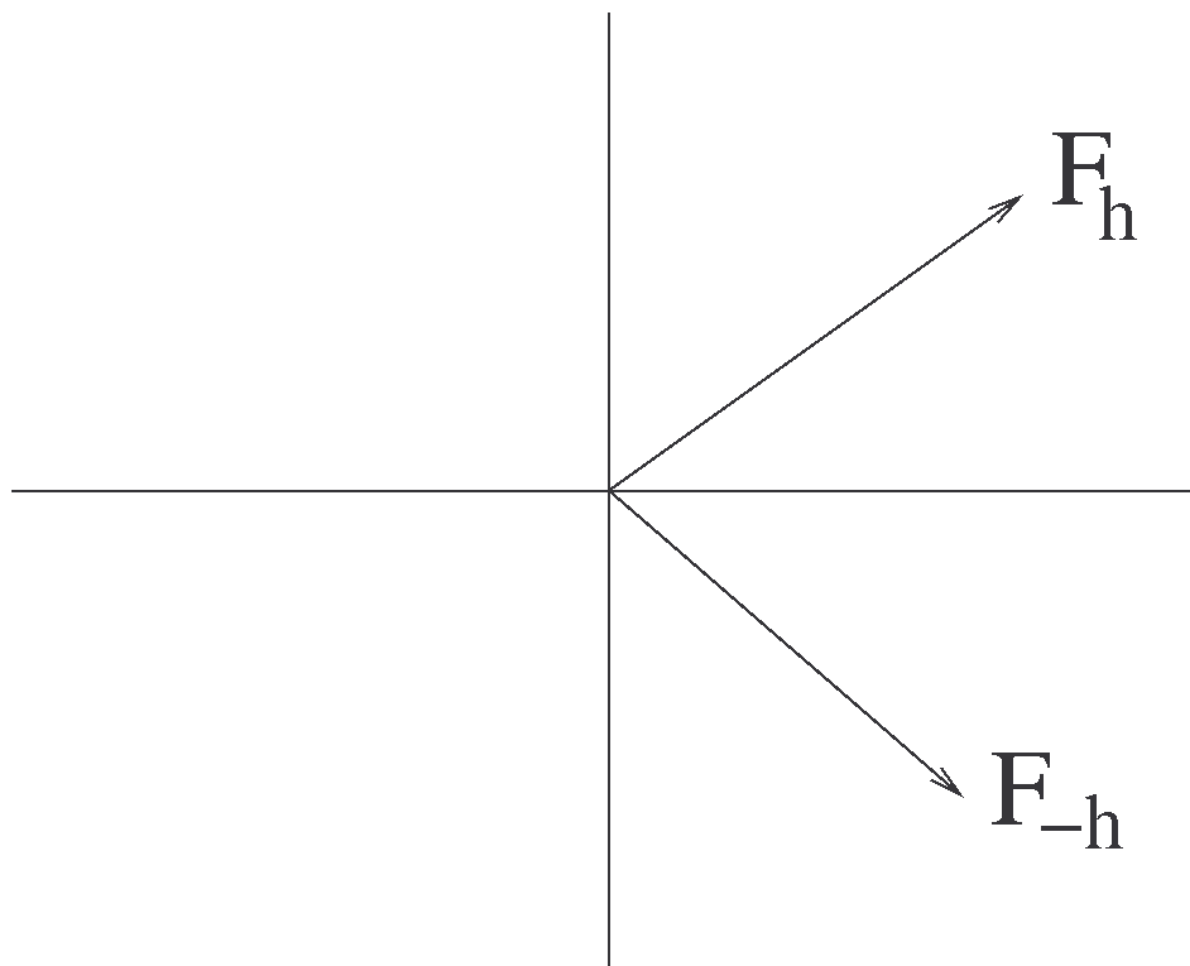# The absolute configuration of pharmaceutically