

PowerPoint File available:

[http://bl831.als.lbl.gov/~jamesh/powerpoint/
UCSF_DQvMI_2017.pptx](http://bl831.als.lbl.gov/~jamesh/powerpoint/UCSF_DQvMI_2017.pptx)

Acknowledgements

Robert Stroud James Fraser John Spence
Chris Nielsen Clemens Schulze-Briese Aina Cohen Ana Gonzalez

UCSF LBNL SLAC

ALS 8.3.1: TomAlberTron

UC Multicampus Research Programs and Initiatives (MRPI)

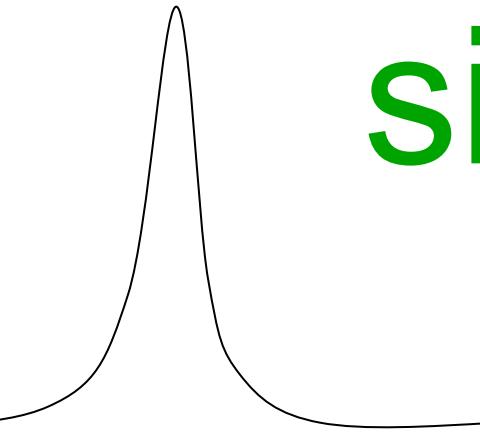
NIH NIGMS R01 GM124149

NIH NIGMS P30 GM124169

Integrated Diffraction Analysis Technologies (IDAT)

Plexxikon, Inc.

Synchrotron Radiation Structural Biology Resource (SLAC)

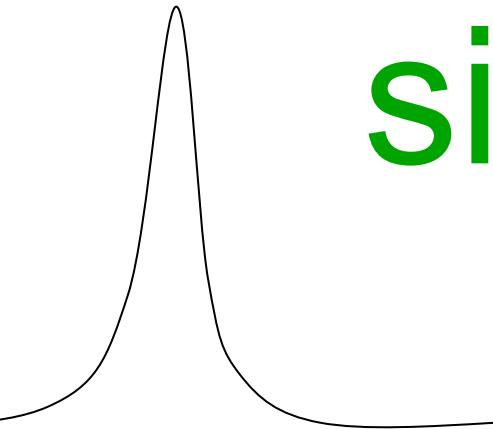


signal

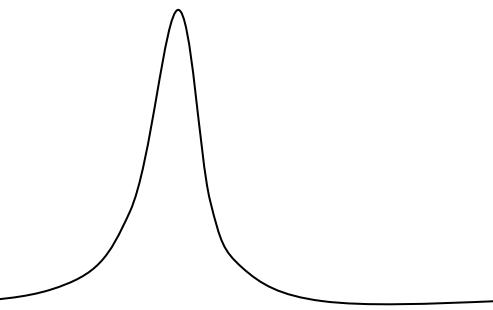
“If you don’t have
good data,
then you have
no data at all.”

-Sung-Hou Kim

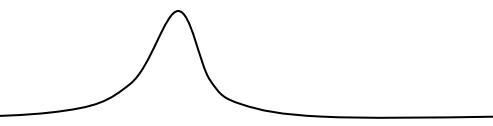
signal vs noise



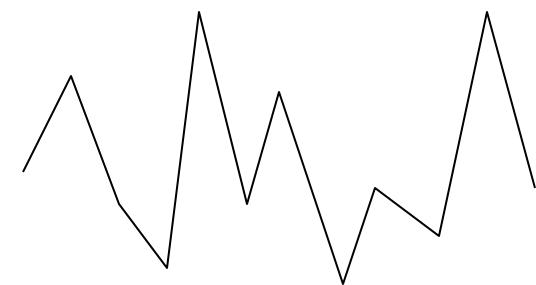
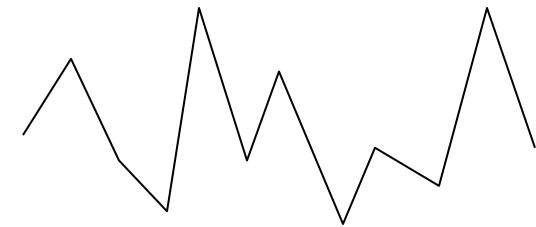
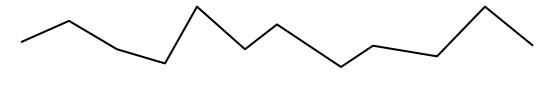
easy



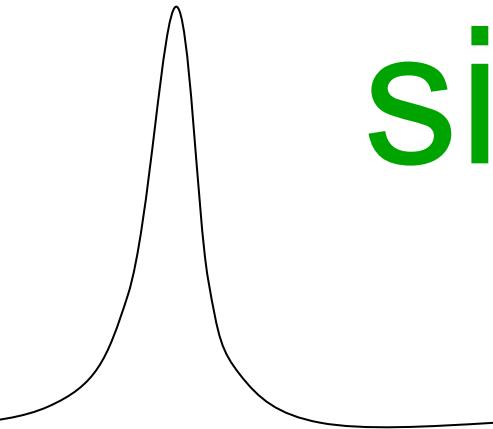
hard



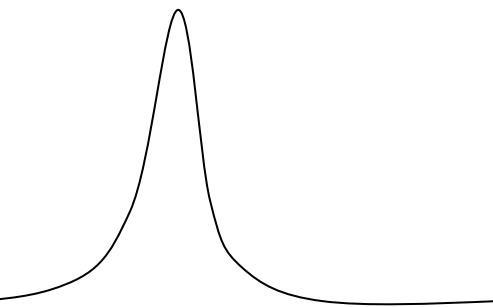
impossible



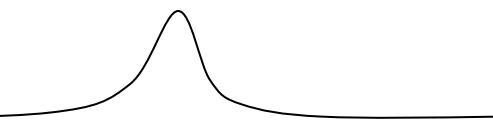
signal vs noise



easy

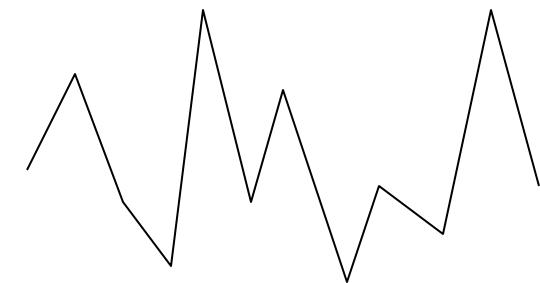
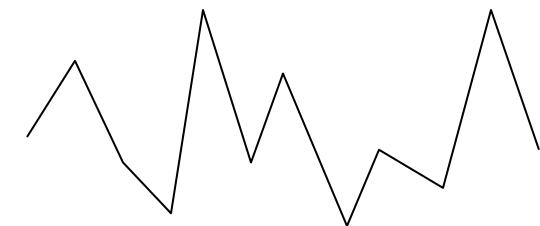
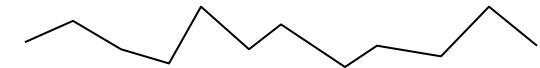


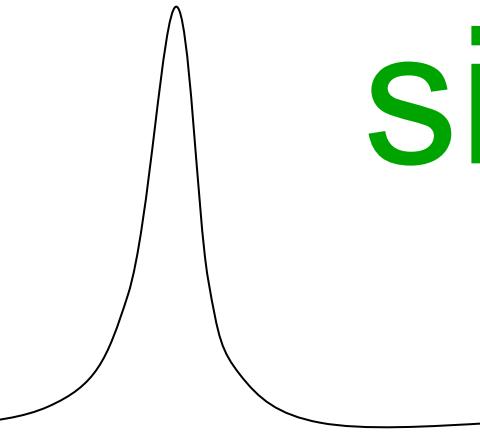
hard



threshold of “solvability”

impossible





signal vs noise



“If you don’t have
good data,
then you must
learn statistics.”

-James Holton

Adding noise

Adding noise

$$1 + 1 = 1.4$$

Adding noise

$$1 + 1 = 1.4$$

$$\sigma_{\text{total}}^2 = \sigma_1^2 + \sigma_2^2$$

Adding noise

$$1^2 + 1^2 = 1.4^2$$

$$\sigma_{\text{total}}^2 = \sigma_1^2 + \sigma_2^2$$

Adding noise

$$1^2 + 1^2 = 1.4^2$$

$$3^2 + 1^2 = 3.2^2$$

$$\sigma_{\text{total}}^2 = \sigma_1^2 + \sigma_2^2$$

Adding noise

$$1^2 + 1^2 = 1.4^2$$

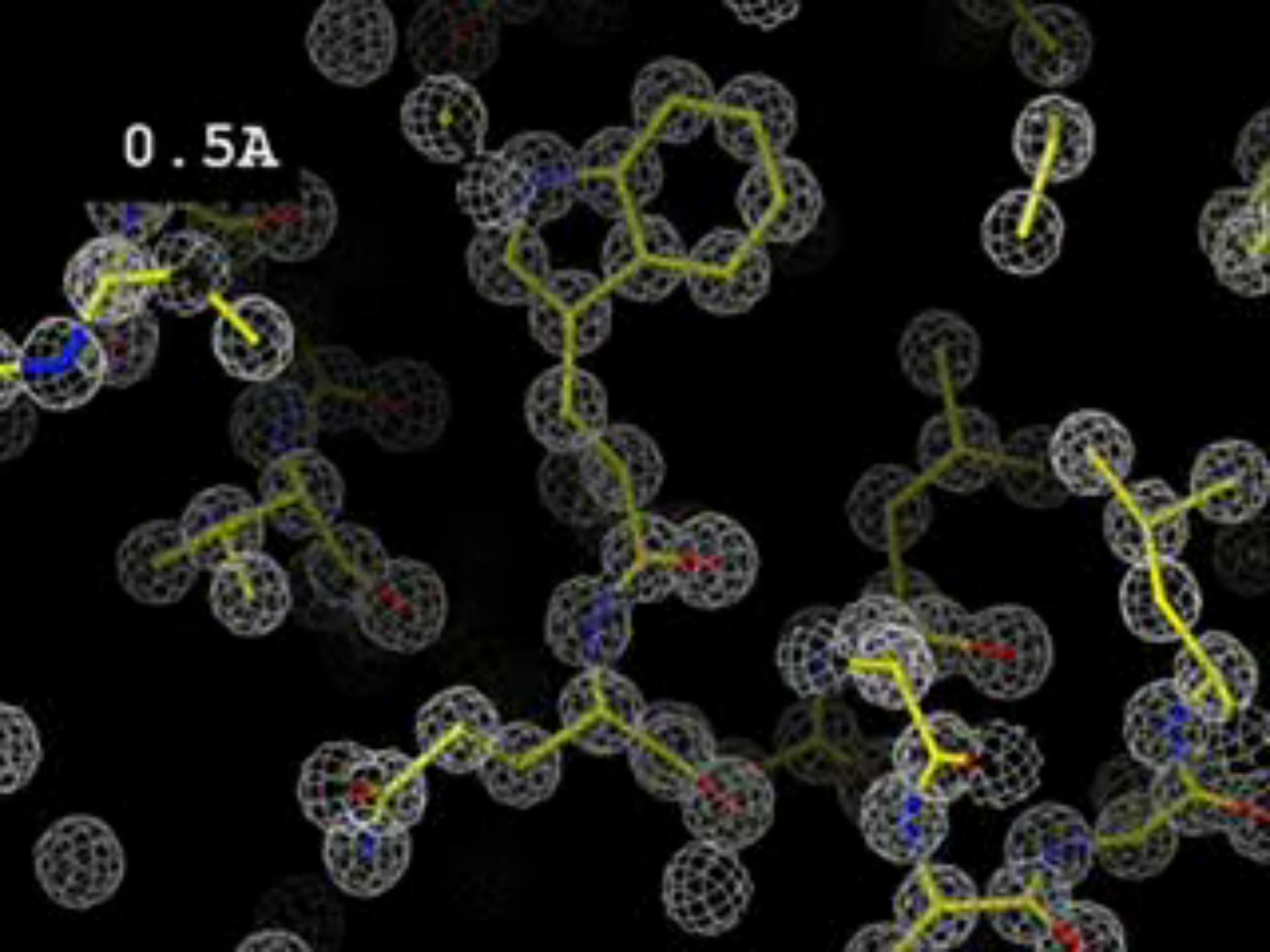
$$3^2 + 1^2 = 3.2^2$$

$$10^2 + 1^2 = 10.05^2$$

What do we want?

1. Resolution
2. Low R factor
3. Mosaicity
4. Nice, round spots
5. No overlaps
6. No twinning
7. Lots of crystals
8. ...

0 . 5 A



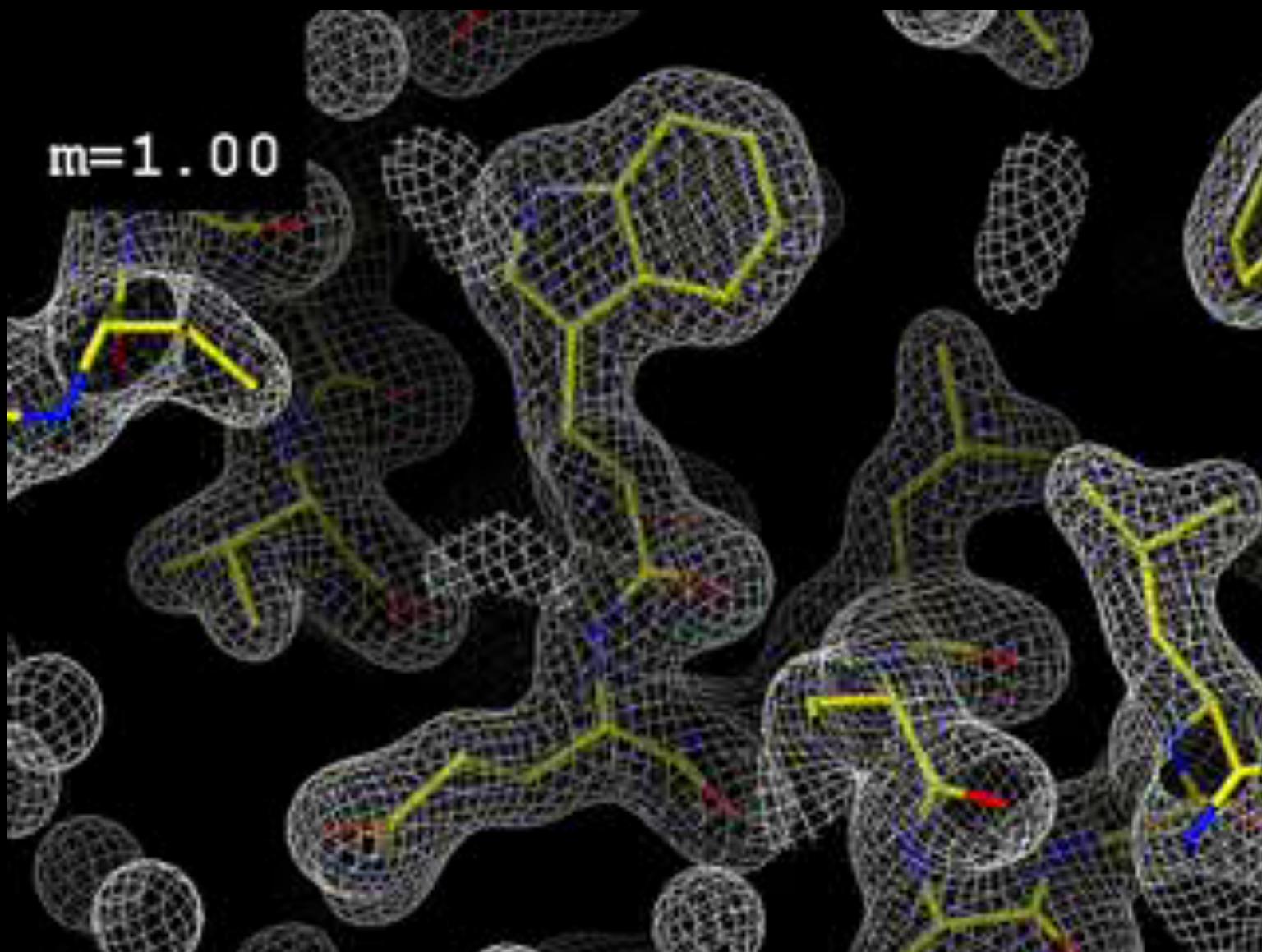
What do we want?

1. Resolution

2. Low R-factor
3. Mosaicity
4. Nice, round spots
5. No overlaps
6. No twinning
7. Lots of crystals
8. ...

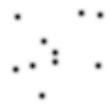
2. Phases

Phases



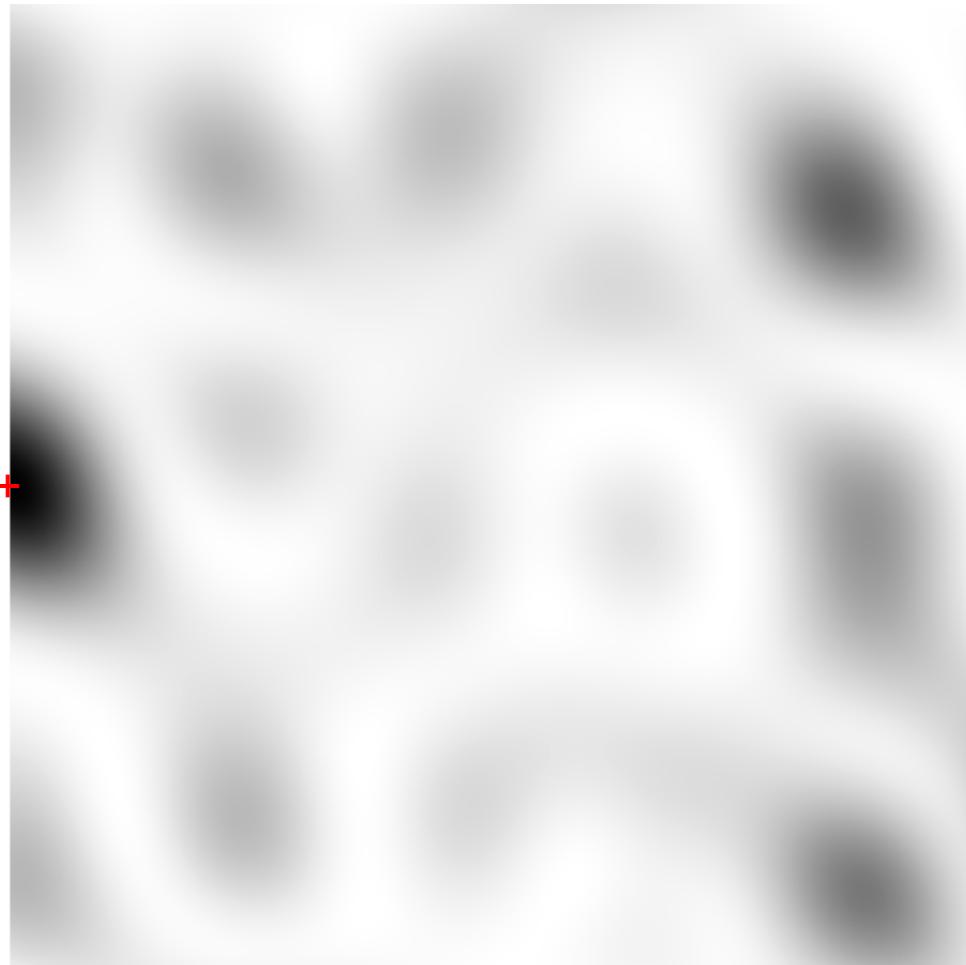
scattering from a structure

sample



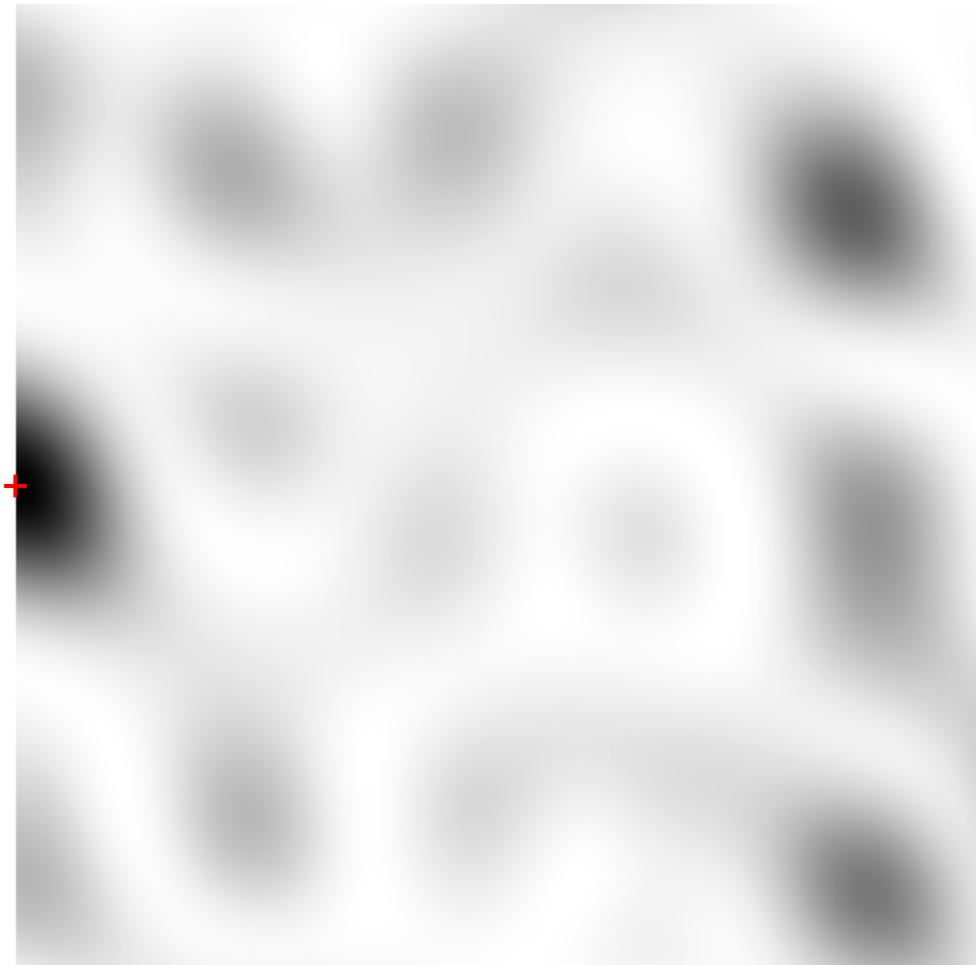
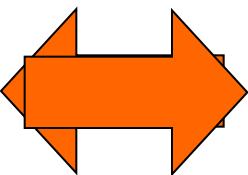
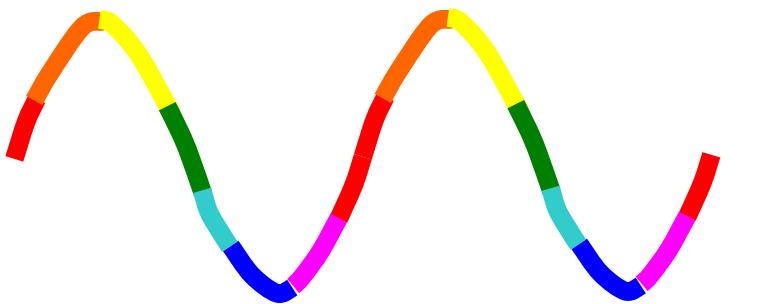
detector

beam center
+



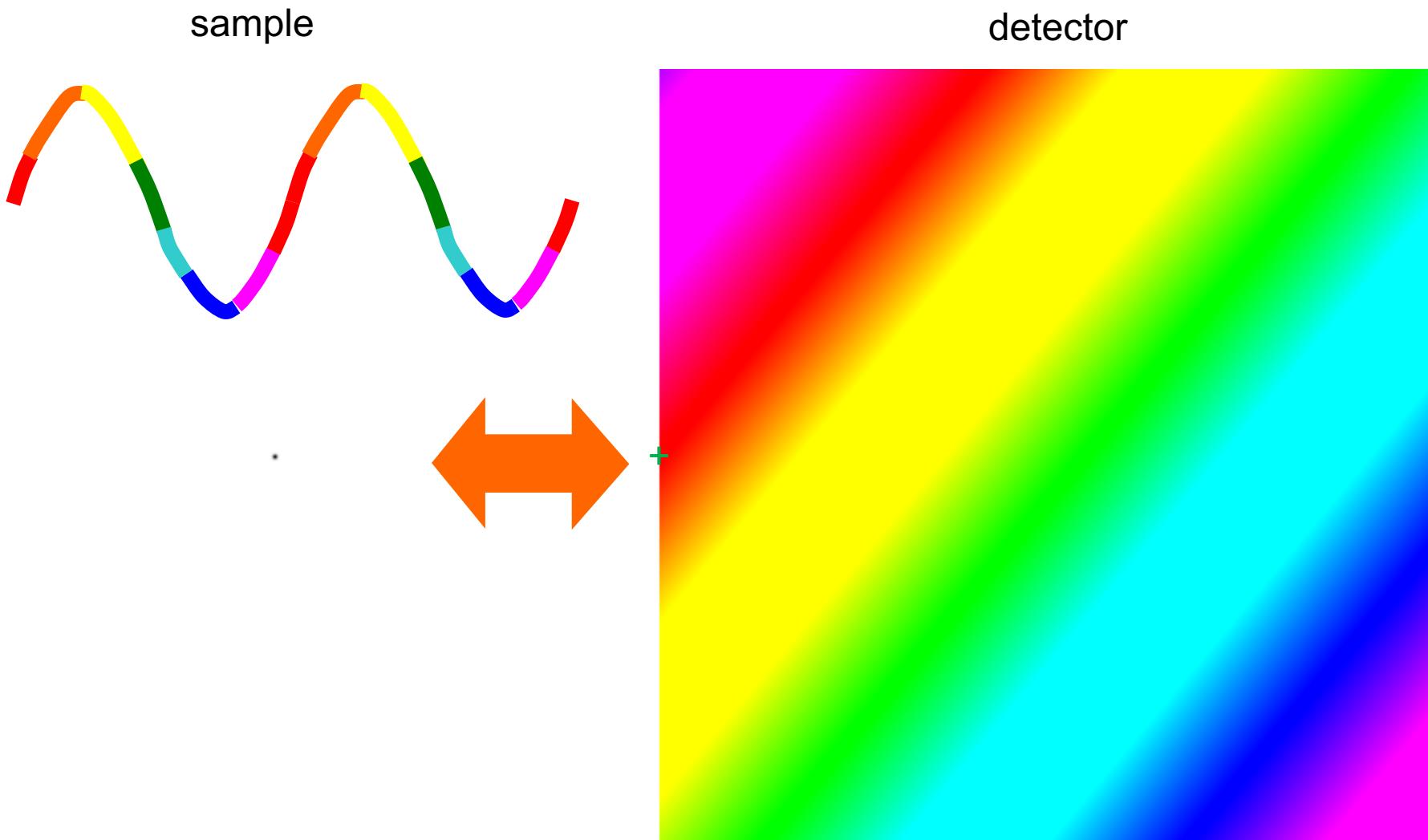
forward Fourier Transform

no phase



scattering from a sample

colored by phase

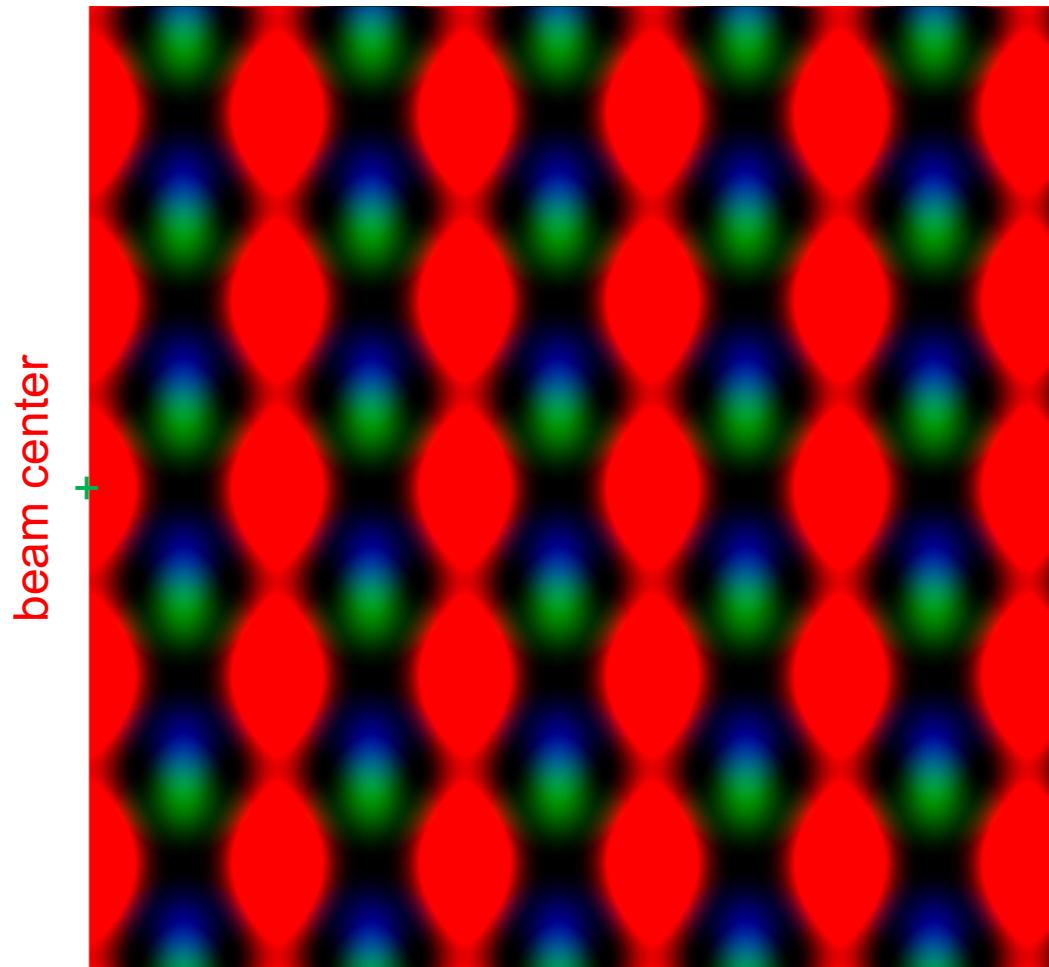


scattering from a lattice

colored by phase

sample

detector

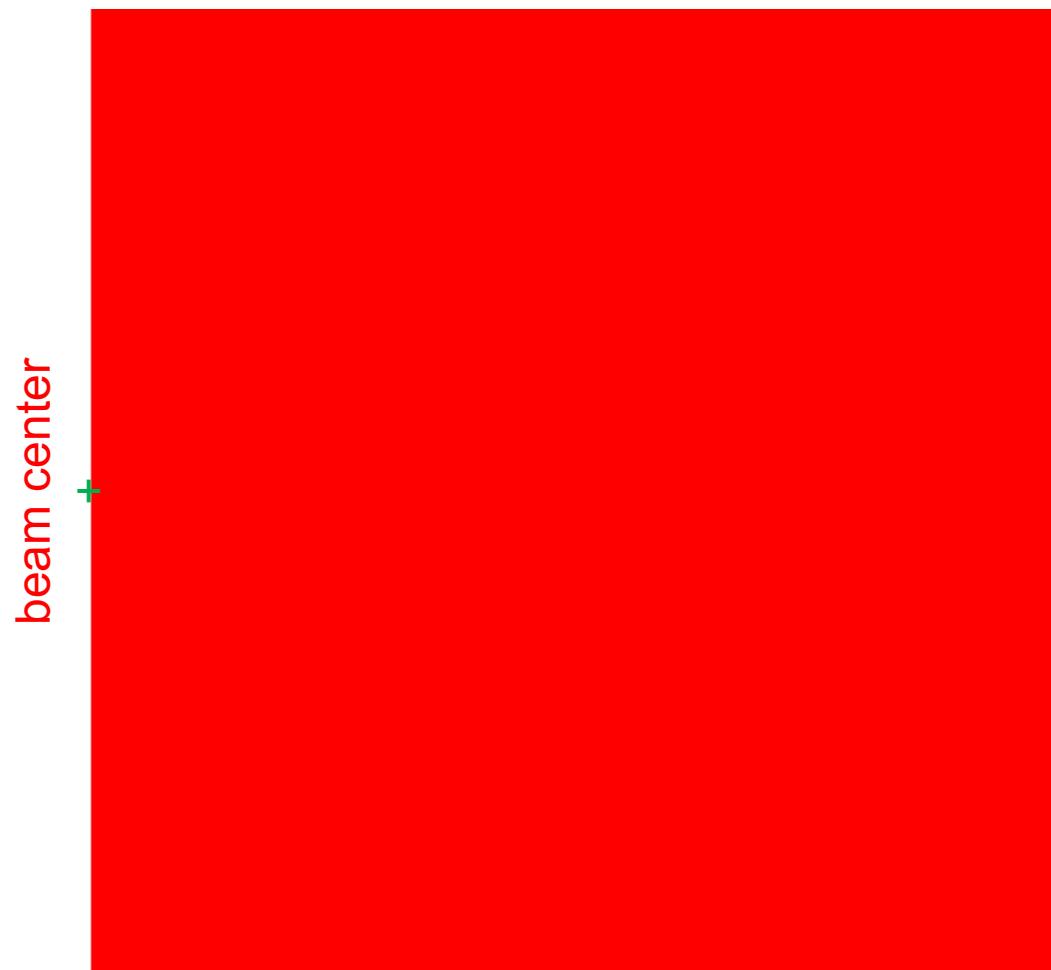


scattering from a lattice

colored by phase

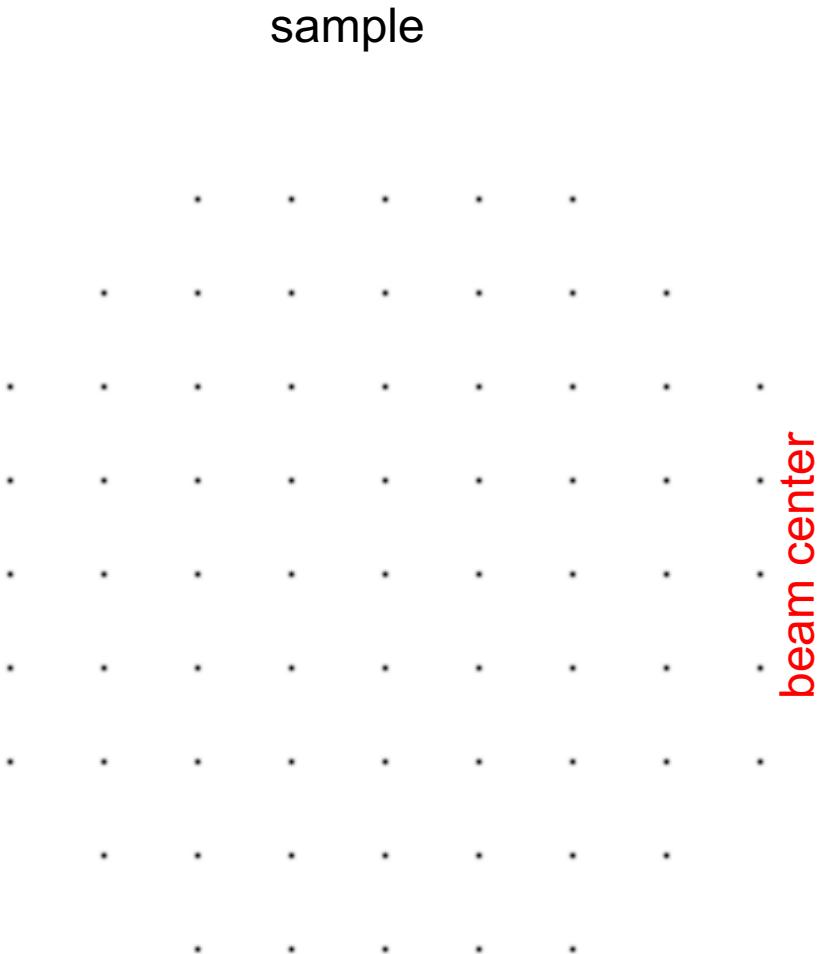
sample

detector

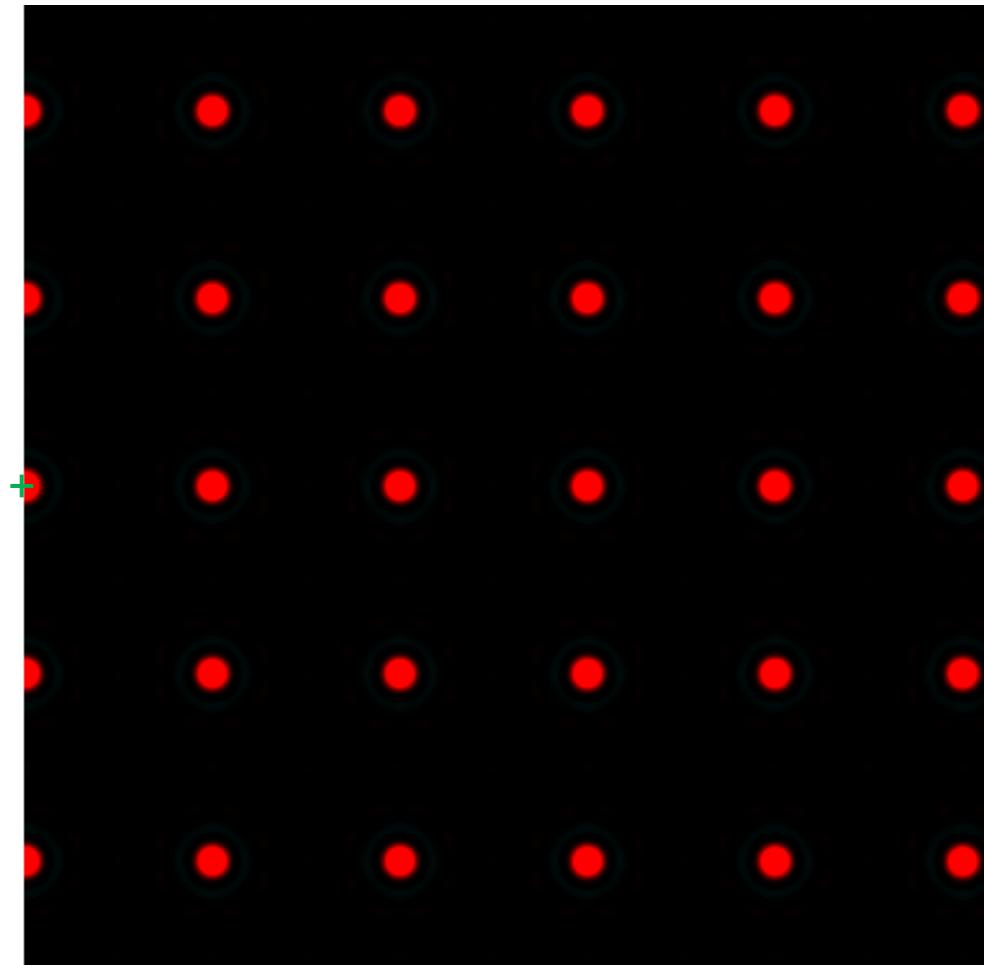


scattering from a lattice

colored by phase



beam center

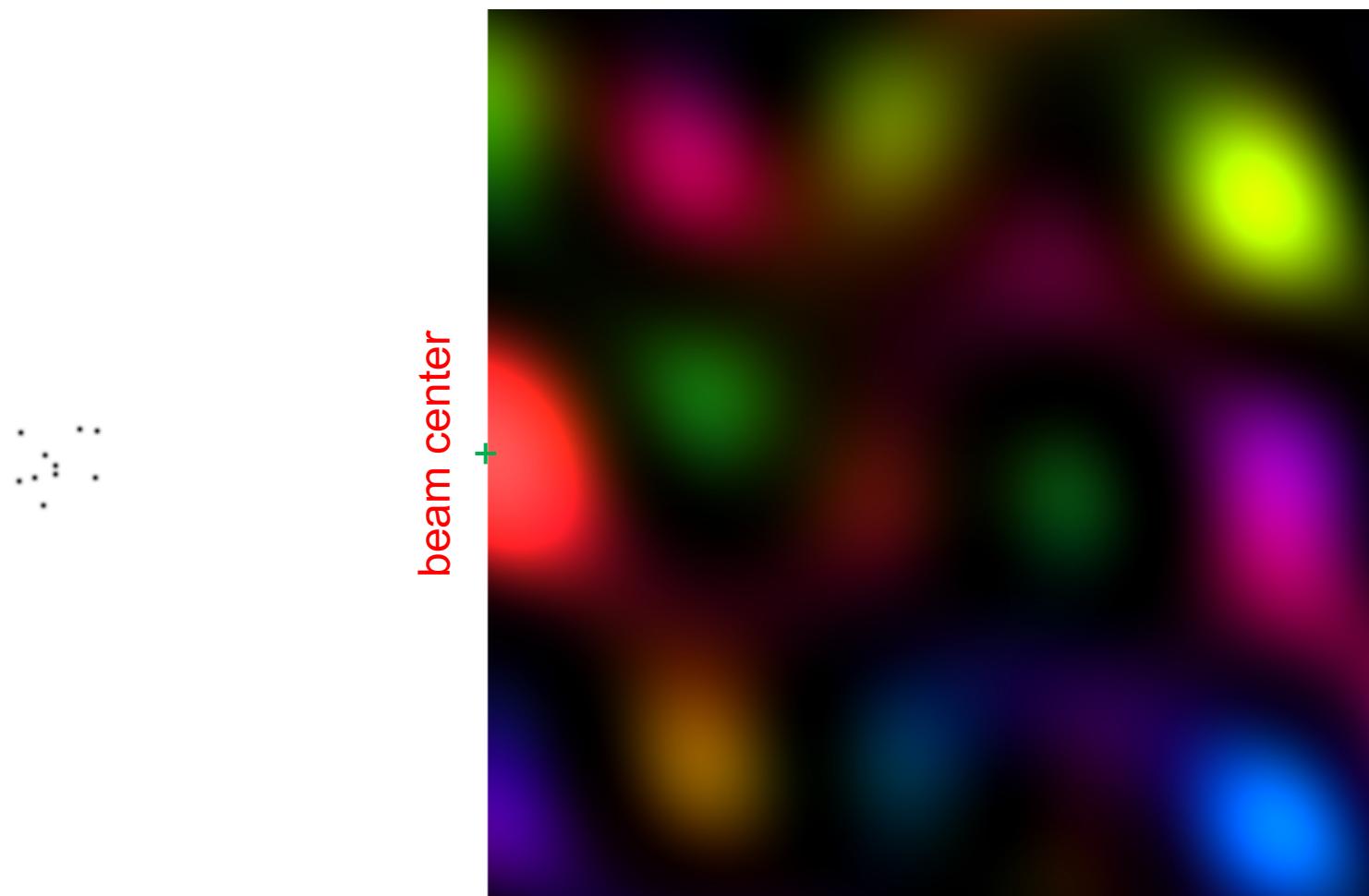


scattering from a crystal structure

colored by phase

sample

detector

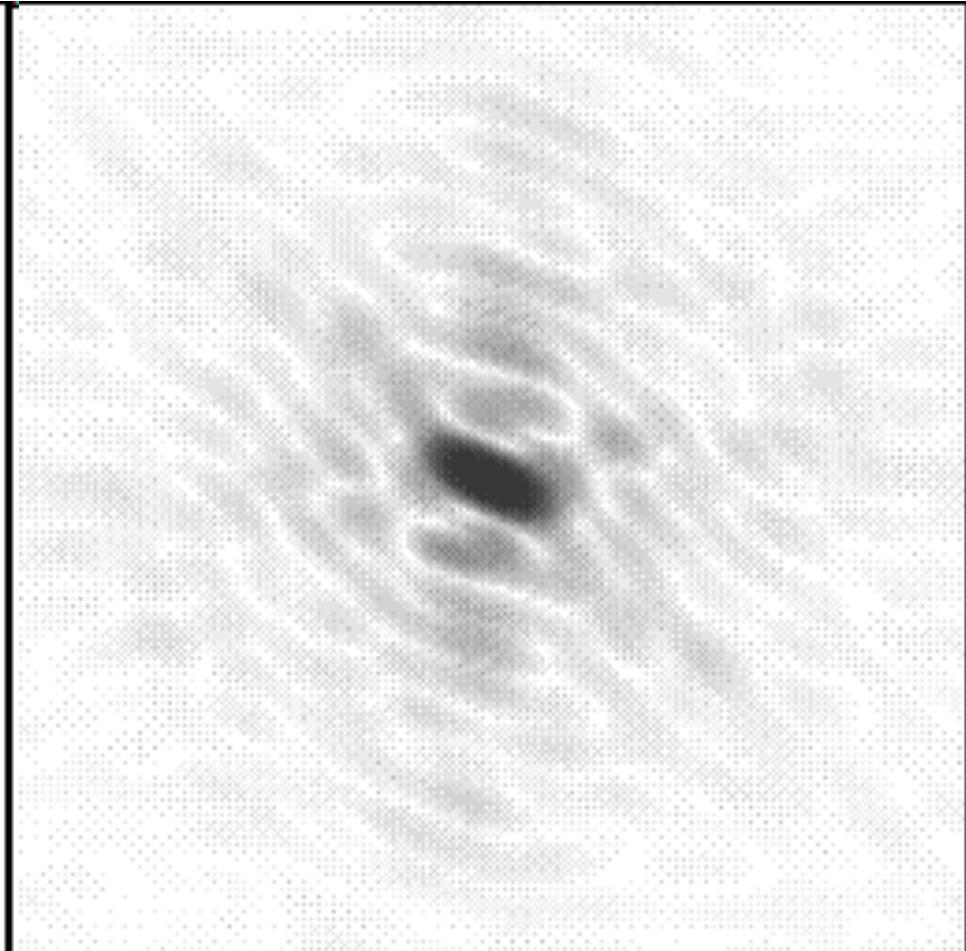


Major Phasing techniques

- Molecular Replacement
- Multiple Isomorphous Replacement
- Multiwavelength Anomalous Diffraction
- Single-wavelength Anomalous Diffraction

