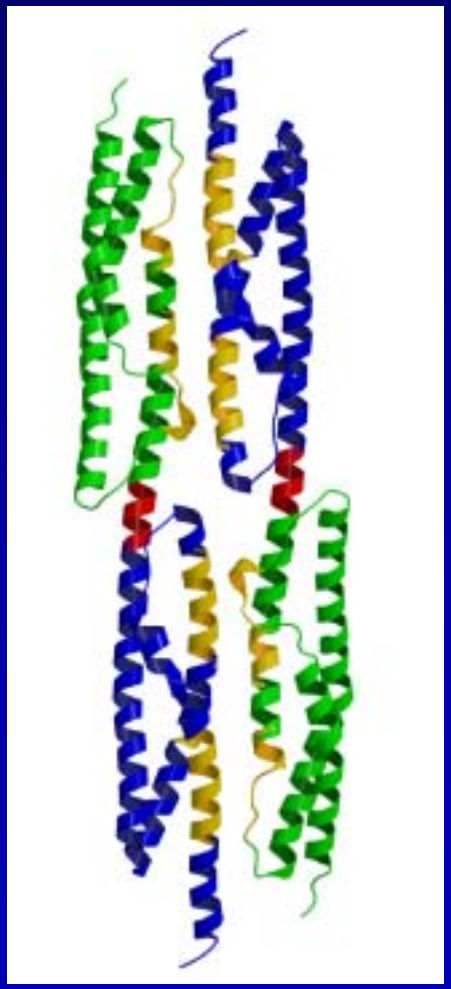
Outline

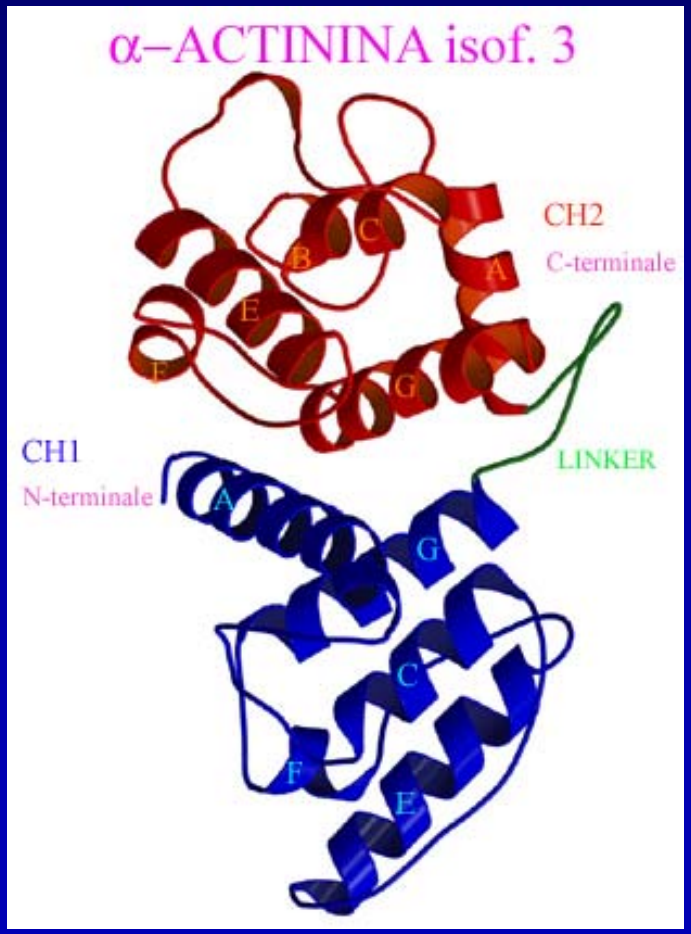
- X-ray sources, why SR
- Crystallisation of proteins
- Structure determination methods
  - multiple isomorphous replacement (MIR)
  - multi-wavelenght anomalous dispersion(MAD)
  - molecular replacement (MR) (only mentioned here)

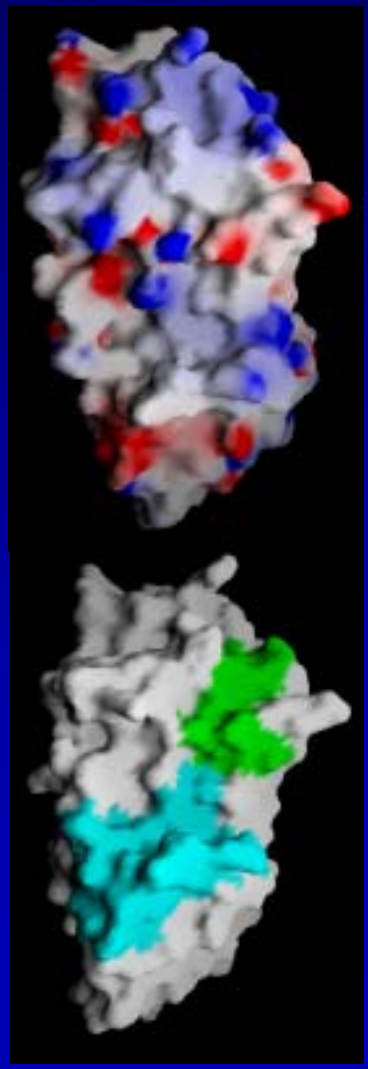
From protein in solution to its 3D structure

- Crystallisation
- X-ray diffraction
- Solution of phase problem (MIR, MAD, MR)
- Building 3D model and refinement

Final result



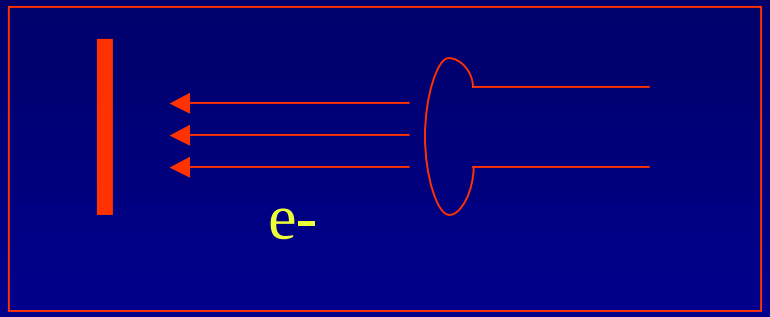




Sources of X-rays

- Wavelength: 1 Å (10^{-10} m)
- Generators - sealed tube or rotating anode
  - targets Cu, Mo
- Synchrotron radiation