interesting molecules
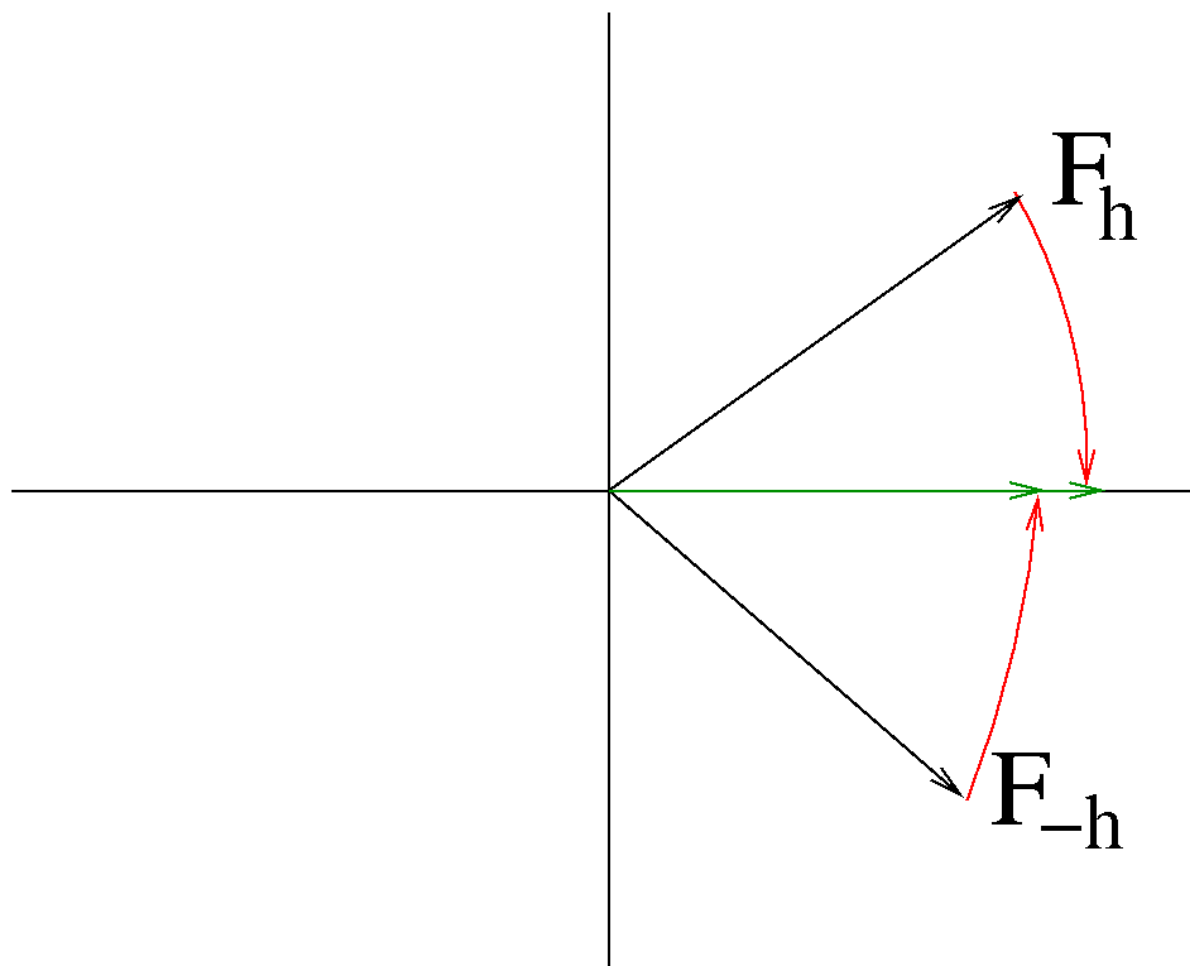
Optimal use of measured data