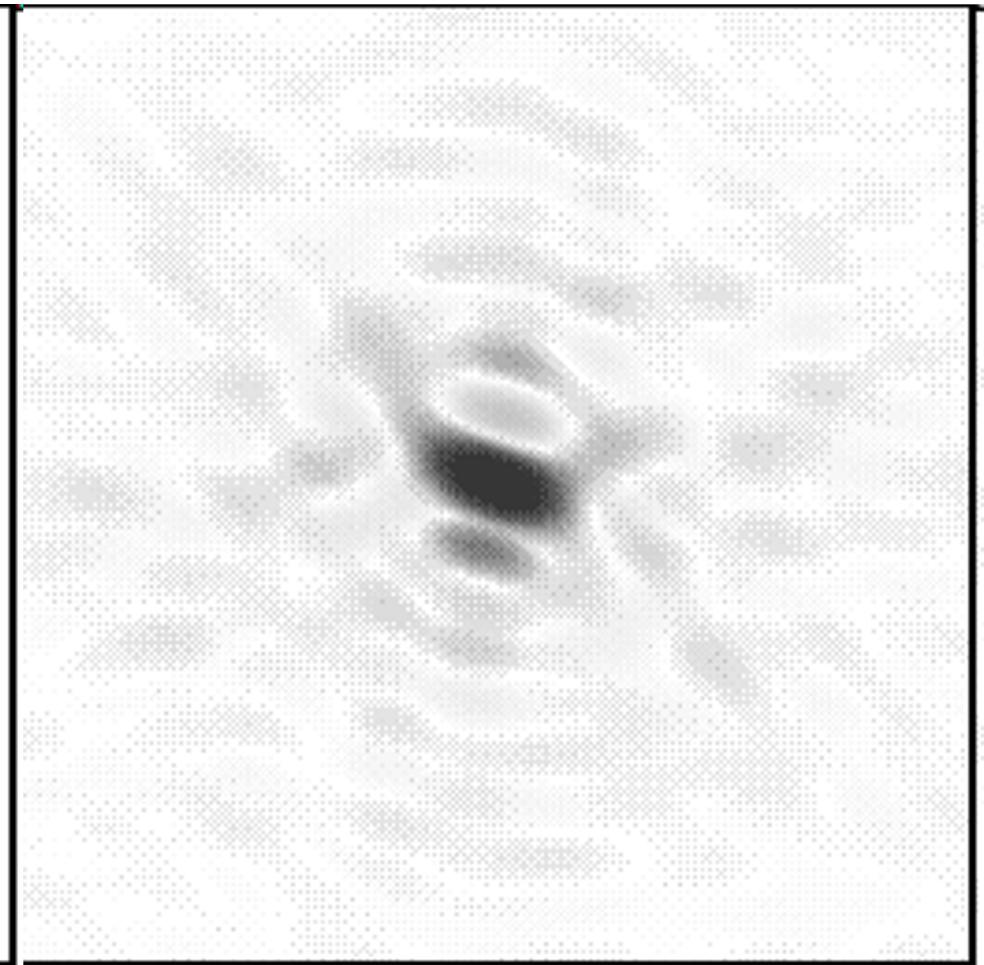
Molecular Replacement

correct structure and intensities



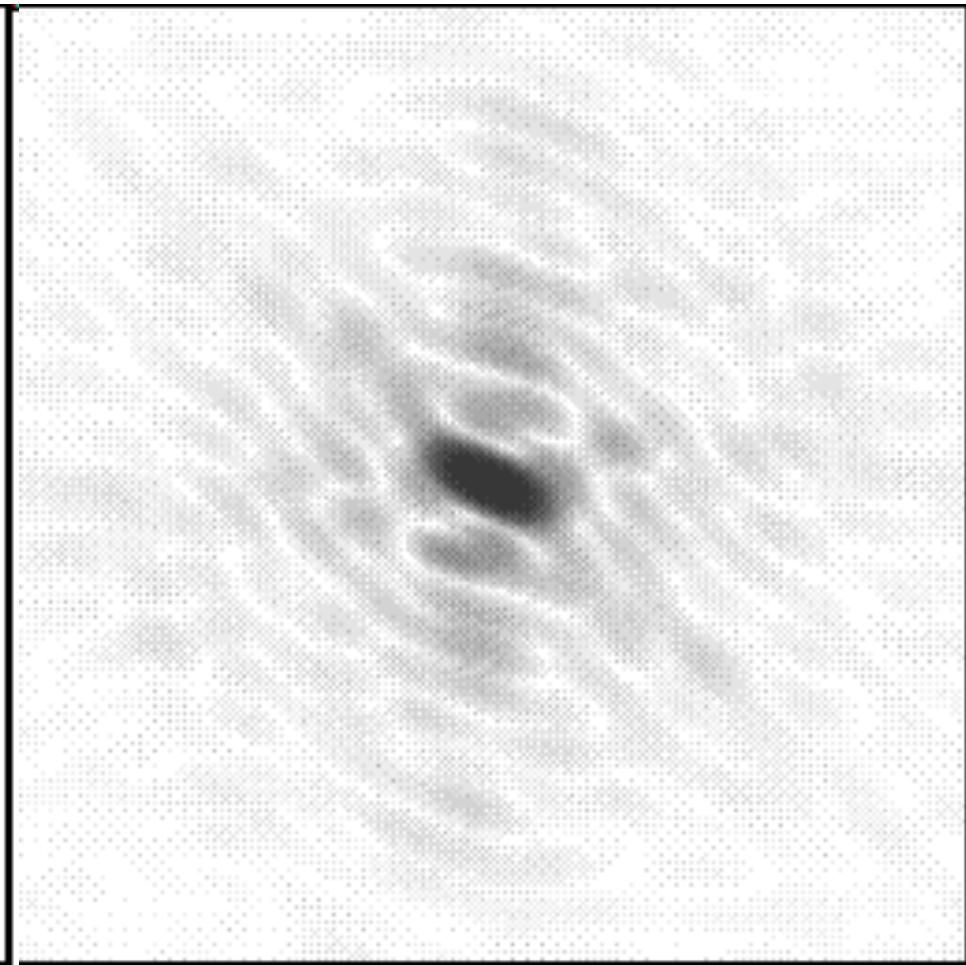
Molecular Replacement

Available model and intensities



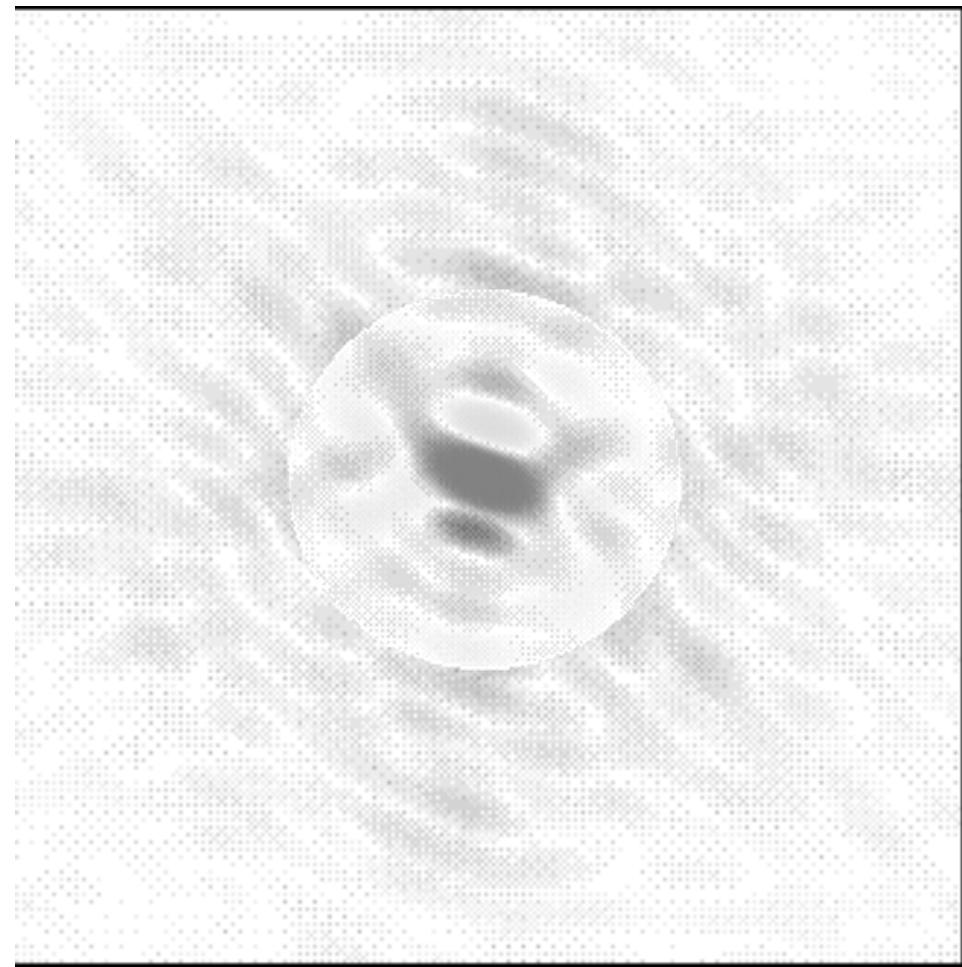
Molecular Replacement

correct structure and intensities



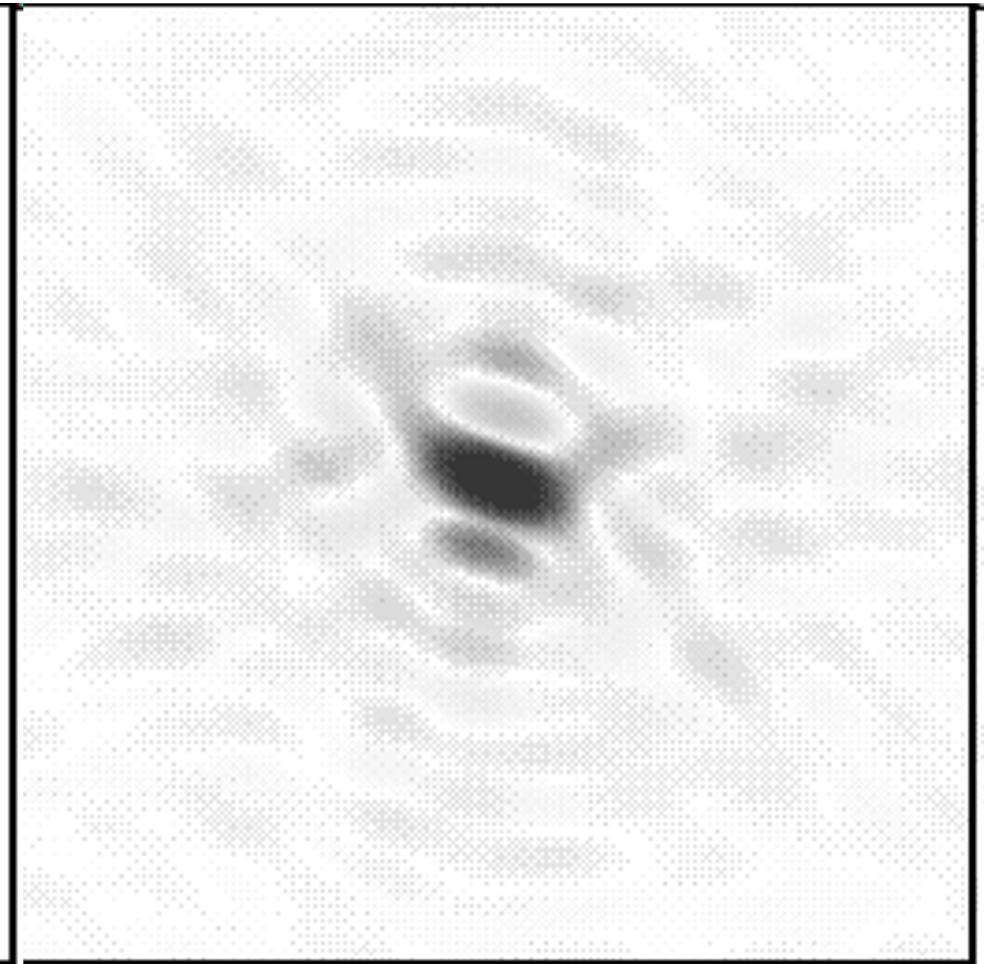
Molecular Replacement

use something similar as a starting model



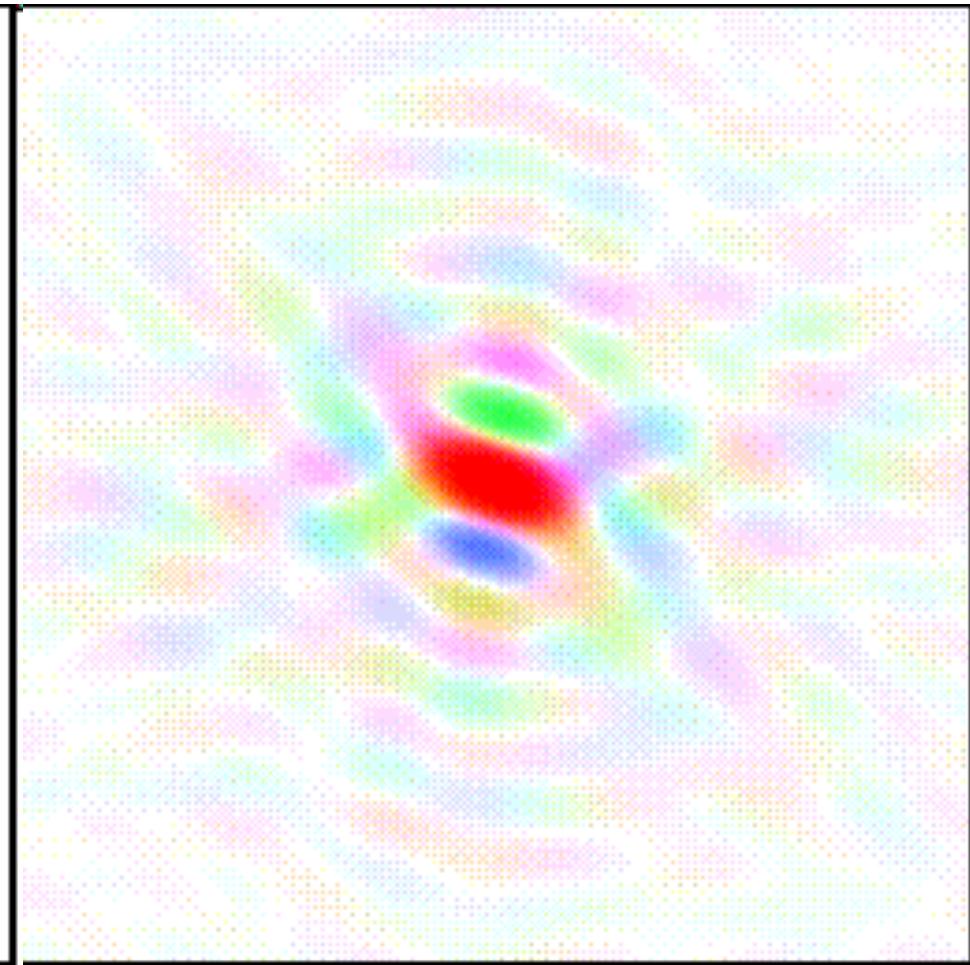
Model Building

current model is missing something



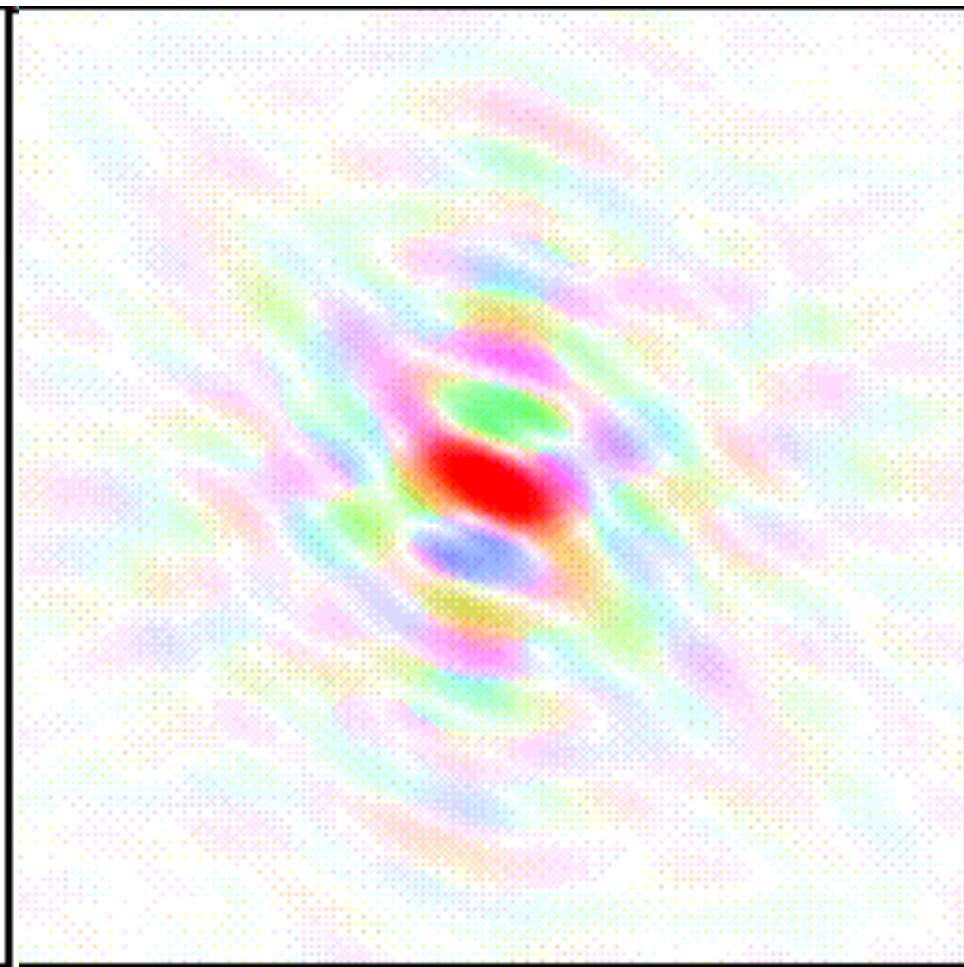
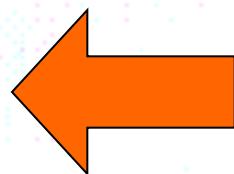
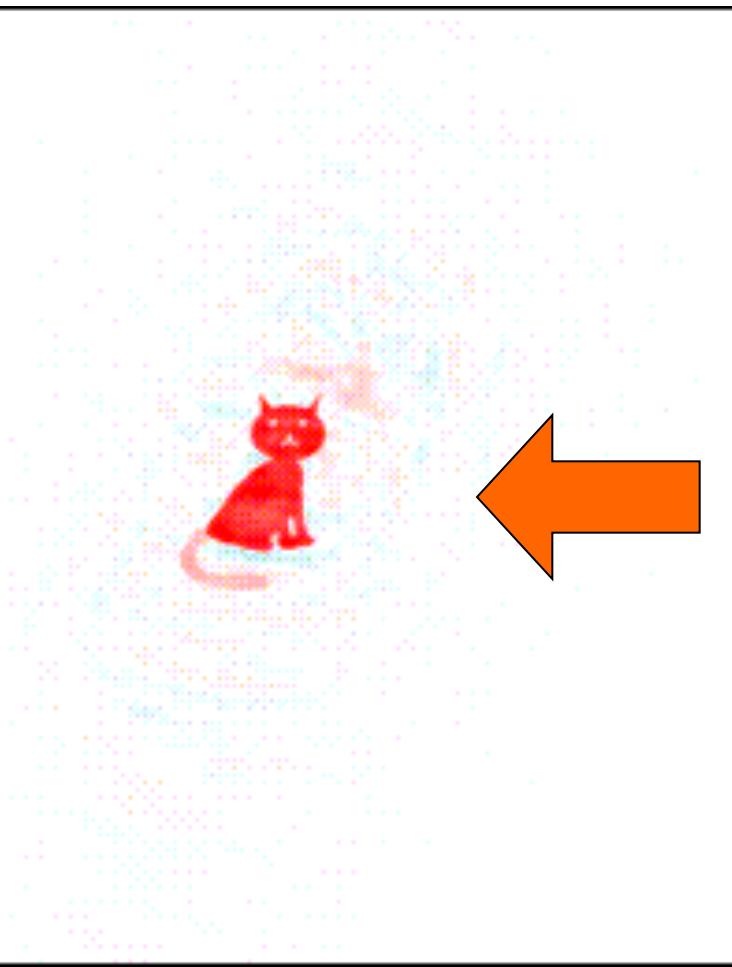
Model Building

phases from available model



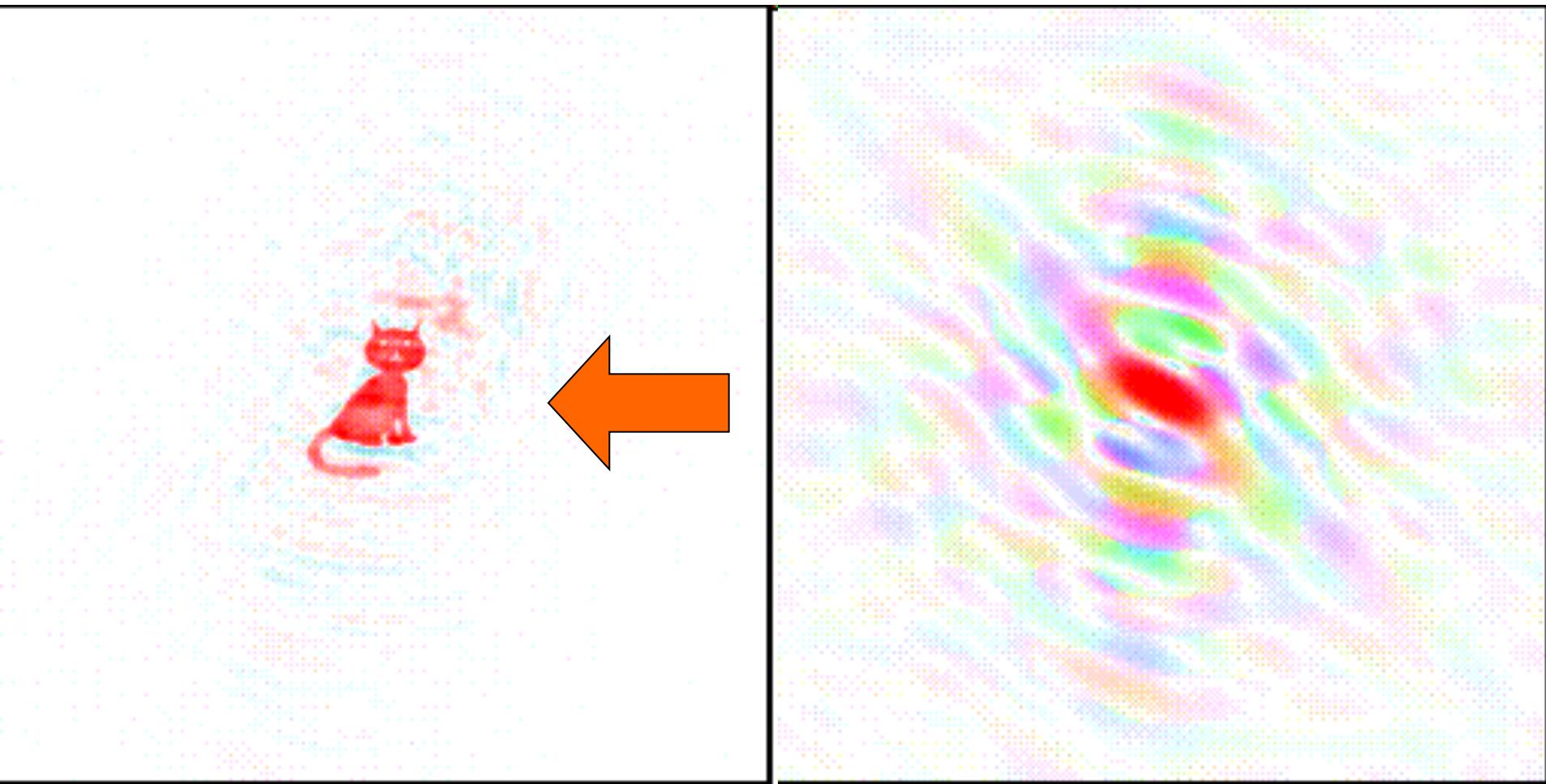
Model Building

missing bits show up, even if not in available model!

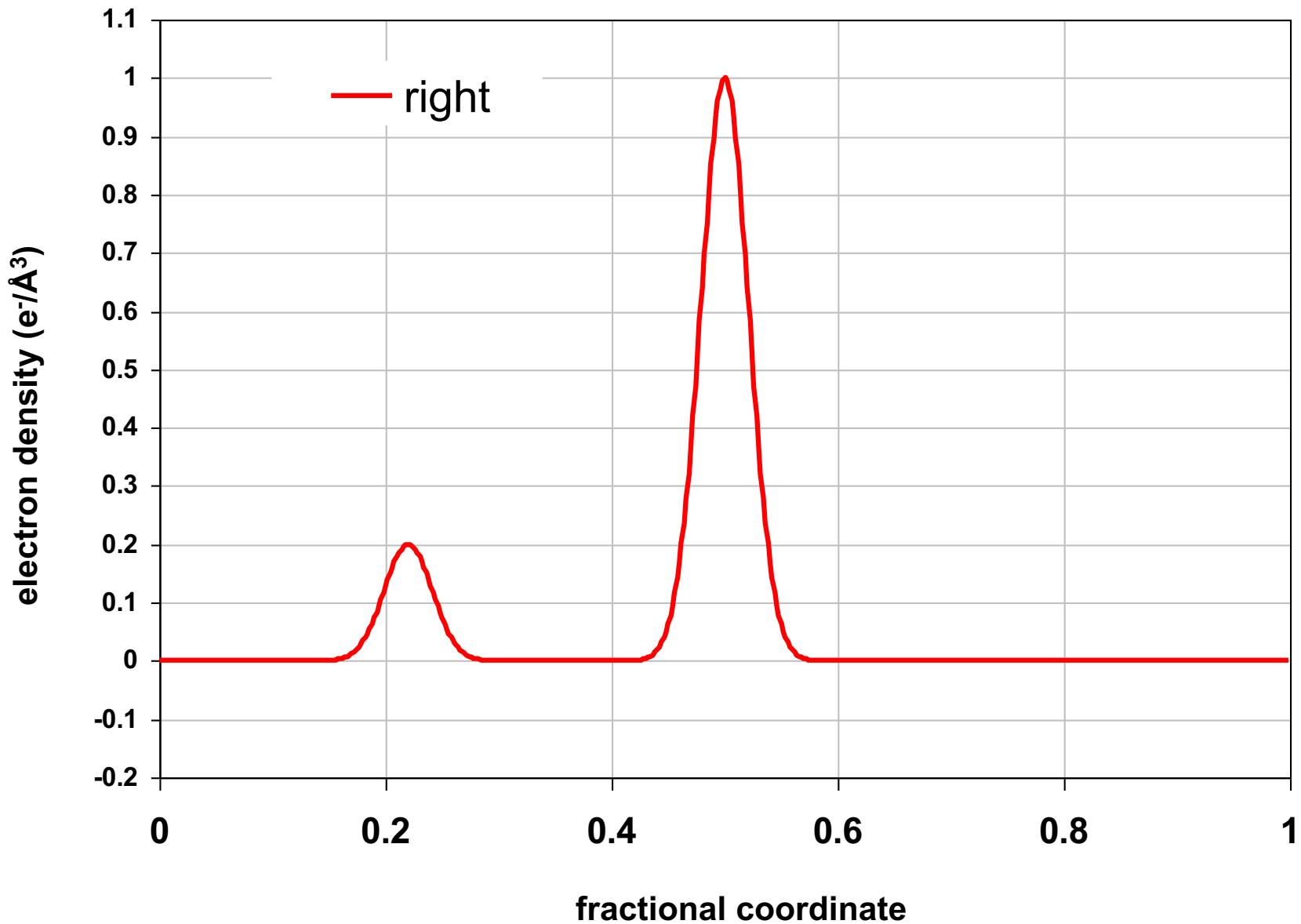


Model Building

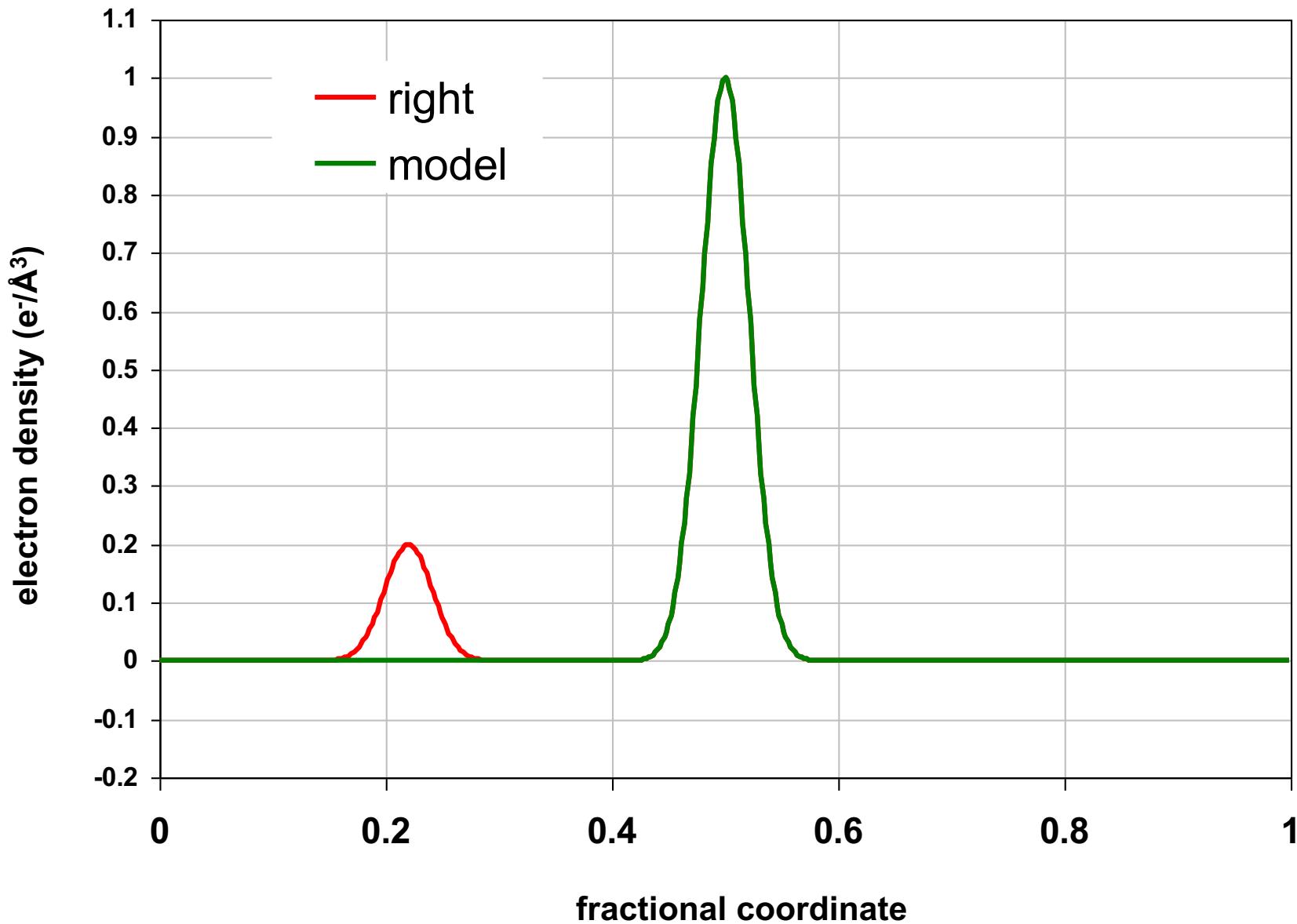
missing bits show up better in $F_{\text{obs}} + (F_{\text{obs}} - F_{\text{calc}})$ map