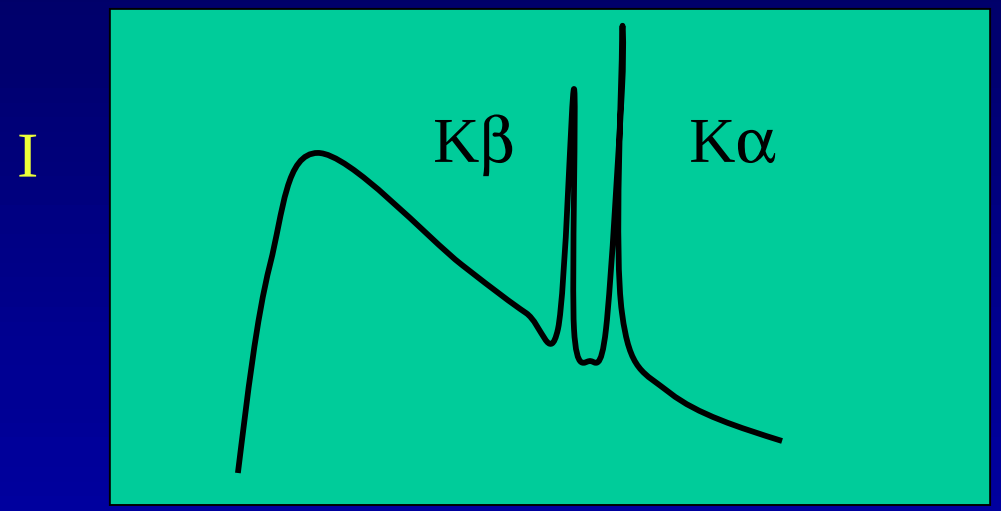
Generators



In high vacuum; working potential and current: 50 mA 60 kV

- Applied power is limited
- Limited intensity of X-rays
  - rotating anode

Generators



Transitions L - K: Kα1, Kα2

Transitions M - K: Kβ1, Kβ2

Synchrotron radiation

- High flux
- High collimation/brilliance
- Tunable
- Time structuire of beam allows time resolved experiments

Synchrotron radiation for structural biology

- Flux: macromolecular crystals are typically weak diffractiors
- High collimation/brilliance: small crystals
- Tunability: solution of phase problem with MAD or SAS
- Time structure of beam: time resolved experiments for study of enzymatic reactions

Crystallisation

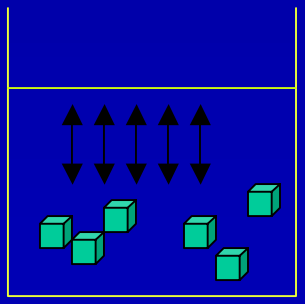
- First Protein crystals
  - 1840: Haemoglobim from earth worm
  - grown from blood between two cover slides
- Modern approach
  - techniques of recombinant DNA
  - pure protein sample

Crystallisation

- Solubility of protein sample depends on
  - pH
  - T
  - ionic strength
- Solubility is precisely defined by the above conditions

Crystallisation - saturated solution

- Protein in solid phase is in equilibrium with the solution
  - Non of the phases is gaining in concentration
  - Thermodynamic force is maintaining the system in equilibrium

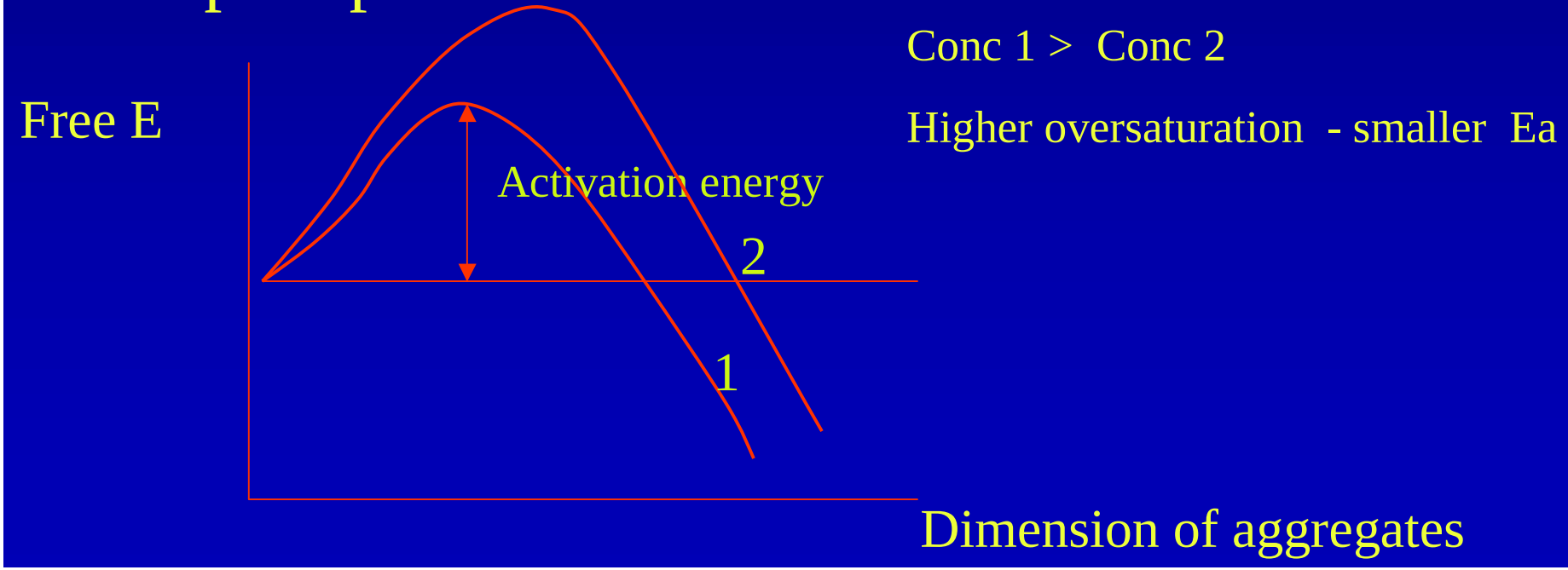


Crystallisation - oversaturated solution

- Concentration of solute is higher that its solubility at given conditions
- Chemical system is not in equilibrium
- Thermodynamics is forcing the system towqrds equilibrium
- Crystals grow from OVERSATURATED solutions

Crystallisation from oversaturated solutions

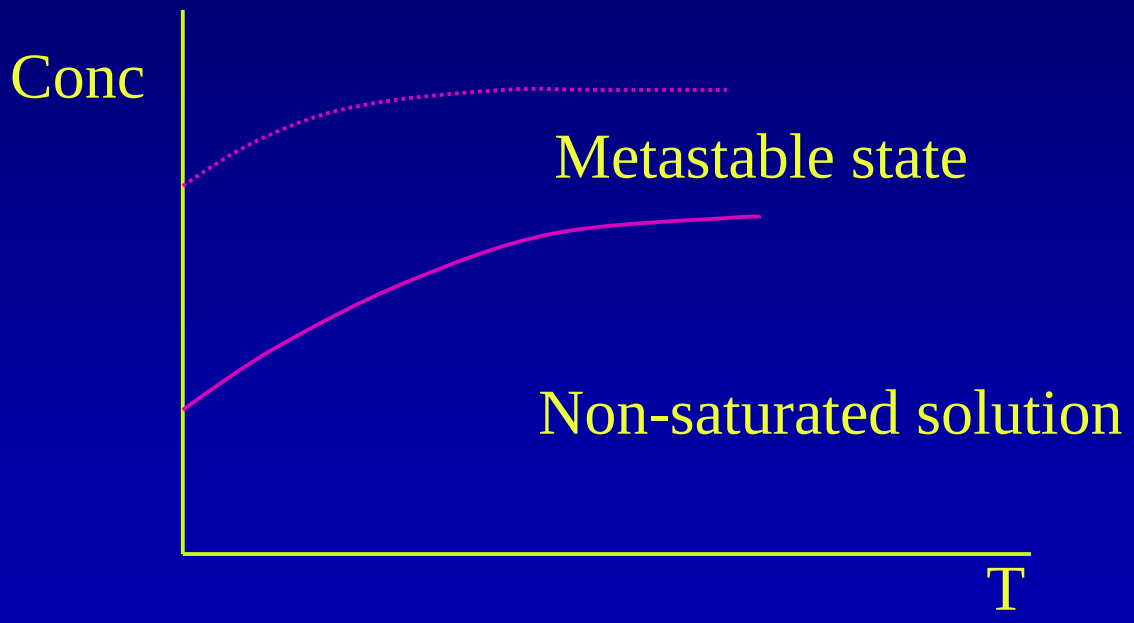
- Oversaturated solution can exist because:
- activation energy is needed for start of precipitate formation