Rob W.W. Hooft, Bruker AXS, Delft
Leo H. Straver, Bruker AXS, Delft
Anthony L. Spek, Utrecht University

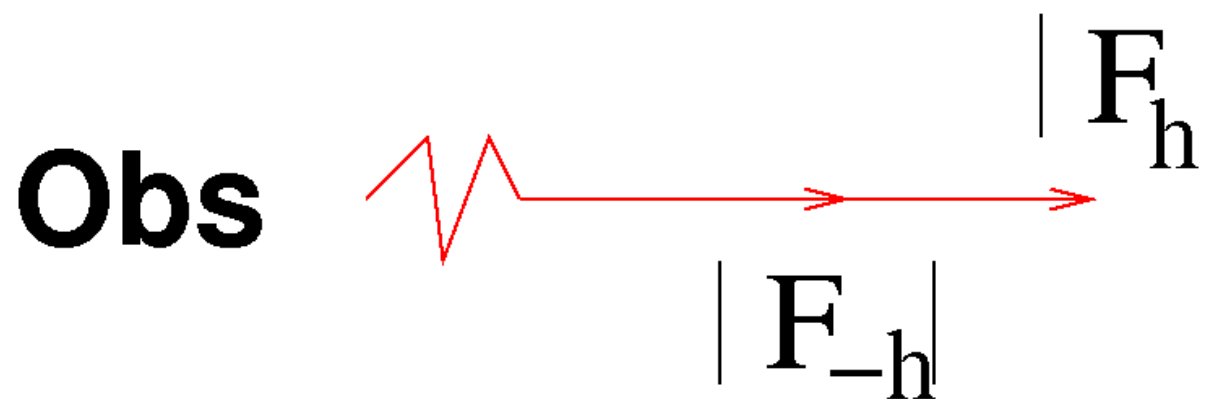Bruker AXS

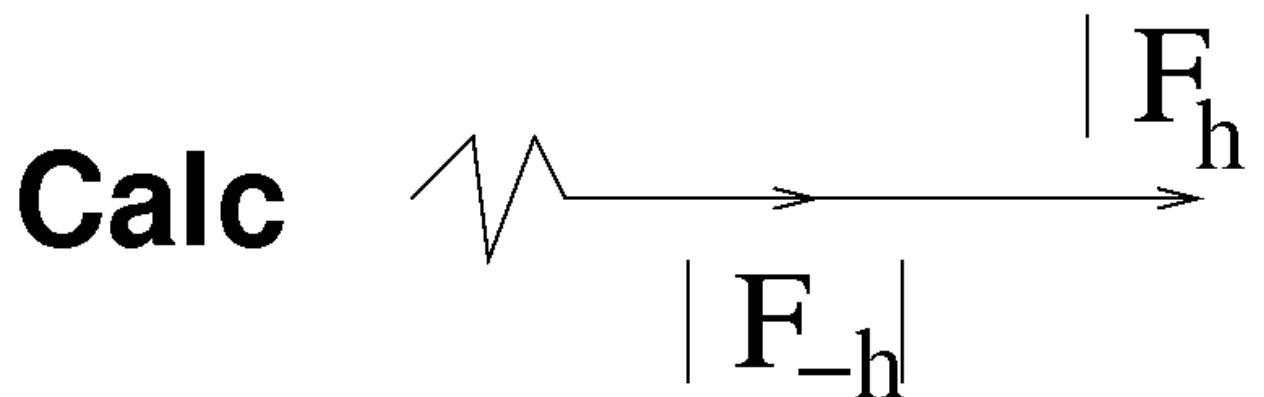# A Bijvoet pair in the complex plane

The Bijvoet difference
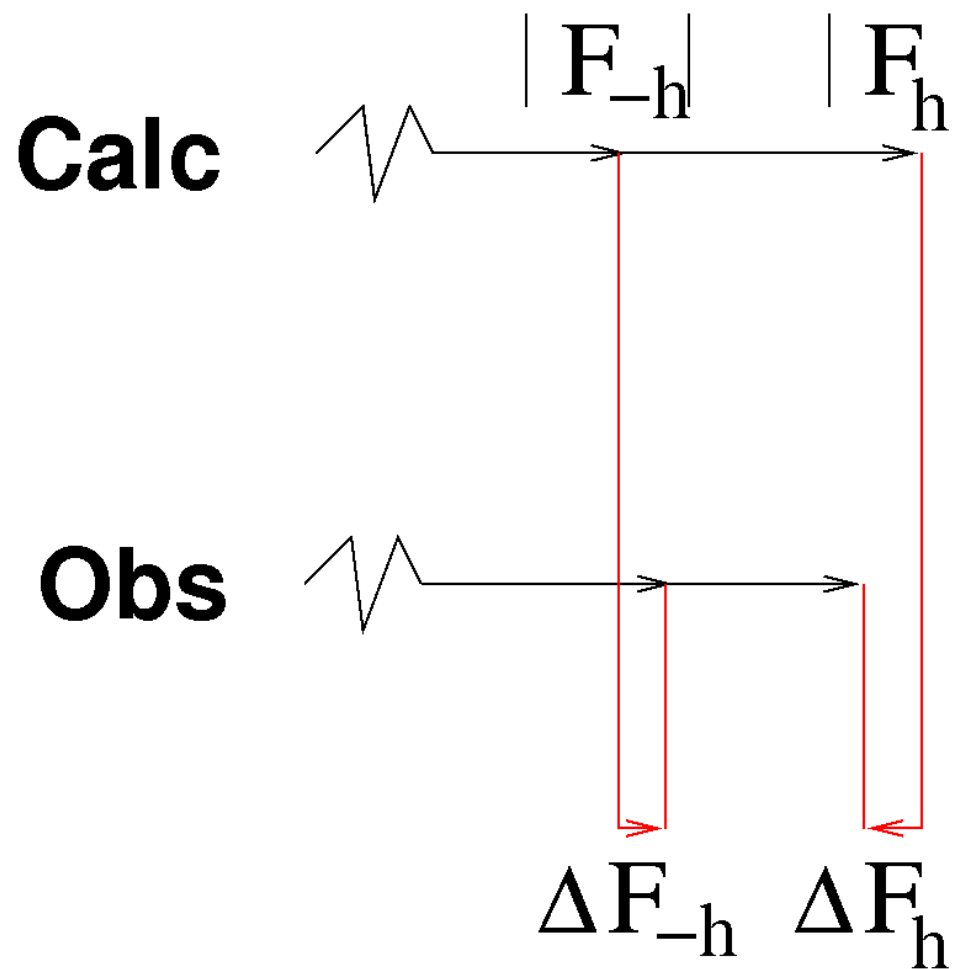
The resonant scattering signal: f"