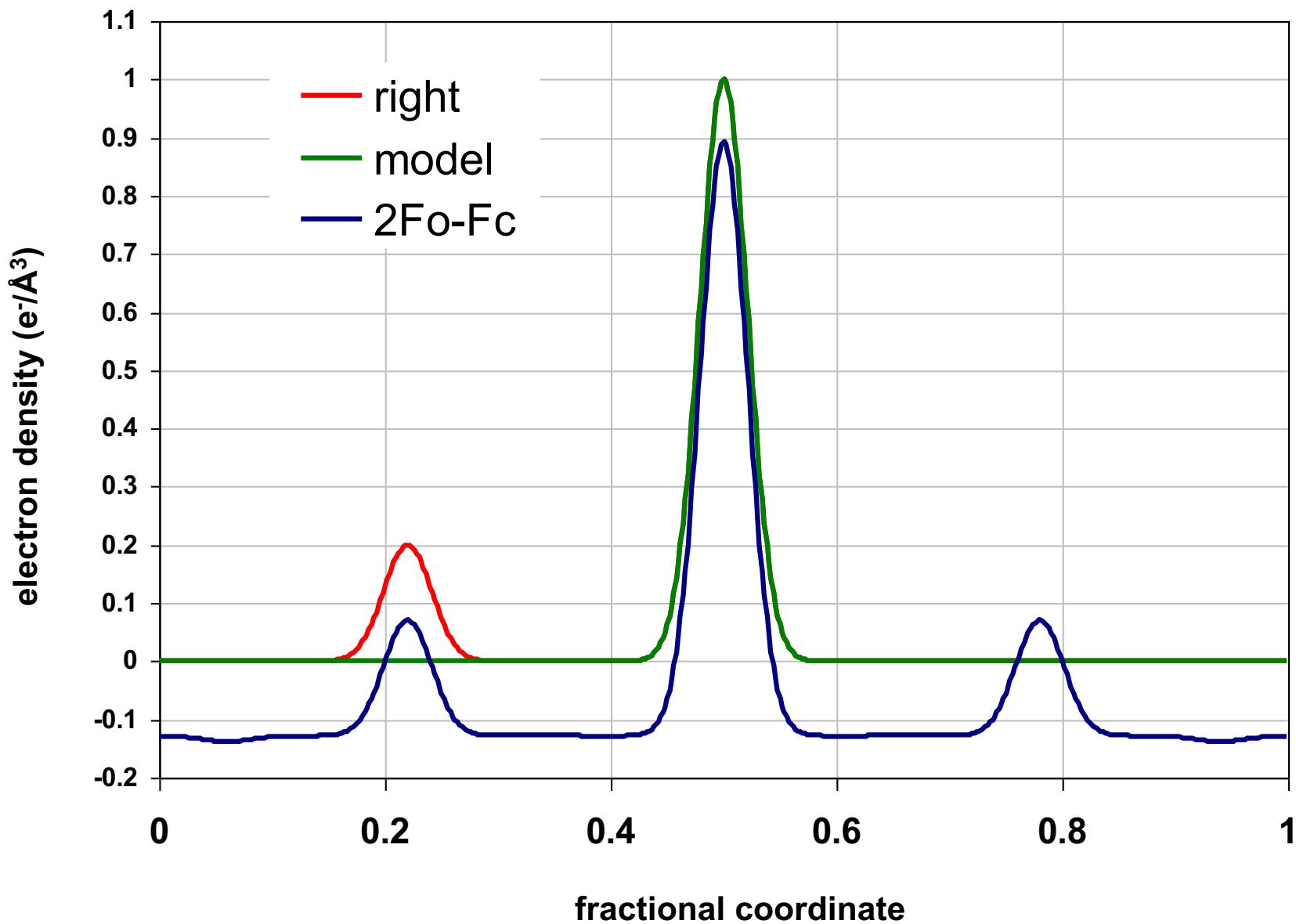
Phase errors: simplest possible case



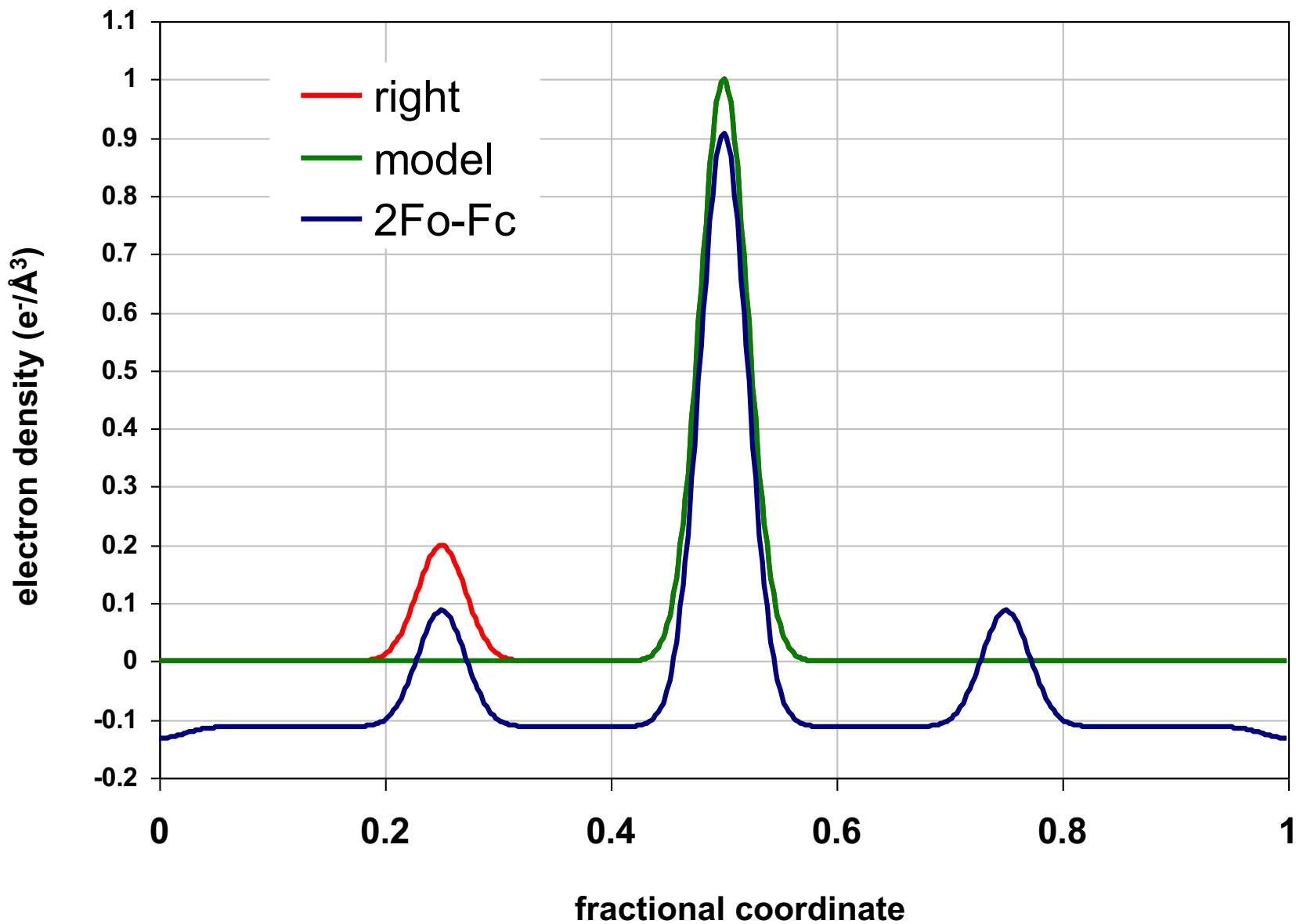
Phase errors: simplest possible case



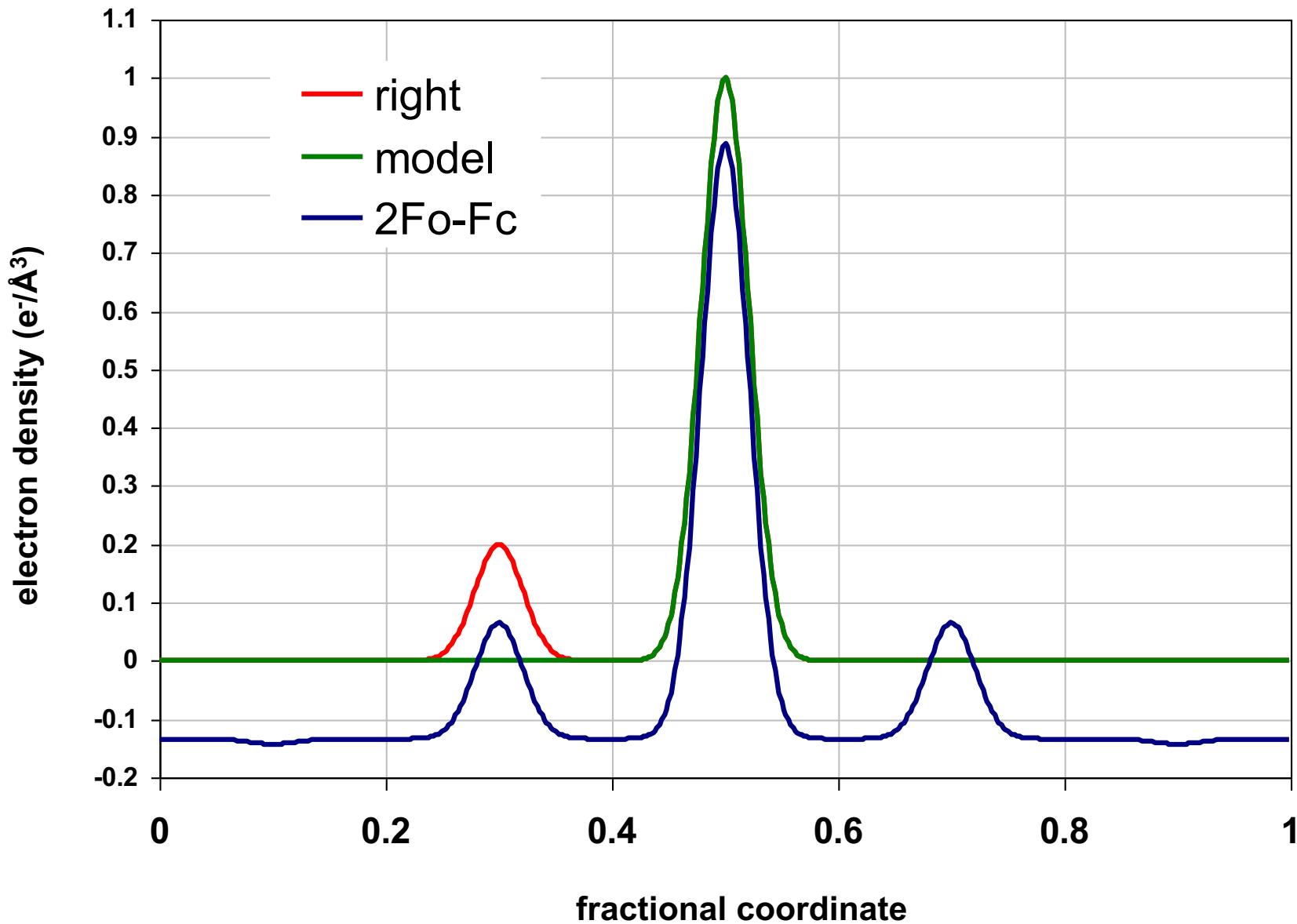
Phase errors: simplest possible case



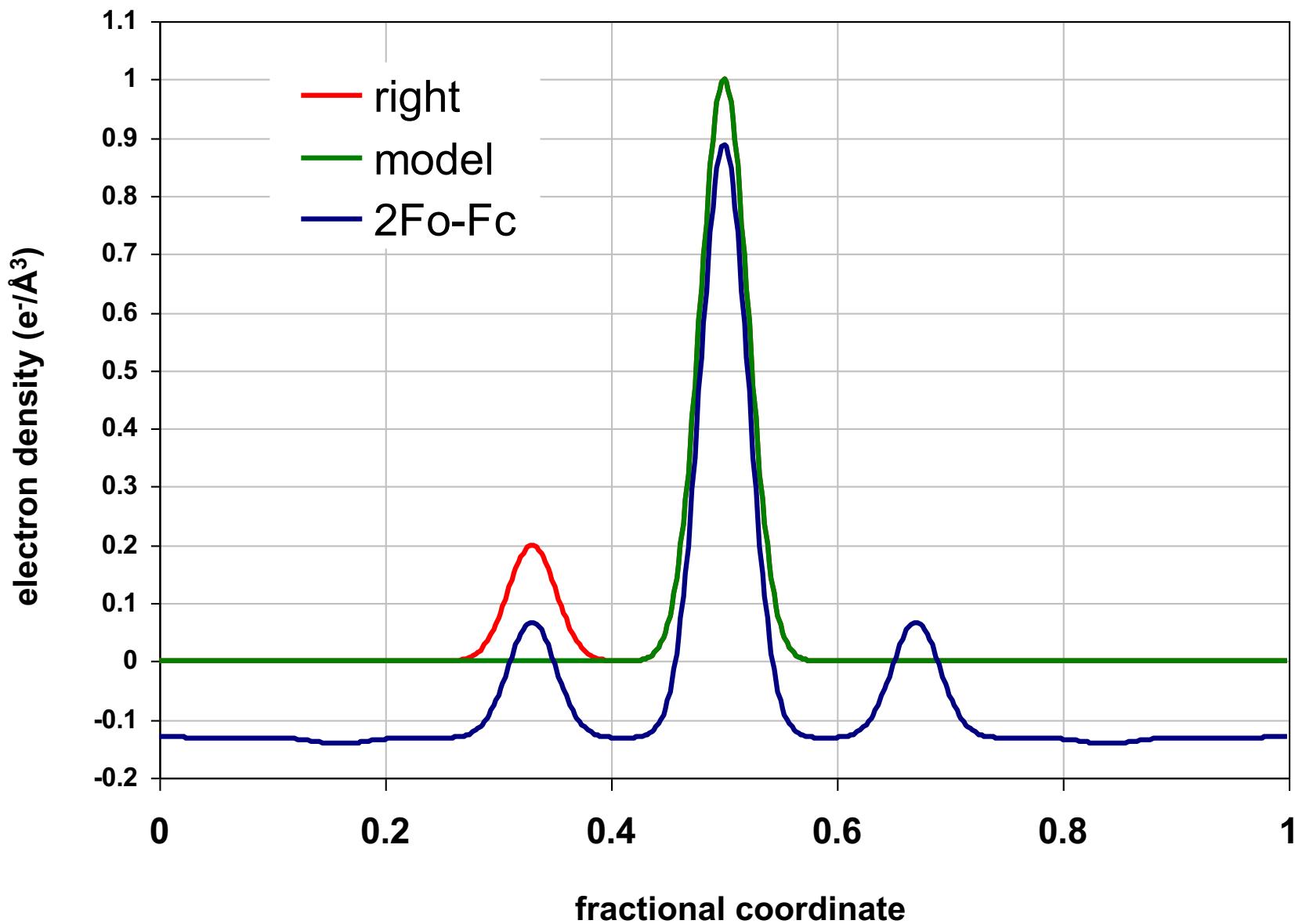
Phase errors: simplest possible case



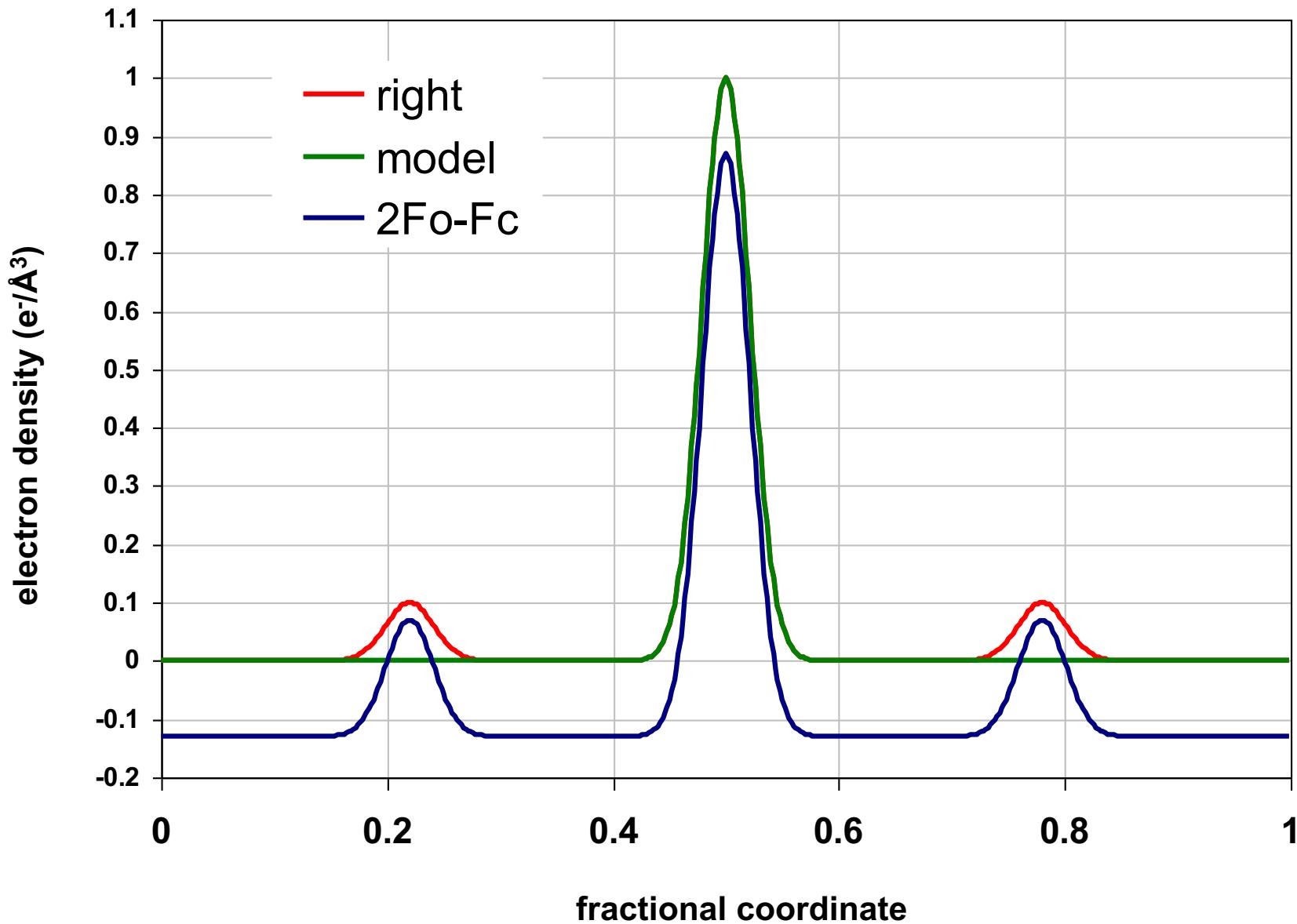
Phase errors: simplest possible case



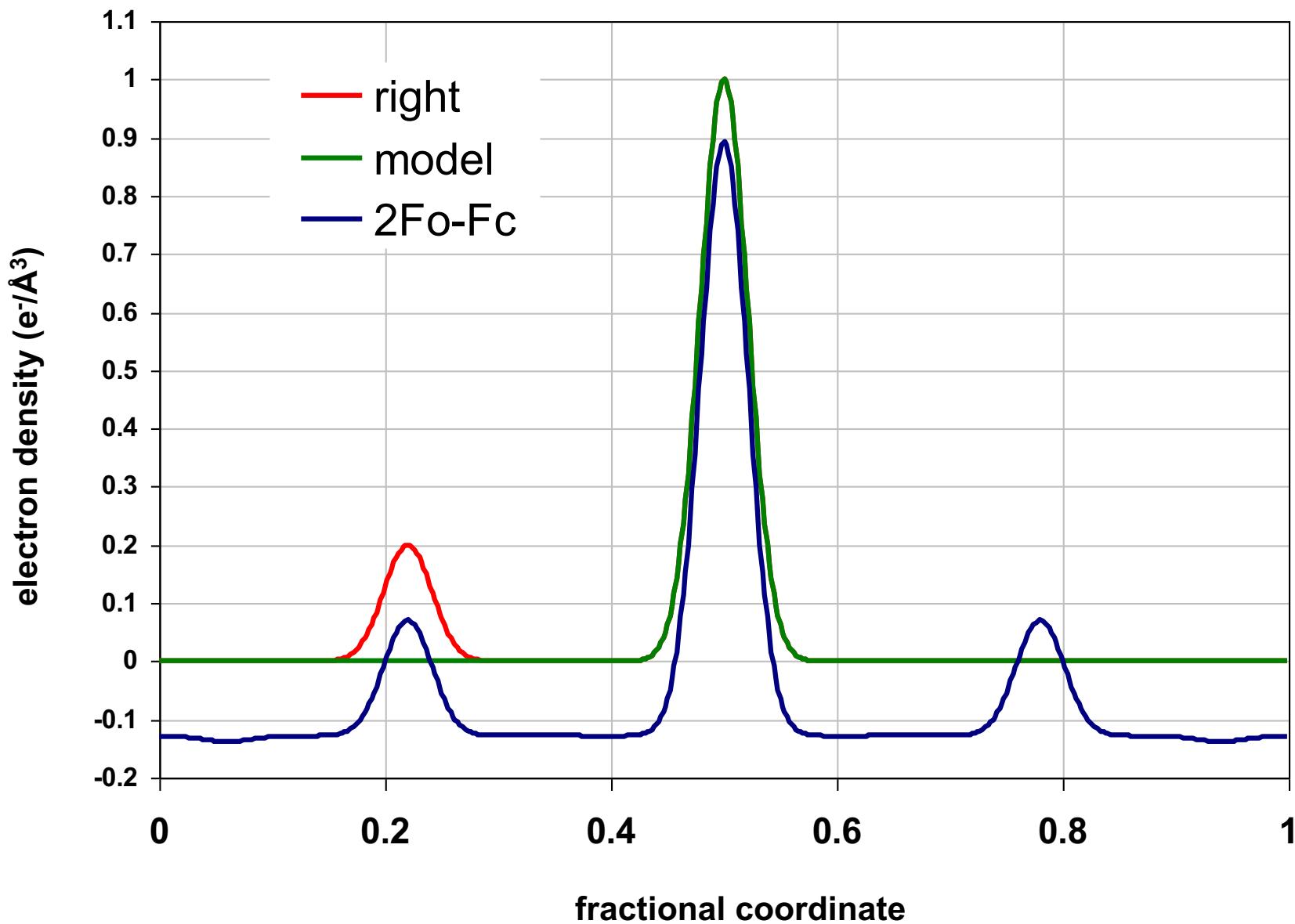
Phase errors: simplest possible case



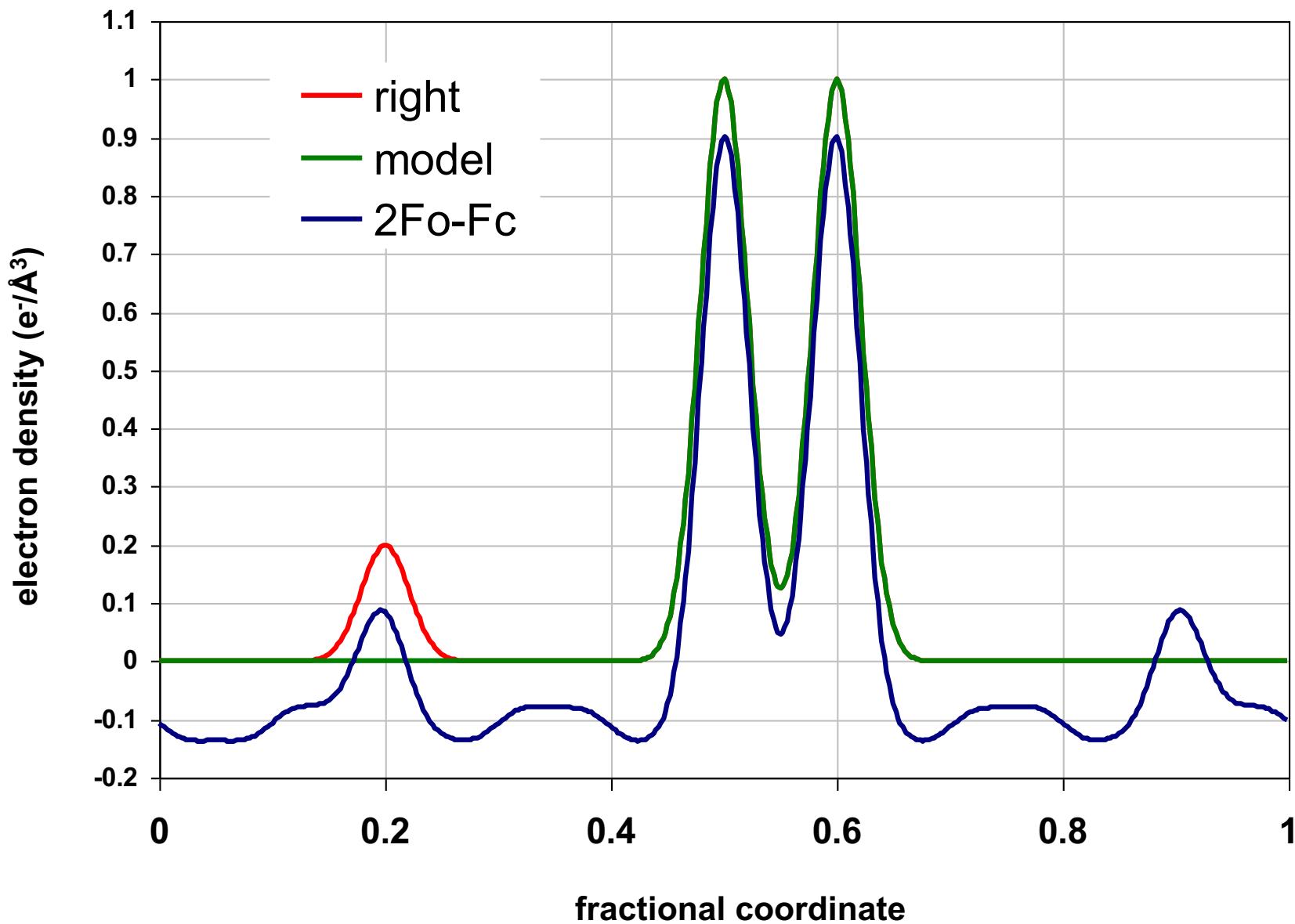
Phase errors: simplest possible case



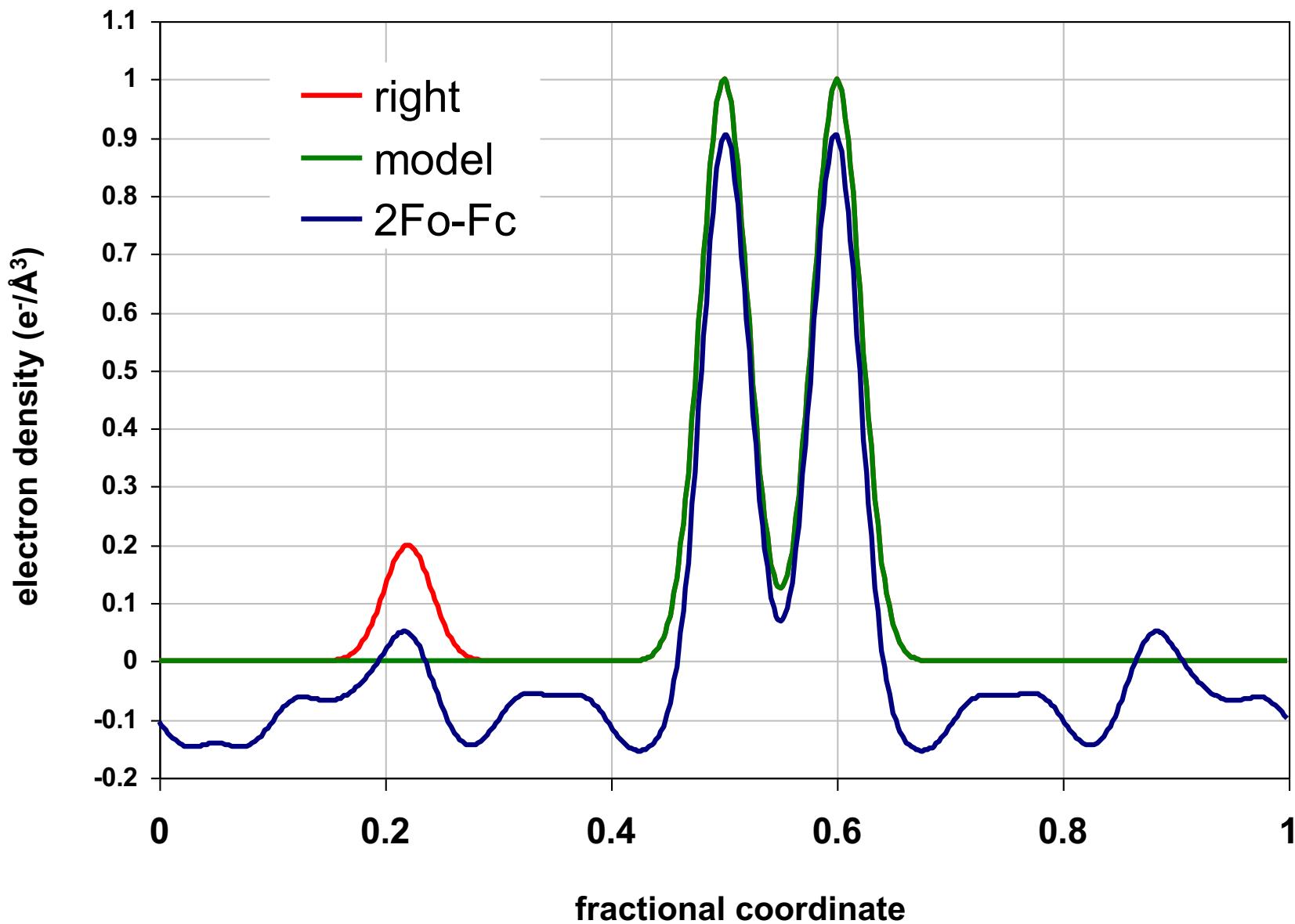
Phase errors: simplest possible case



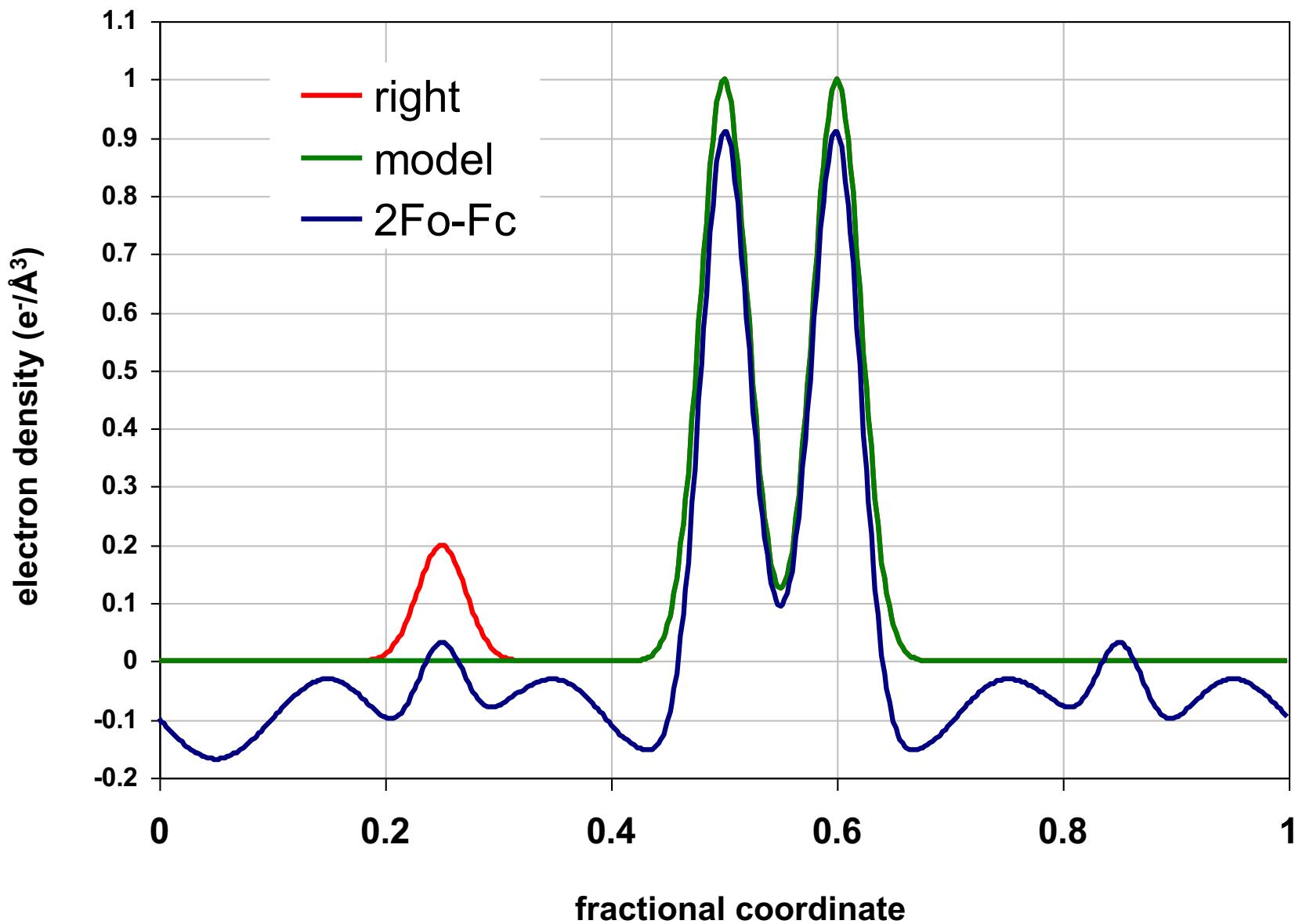
Phase errors: 3-atom case



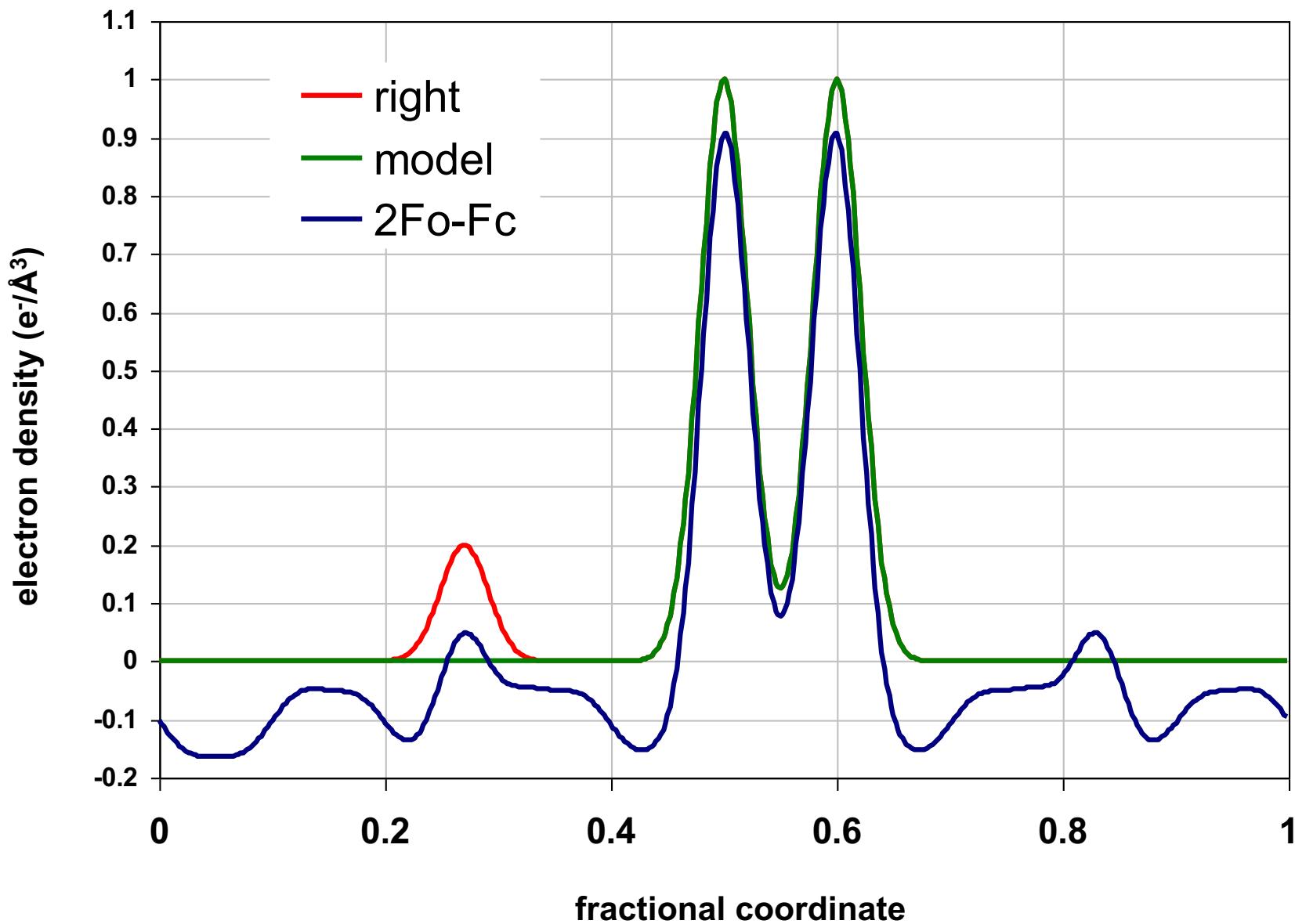
Phase errors: 3-atom case



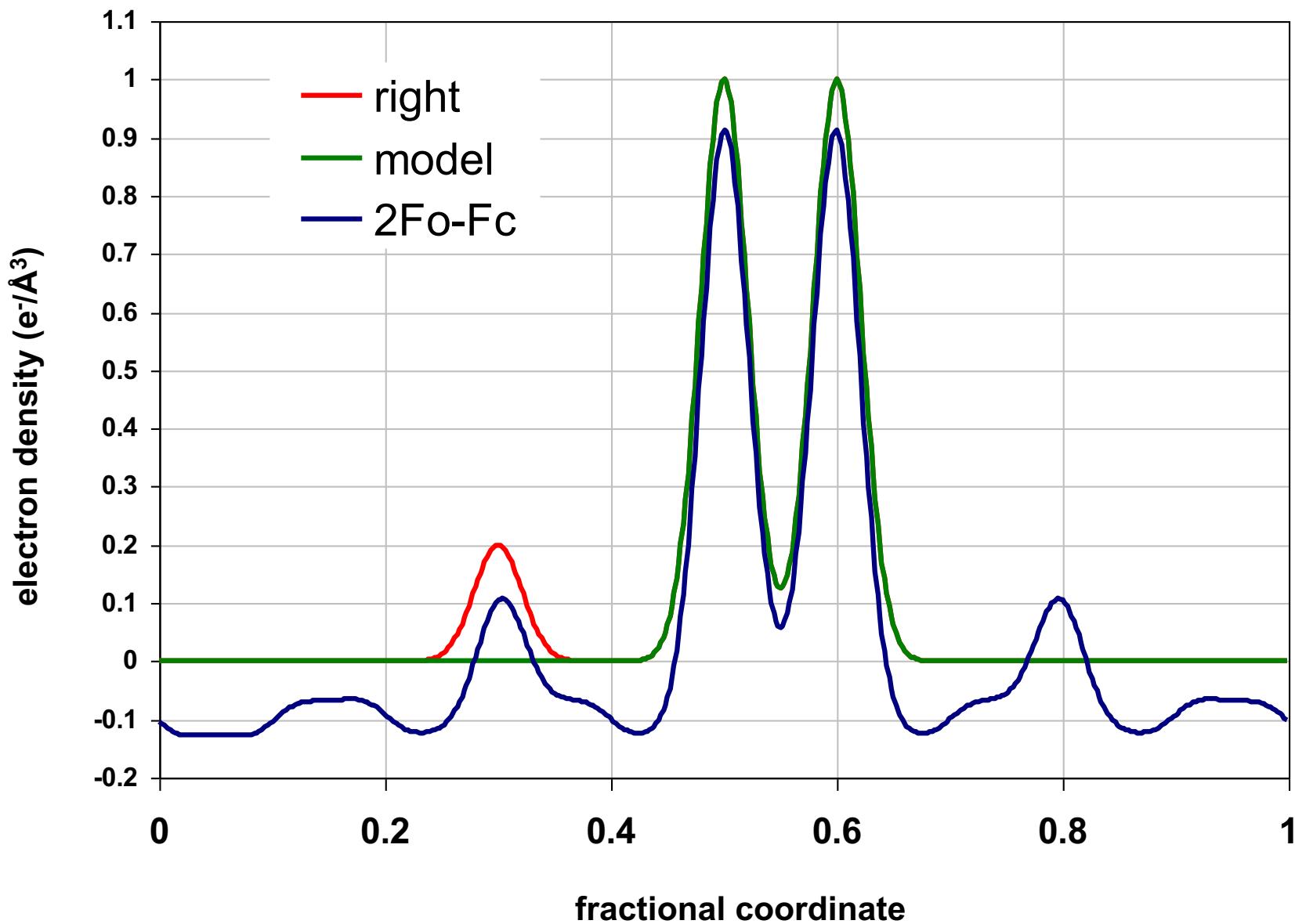
Phase errors: 3-atom case



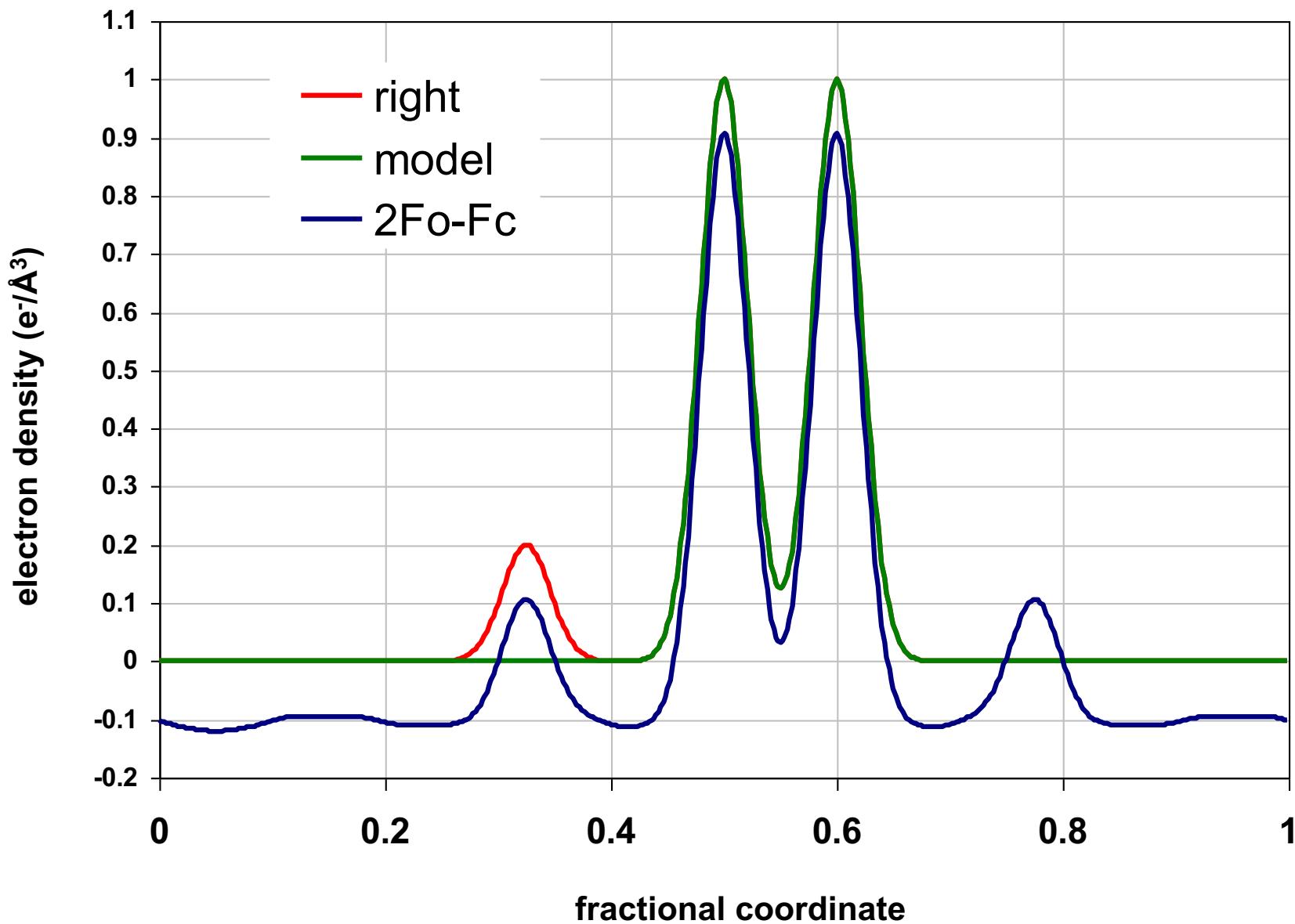
Phase errors: 3-atom case



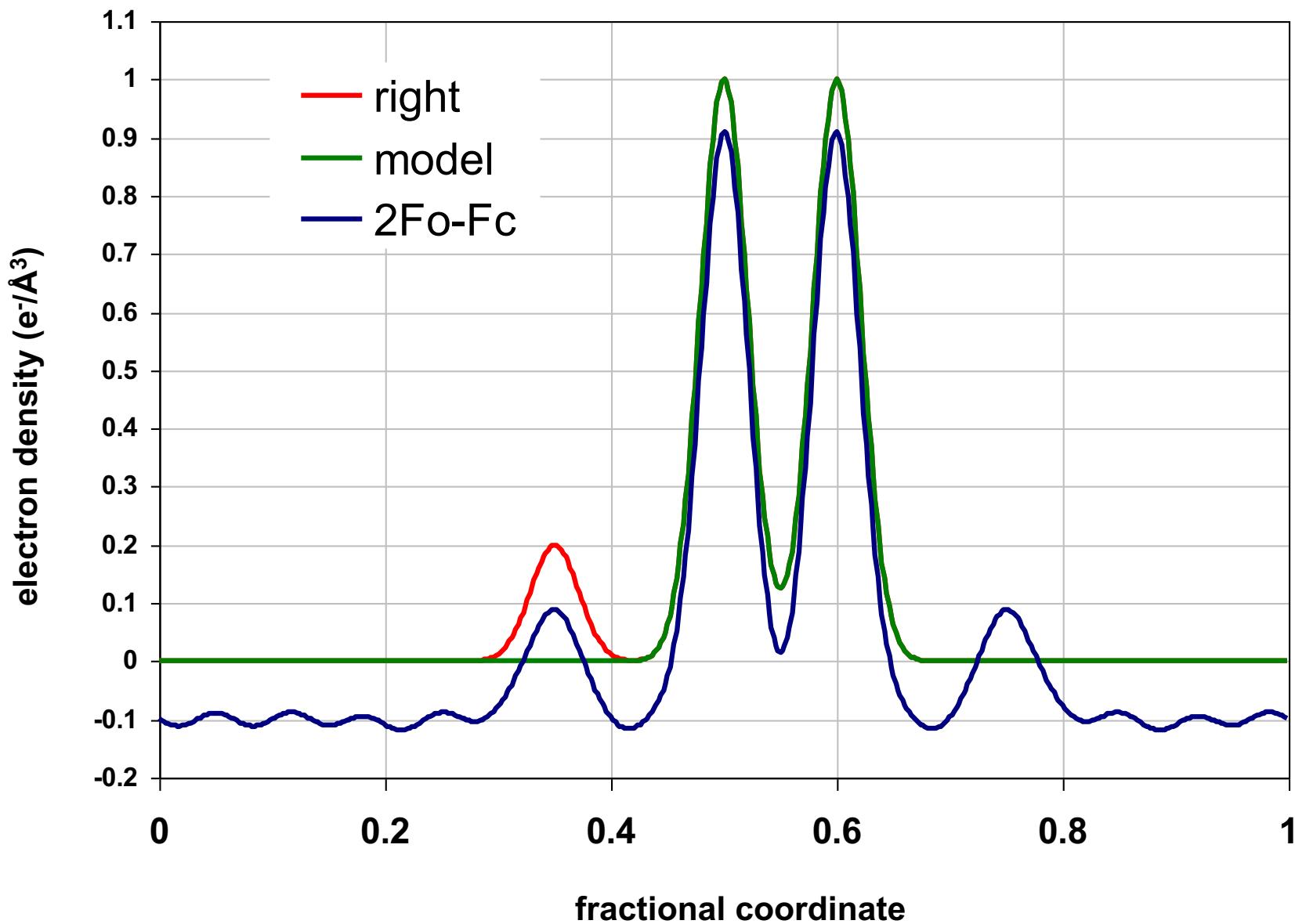