Conc 1 > Conc 2

Higher oversaturation - smaller Ea

Crystallisation - Phase diagram



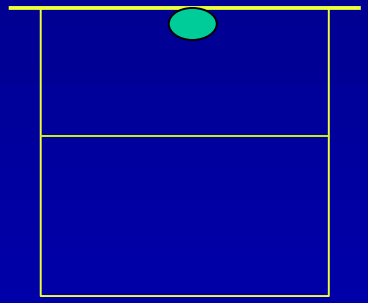
Similar diagram for Conc/pH, Conc/ionic strength

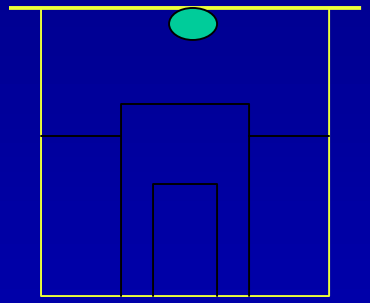
Crystallisation strategies

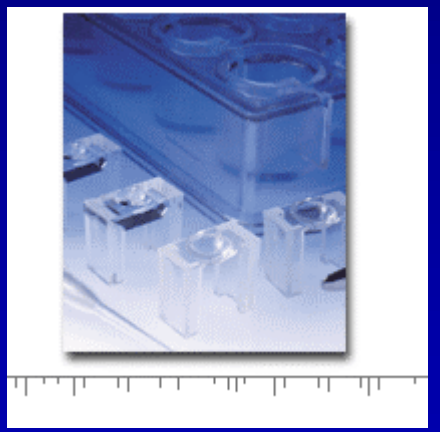
- high number of parameters - multidimensional space
  - systematic search
  - sparse-matrix
    - investigation of 100-200 different conditions at different temperatures (20C, 4C)

Crystallisation - vapour diffusion

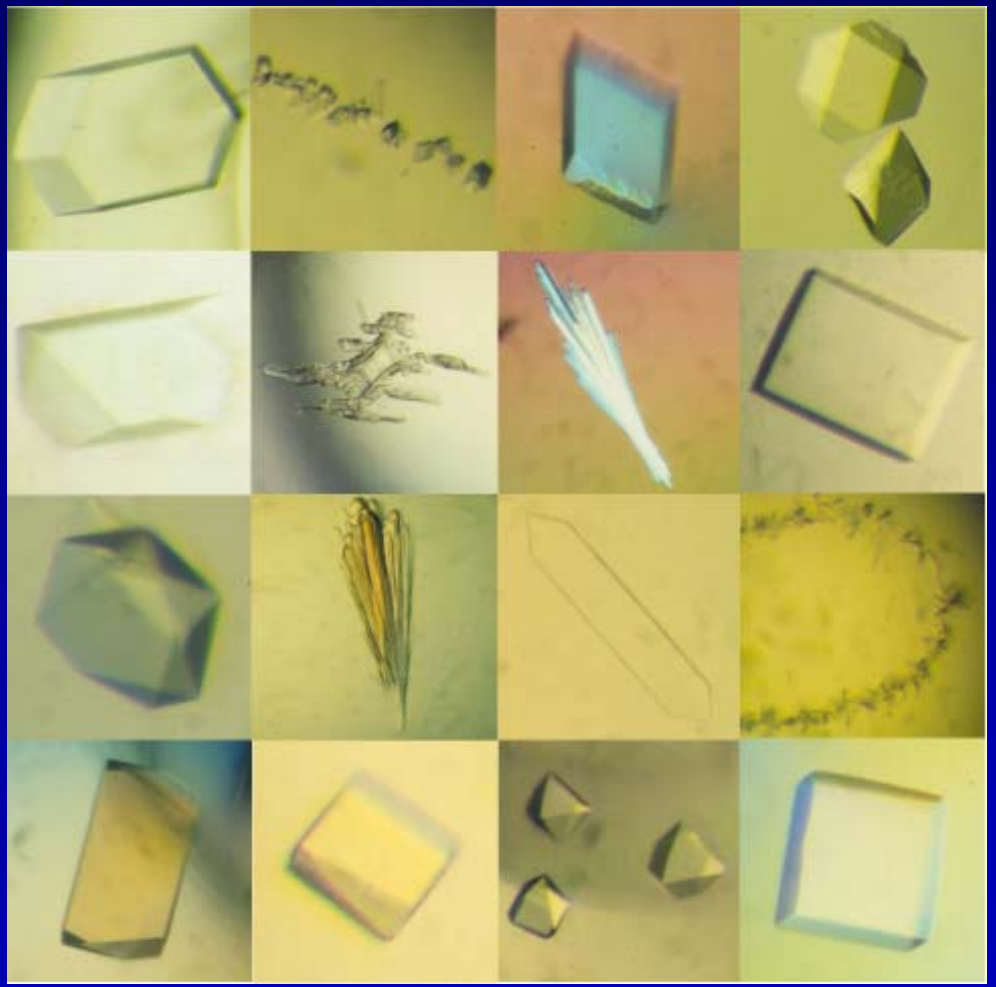
- Hanging drop
- Sitting drop







Crystal Gallery



MIR

- **Multiple isomorphous replacement**
- measure apmpludes of diffracted beams
- determination of phases indirectly
- MIR is very often used *ab initio* method for determination of phases for macromolecules

MIR

- 1954 Perutz *et al*. solve structure of heamoglobin
- For macromolecules: proteins, DNA, RNA
- **Basic principle**: binding of heavy atom to (Hg, Pt, U, Pb, etc) the macromolecule which does not change structure upon this treatment

MIR -isomorphous derivative

- ideal isomorphous derivative: the only difference between the native and derivative is the bound heavy metals atoms(s)
- Approach: ‘soak’ of protein crystal in solution containing heavy metals