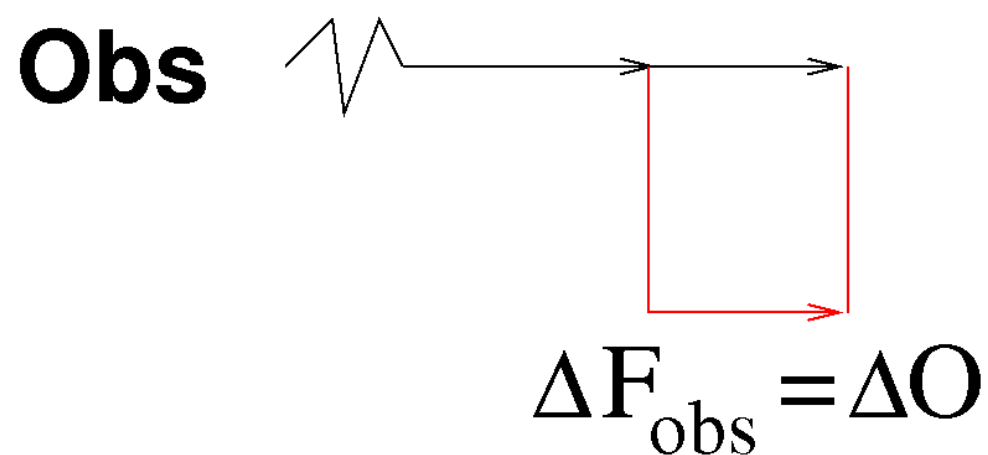|	f"(CuKα)	f"(MoKα)
Se	1.14	2.23
Cl	0.70	0.16
S	0.56	0.12
O	0.032	0.006

Model and observation

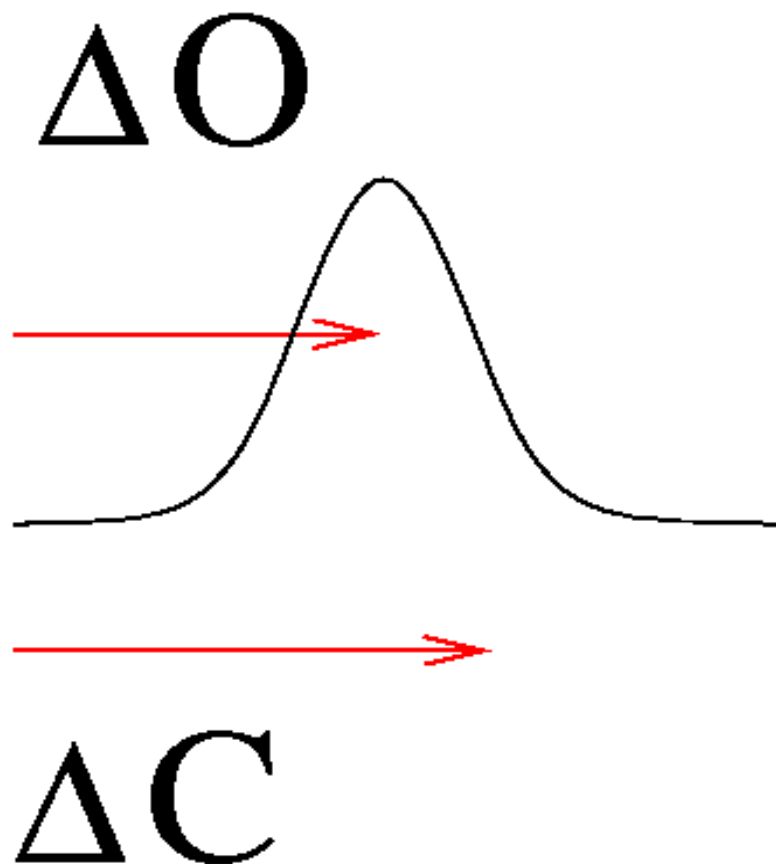# Difference between observations and model