Phase errors: 3-atom case



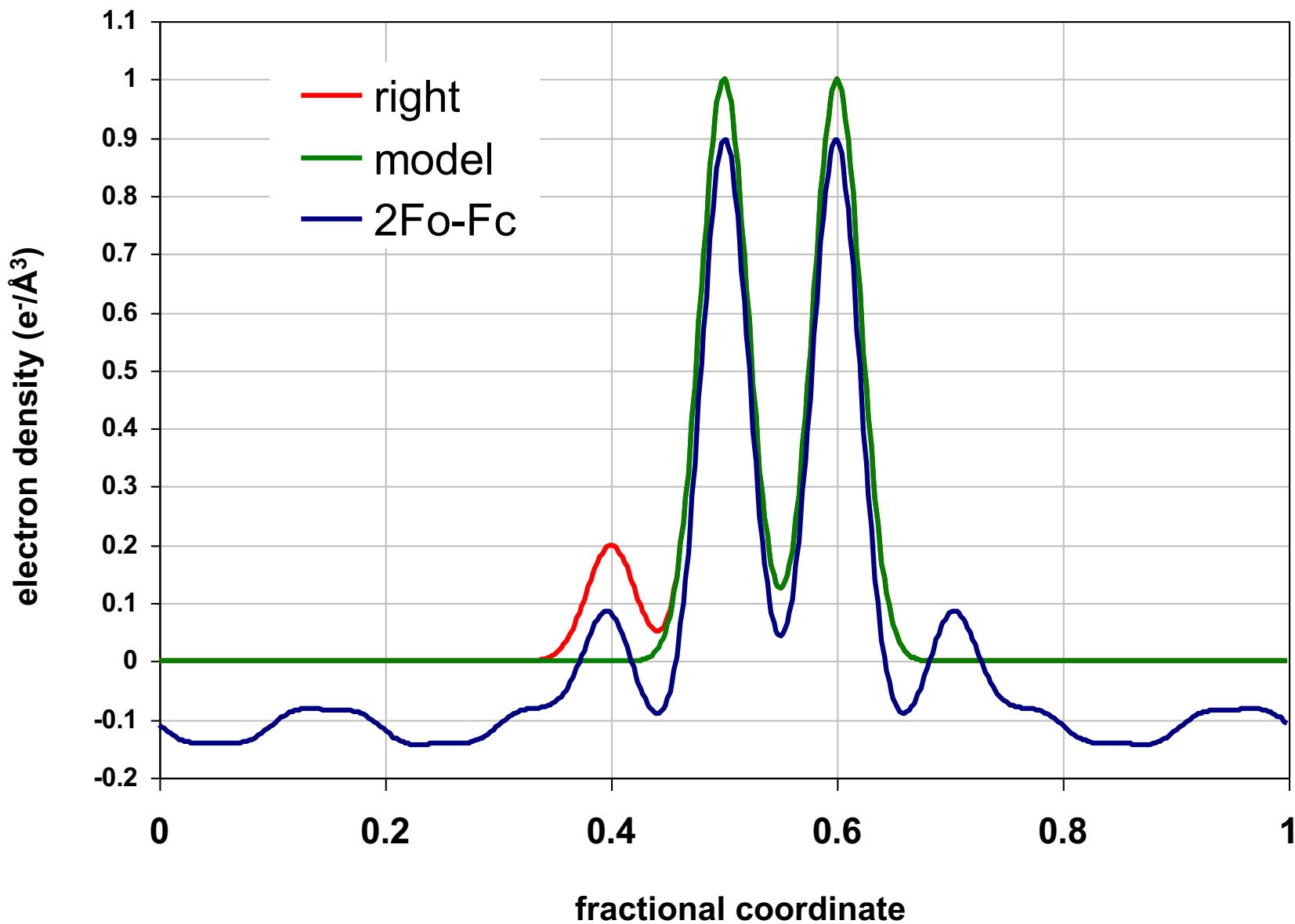
Phase errors: 3-atom case



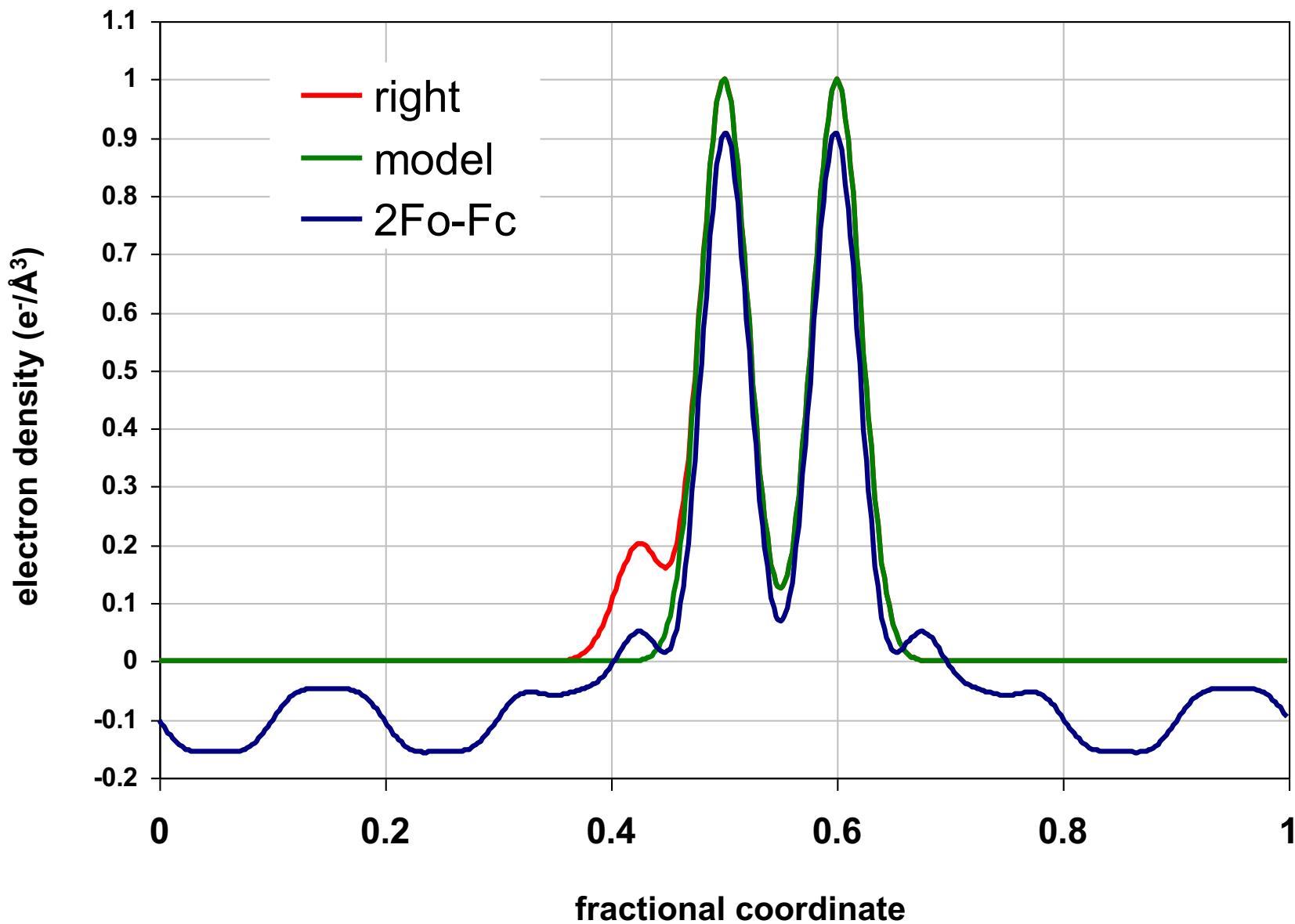
Phase errors: 3-atom case



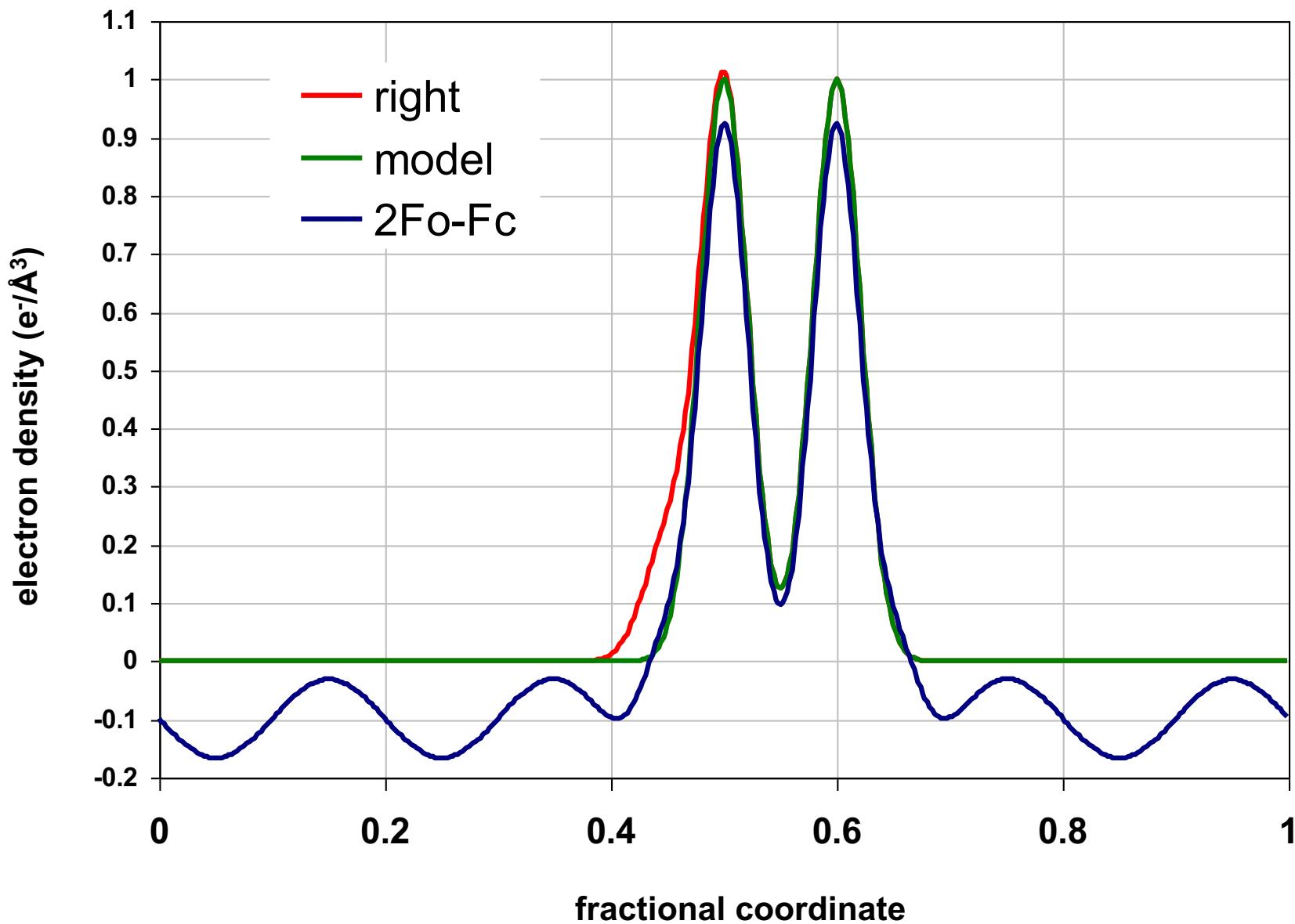
Phase errors: 3-atom case



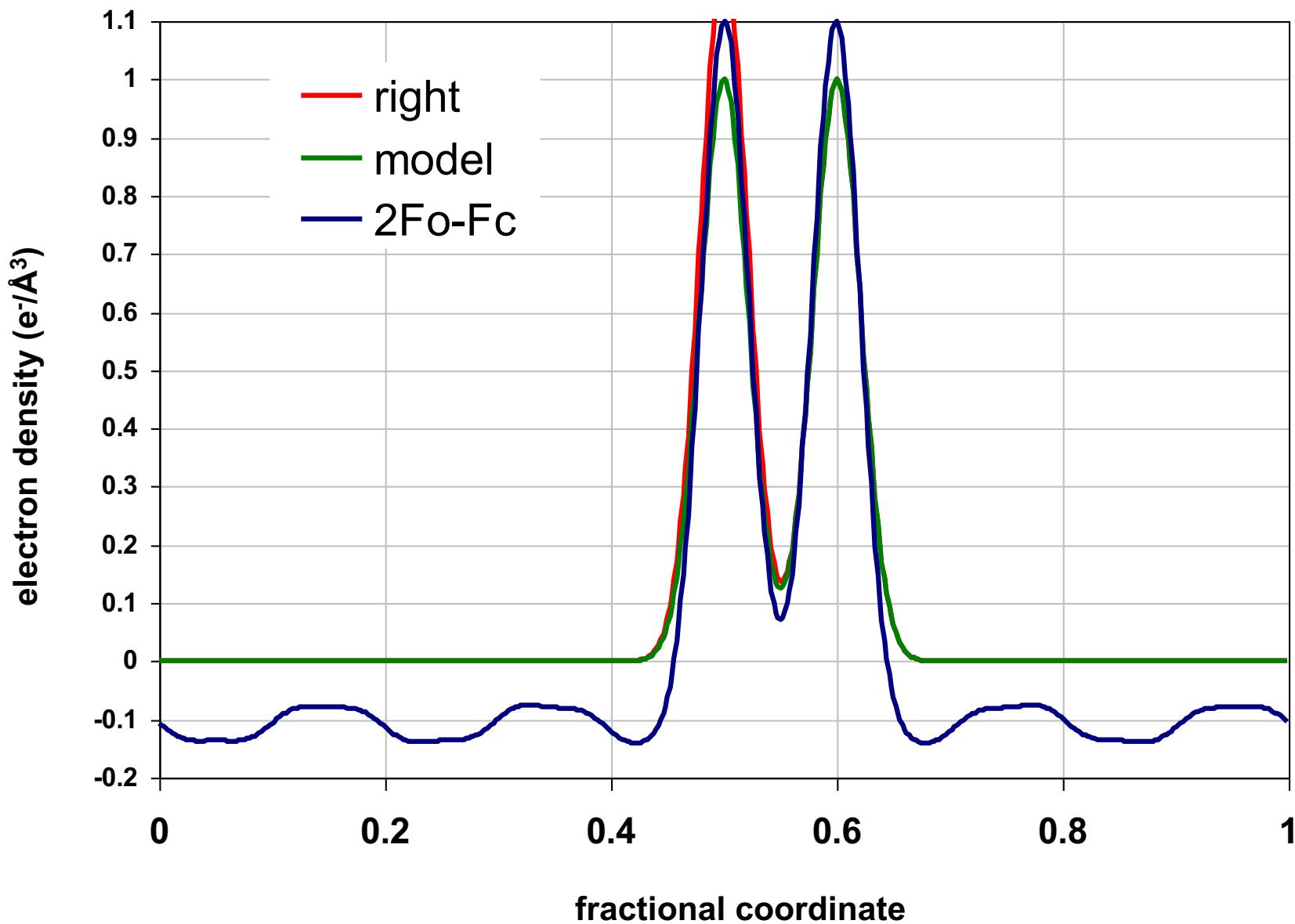
Phase errors: 3-atom case



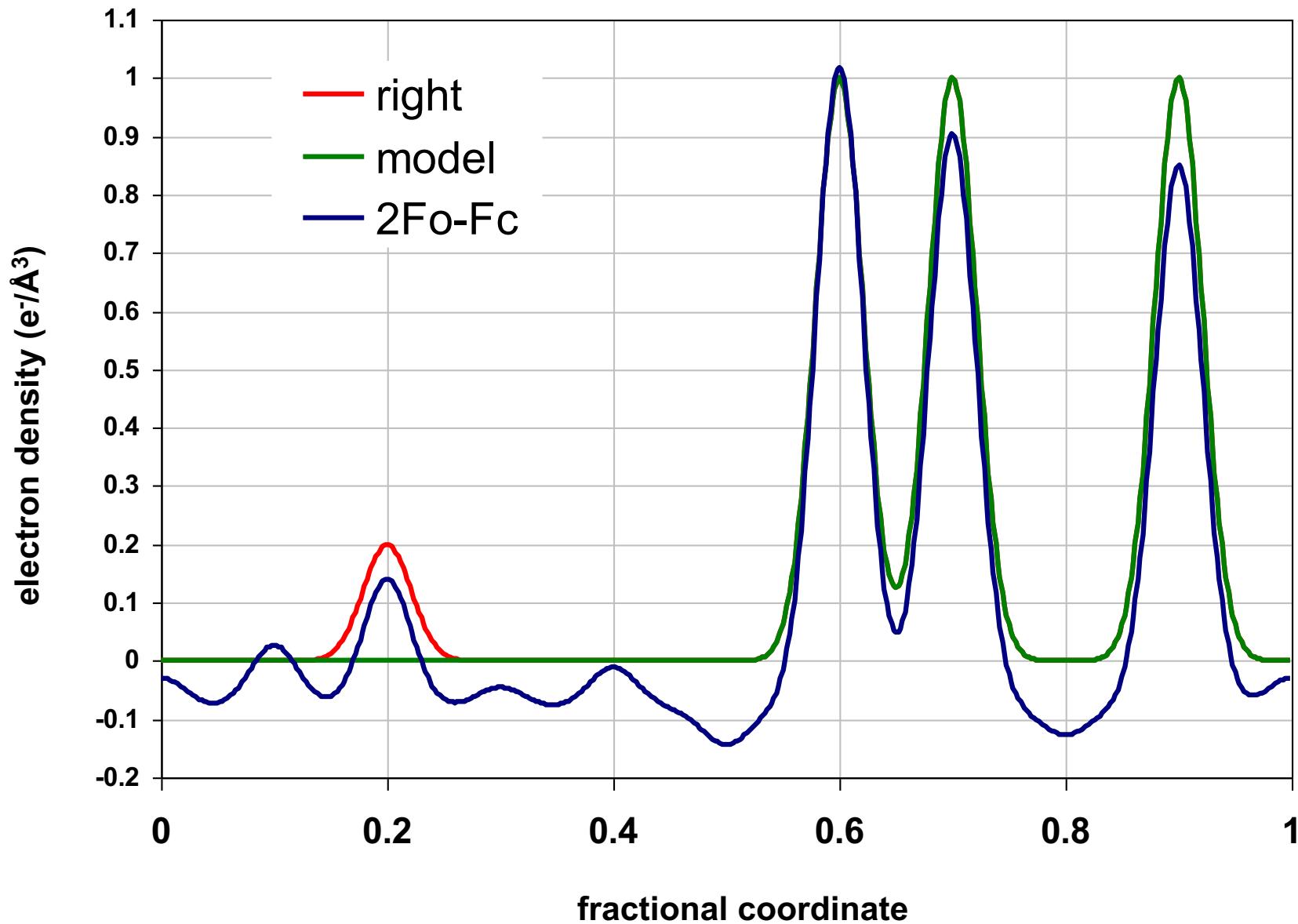
Phase errors: 3-atom case



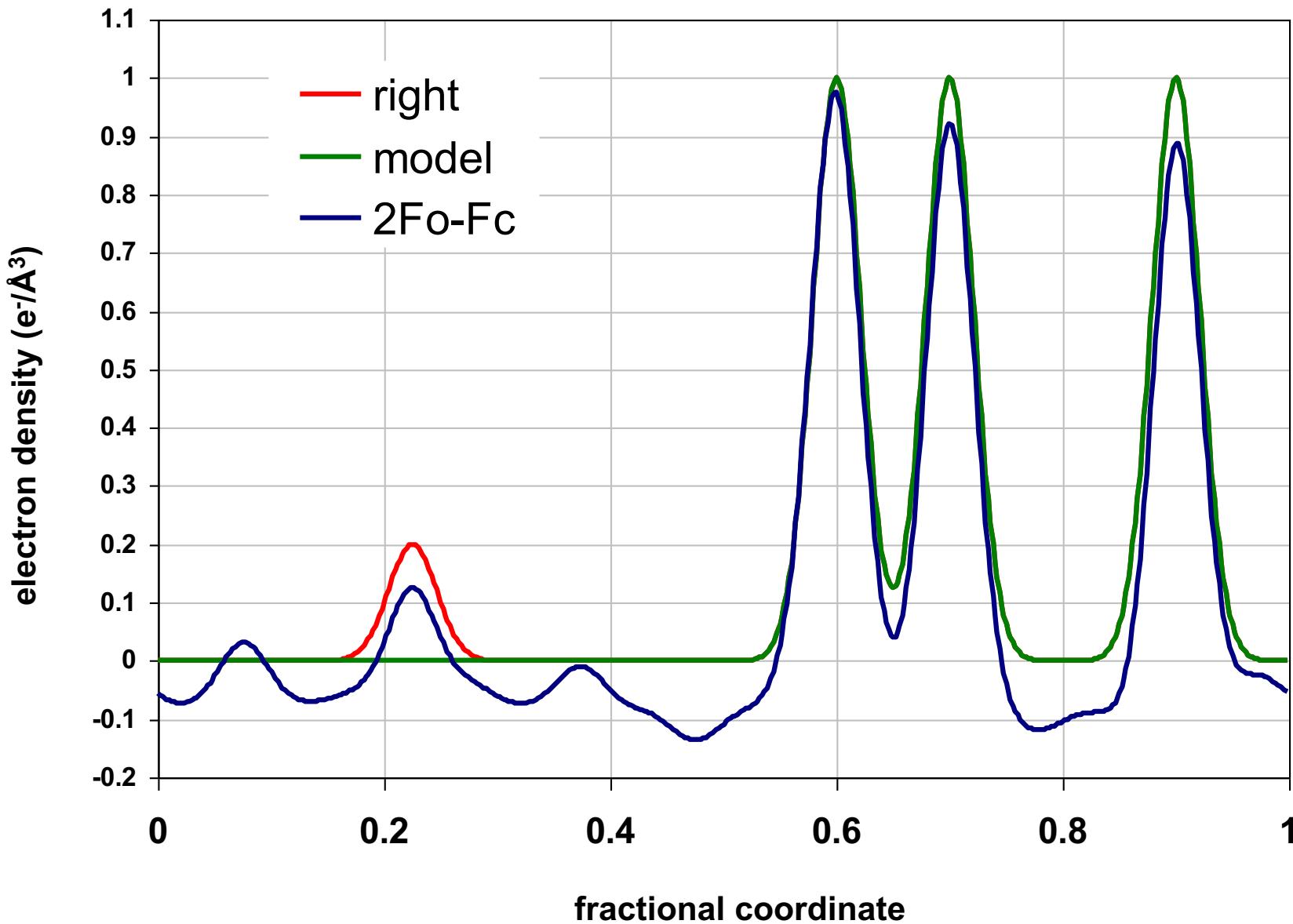
Phase errors: 3-atom case



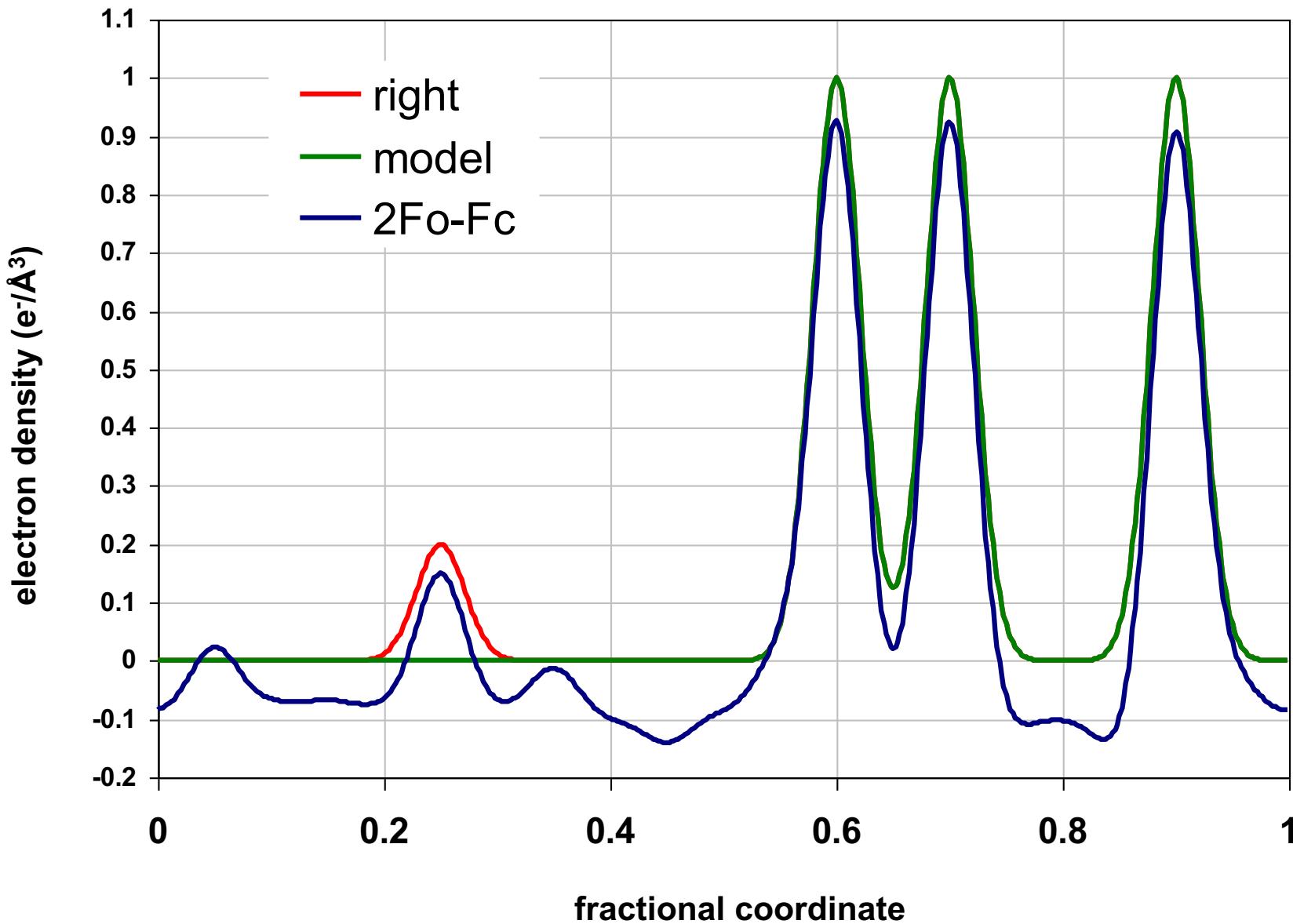
Phase errors: asymmetric case



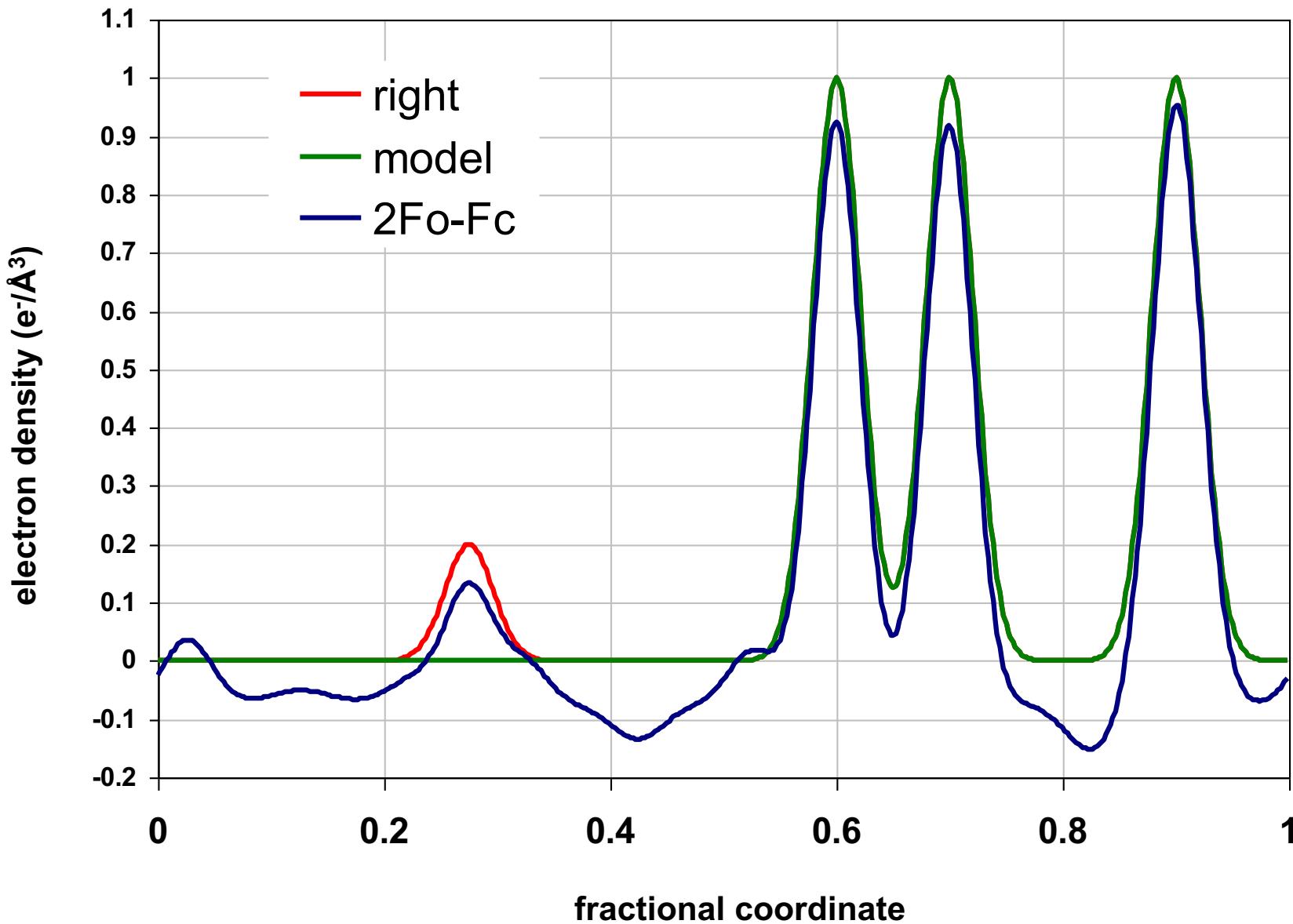
Phase errors: asymmetric case



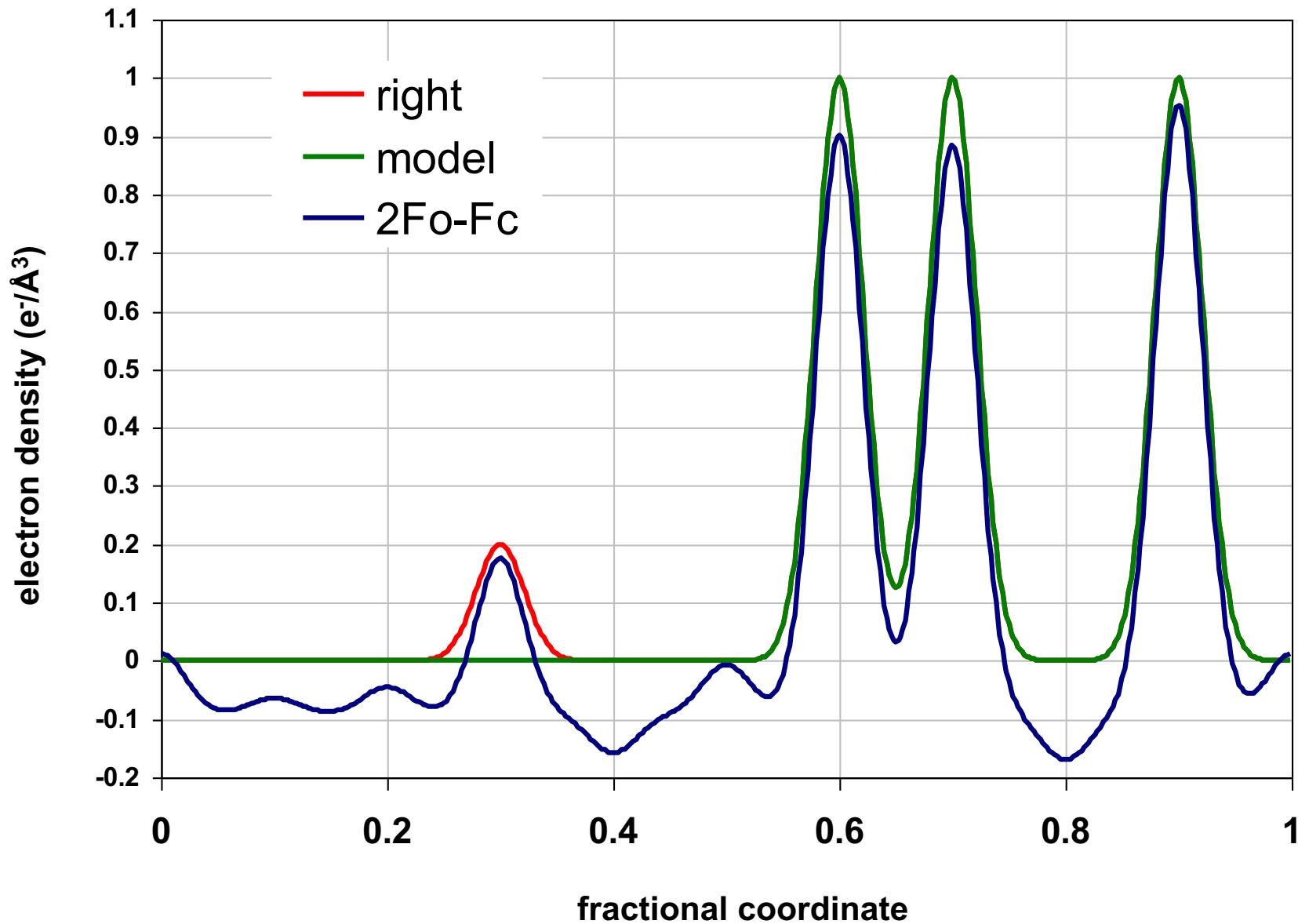
Phase errors: asymmetric case



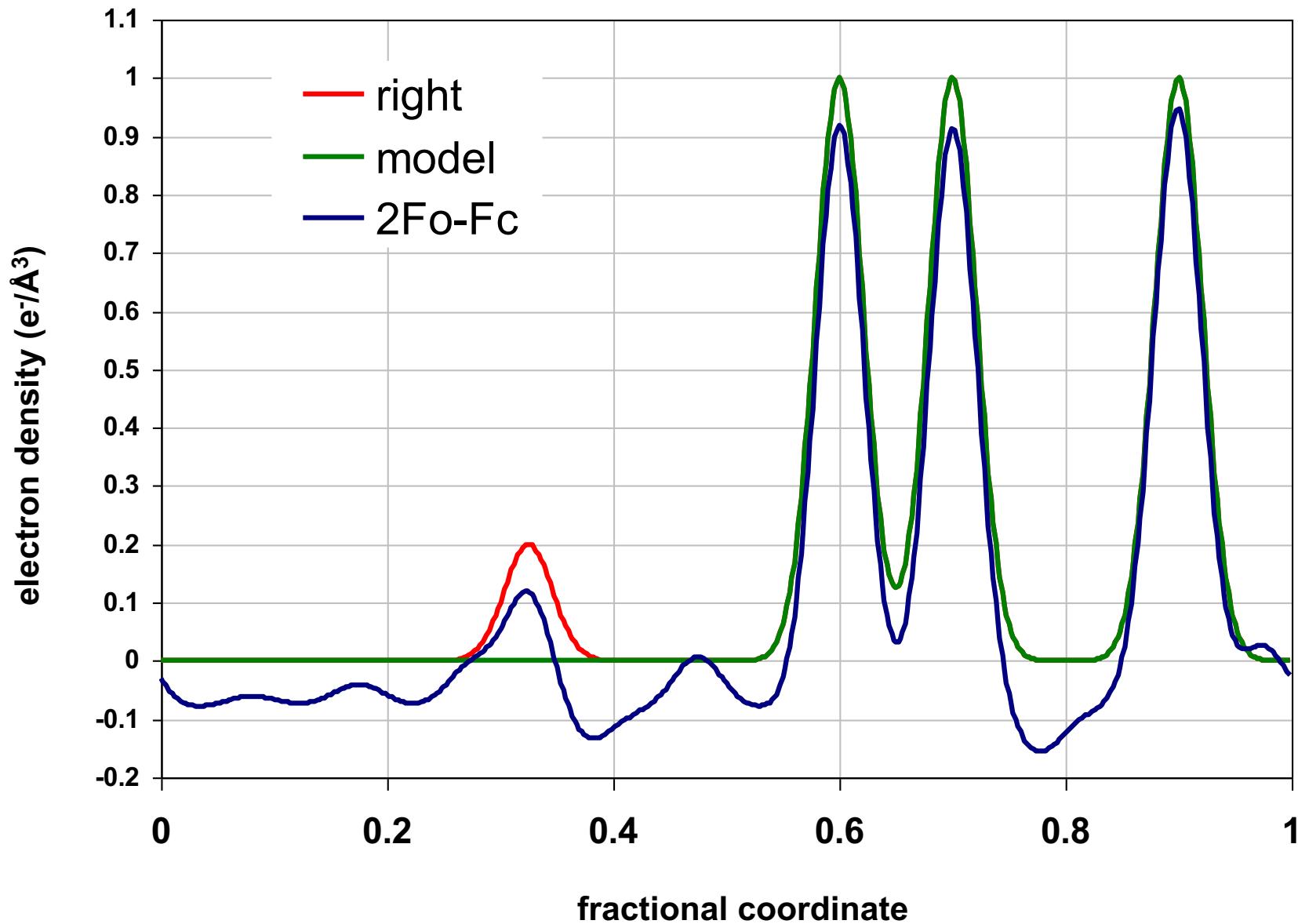
Phase errors: asymmetric case



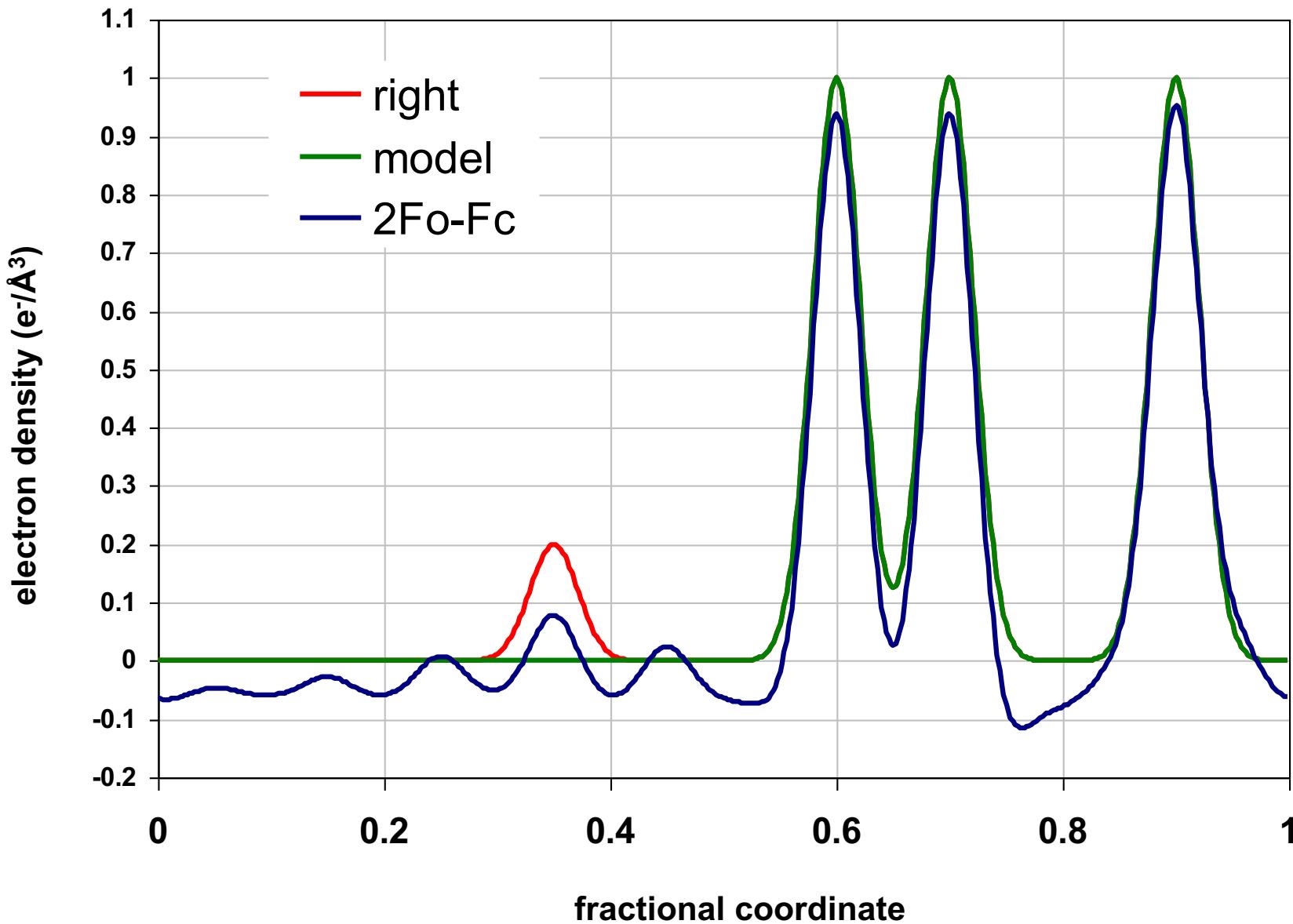
Phase errors: asymmetric case



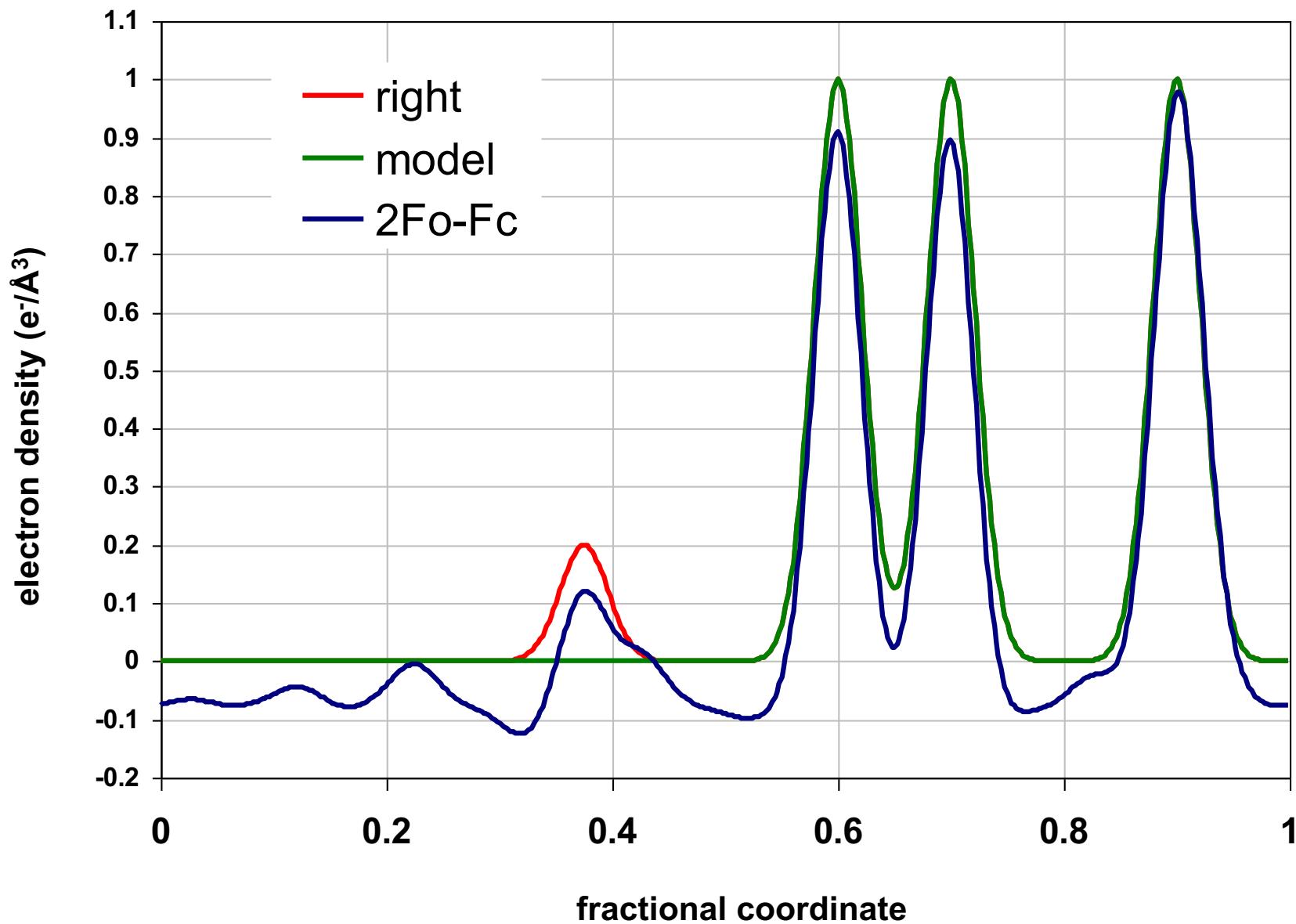
Phase errors: asymmetric case



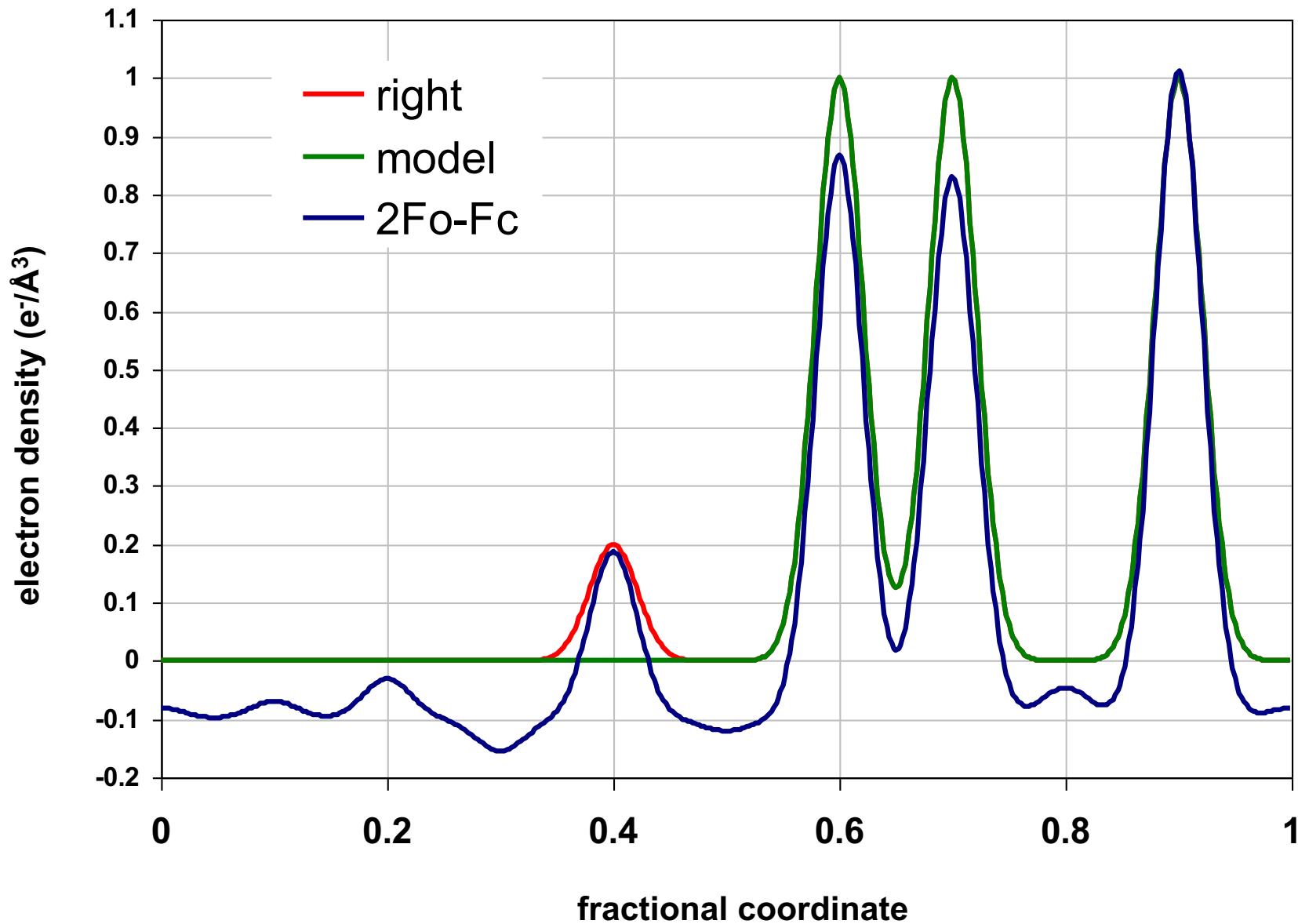
Phase errors: asymmetric case