MIR -isomorphous derivative

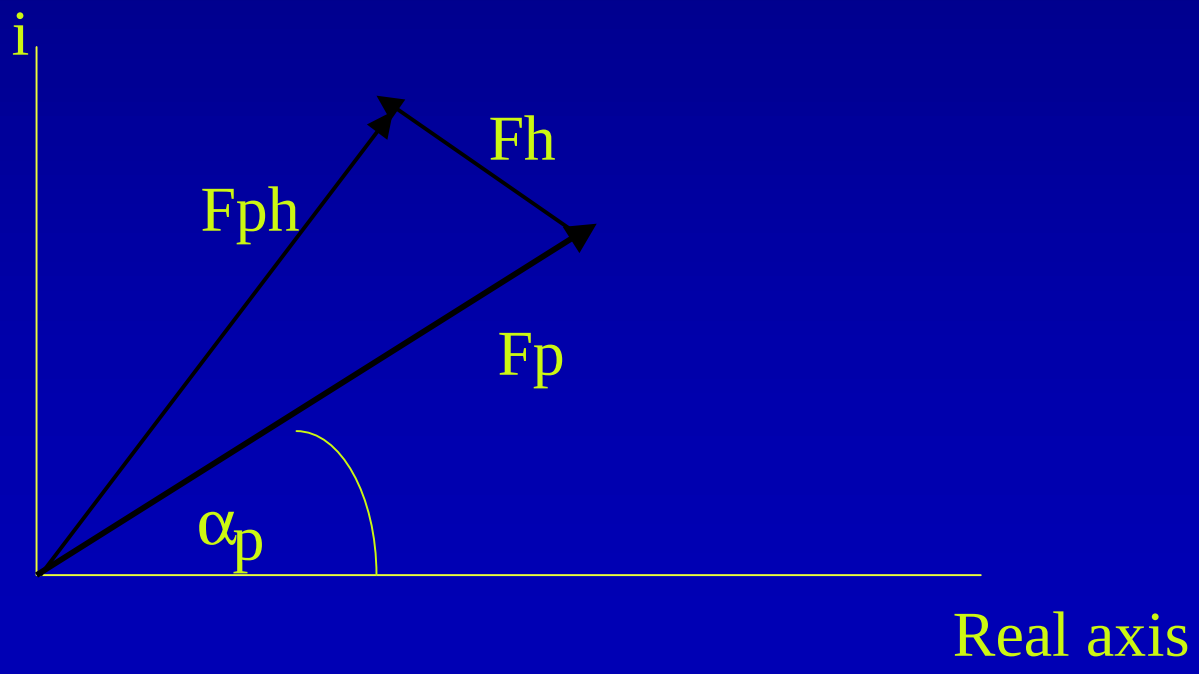
- presence of heavy atom(s) introduces differece to the diffraction pattern with respect to the diffraction pattern of the native crystal
  - differnecs are in INTENSITIES of diffcìratde x-rays
- when heavy atom binds isomorphously the dirrerenced between the two samples are

MIR - approach

- X-ray data collection of native crystals
- preparation of heavy atom derivative
- X-ray data collection of derivative
- determination of heavy atom positions (Patterson)
- refinement of heavy atom coordinates and phase calc.
- calc. Of electron density
- building and refinement of atomic model

MIR - why does it work

- Fph, Fp, Fh: structure factors
  - derivative, protein, heavy atom



MIR - why does it work

- Fph = Fp + Fh
- knowing the position of heavy atom(s) we can calculate:
  - amplitude Fh
  - faze αh
- $F = A + iB$; $A(hkl) = \Sigma f_i \cos[2\pi(hx_i + ky_i + lz_i)]$; $B(hkl) = \Sigma f_i \sin[2\pi(hx_i + ky_i + lz_i)]$

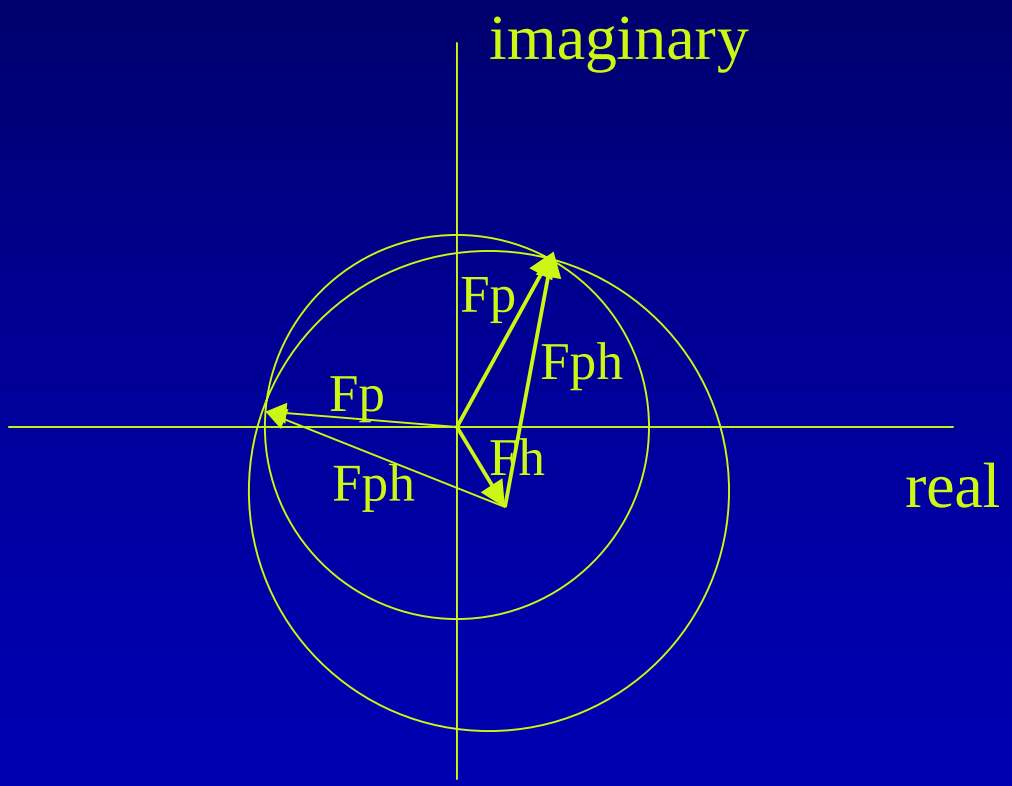
MIR - how does it work

- Measured:
  - amplitudes Fph
  - amplitudes Fp
- Calculated:
  - heavy atom positions and from them
    - amplitude: Fh
    - faze αh

MIR - how does it work

- $Fph^2 = Fp^2 + Fh^2 + 2FpFhcos(\alpha p - \alpha h)$
  - amplitudes: Fp, Fh, Fph
  - fazes: αh, αp
- $\alpha p = \alpha h + \arccos[(Fph^2 - Fp^2 - Fh^2)/2FpFh]$
- $\alpha p = \alpha h$ +/- α’