Observed and calculated differences

$$\text{Calc} \qquad |F_{-h}| \qquad |F_h|$$

$$\Delta F_{calc} = \Delta C$$

$$\text{Obs}$$

$$\Delta F_{obs} = \Delta O$$
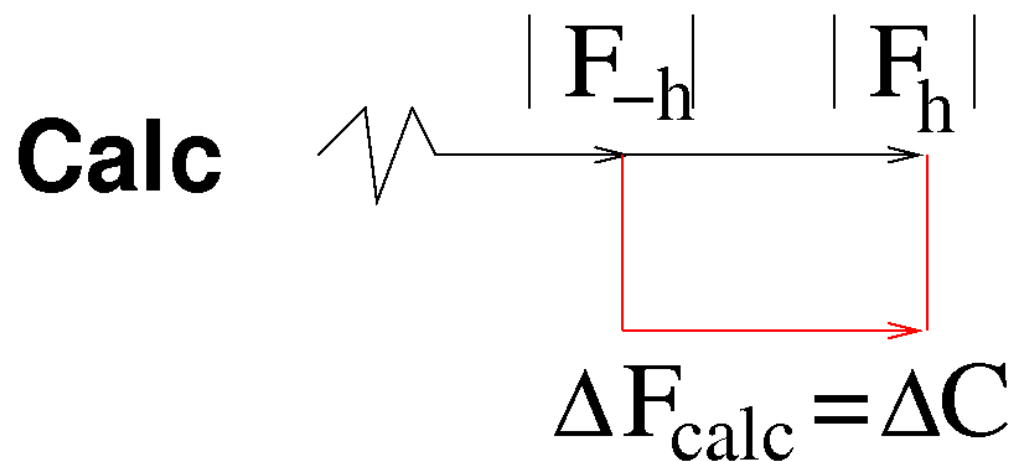
Standard deviations