Phase errors: asymmetric case



Phase errors: asymmetric case

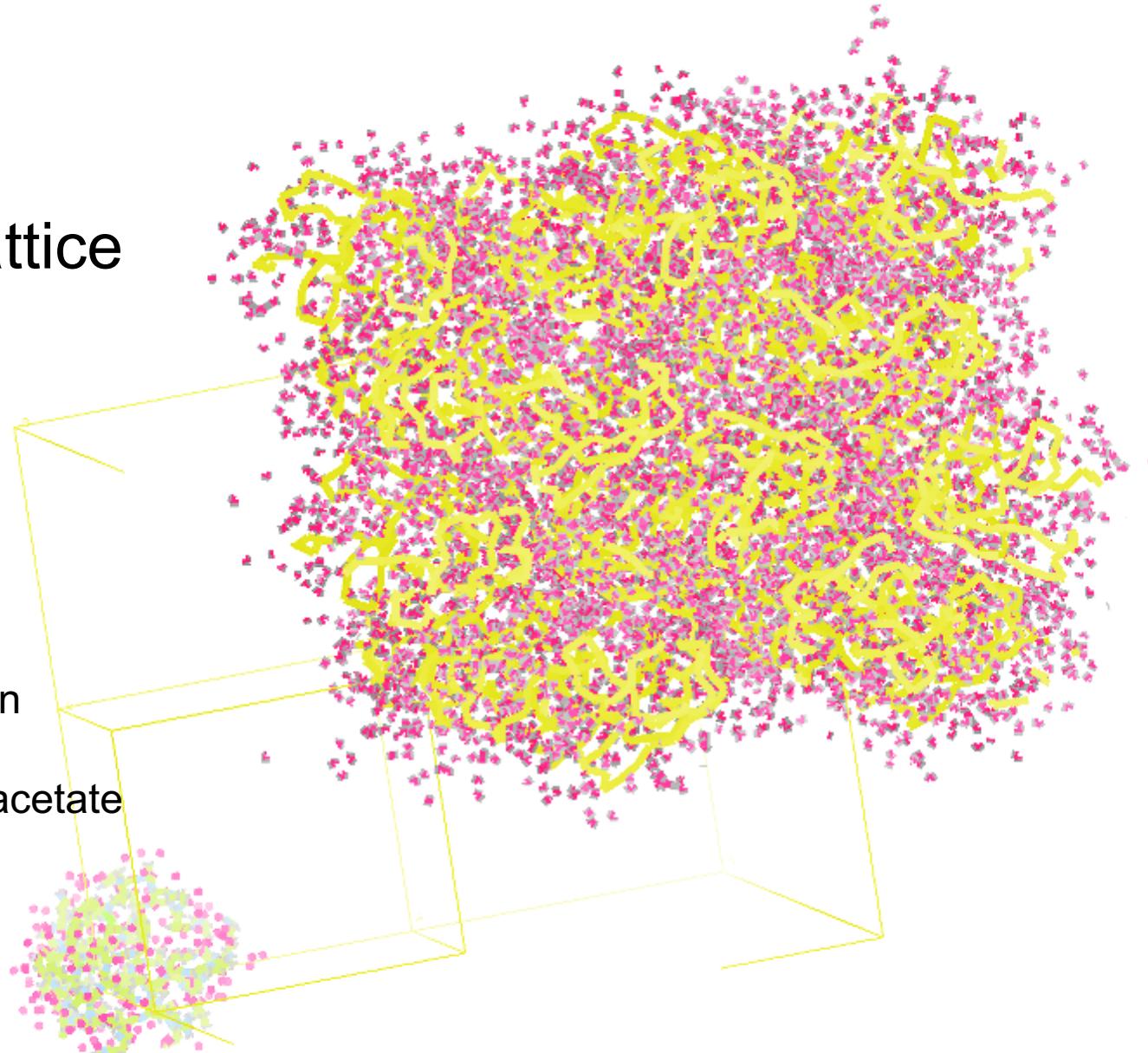


Molecular Dynamics Simulation

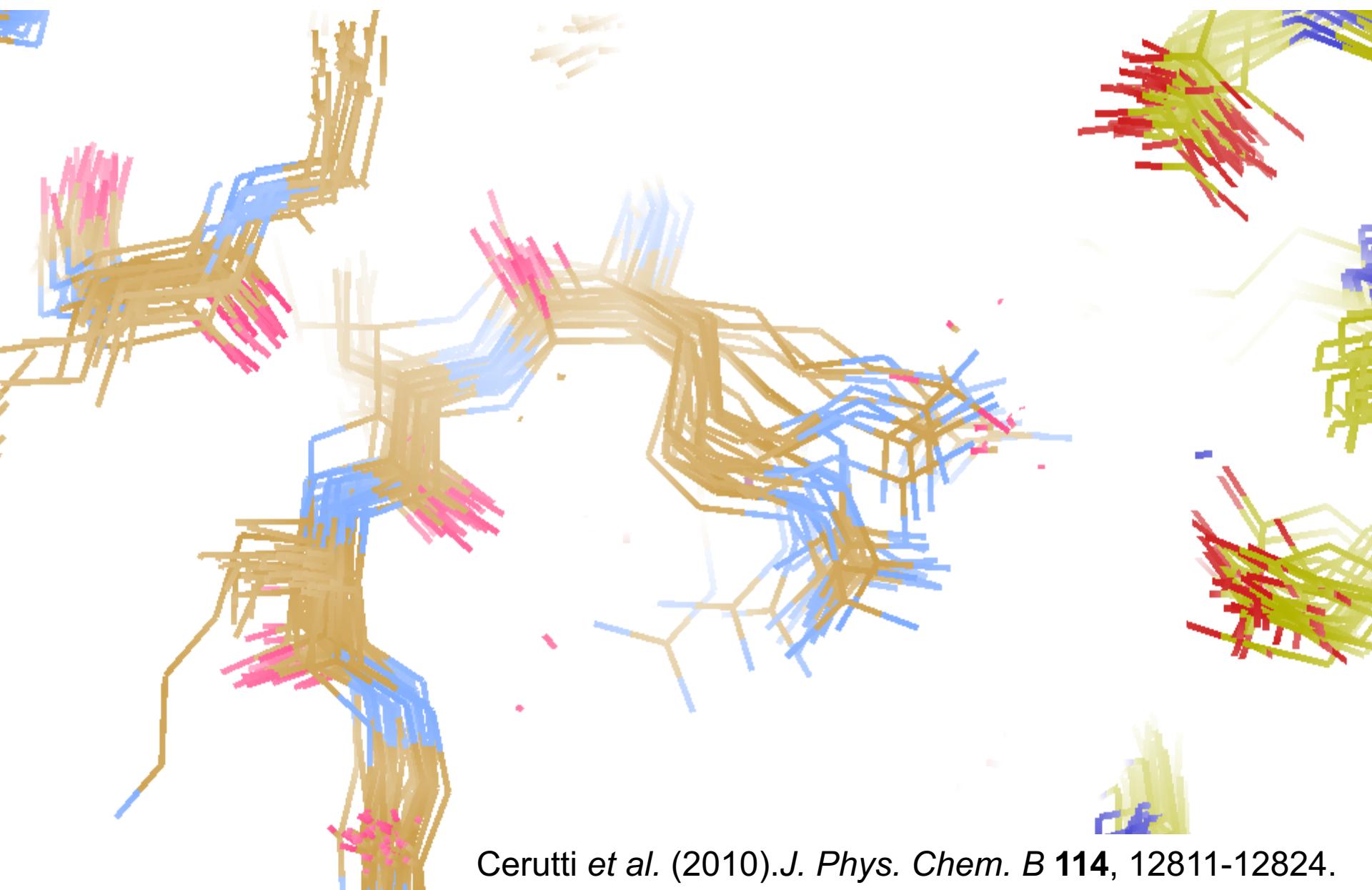
using **real**
crystal's lattice

1aho
Scorpion toxin

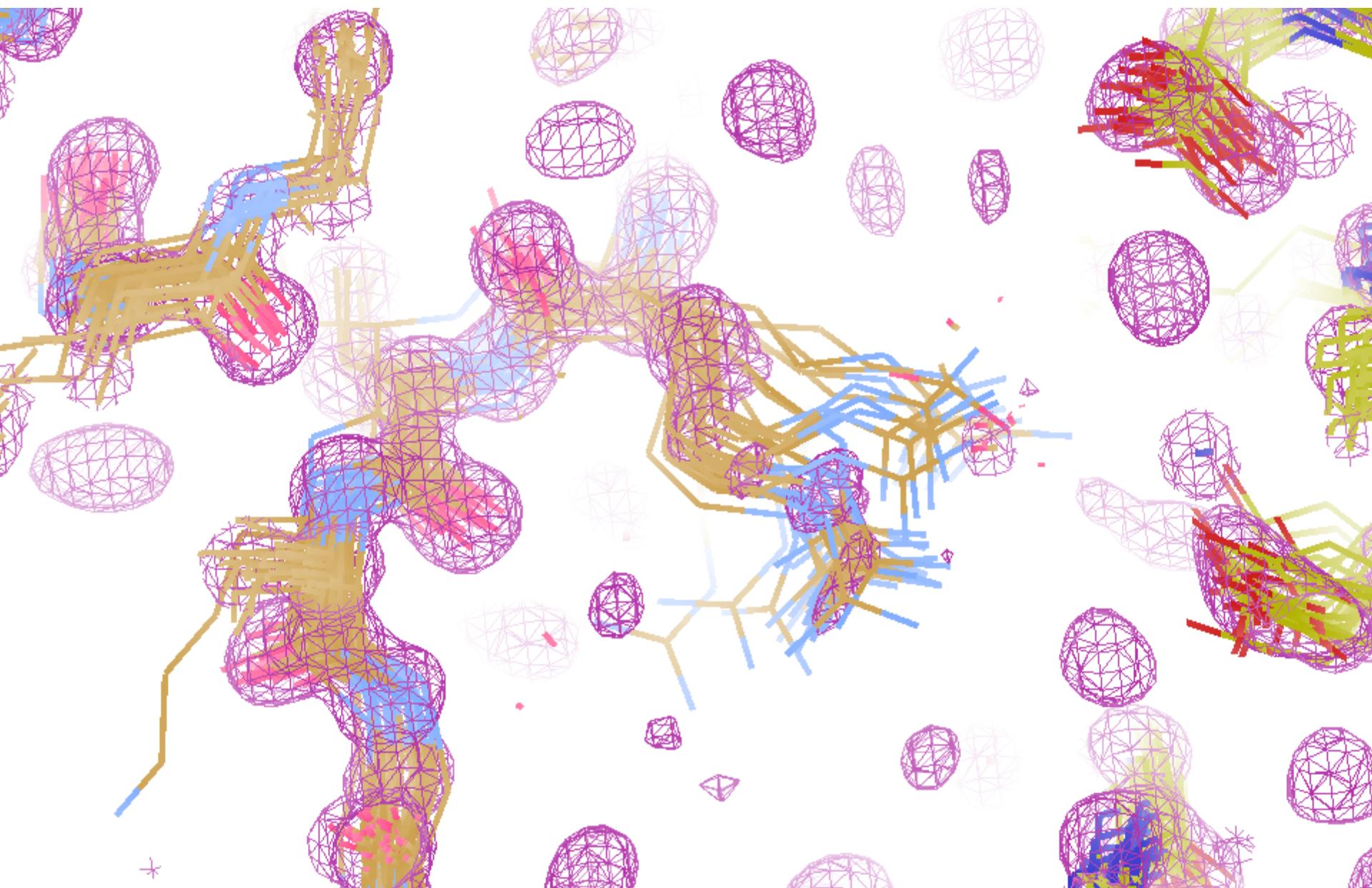
0.96 Å resolution
64 residues
Solvent: H₂O + acetate



MD simulation: 30 conformers from 24,000

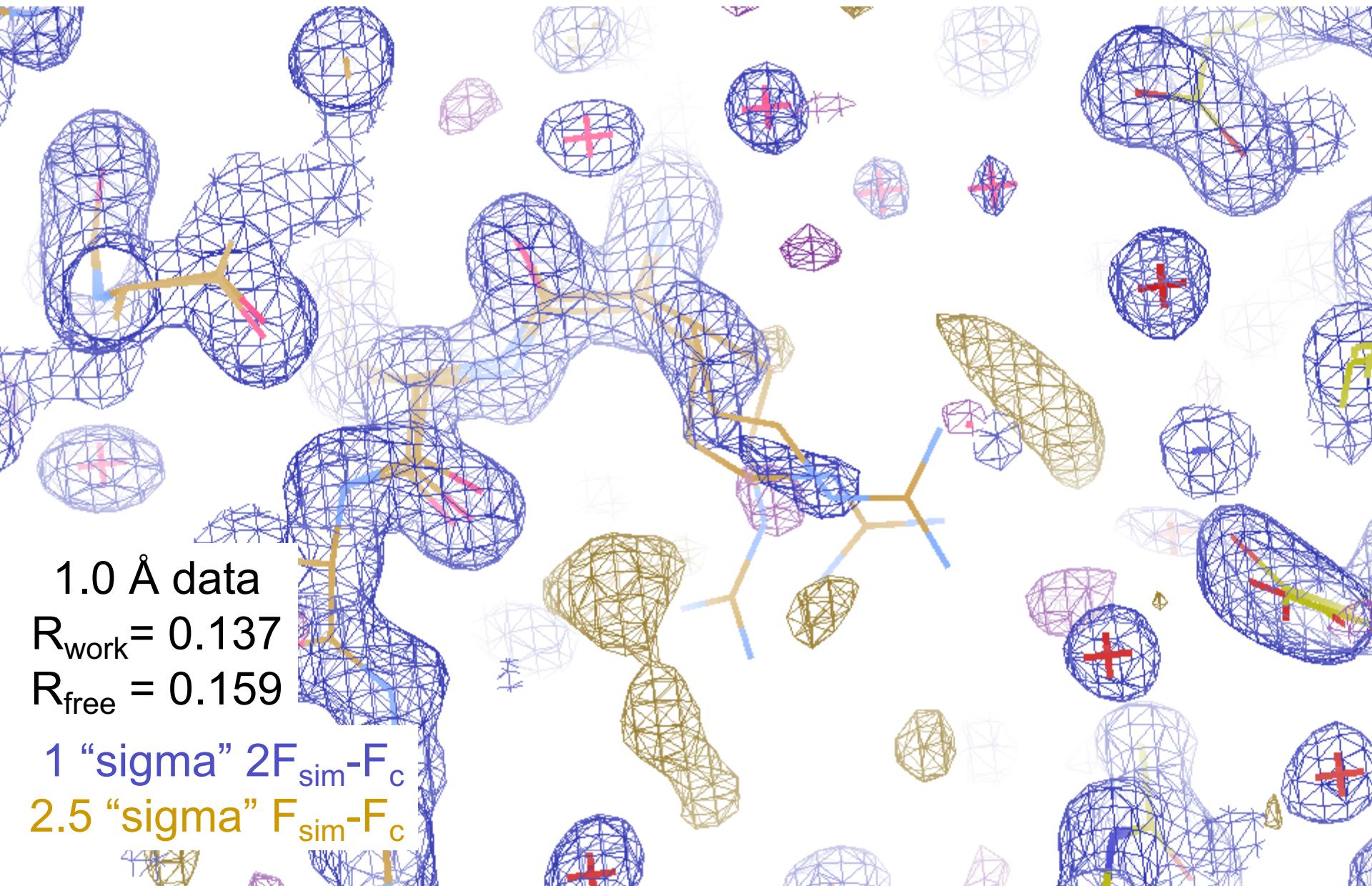


Electron density from 24,000 conformers

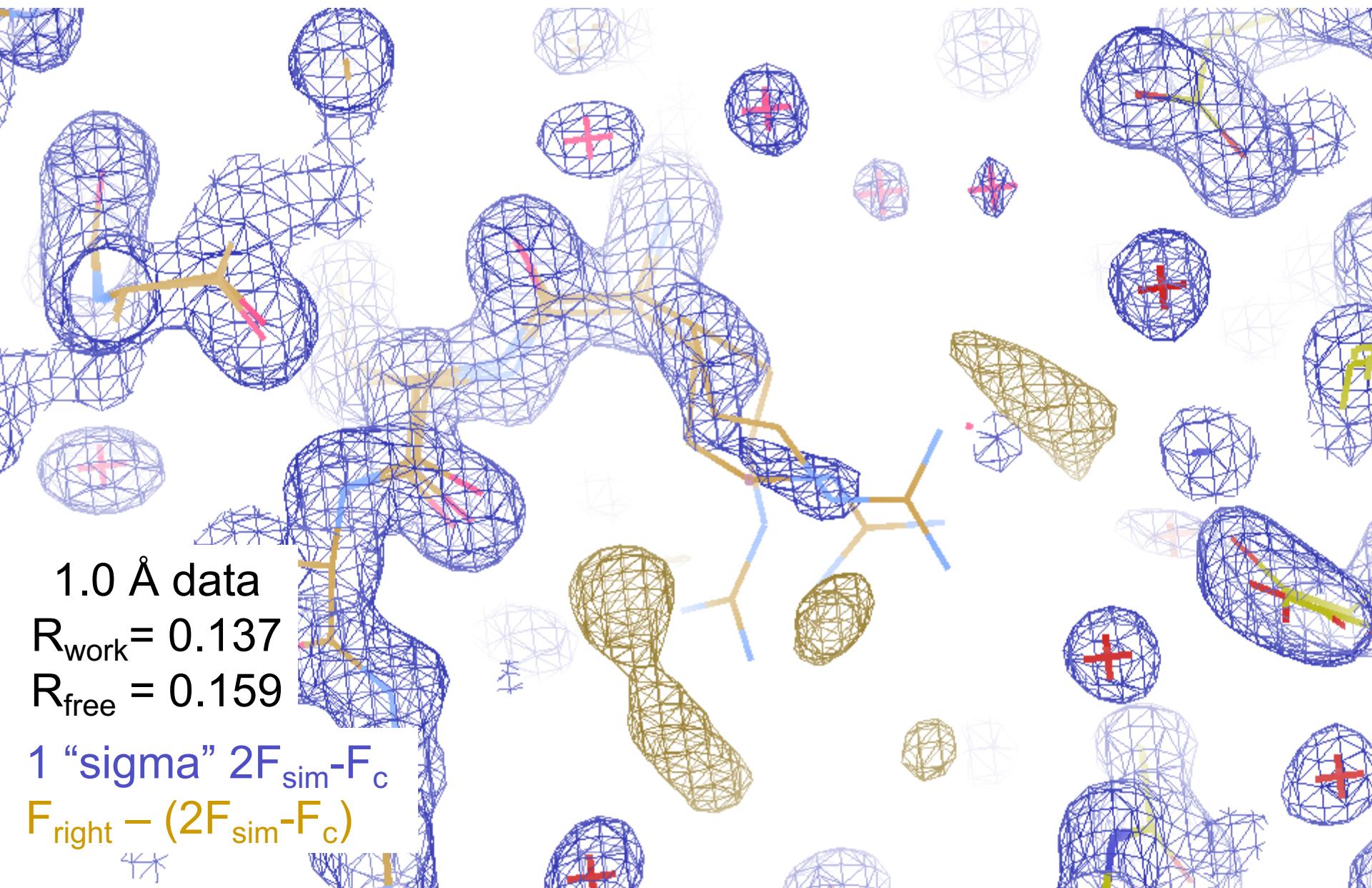


MLFSOM simulation

$2F_{\text{sim}} - F_{\text{calc}}$ and $F_{\text{sim}} - F_{\text{calc}}$ maps



$2F_{\text{sim}} - F_{\text{calc}}$ and “true difference” maps



1.0 Å data

$R_{\text{work}} = 0.137$

$R_{\text{free}} = 0.159$

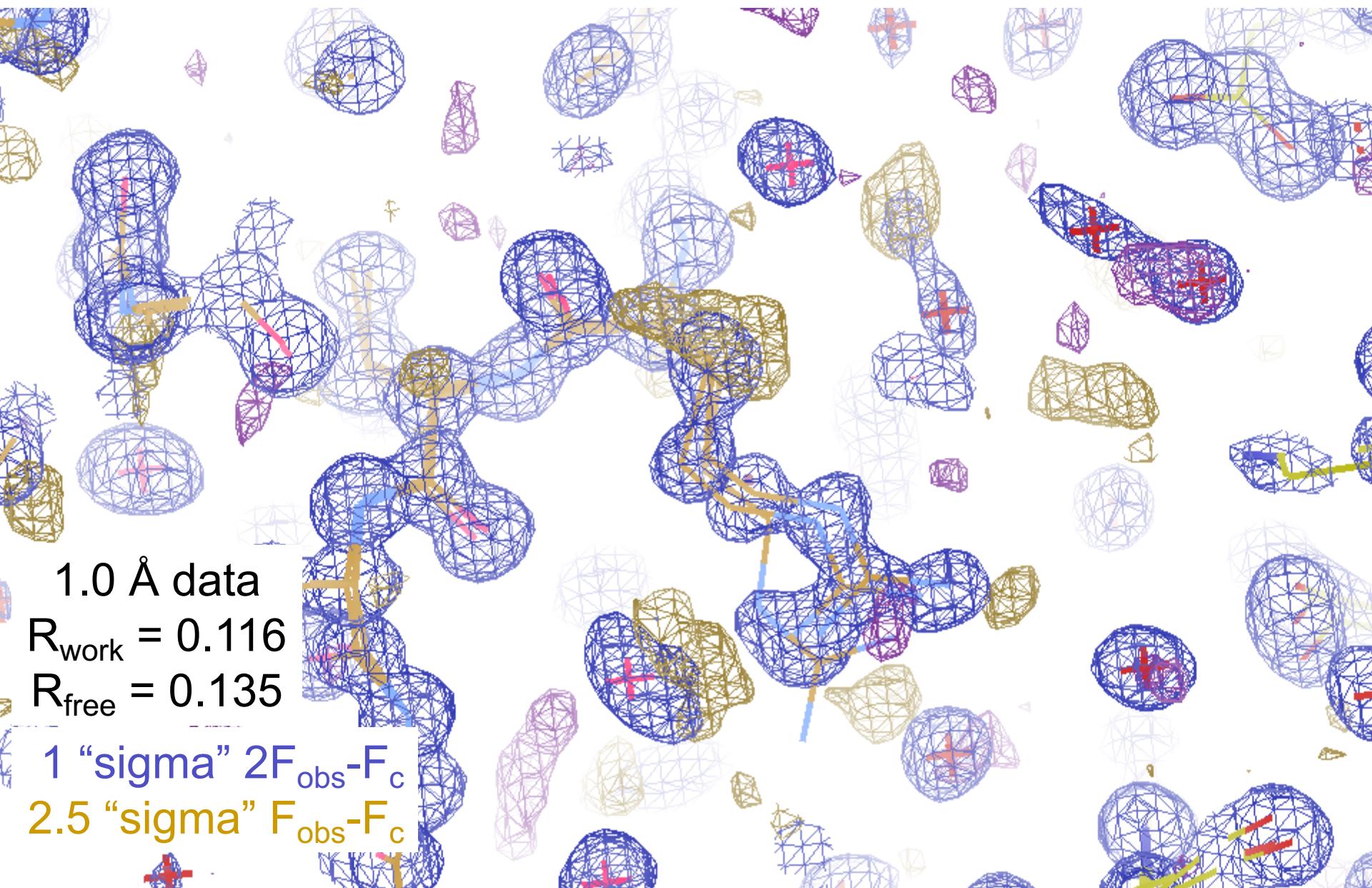
1 “sigma” $2F_{\text{sim}} - F_{\text{c}}$

$F_{\text{right}} - (2F_{\text{sim}} - F_{\text{c}})$

30 conformers from 24,000



Conventional model vs X-ray data



How many conformers are needed?