MIR - how does it work

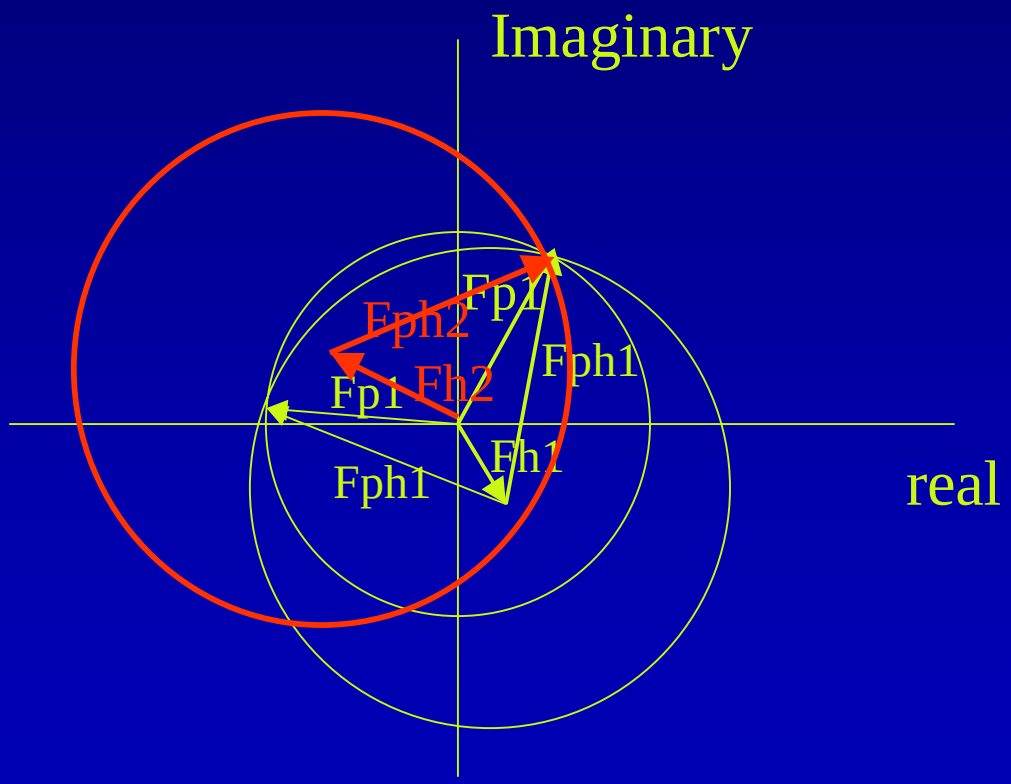


MIR - how does it work

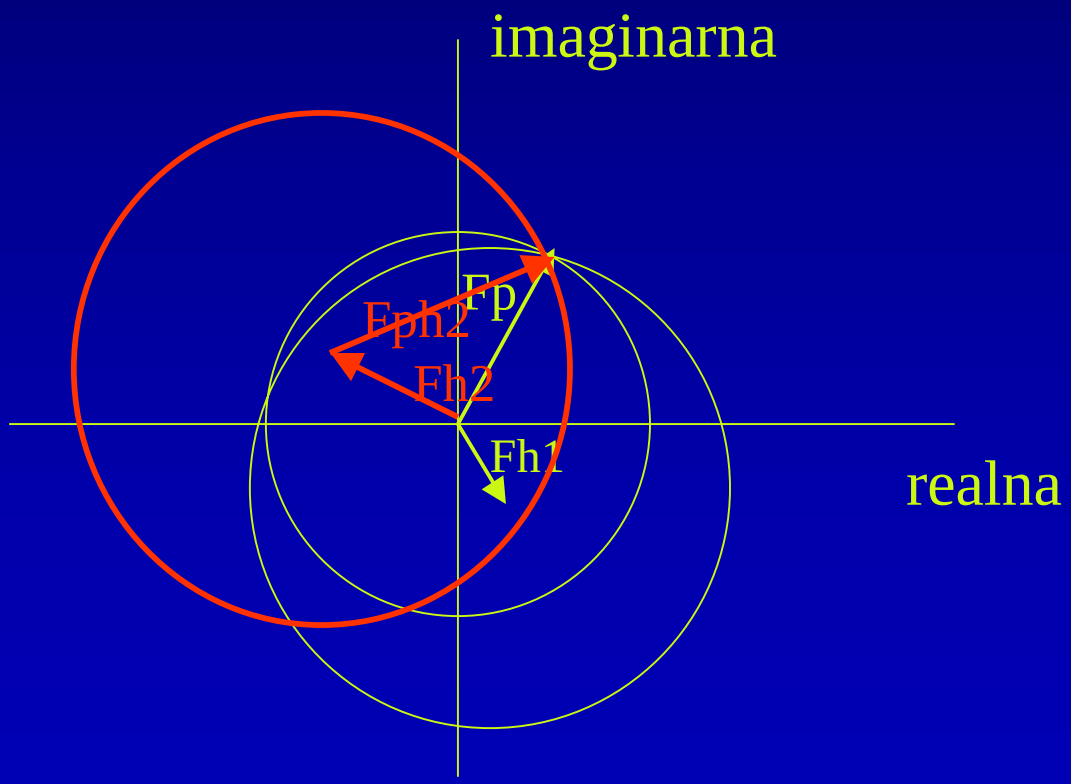
- we can not determine the phase un-ambigously with only one derivative:
- $\alpha pa = \alpha h + \alpha'$
- $\alpha pb = \alpha h - \alpha'$
- αpa in αpb represent structures A and B
  - A is a mirror image of B
- proteins are chiral molecules (L-amino amino acids)

MIR - how does it work

- we need at least two derivatives for correct determination of phases



MIR - how does it work



MIR - how does it work

- having determined αp for each diffracted beam hkl we can calculate electron density

$$\rho(xyz) = (1/V_c)\Sigma\Sigma\Sigma Fp(hkl)\exp[-2\pi i(hx + ky + lz)]$$

- $Fp = Fp \exp(i\alpha p) = Fp \cos\alpha p + iFp \sin\alpha p$

MIR - often used havy atoms

- Pt, Au, Hg, Pb, Th, U, Re, Os, Ir, Xe
- Pd, Ag (small atomic number) - for small proteins
- J - iodinated tyrosins, modified nucleic acid bases (J, Br)
- lantanides (La-Lu) - can substitute Mg or Ca
- noble gasses (Xe, Kr

Determination of heavy atom positions:Patterson

- Patterson 1934
- Fourier synthesis with
  - coefficients: AMPLITUDEs^2
  - fazes: 0.0
- $P(u,v,w) = (1/V_c)\ \Sigma\Sigma\Sigma|F(hkl)|^2 \exp[-2\pi i(hu + kv + lw)]$
- $P(u,v,w) = (1/V_c)\ \Sigma\Sigma\Sigma|F(hkl)|^2 \cos[2\ \pi(hu + kv + lw\ )]$
- Convolution of electron density
  - $P(u,v,w) = \rho(xyz)\ \rho(x+u,y+v,z+w)dxdydz$

Patterson function

- has N^2 -N maxima + maximum at the origin
- distance of maxima from the origin depends on the length of the inter-atomic vector
- intensity of maxima depends on the priduct of the atomic numbers of t atoms i and j: $Z_i Z_j$

Patterson function

- Symmetry: centrosymmetric
  - vectors AB, BA
- possible only 24 space groups: loss of information when going from F(hkl) to |F(hkl)| without using phase