Obs

ΔO
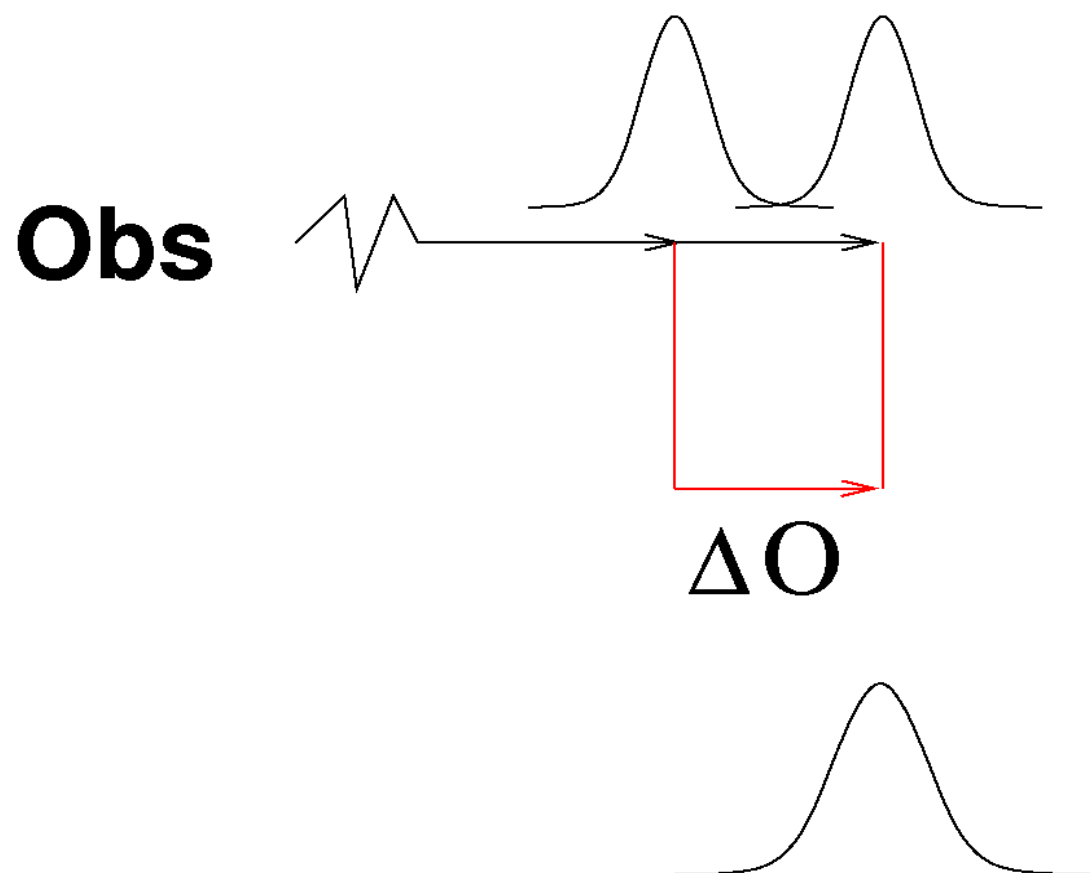
Likelihood of correspondence

Absolute structure sensitivity

	Strong signal		Weak signal		Strong signal
$\Delta C/\sigma \gg 1$	--	--	--	-	++ or --
$\Delta C/\sigma \approx 1$	--	-	-	+	--
$\Delta C/\sigma \approx 0$	--	-	+	-	--
$\Delta C/\sigma \approx -1$	--	+	-	-	--
$\Delta C/\sigma \ll -1$	++ or --	-	--	--	--
	$\Delta O/\sigma \ll -1$	$\Delta O/\sigma \approx -1$	$\Delta O/\sigma \approx 0$	$\Delta O/\sigma \approx 1$	$\Delta O/\sigma \gg 1$

Normal distribution for a single model

$$z = \frac{\Delta C - \Delta O}{\sigma_{\Delta O}}$$

$$\mathrm{p}(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

Generalized model

Mix between the opposite structures at arbitrary ratios:

$$(1-y)\Delta C + y(-\Delta C) = (1-2y)\Delta C \equiv \gamma \Delta C$$

This gives:

$$x(\gamma) = \frac{\gamma \Delta C - \Delta O}{\sigma_{\Delta O}} \qquad\qquad \mathrm{p}(x(\gamma)) = \frac{1}{\sqrt{2\pi}} e^{-x(\gamma)^2/2}$$

What is p?

$$x(\gamma) = \frac{\gamma \Delta C - \Delta O}{\sigma_{\Delta O}} \qquad p(x(\gamma)) = \frac{1}{\sqrt{2\pi}} e^{-x(\gamma)^2/2}$$

p is the chance of observing the given intensities, given the structural model and given the value of γ

Bayesian statistics

$$p(\gamma / \text{observations}) = \frac{p(\text{observations} / \gamma)p(\gamma)}{p(\text{observations})}$$

Delayed normalization

We delay normalization until we have determined the complete probability distribution:

$$\mathrm{p}(\gamma/\mathrm{observations}) = \frac{\mathrm{p}(\mathrm{observations}/\gamma)\mathrm{p}(\gamma)}{\sum p_u(\gamma/\mathrm{observations})}$$

Prior selection: enantiopure compound

In the case of enantiopure pharmaceuticals there are only two models: the proposed model and its opposite.

- P for the proposed model we call P2(OK)
- P for the opposite model we call P2(wrong)
- The prior p(γ) is ½ if $\gamma=1$ or $\gamma=-1$, and 0 otherwise.

Prior selection: racemic twinning

In the case of a potential racemic twin, γ becomes a continuous function

- Physics restrains values to $\gamma \geq -1$ and $\gamma \leq 1$
- Prior: $p(\gamma)=1$ if $\gamma \geq -1$ and $\gamma \leq 1$, $p(\gamma)=0$ elsewhere

This makes sense from a statistical point of view, but it does not give any warning if there would be a serious aberration

Alternative:

- Prior definition $p(\gamma)=1$ everywhere
- i.e. no prior knowledge is used

Prior selection: racemic twinning (2)

Prior definition p(γ)=1 everywhere, i.e. no prior knowledge

p_u(γ|observations) looks like a Gaussian. We can calculate a mean, G, and associated standard deviation:

$$G = \frac{\int \gamma p_u(\gamma)d\gamma}{\int p_u(\gamma)d\gamma} \qquad \sigma^2(G) = \frac{\int (\gamma - G)^2 p_u(\gamma)d\gamma}{\int p_u(\gamma)d\gamma}$$