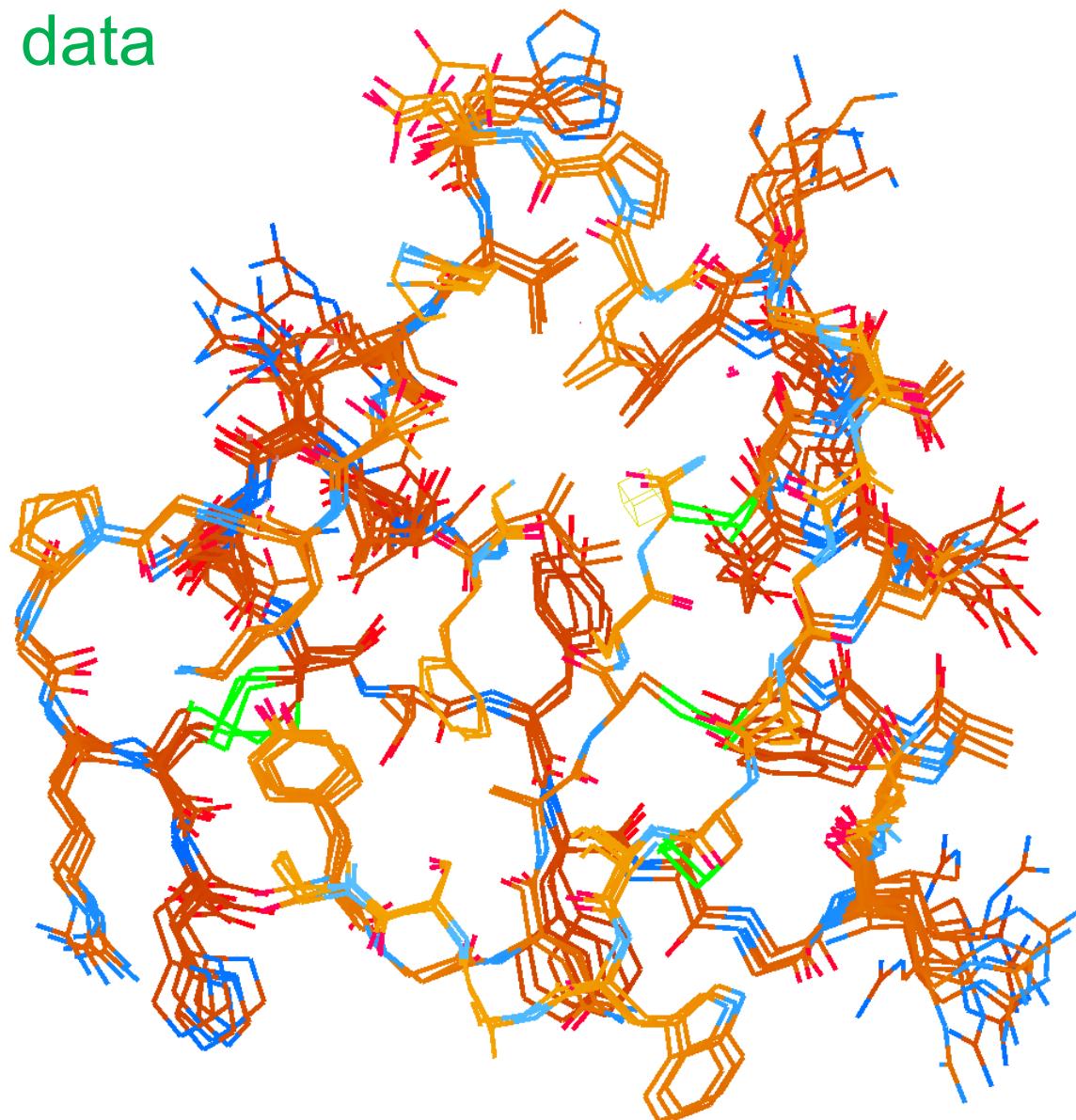
MD-derived fake data

Answer: 2-14

$$R_{\text{work}} = 0.0441$$

$$R_{\text{free}} = 0.0516$$

vs F_{sim}



How many conformers are needed?

Observed x-ray data

Answer: 1-7

$$R_{\text{work}} = 0.135$$

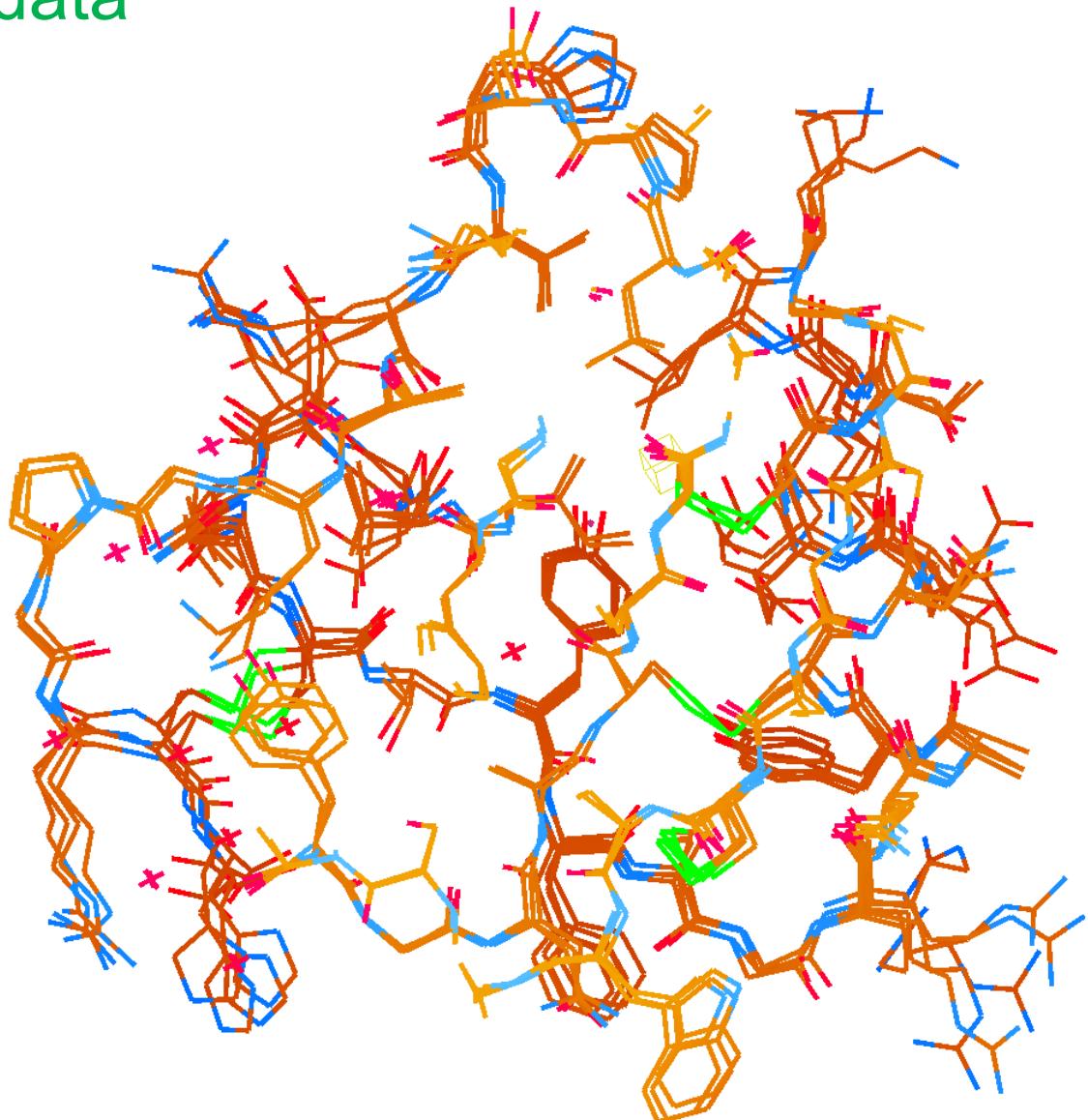
$$R_{\text{free}} = 0.158$$

vs F_{obs}

1aho:

$$R_{\text{work}} = 0.1475$$

$$R_{\text{free}} = 0.1685$$



Adding noise

$$12\%^2 + 12\%^2 = 17\%^2$$

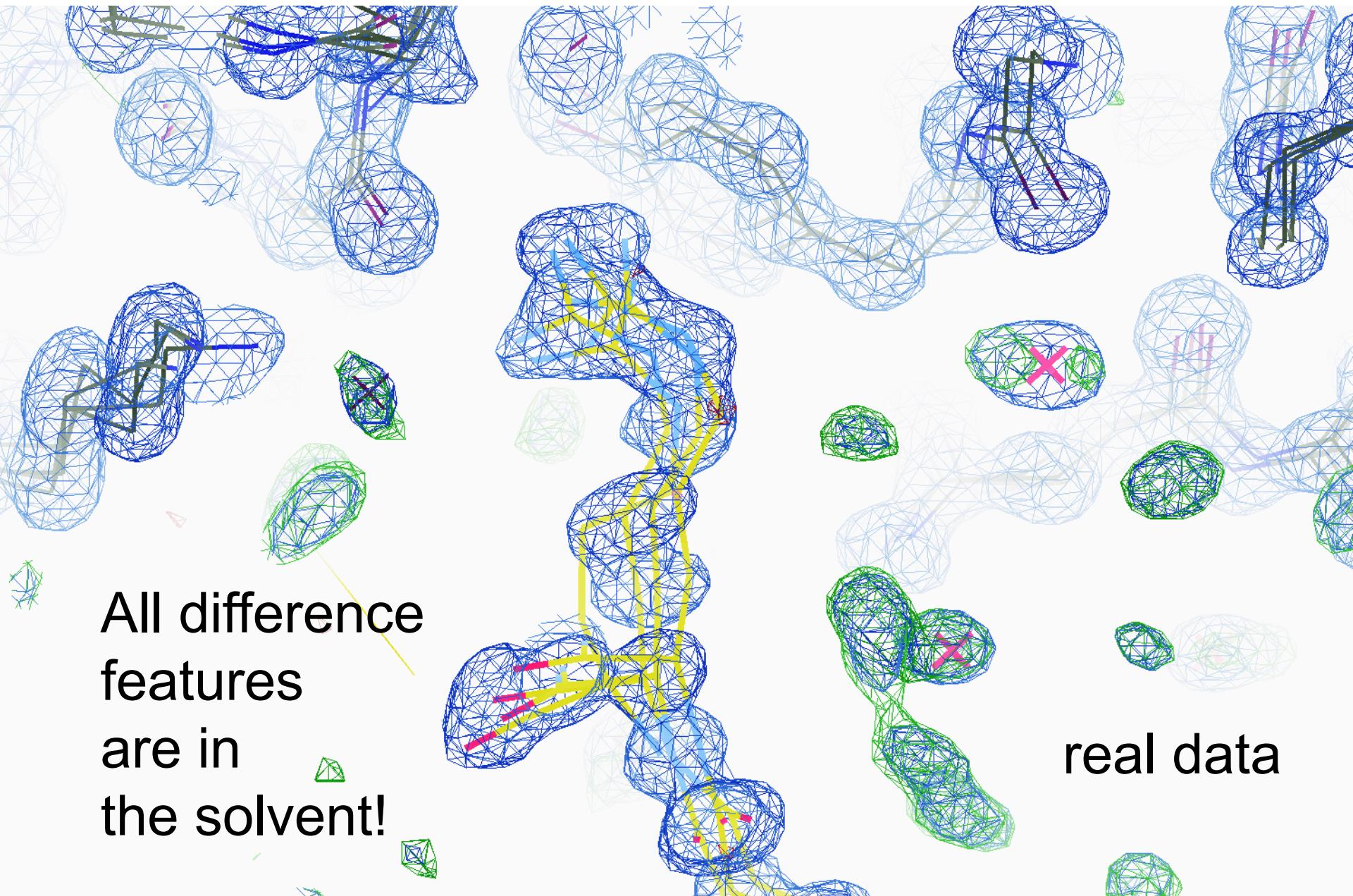
$$R_{\text{protein}} + R_{\text{solvent}} \approx R_{\text{work}}$$

Adding noise

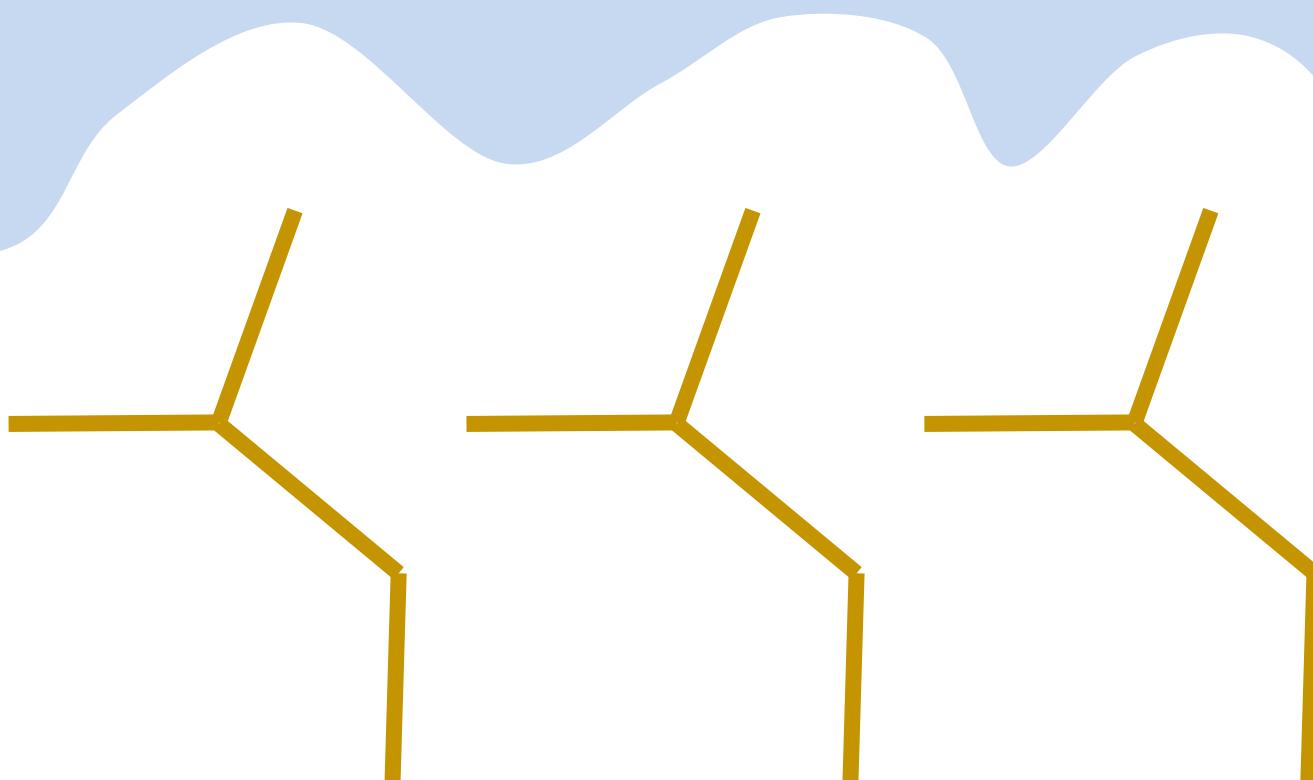
$$4\%^2 + 12\%^2 = 12.6\%^2$$

$$R_{\text{protein}} + R_{\text{solvent}} \approx R_{\text{work}}$$

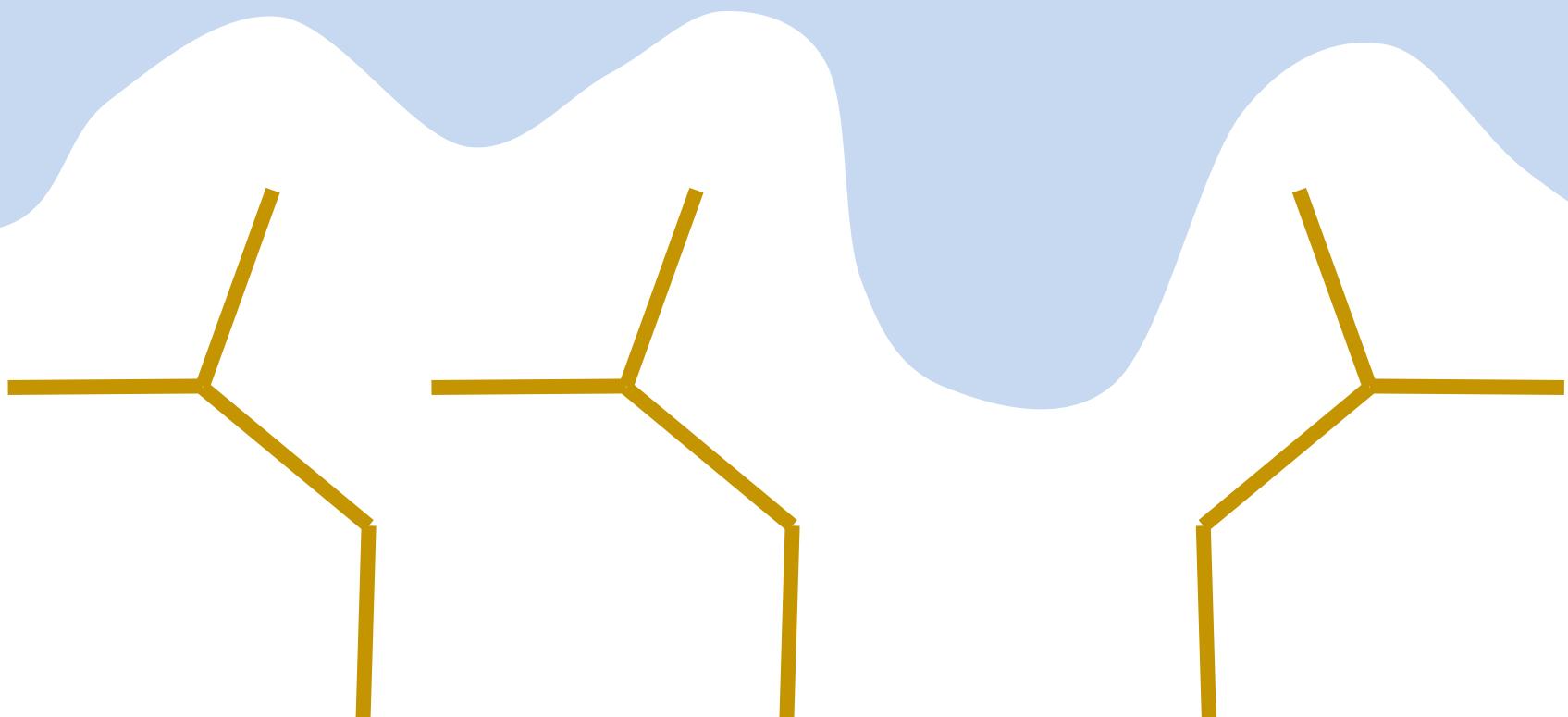
How many conformers are there?



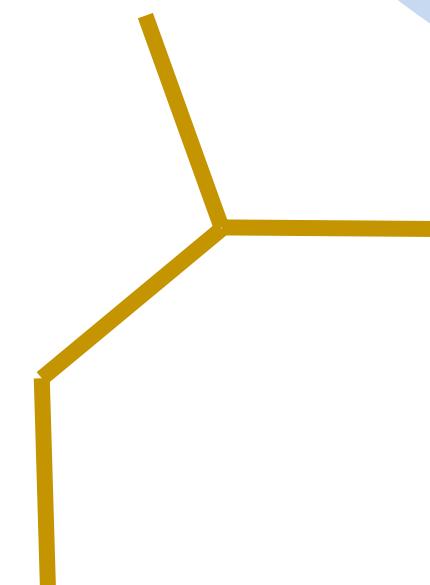
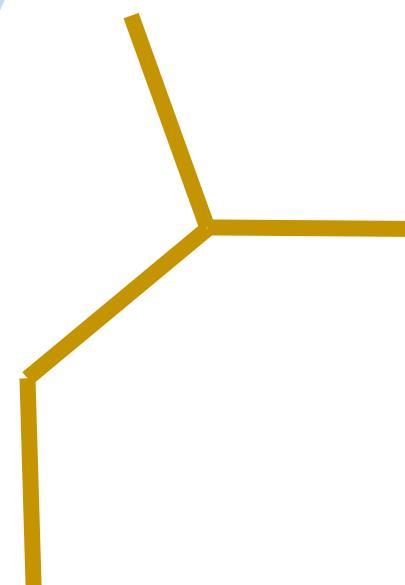
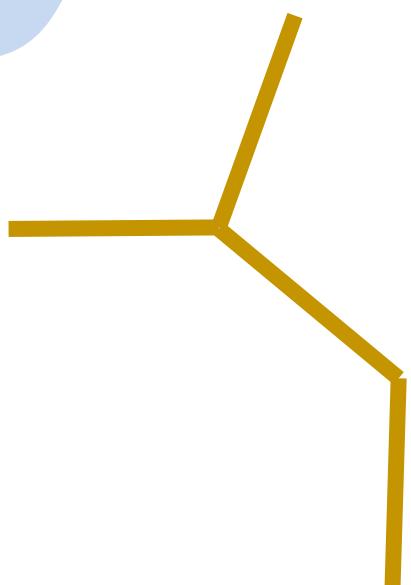
Solvent boundary problem



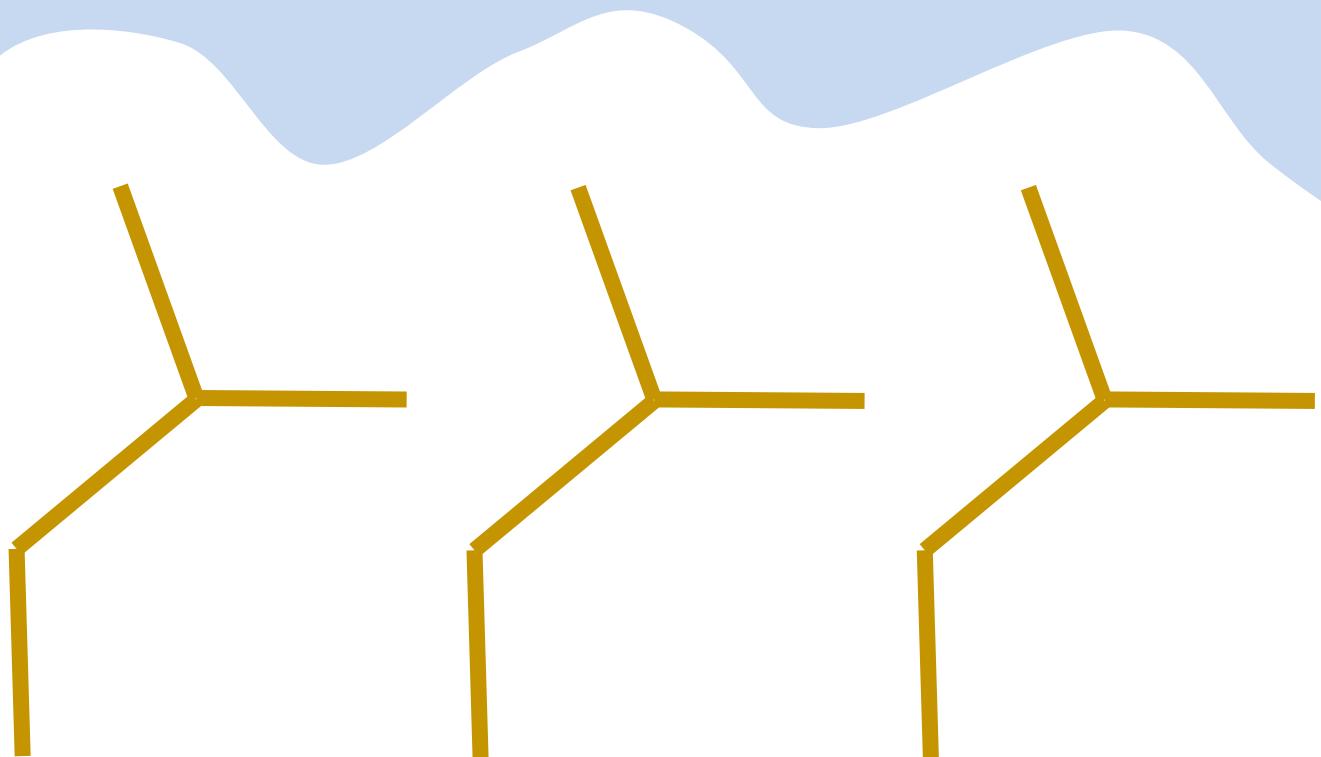
Solvent boundary problem



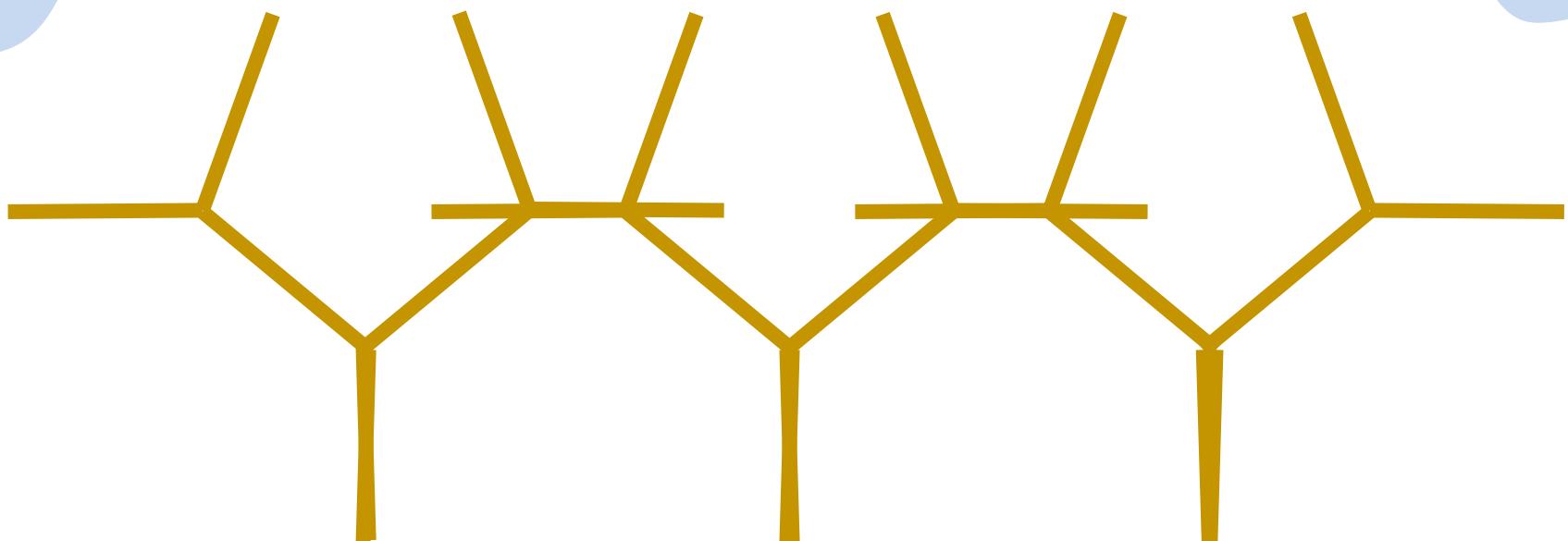
Solvent boundary problem



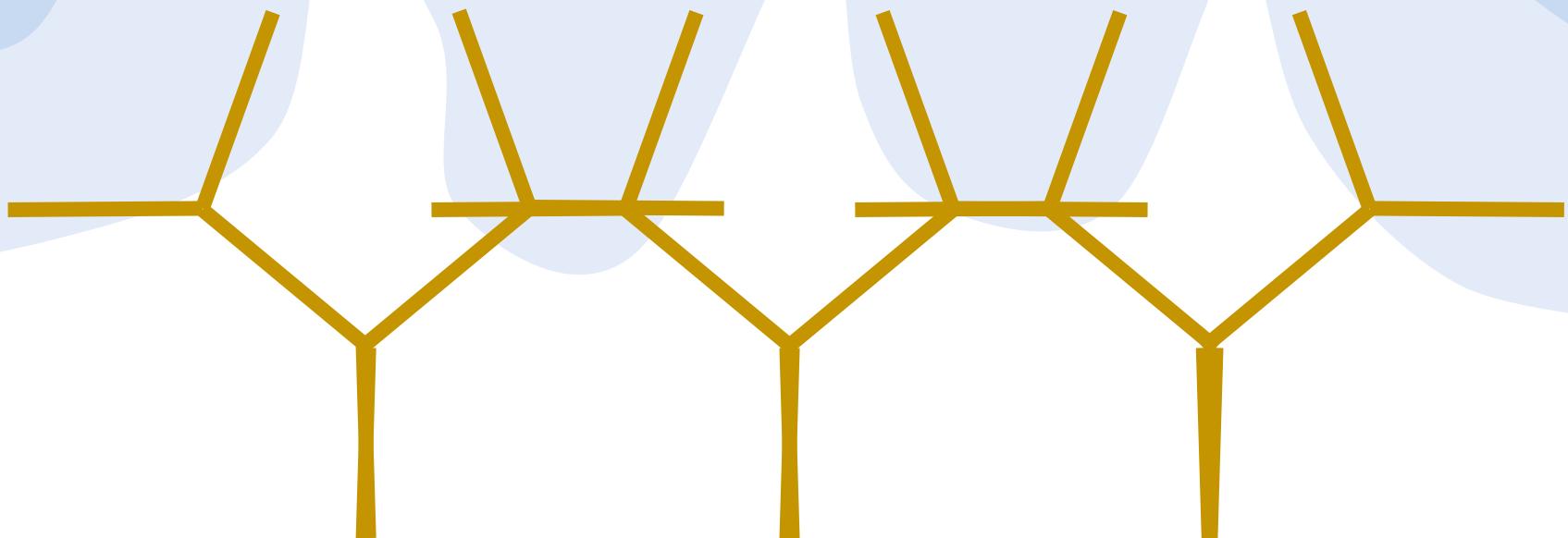
Solvent boundary problem



Current implementation



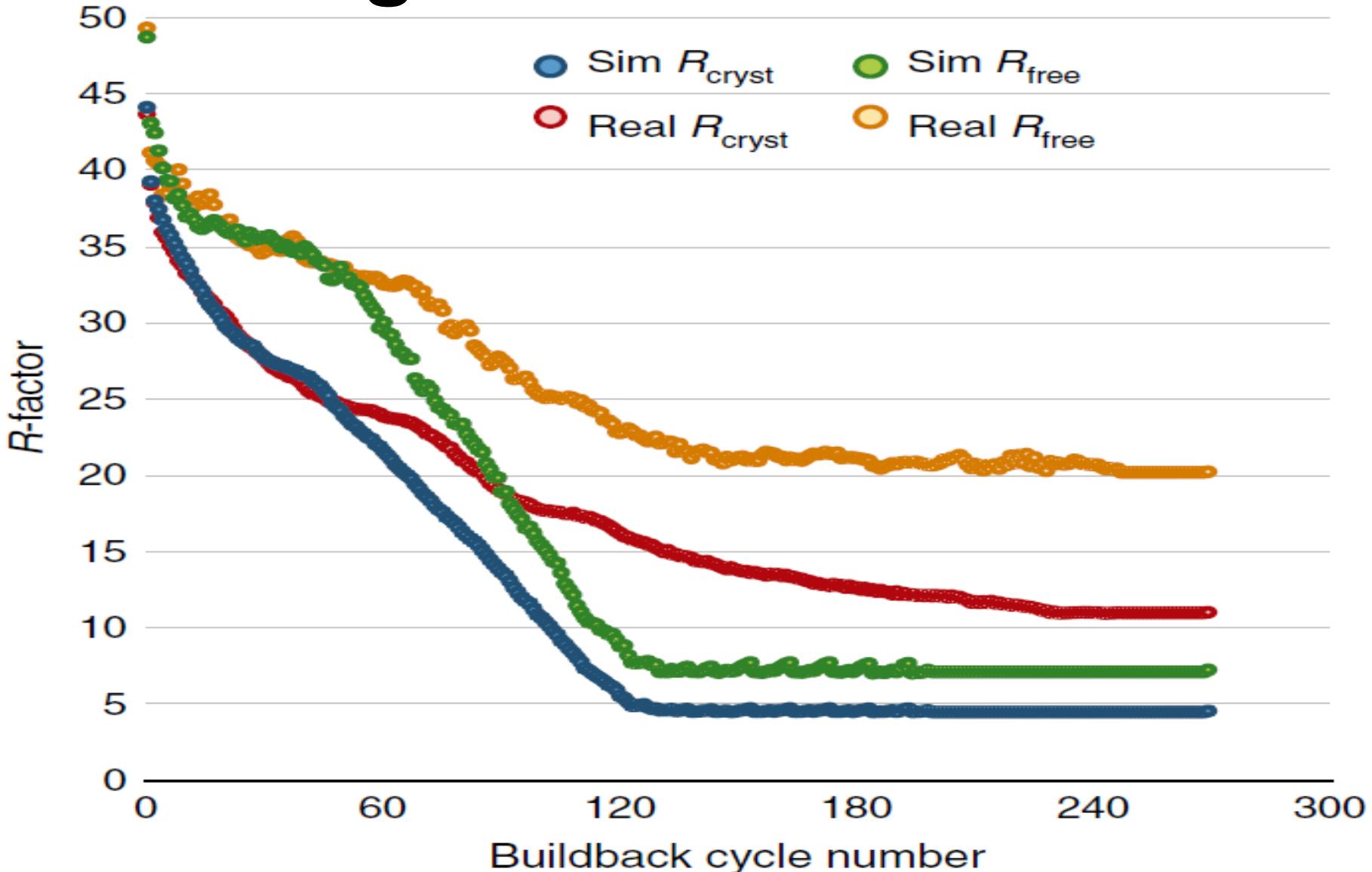
Better implementation? (not available yet)



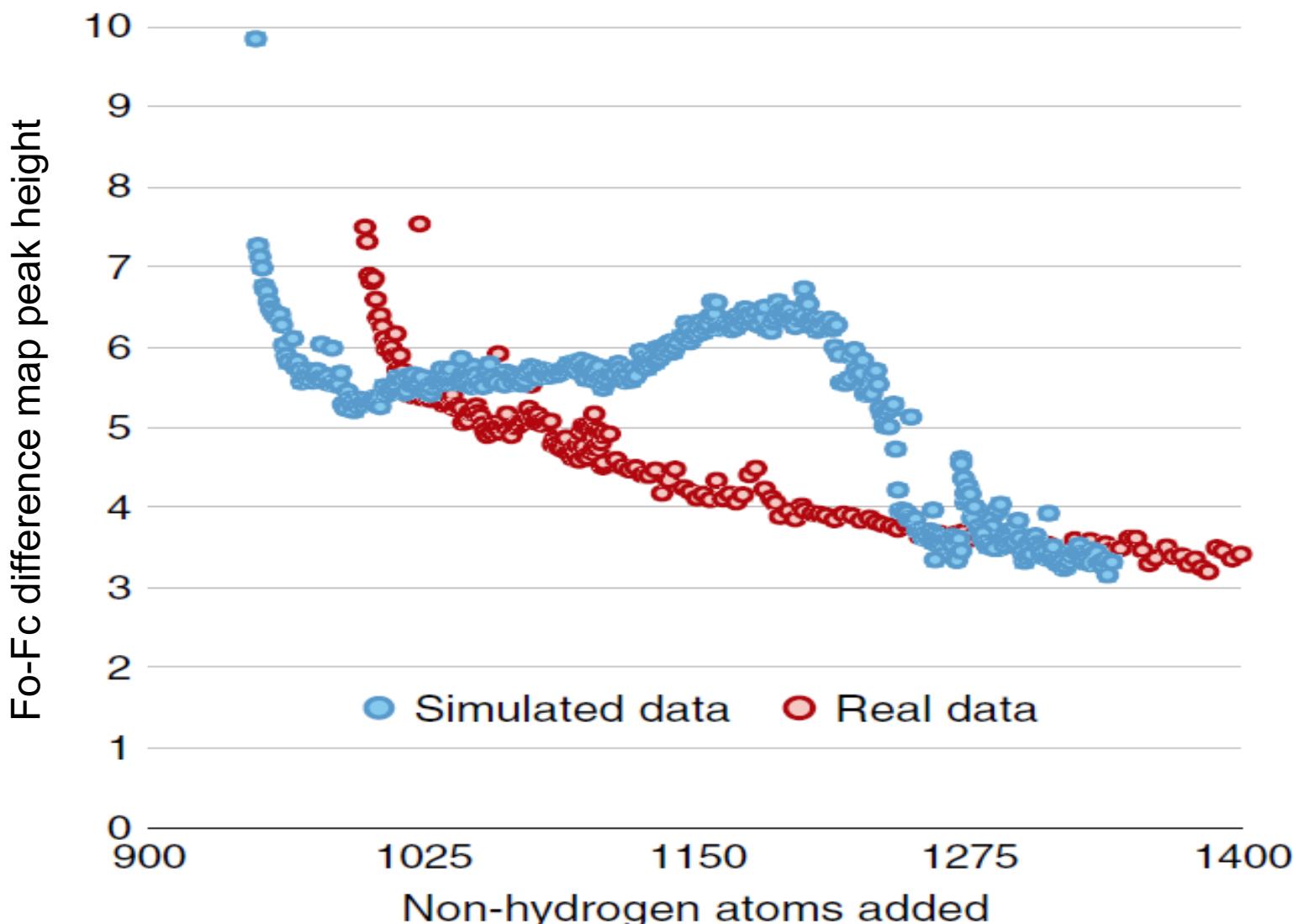
Getting rid of “bias”

1. Throw out everything not 100% sure about
 - Alt confs? Side chains? Loops?
 - Can often cut down to 50% or more
2. Build **bigest** $F_{\text{obs}} - F_{\text{calc}}$ feature
3. Refine to **convergence**
 - Atoms stop moving, not just R settling
4. Goto 2.
 - Note: Repeating MR = rigid-body shift
 - Solvent-flattened SAD phases not better

Building back from mainchain



Building back from mainchain



Take-home lesson:

Always build into the
biggest

$F_{\text{obs}} - F_{\text{calc}}$ difference feature

(least likely to be noise)

If biggest feature is **negative**:
adjust bulk solvent parameters

“R” factors

R = % error

“R” factors

$$R = \frac{\sum |F_{obs} - F_{calc}|}{\sum F_{obs}}$$

completely random:	0.59
starting MR solution:	0.4-0.55
something still wrong?:	> 0.3
correct chain trace:	< 0.2
small molecule:	~ 0.05

“R” factors

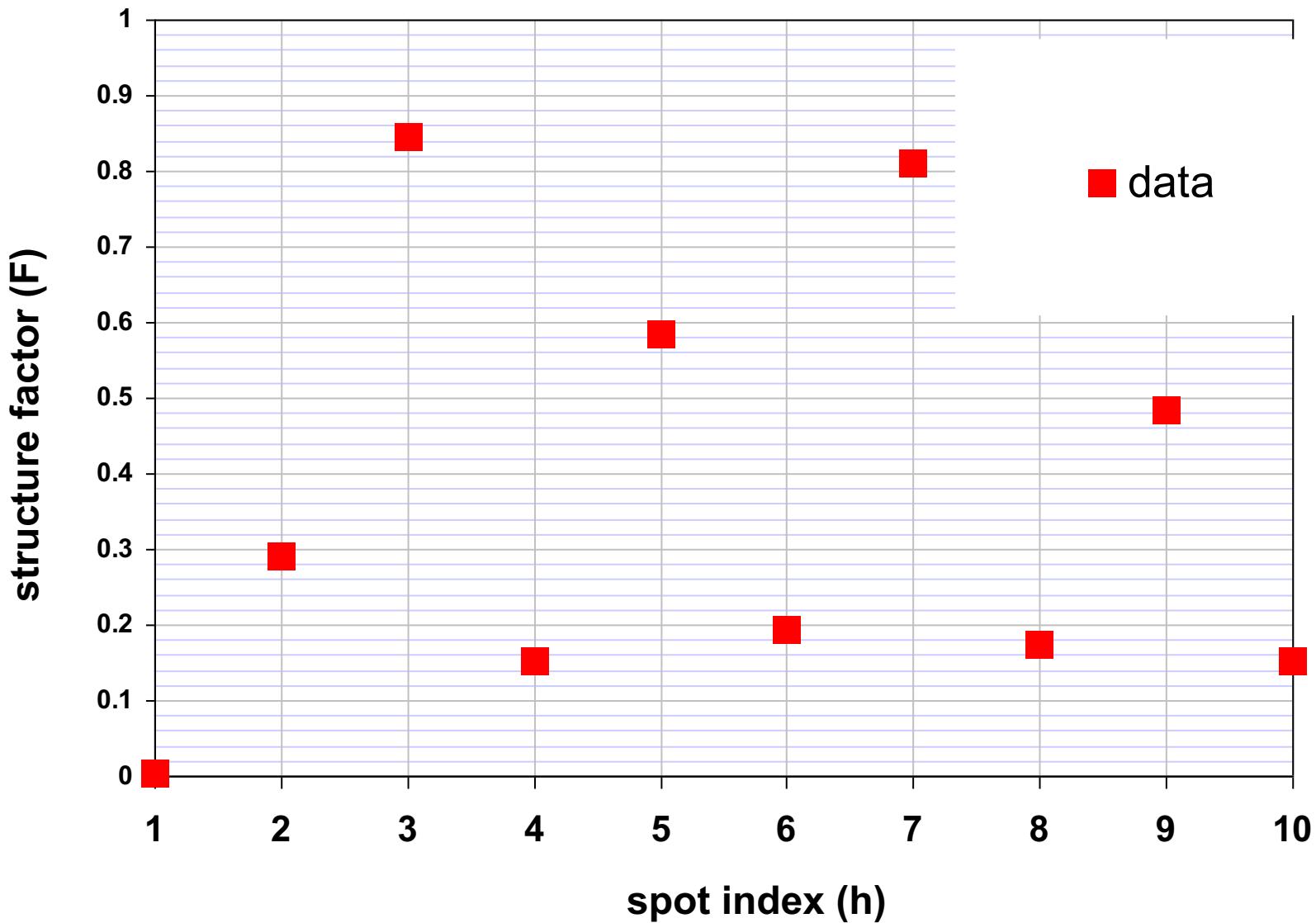
R_{cryst} (or just “R”)

observed vs calculated data (Fs)