Patterson function with one heavy atomom

- $P2_1$ space group
- x,y,z; -x, y+0.5, -z
- Maxima in Pattreson maps corresponding to the vector between two symmetry related atoms will be at:
  - u = x - (-x) = 2x;
  - v = y - (y + 0.5) = -0.5 ;
  - w = z - (-z) = 2z

Patterson function with one heavy atomom

- u = 0.25; y = 0.5; w = 0.44
- x = u/2 = 0.25/2 = 0.125
- z = w/2 = 0.44/2 = 0.22
- Heavy atom position:
  - x = 0.125
  - y = anything
  - z = 0.22

MIR - determination of heavy atom positions

- $P(u,v,w) = (1/V_c)\ \Sigma\Sigma\Sigma|F(hkl)|^2 \exp[-2\pi i(hu + kv + lw)]$
- $P(u,v,w) = (1/V_c)\ \Sigma\Sigma\Sigma|F(hkl)|^2 \cos[2\ \pi(hu + kv + lw\ )]$
- difference Pattersonova method
- Coefficients:
- $(|\ Fp\ |\ -\ |\ Fph\ |)^{\ 2}$
- most intense maxima will correspond to interatomic vectors bewteen heavy atoms

MIR - determination of heavy atom positions

- atom 1: 0.25 0.11 0.32
- atom 2: 0.1 0.35 0.15
- maximum v Pattersonovi map at:
- 0.25 - 0.1; 0.11-0.35; 0.32-0.15
- 0.15 -0.24 0.17

MIR - final step - calc. Of electron density

- having determined αp for each diffracted beam hkl we can calculate electron density

$$\rho(xyz) = (1/V_c)\Sigma\Sigma\Sigma Fp(hkl)\exp[-2\pi i(hx + ky + lz)]$$

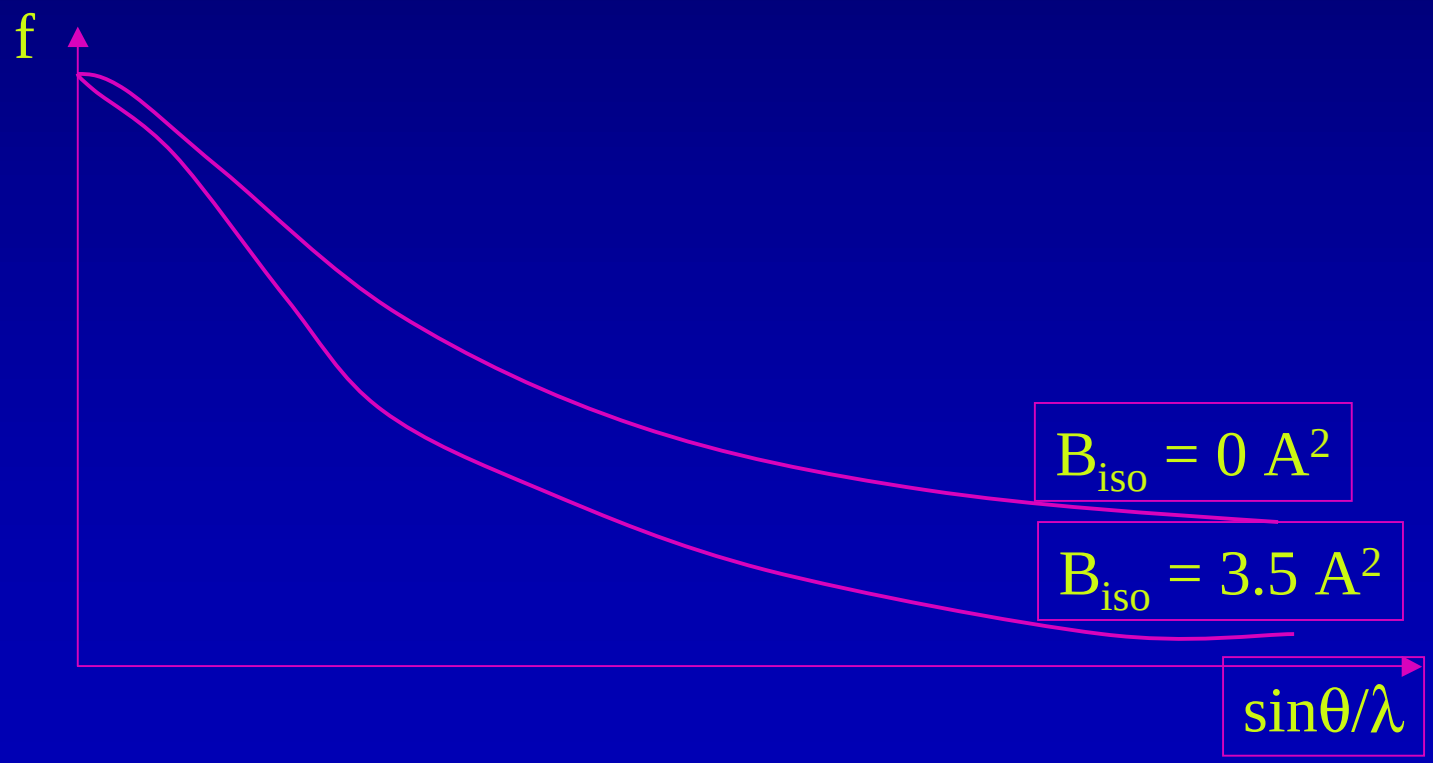
- $Fp = Fp \exp(i\alpha p) = Fp \cos\alpha p + iFp \sin\alpha p$

Structure, scattering factor

- structure factor F(h) is a vector sum of the contribution to that reflection of diffracted amplitude from every atom in the unit cell
- intensity of diffraction from each type if atom is determided by scattering factor *f*
- scattering factor *f*
- is proportional to the number of electrons and inversely proportional to the angle of diffraction 2θ

Scattering factor

- calculated



Normal scattering

- when incident λ is distant from the natural frequency of oscillation (ω) of the electron