Example

OptionMenus
Stereo Opts
Incl-HAtoms
Solid-Style
Rod -Style
CPK +Stick
Straw-Style
Stick-Style
BWC Res ARU
ViewOptions
GeomCalc
LabelSize)
UnitCellBox
Resd012..
H-Bonds-X
PackRange
Label -Hat+
CRotY >>
<<-RotZ+>>
<<-RotY+>>
<<-RotX+>>
Col Res ARU
Decoration
EPS Pov Ras
Reset End
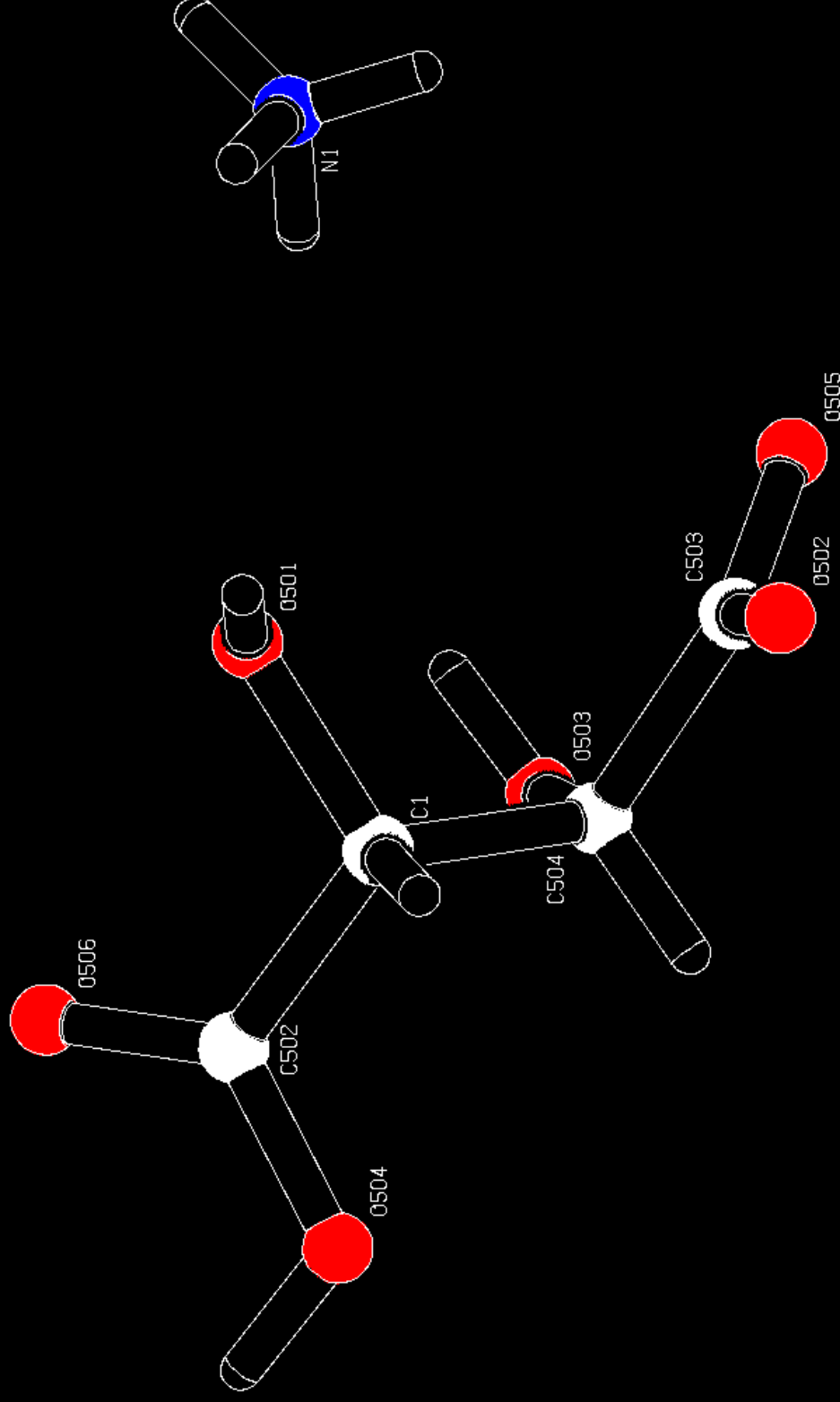
Exit

MenuActive

NOMOVE FORCED RES=0

N1

O501
C503
O505
O502
C1
O503
C504
C502
O506
O504

-90 X

R = 0.02

-148 ambi

Z -148

PLATON-JUL 20 13:54:17 2006 - (170706)

Y
9 -3

INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)

>>

BIJVOET-PAIR SCATTER PLOT

PLATON-Jul 20 13:49:10 2006 - (170706)