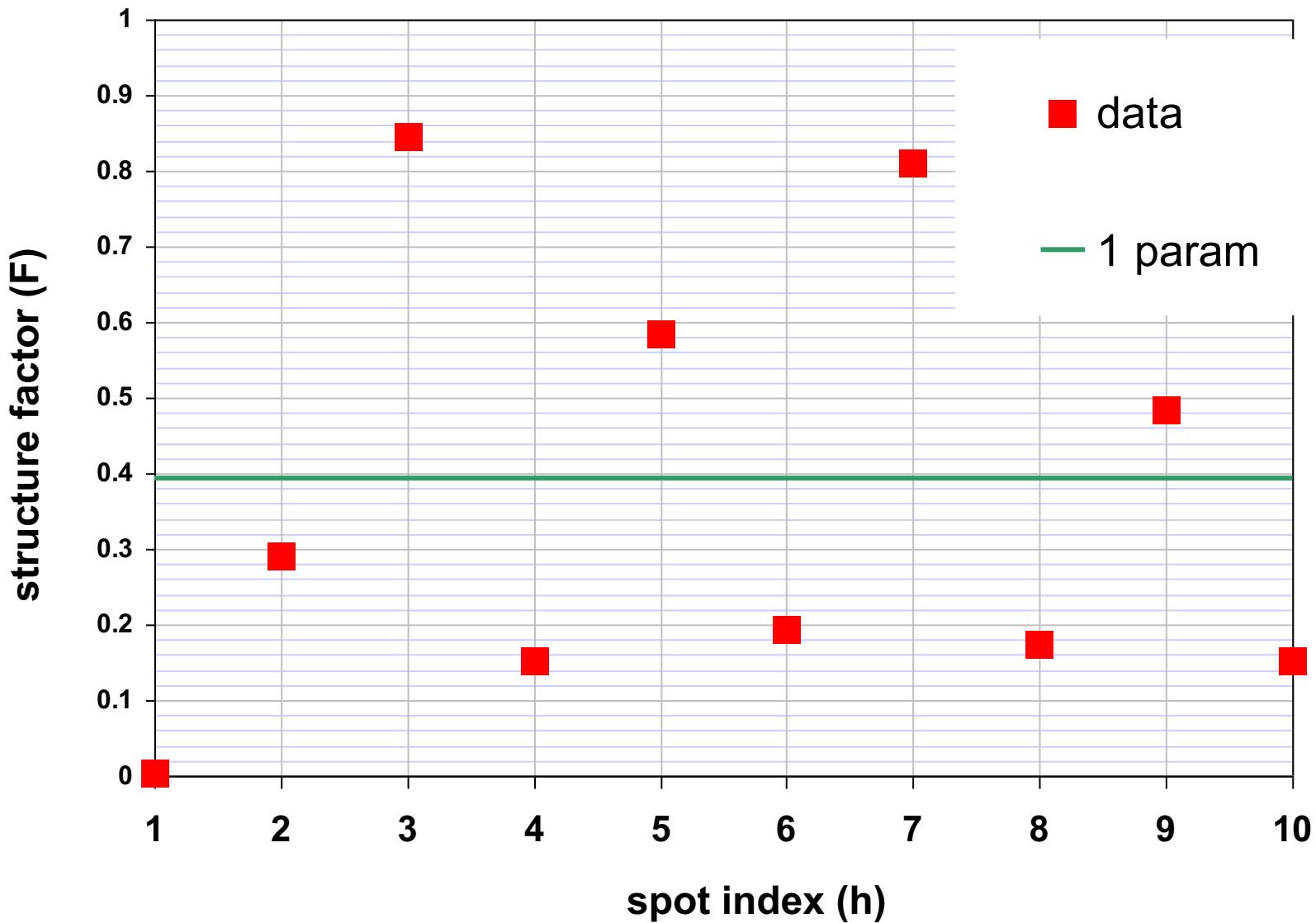
R_{free}

cross-check with “random” subset of data
should be < 0.3 and < $R_{\text{cryst}} + 0.1$

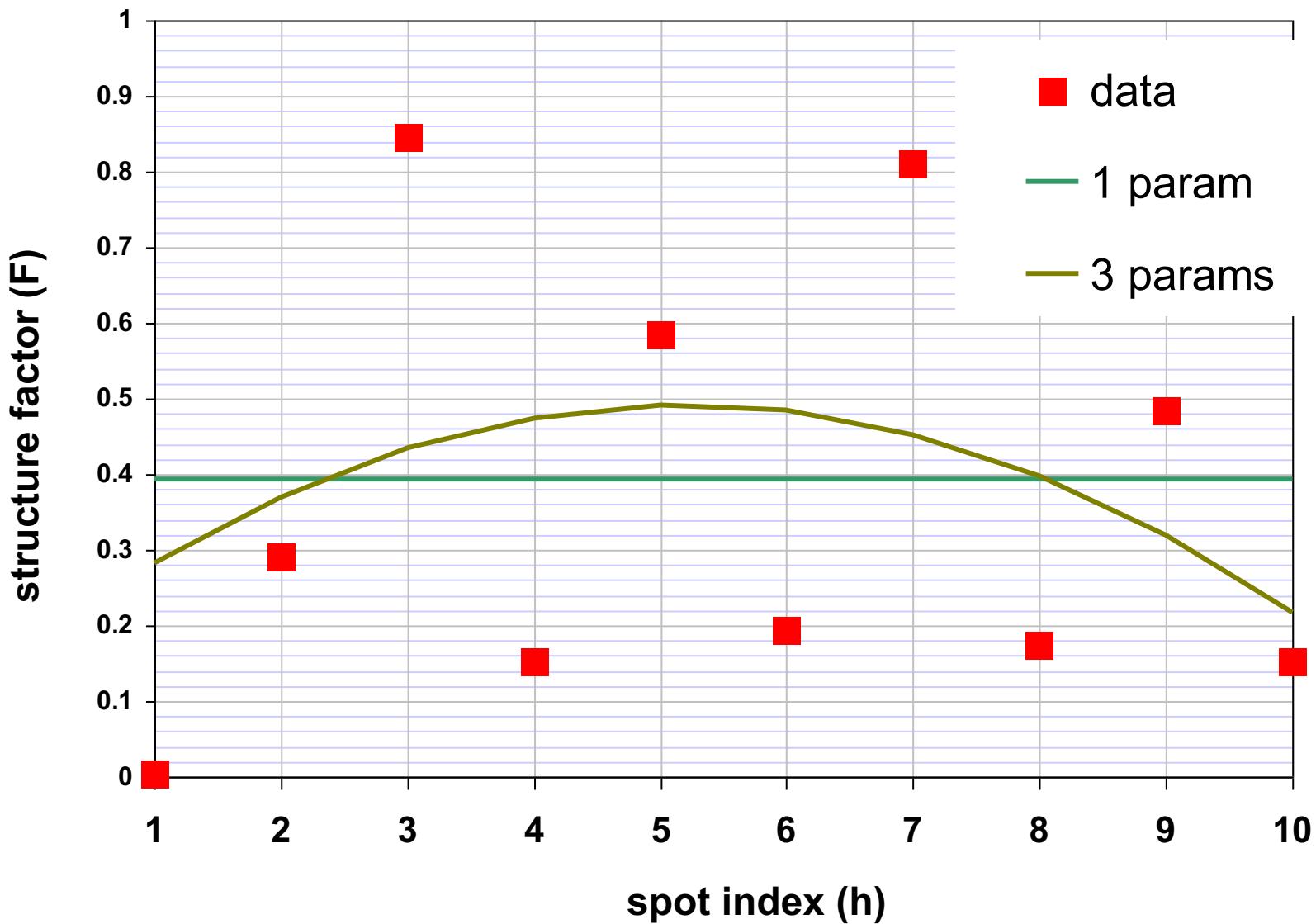
Fitting data



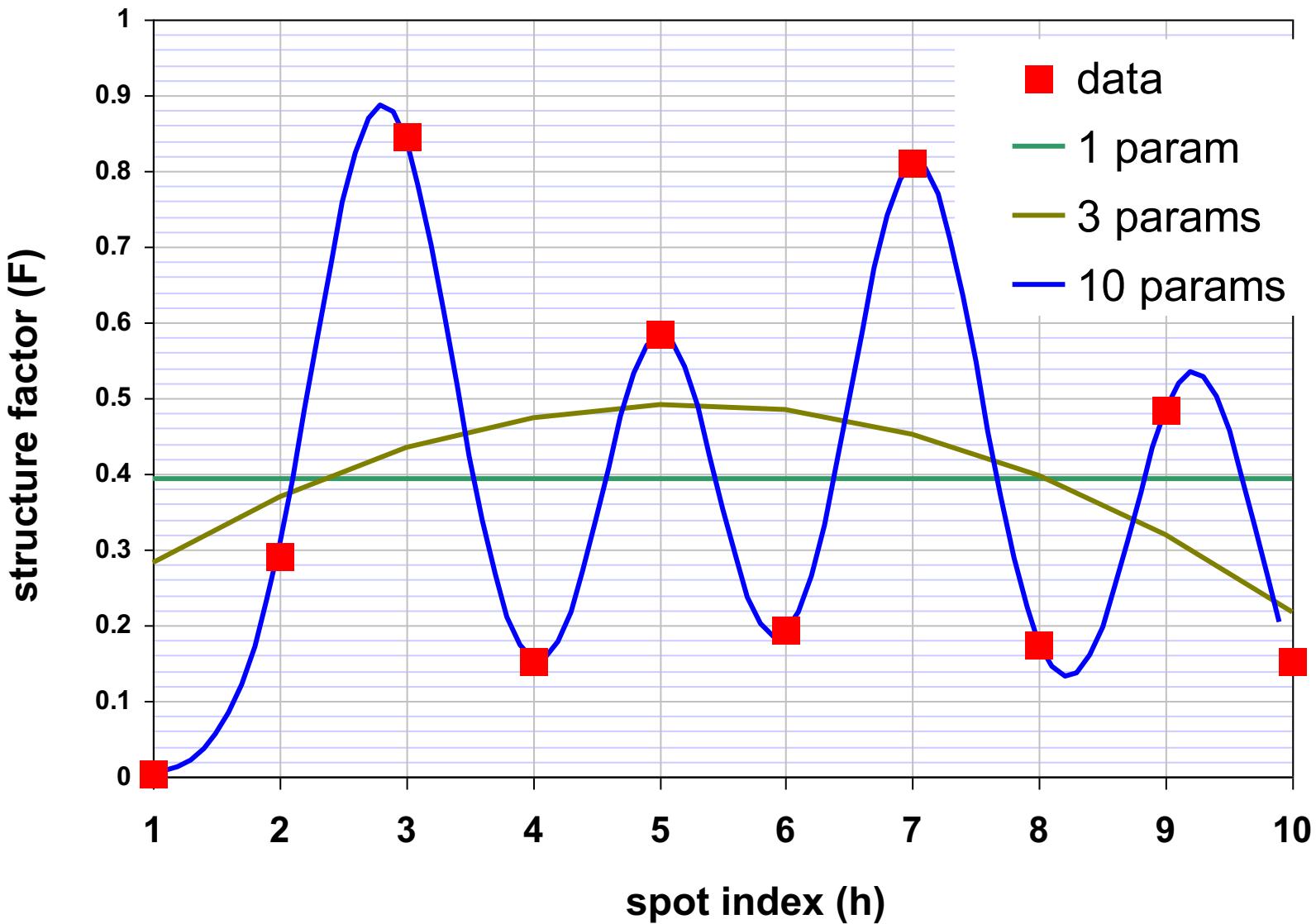
Fitting data



Fitting data



Fitting data



“R” factors

R_{cryst} (or just “R”)

observed vs calculated data (Fs)

R_{free}

cross-check with “random” subset of data
should be < 0.3 and < $R_{\text{cryst}} + 0.1$

$R_{\text{sym}} = R_{\text{merge}}$ (self-consistency of data: ls)

R_{merge}

$$R_{\text{merge}} = \frac{\sum |I_{\text{obs}} - \langle I \rangle|}{\sum I_{\text{obs}}}$$

blows up
as $I_{\text{obs}} \rightarrow 0$

completely random:	0.59
weak data (high angle):	0.7- ∞
wrong symmetry choice?:	~0.2-0.55
small or disordered crystal:	~0.1-0.2
typical:	~ 0.05

“R” factors

R_{cryst} (or just “R”)

observed vs calculated data (F_s)

R_{free}

cross-check with “random” subset of data
should be < 0.3 and $< R_{\text{cryst}} + 0.1$

$R_{\text{sym}} = R_{\text{merge}}$ (self-consistency of data: ls)

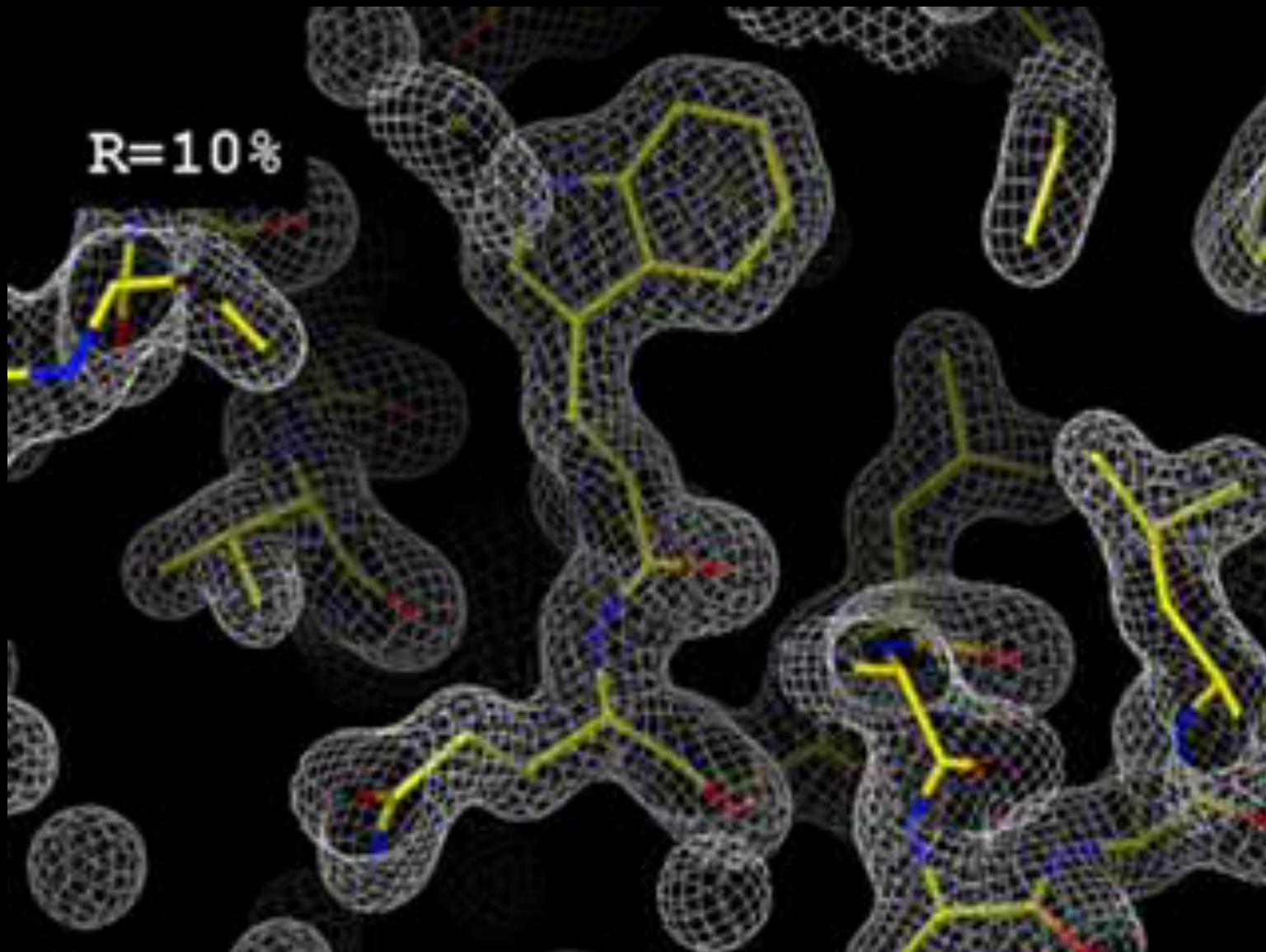
R_{rim} R_{pim} R_{meas}

R_{iso} R_{anom}

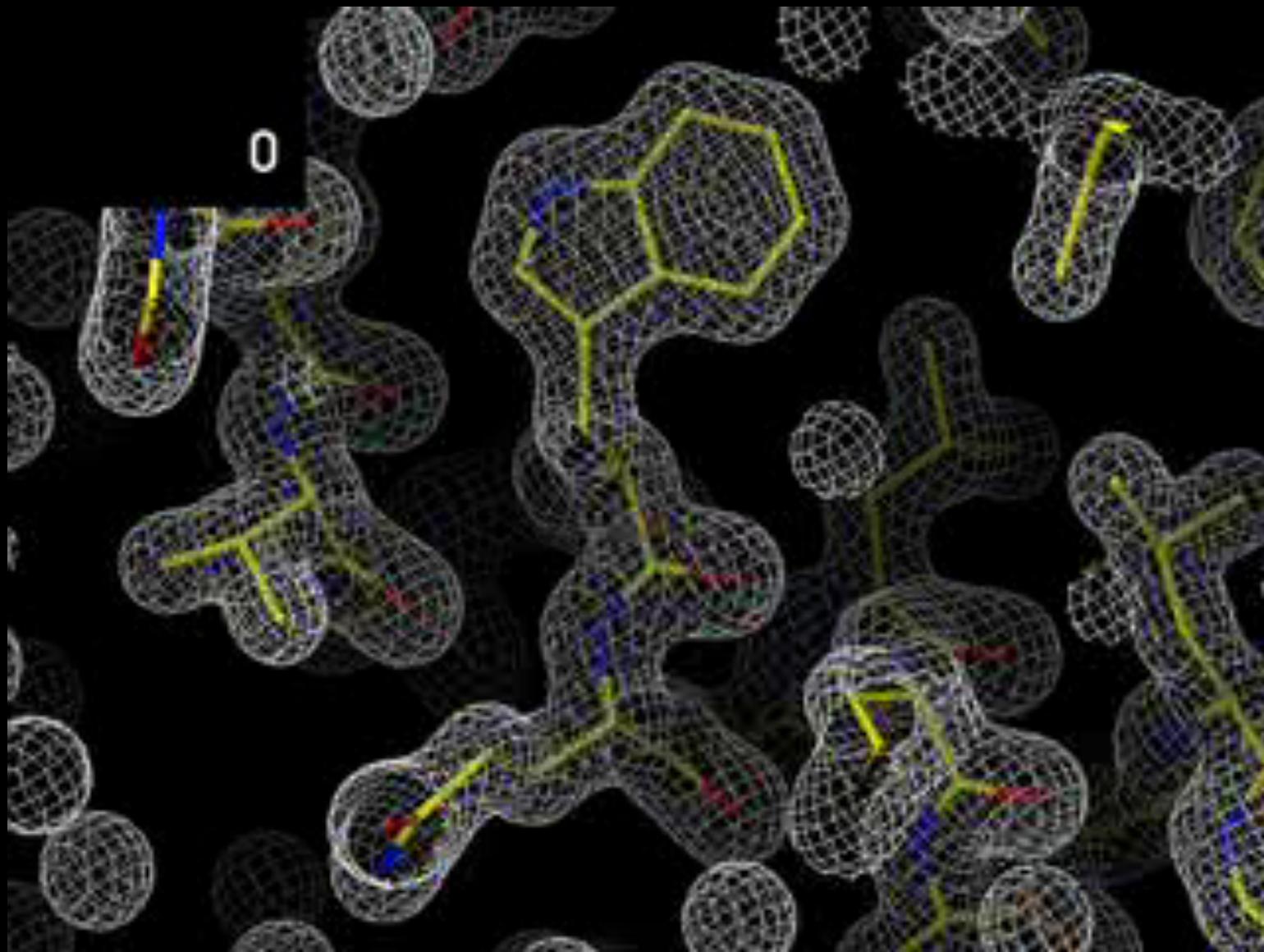
R_{diff}

R_{sym} is “unfair” to
high multiplicity

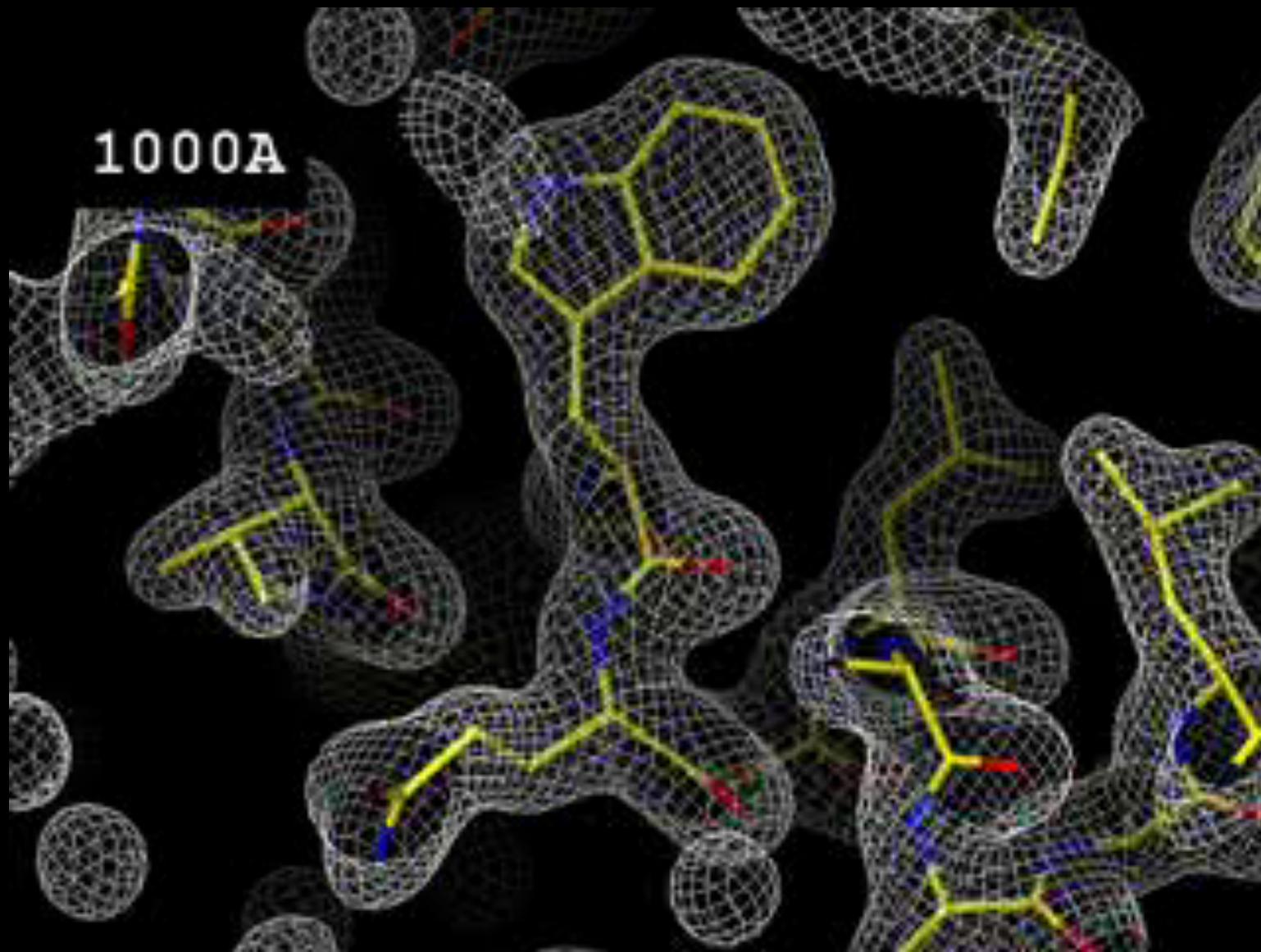
R-factor



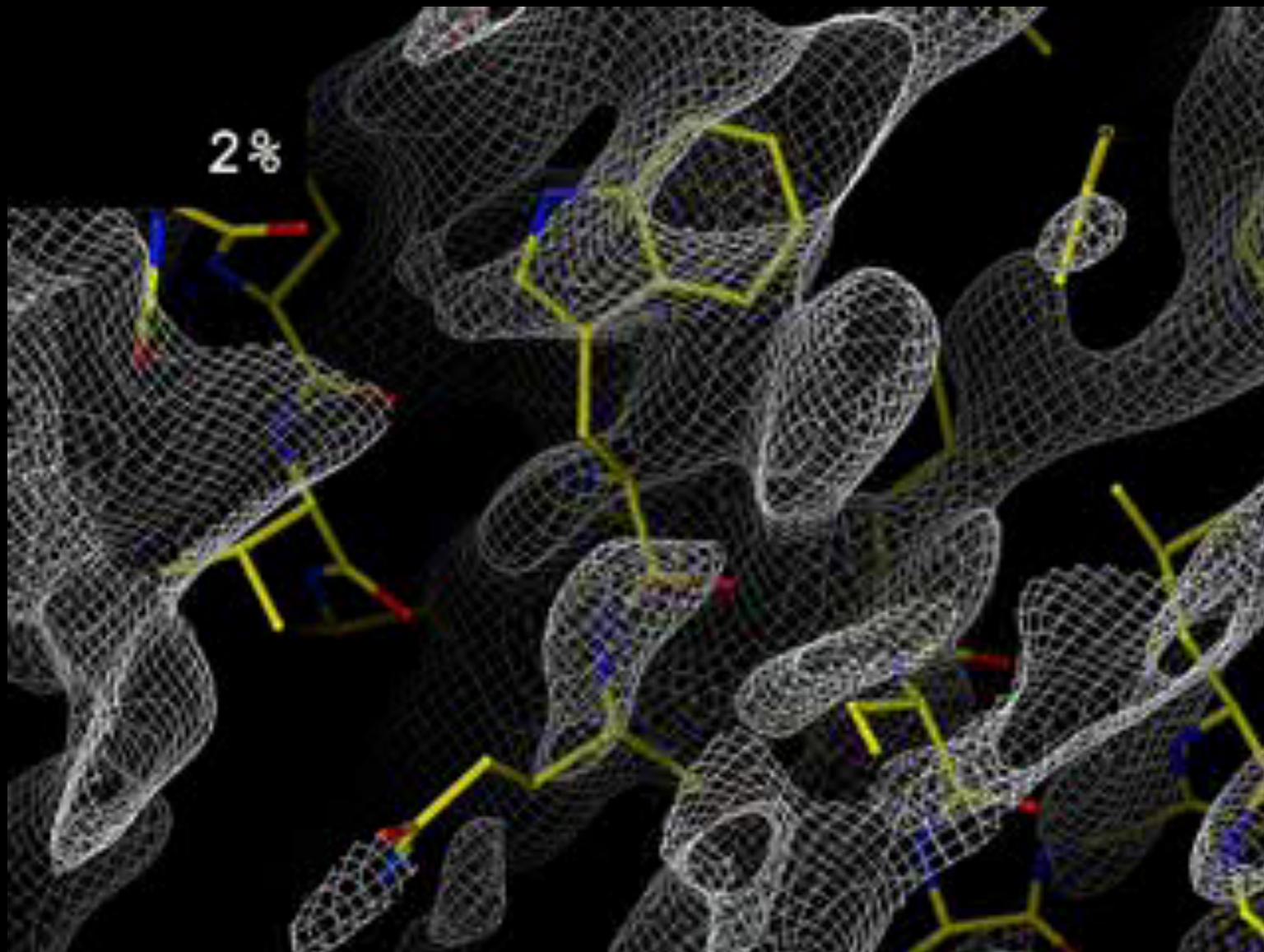
Overloads



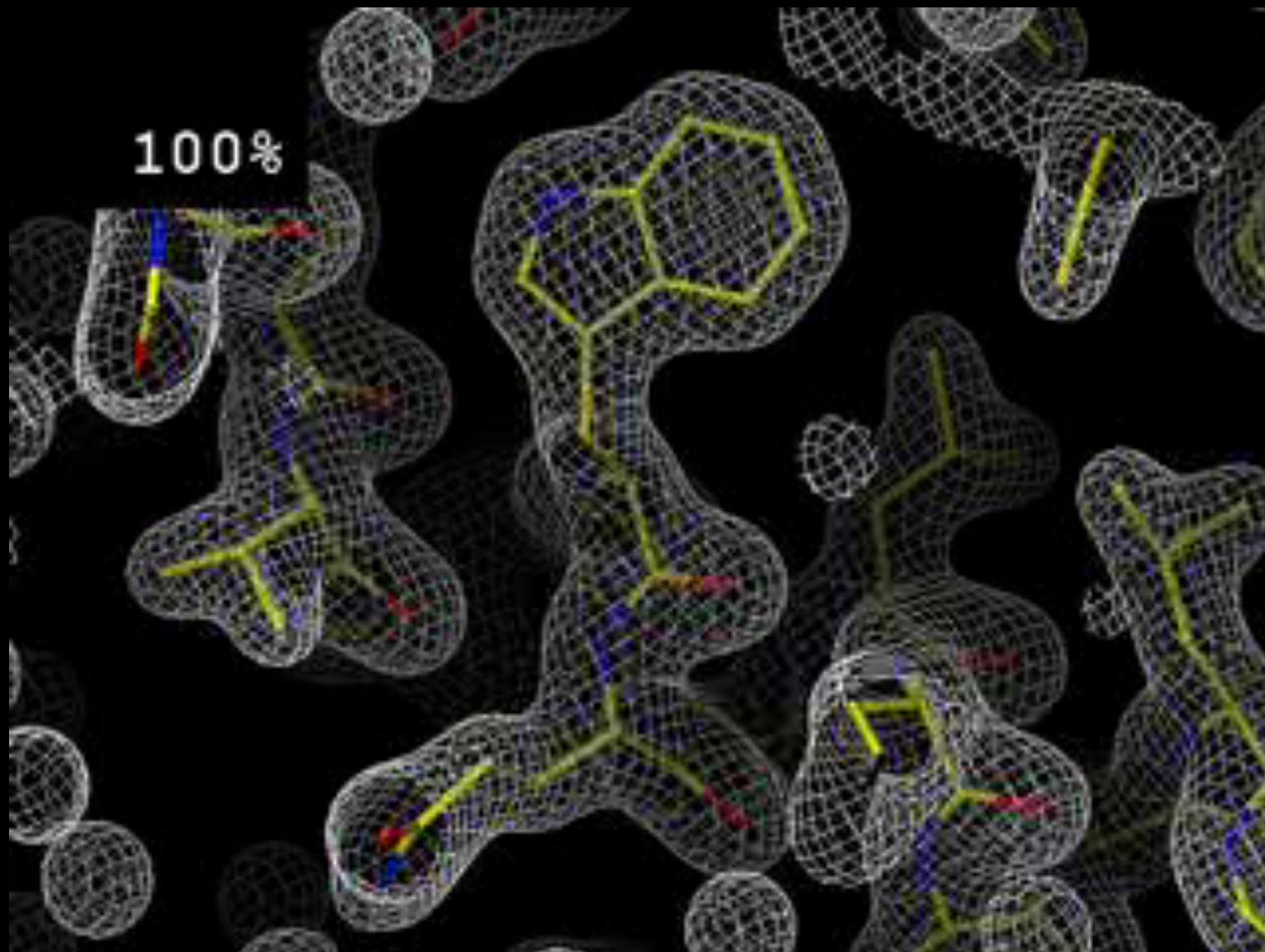
Resolution: low-angle cutoff



Completeness: missing wedge



Completeness: random deletion



And the dominant source of error is...

- Resolution

Disorder (including Radiation Damage)

Background

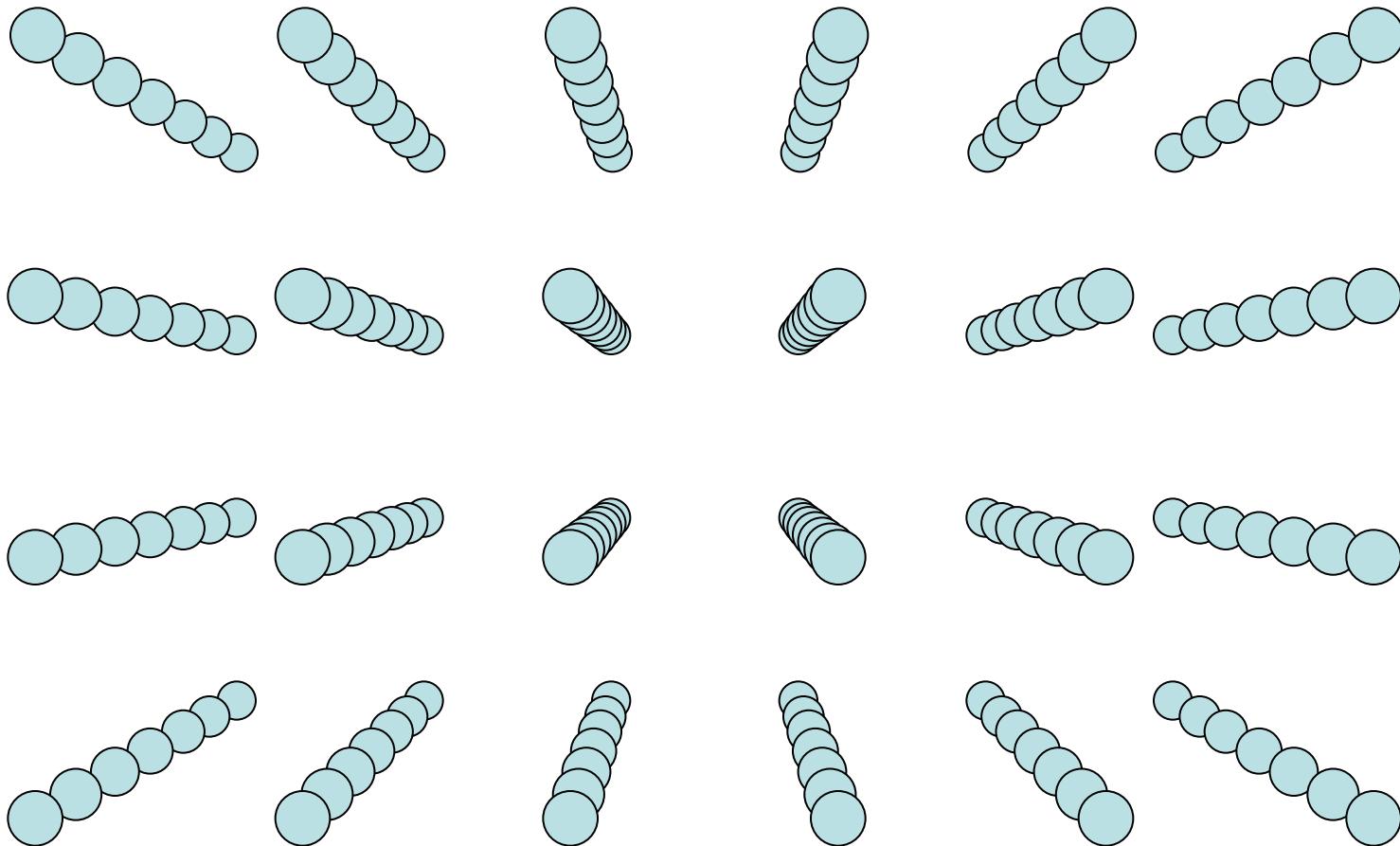
- Phases

Non-isomorphism (including Radiation Damage)

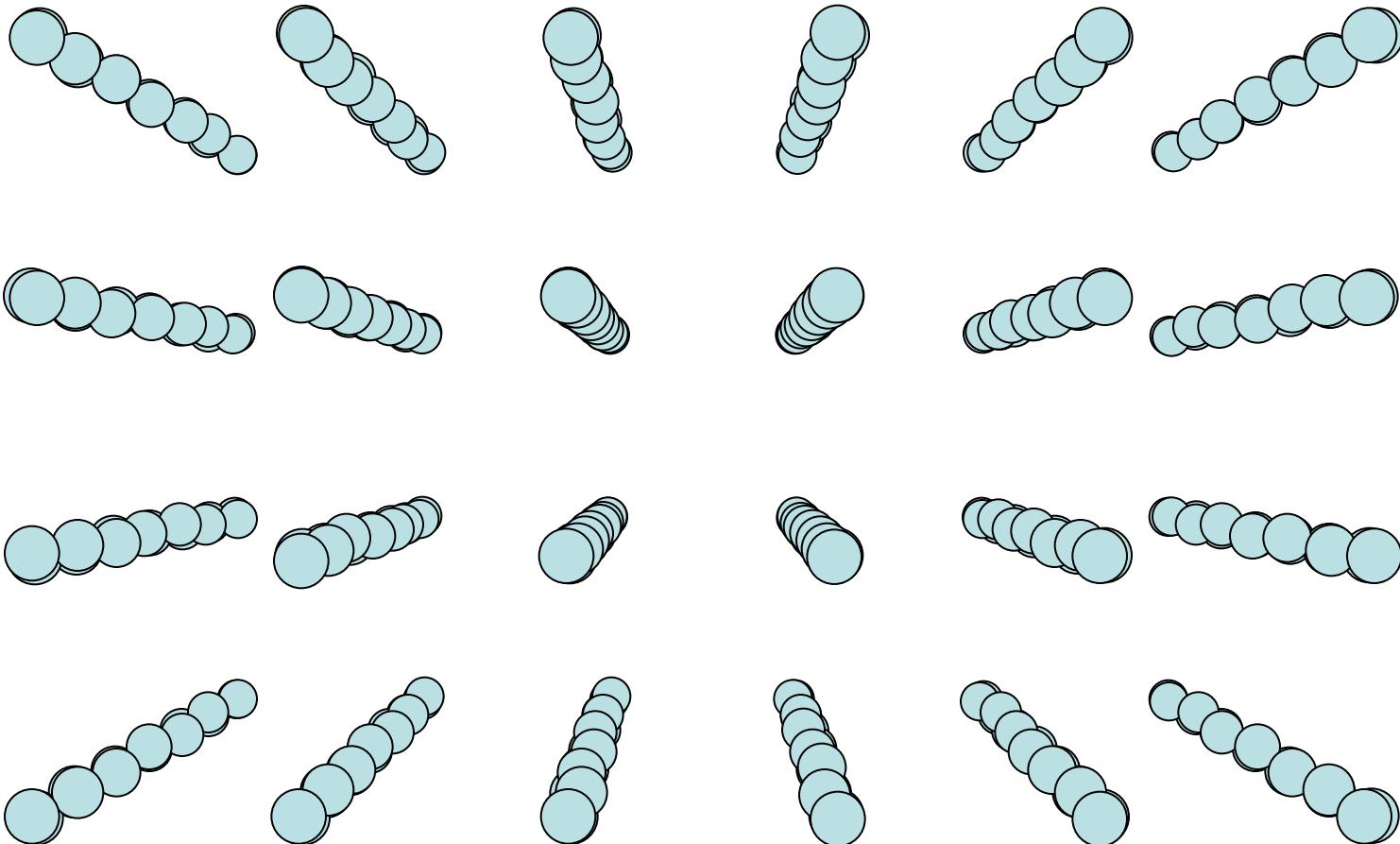
Vibration (beam, xtal, spindle, etc.)