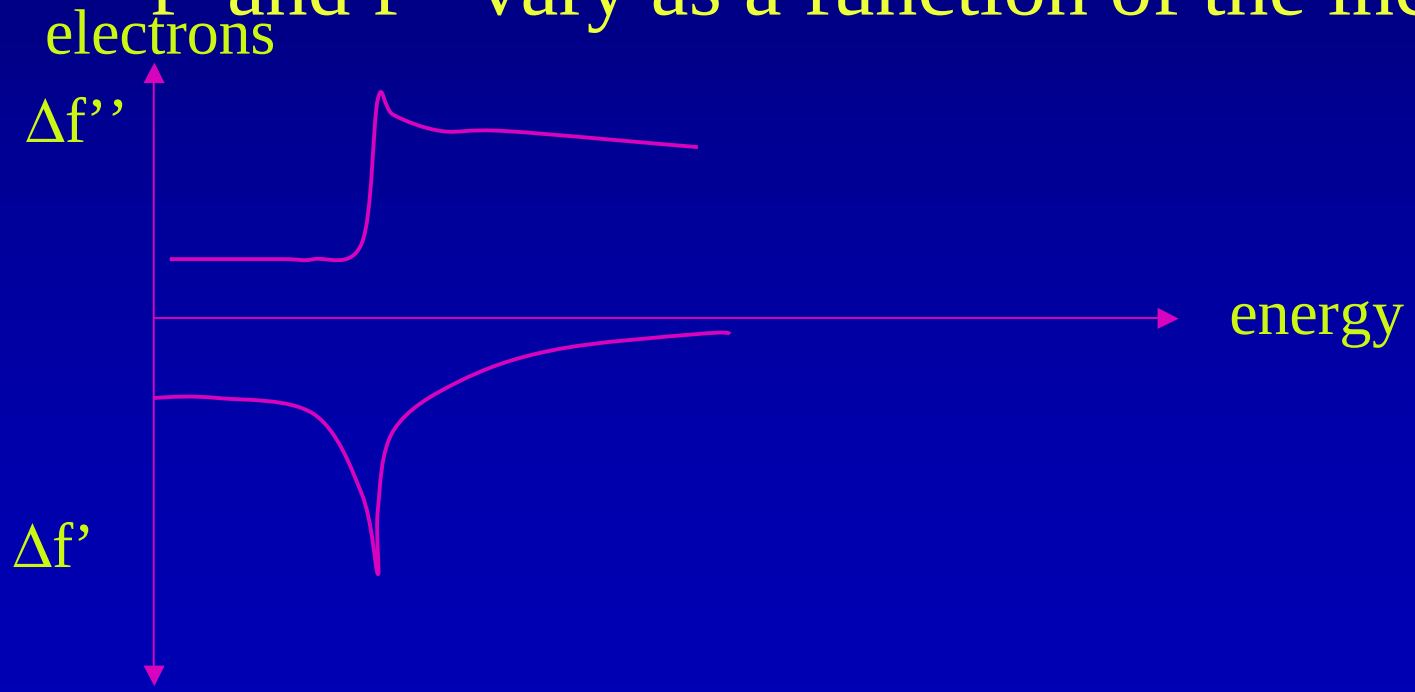
- elastic (Thompson) scattering
  - *f* is real and positive quantity

Anomalous scattering

- when incident λ is close to the natural frequency of oscillation (ω) of the electron
  - electrons execute resonat vibrations
  - diffraction occurs from such a resonant entity
- scattering factor *f* is a complex quantity
  - $\Delta f'$ - dispersive component (real)
  - $i\Delta f''$ - Bijvoet component (complex)
- **anomalous scattering**

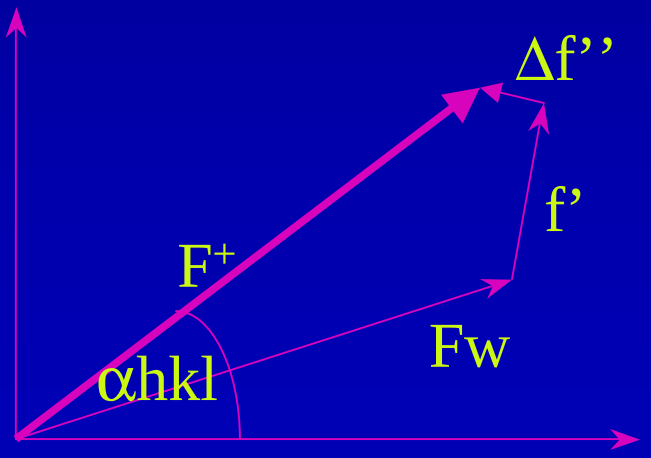
Anomalous scattering

- $f^{anom} = f_o + \Delta f' + i\Delta f'' = f' + i\Delta f''$
- f’ and f’’ vary as a function of the incident λ



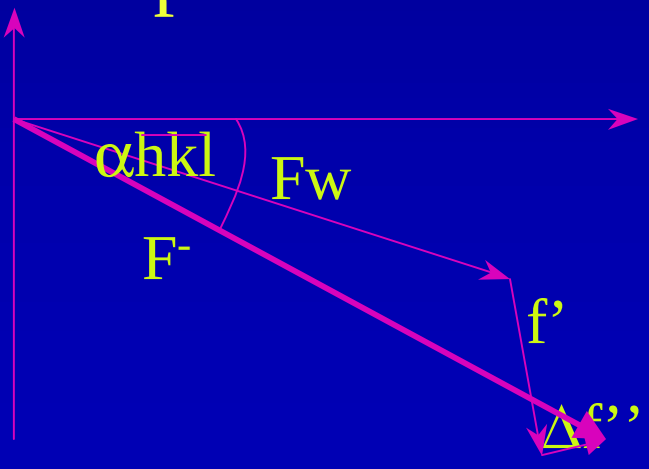
Anomalous scattering: Fhkl

- Fw = resultant of scattering from atoms without dispersion
- f’ and Δf’’ scattering factors from each atom with dispersion



Anomalous scattering: Fhkl

- Fw = resultant of scattering from atoms without dispersion
- f’ and Δf’’ scattering factors from each atom with dispersion



Anomalous scattering: Friedel law

- F^+ and F^- would be equal in magnitude in the absence of $\Delta f''$ vector
- although magnitudes of $\Delta f''$ are equal for Friedel mates
- phase of the $\Delta f''$ is always $\pi/2$ ahead of all other diffracted vectors
- **Friedel mates are not equal**

Bijvoet and dispersive signals

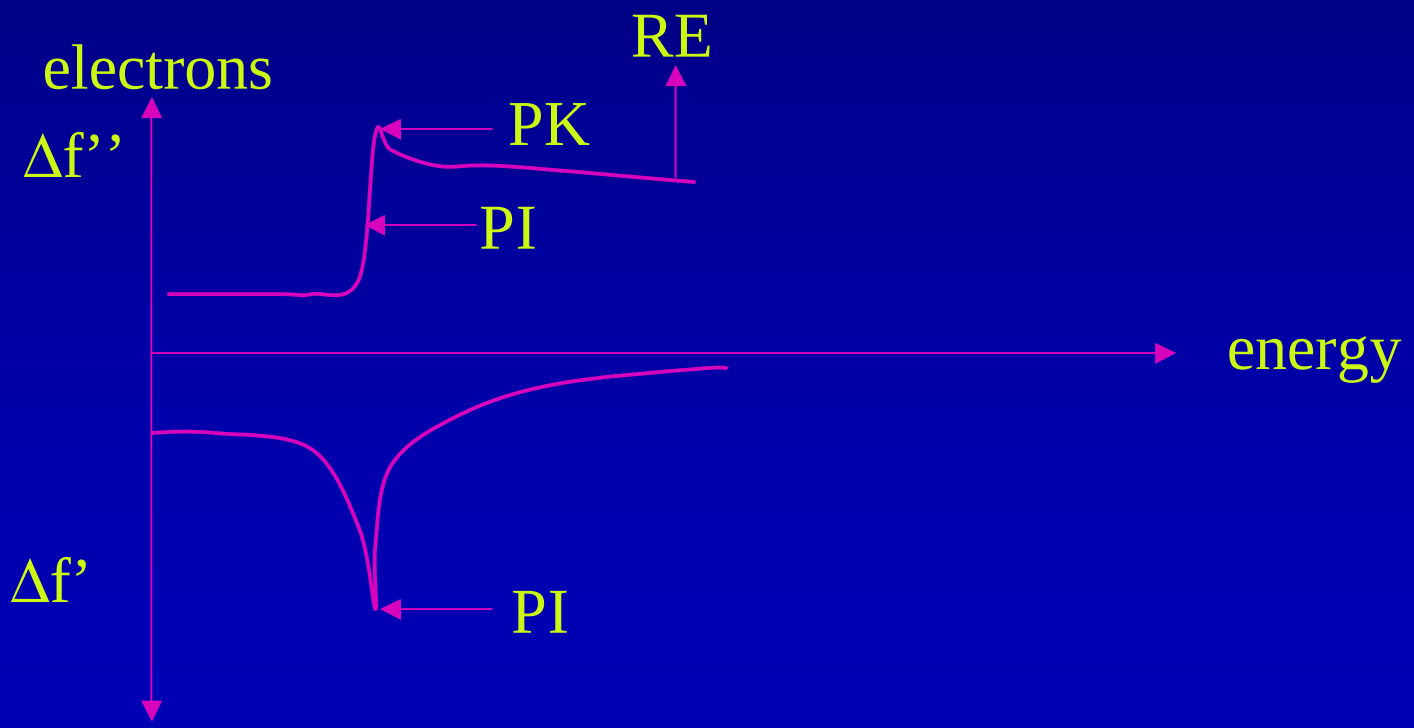
- **Bijvoet signal** can be derived by measuring the differences in intensities of Friedel mates at suitably chosen λ
- magnitude f’ is the same for Friedel mates
- f’ signal can be used by measuring intensities of a given replection at suitably chosen pair of λ and determining the differences between those measurments