R = 0.02

DELTA Fcalc**2

DELTA Fobs**2

1.20

0.91

ambi

BIJVOETMENU
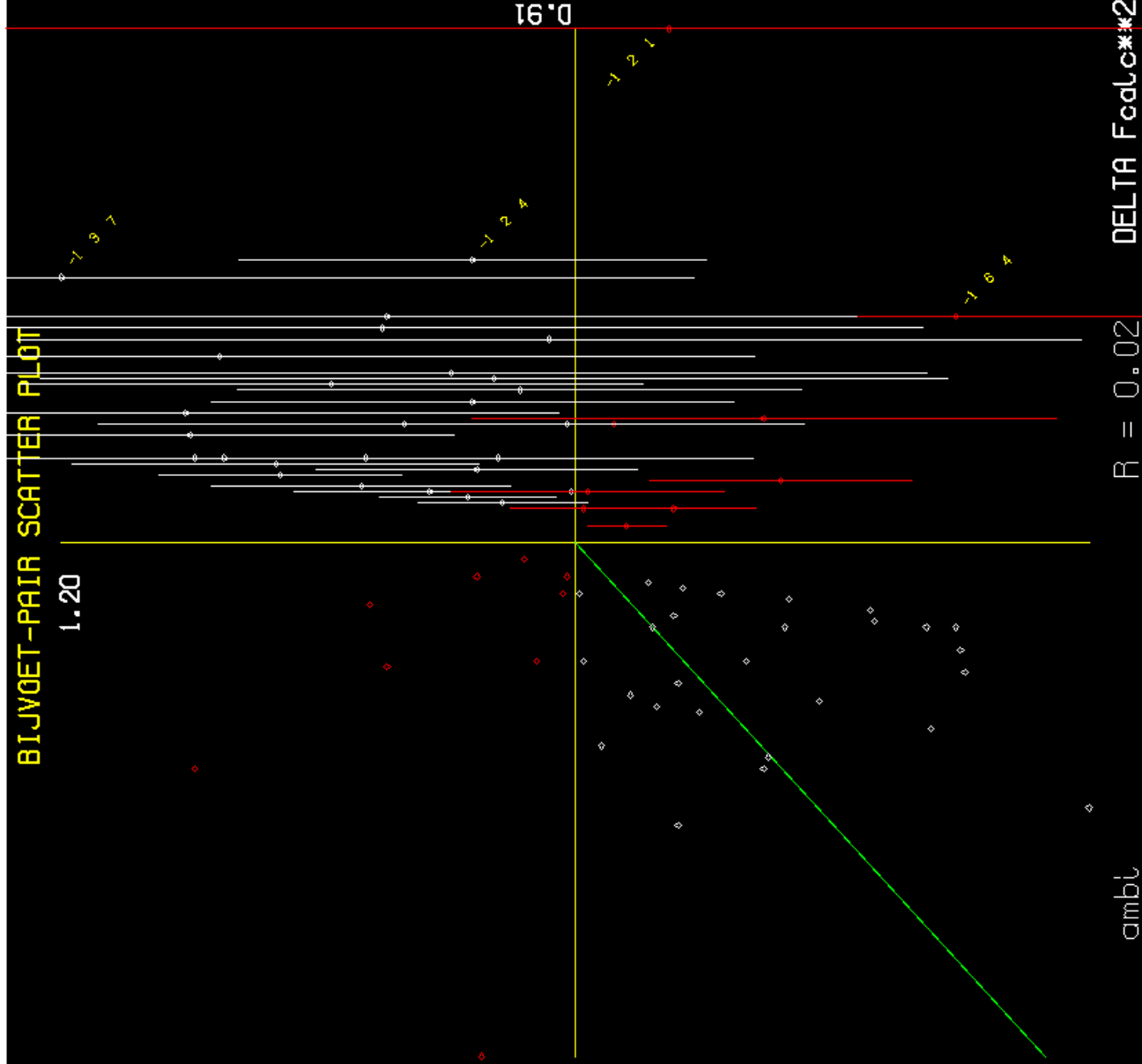
Space Group P212121	
Wavelength	0.71073	SigmaCriter
Flack Param	-0.100	
Flack S.U.	0.900	

Bijvoet Pairs:

Sigma Crit..	0.25
Selected ...	38
Out of	612
Number Plus	29
Number Minus	9
Aver. Ratio	1.391
RC	1.206

Normal Prob. Plot

Corr. Coeff.	0.991
Intercept ..	-0.022
Slope	0.840	ApplySlope

Bayesian Statistics

Bijvoet Pairs	612
P2(true)....	1.000
P3(true)....	0.998
P3(rac-twin)	0.002
P3(false) ..	0.2E-09
G	1.0936
G S.U.	0.3116
FLEQ	-0.047	EPS-File
FLEQ S.U. ..	0.156	End

Exit

MenuActive

INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)

>>

Experiences

σ_y is about half of σ_x sometimes even less

y clusters very well around 0.0

absolute structure of $C_xH_yO_z$ can be determined reliably using CuKα radiation

....sometimes even using MoKα radiation

When Bijvoet pairs are measured explicitly, the least-squares refinement of the Flack x parameter is not correlated with atomic coordinates

Conclusions

A new method to determine absolute structure

Pharmaceutical compounds, known to be enantiopure.

Explicit calculation of the likelihood of an inversion (P2) without assumption about the Gaussian shape of the probability curve

The method can be used for very weak anomalous scatterers

But:

Bijvoet pairs must be measured explicitly

This absolute structure determination can not take place inside the full matrix refinement.