Phasing with MAD

- **Karle formulation**
  - algebraic approach
- **MAD as a special case of MIR**

Phasing (as special case of MIR)

- native (reference dataset), e.g. PI
- PK, RE considered as derivatives



Phasing (as a special case of MIR)

- **for each walevength *i***
- dispersive term $\Delta f_i' = f_i' - f_1'$ is the source of isomorphous differences between $F(\lambda_i)$ and $F(\lambda_1)$
- **within one wavelenght**
- f_i'' gives rise to Bijvoet differences between F(hkl) and F(-h, -k, -l)

Anomalous scatterers for phasing

- primary choice based on wavelengths emitted by synchrotrons
- 0.35 - 3.5 A
  - subset of atoms in the range 29 (Ca) - 47 (Ag), and 50(Sn) to 92(U)
- effectively 0.7 - 1.7 A due to experimental difficulties

Anomalous scatterers for phasing

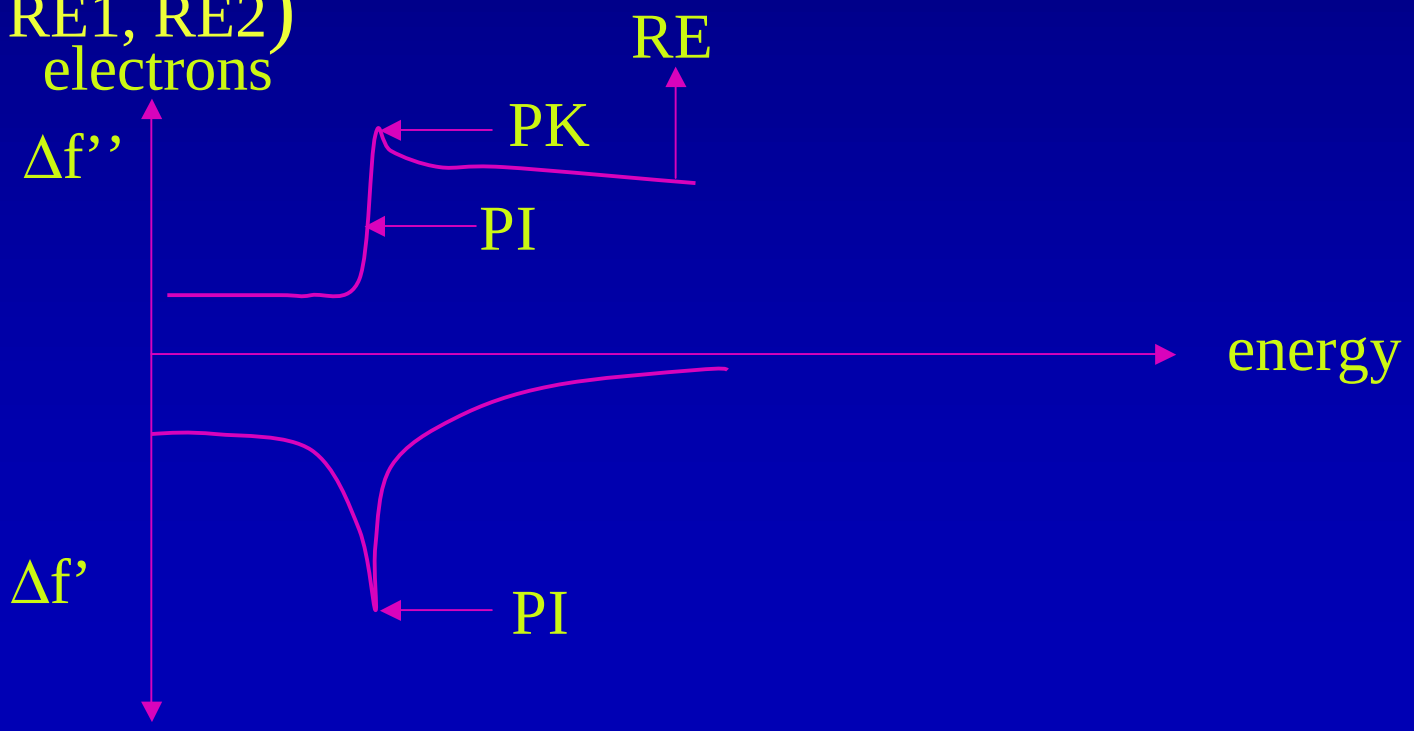
- heavy atoms
  - metals (either introduced or part of the native enzyme)
  - Br, J, Kr, Xe
- Se intruduced into the overexpressed protein via Se-Met
  - auxotroph strains of *E.coli*
  - extensively used method with high rate of success

WHY MAD

- excellent phasing vehicle in combination of
  - use of third generation syncrotrons and
  - techniques of molecular biology for preparation of Se-met labelled samples
- theoretically no problems with non-isomorphism

Choice of Wavelenghts

- maximizing the Bijvoet signal (PK)
- maximizing the dispersive signal (PI, RE, RE1, RE2)



Experimental requirements

- tunable X-ray source
- energy resolution (bandpass) < 10^{-4}
  - Wavelengths of PK and PI
  - precise position of the absorption edge depends on the chemical environment of the element
  - the position of the absorption edge must be determined on the crystal itself
  - by measuring the fluorescence of the element

Diffraction experiment

- differences due to anomalous scatterer are often small
- carefull measurments of intensities
  - set the counting statistics in function of expected anomalous signal