Detector calibration

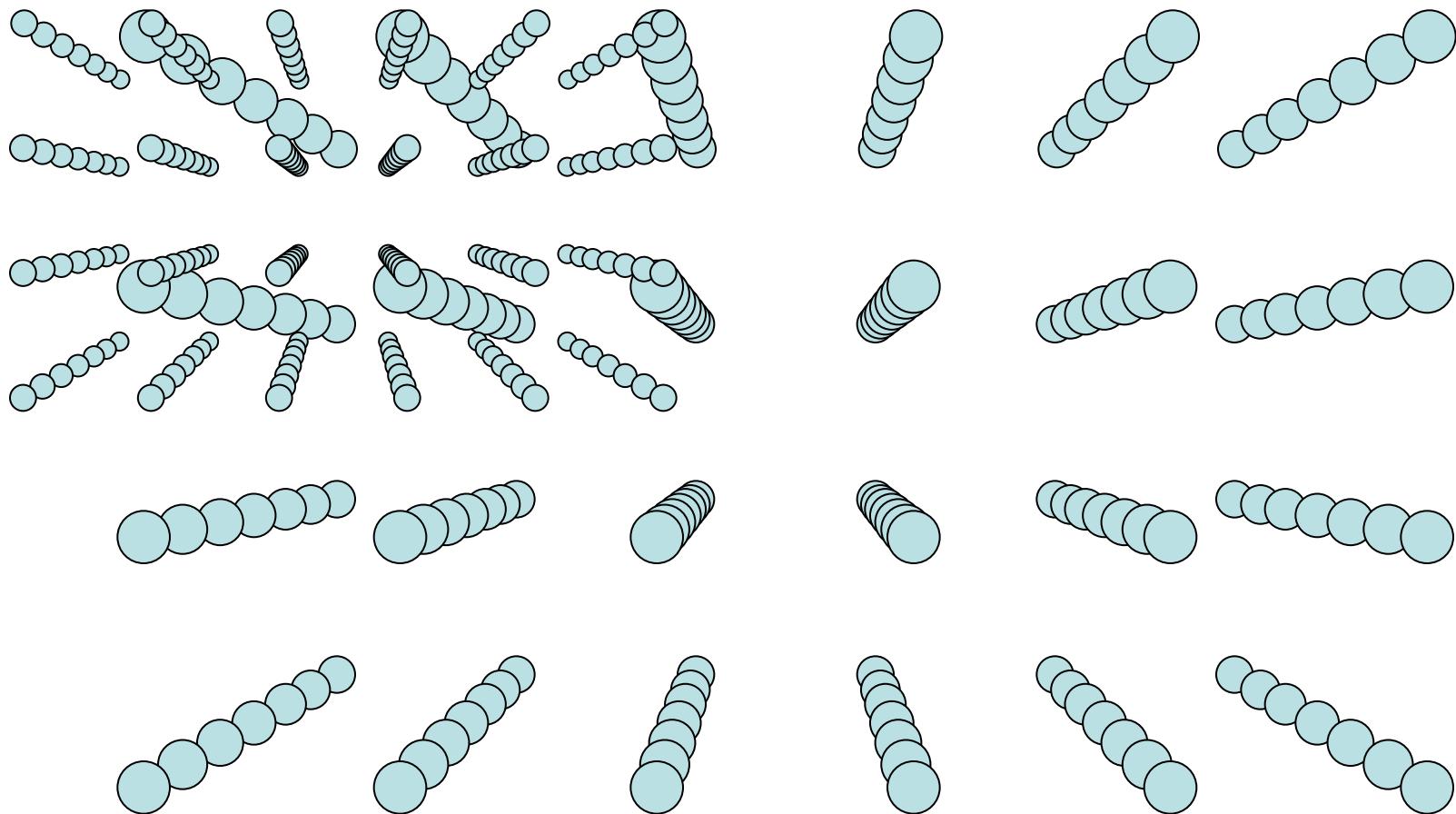
What is “disorder”?



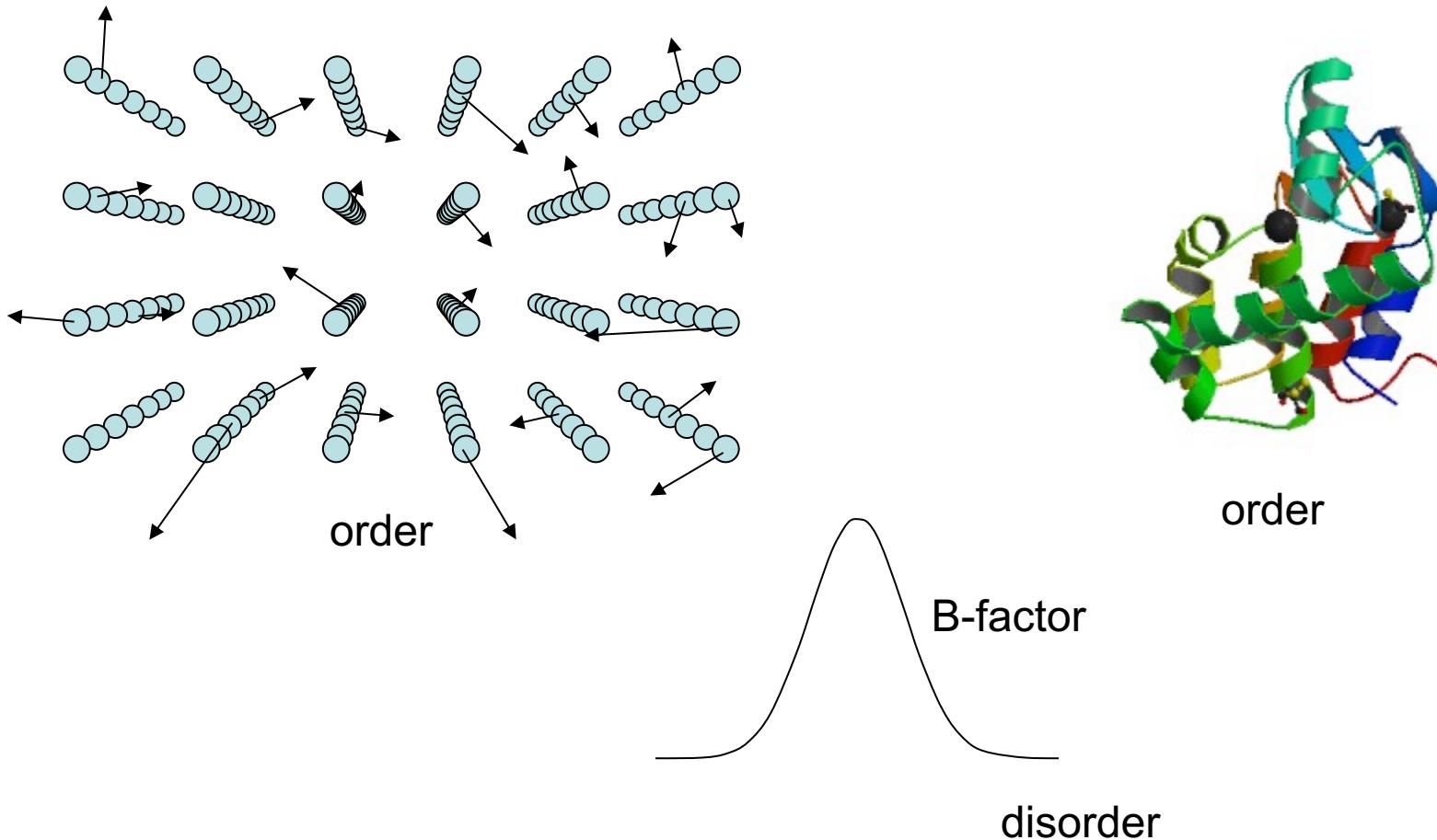
What is “disorder”?



What is “disorder”?



What is “disorder”?



“B” factors

ATOM	122	N	LEU	A	13	-3.244	25.808	19.998	1.00	16.96
ATOM	123	CA	LEU	A	13	-2.877	25.448	21.355	1.00	15.29
ATOM	124	C	LEU	A	13	-2.792	23.966	21.561	1.00	17.54
ATOM	125	O	LEU	A	13	-1.814	23.493	22.143	1.00	16.35
ATOM	126	CB	LEU	A	13	-3.907	26.164	22.268	1.00	18.72
ATOM	127	CG	LEU	A	13	-3.577	25.982	23.738	1.00	21.19
ATOM	128	CD1	LEU	A	13	-2.283	26.820	24.019	1.00	19.43
ATOM	129	CD2	LEU	A	13	-4.702	26.474	24.639	1.00	24.65
ATOM	130	N	SER	A	14	-3.677	23.149	20.979	1.00	15.96
ATOM	131	CA	SER	A	14	-3.646	21.711	21.061	1.00	18.26
ATOM	132	C	SER	A	14	-2.373	21.203	20.360	1.00	18.71
ATOM	133	O	SER	A	14	-1.747	20.315	20.930	1.00	17.47
ATOM	134	CB	SER	A	14	-4.875	21.077	20.419	1.00	17.62
ATOM	135	OG	ASER	A	14	-4.825	19.665	20.388	0.50	20.89
ATOM	136	OG	BSER	A	14	-6.027	21.408	21.164	0.50	18.67
ATOM	137	N	LYS	A	15	-2.045	21.772	19.215	1.00	18.03
ATOM	138	CA	LYS	A	15	-0.799	21.361	18.555	1.00	18.12
ATOM	139	C	LYS	A	15	0.446	21.707	19.351	1.00	18.81
ATOM	140	O	LYS	A	15	1.400	20.948	19.411	1.00	17.77
ATOM	141	CB	LYS	A	15	-0.700	22.034	17.177	1.00	14.49
ATOM	142	CG	LYS	A	15	-1.727	21.368	16.256	1.00	16.12
ATOM	143	CD	LYS	A	15	-1.663	22.147	14.936	1.00	19.40
ATOM	144	CE	ALYS	A	15	-2.725	21.614	13.986	0.50	17.42

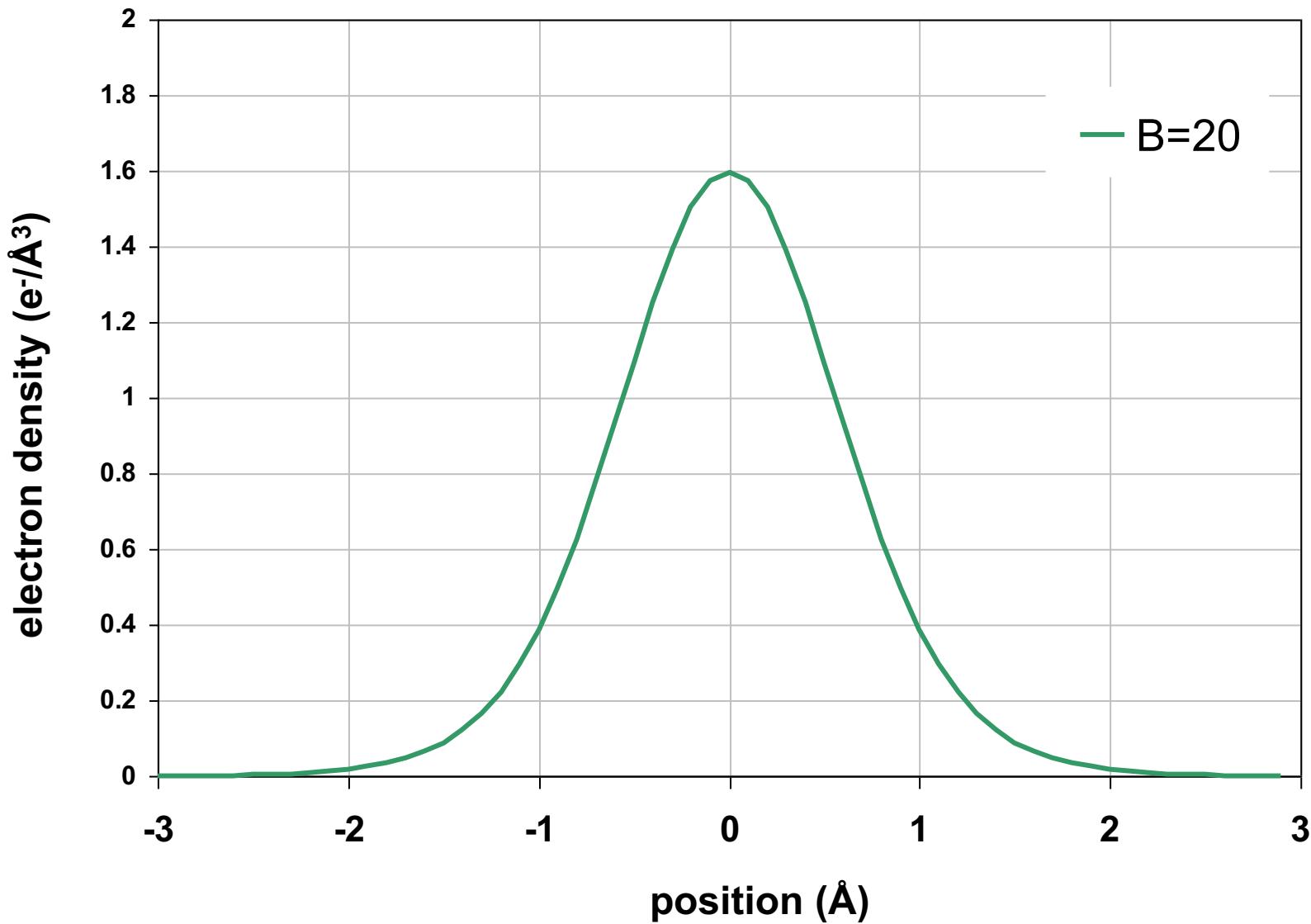
“B” factors

$$B \approx 4d^2 + 12$$

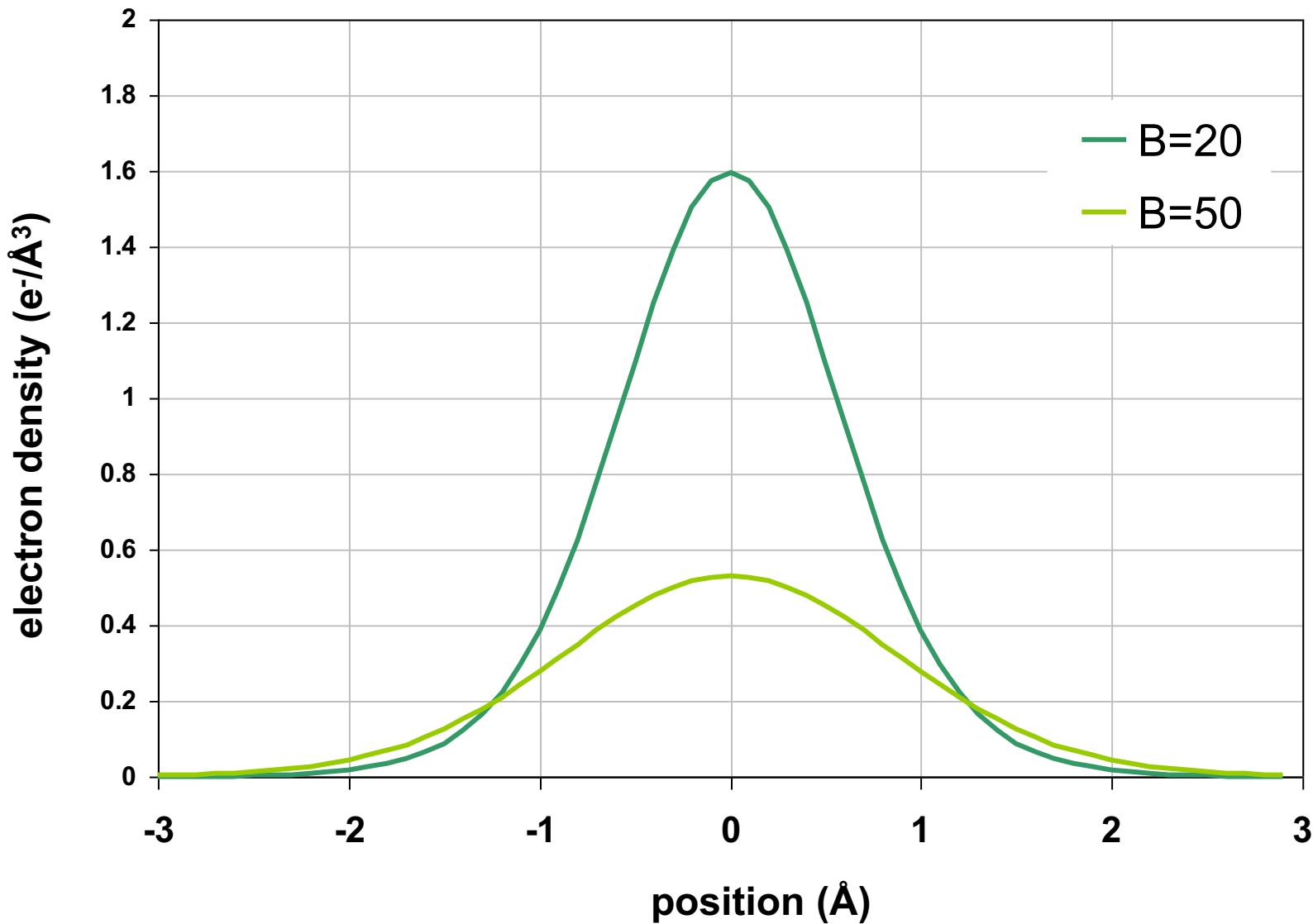
d = resolution in Å

essentially, the “resolution” of an atom

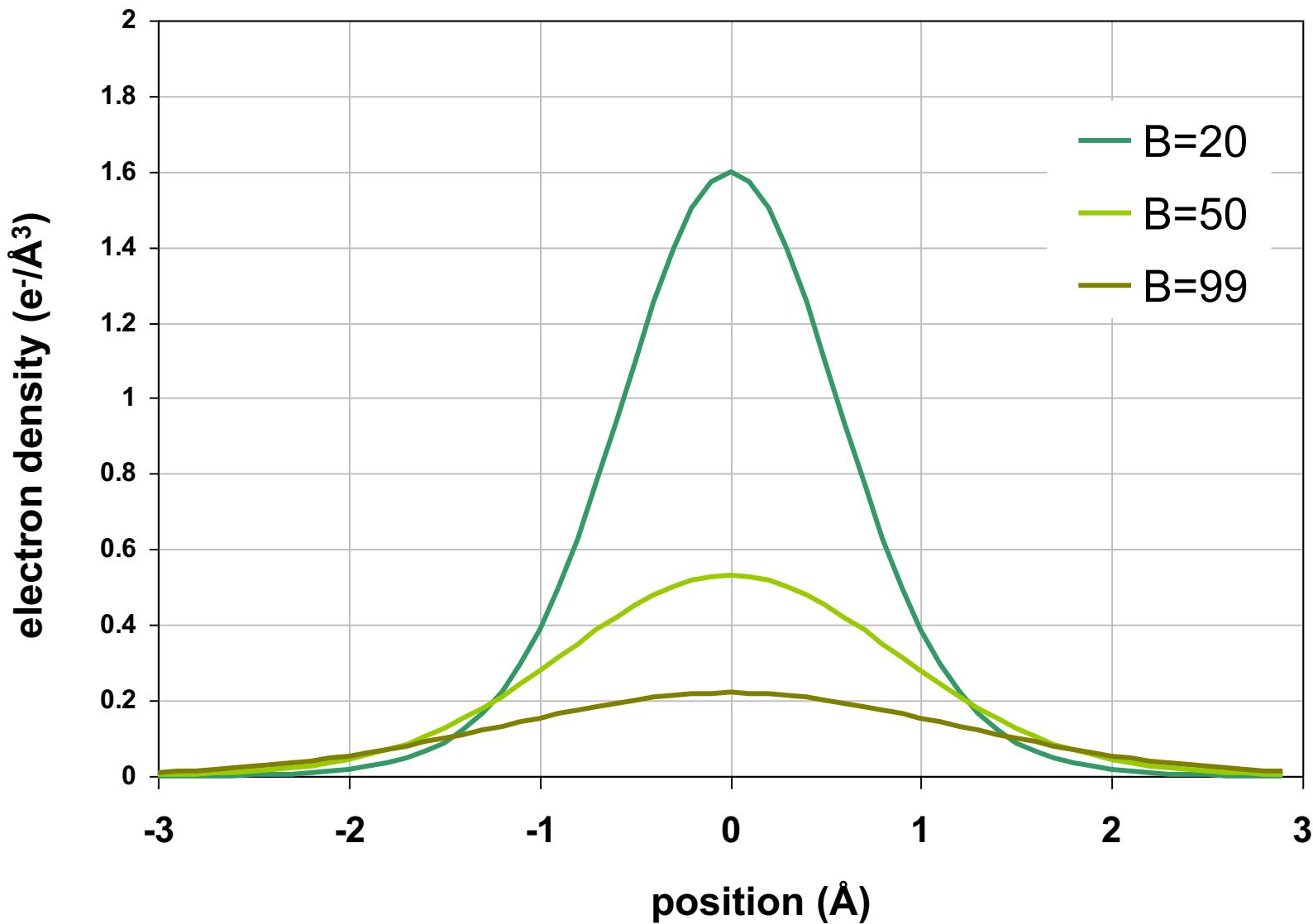
“B” factors



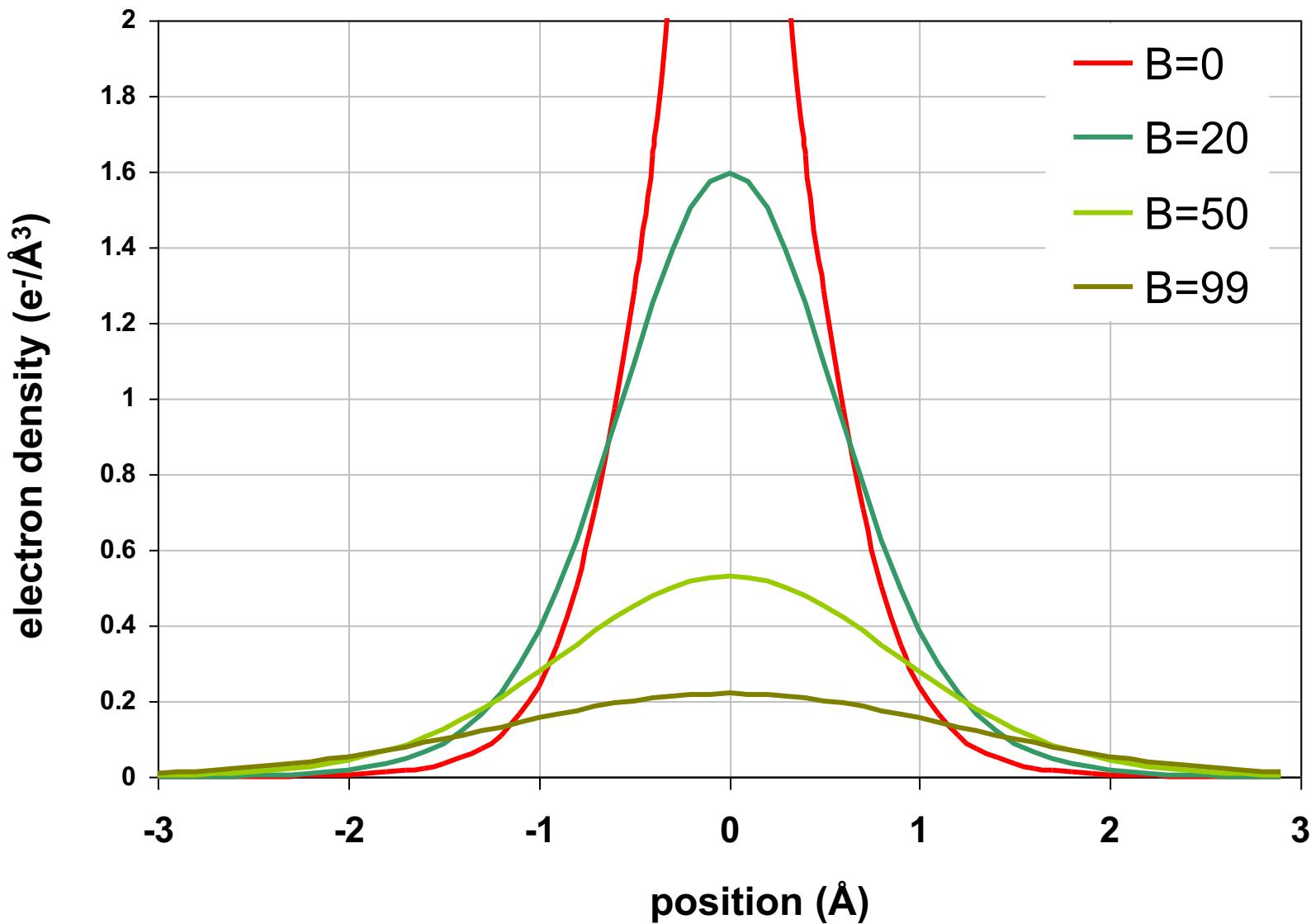
“B” factors



“B” factors



“B” factors



The B factor

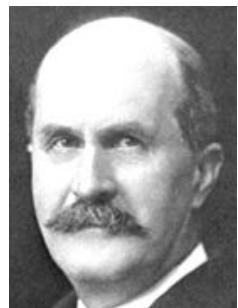
The effects of temperature have been foreseen and calculated by Debije *, and the figures of the table may be looked on as measurements of the Debije effect.

It is interesting to compare the results given above with the corresponding figures calculated from Debije's formula. The latter may be written in the form

$$I_\theta = \frac{A}{\sin^2 \theta} \cdot e^{-B \sin^2 \theta},$$

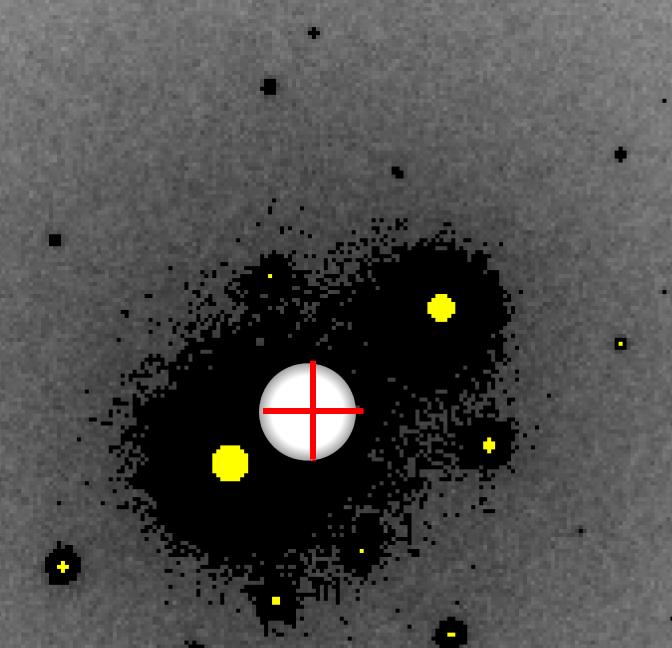
where A and B are constants, given the crystal, the wavelength, and the temperature, and θ is the glancing angle at which the pencil is reflected. I_θ is the intensity of the reflected pencil.

The constant B contains a quantity which Debije calls the characteristic temperature of the crystal. Its value for rocksalt is not known with certainty, but is believed to be



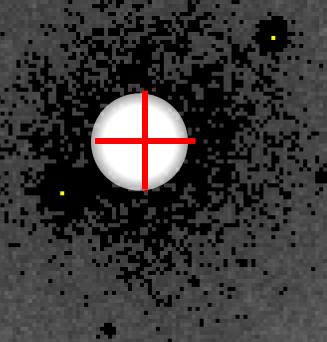
B factor from image analysis

$B = 500$



B factor from image analysis

B = 20



The real problem is:



Averaging
+
Dynamics
=

Poor
resolution

a grand challenge for structural biology in the 21st century

What Causes Disorder?

- Function
- Radiation damage
- Mosaicity
- Strain
- Xtal abuse
- Dehydration

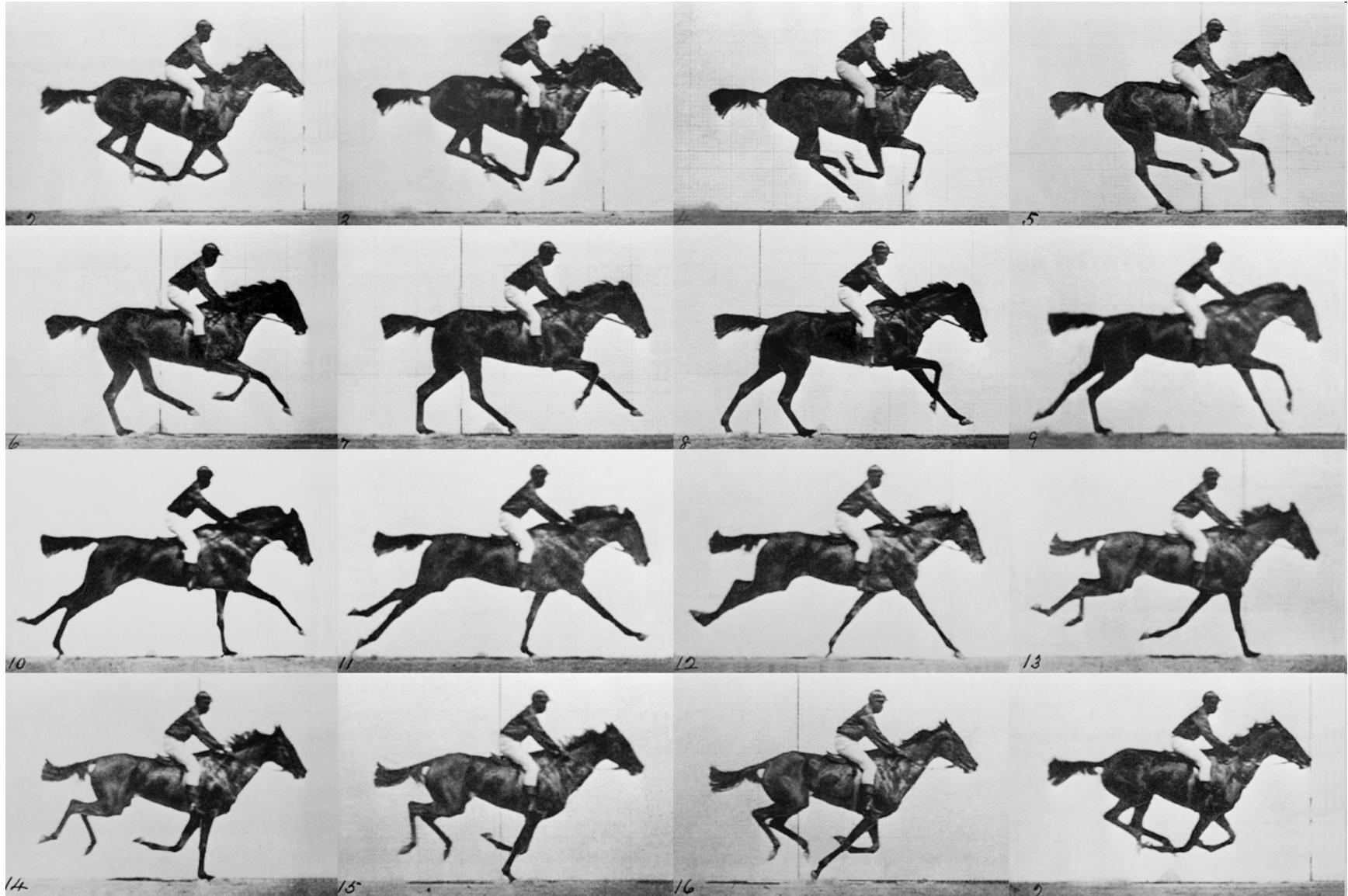
What Causes Disorder?

- Function
- Radiation damage
- Mosaicity
- Strain
- Xtal abuse
- Dehydration

Muybridge's galloping horse (1878)



realistic “crystal” of horses



average structure: galloping horse



Ways to improve order

- New crystal form
 - Purify
 - add a column
 - heat shock
 - Bind something
 - Known ligand
 - Silver bullets
 - Change the clone
 - chop off floppy bits
 - homolog
 - Ala -> Lys mutants
 - Make Protein Sit Still!!!
- 8000 cases in PDB of resolution improvement
 - 80% of these are different crystal forms
 - > 1 Å improvement → 94% different forms



What Causes Disorder?

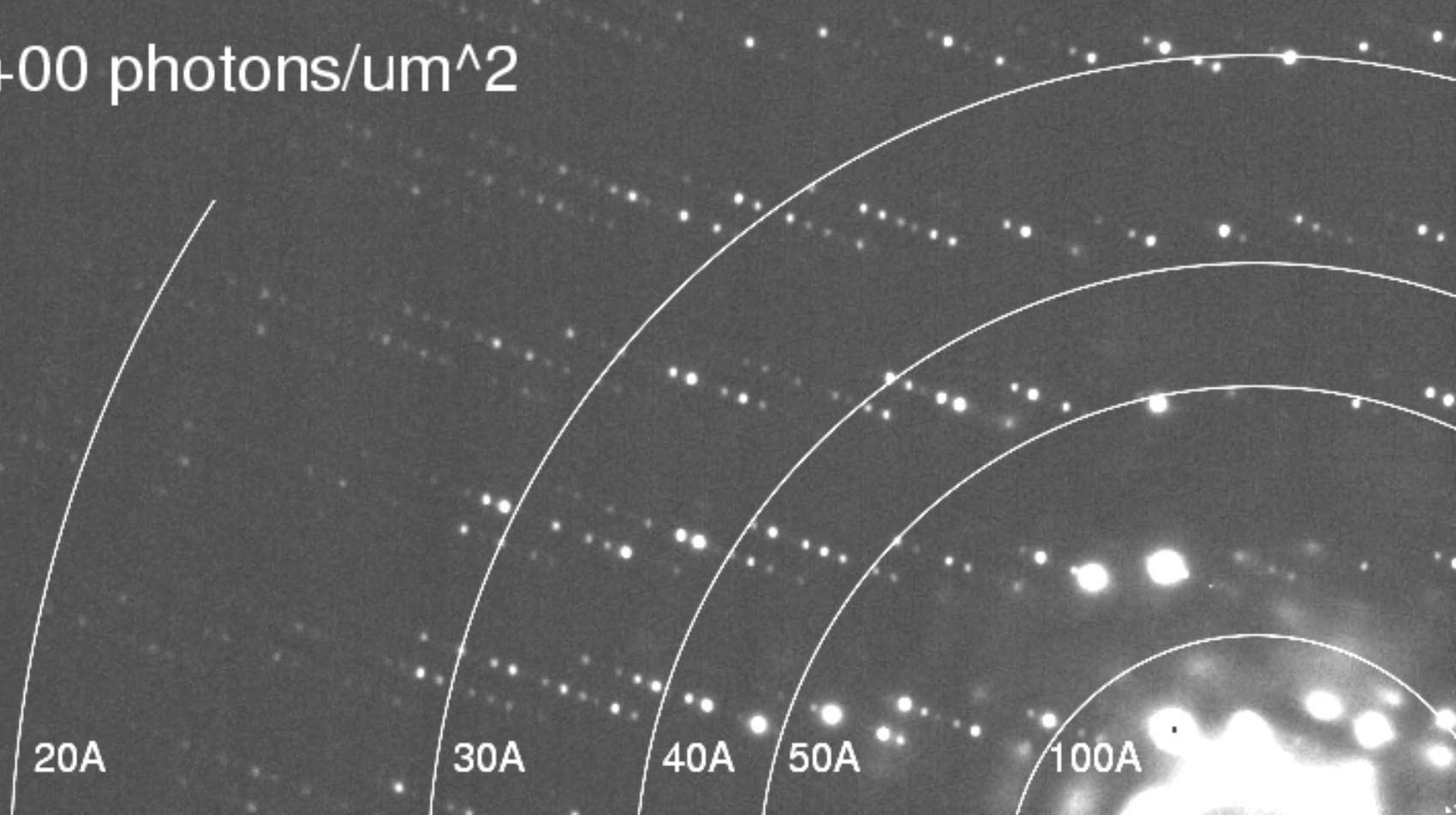
- Function
- **Radiation damage**
- Mosaicity
- Strain
- Xtal abuse
- Dehydration

Types of radiation damage

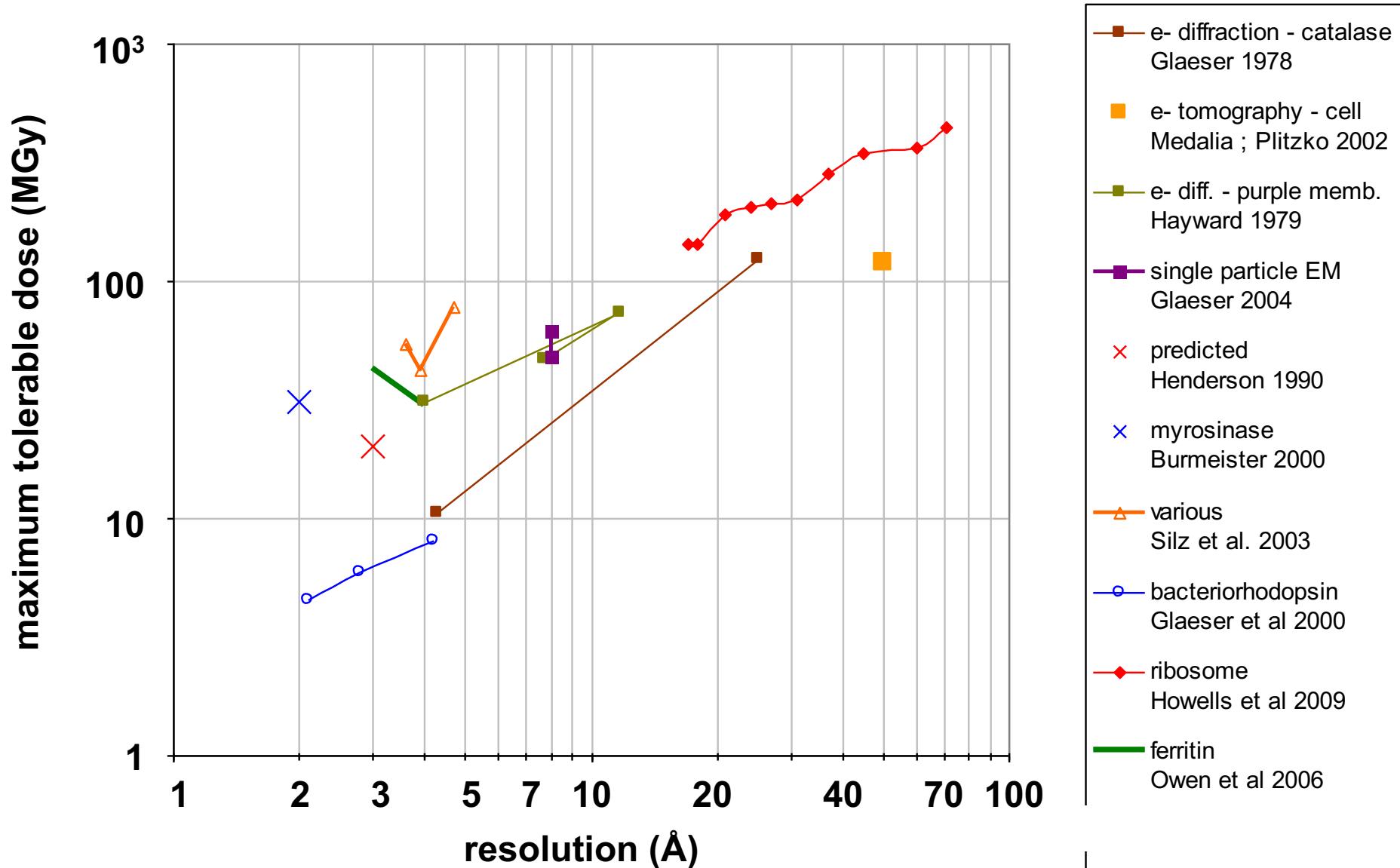
- Global
- Specific

Global damage

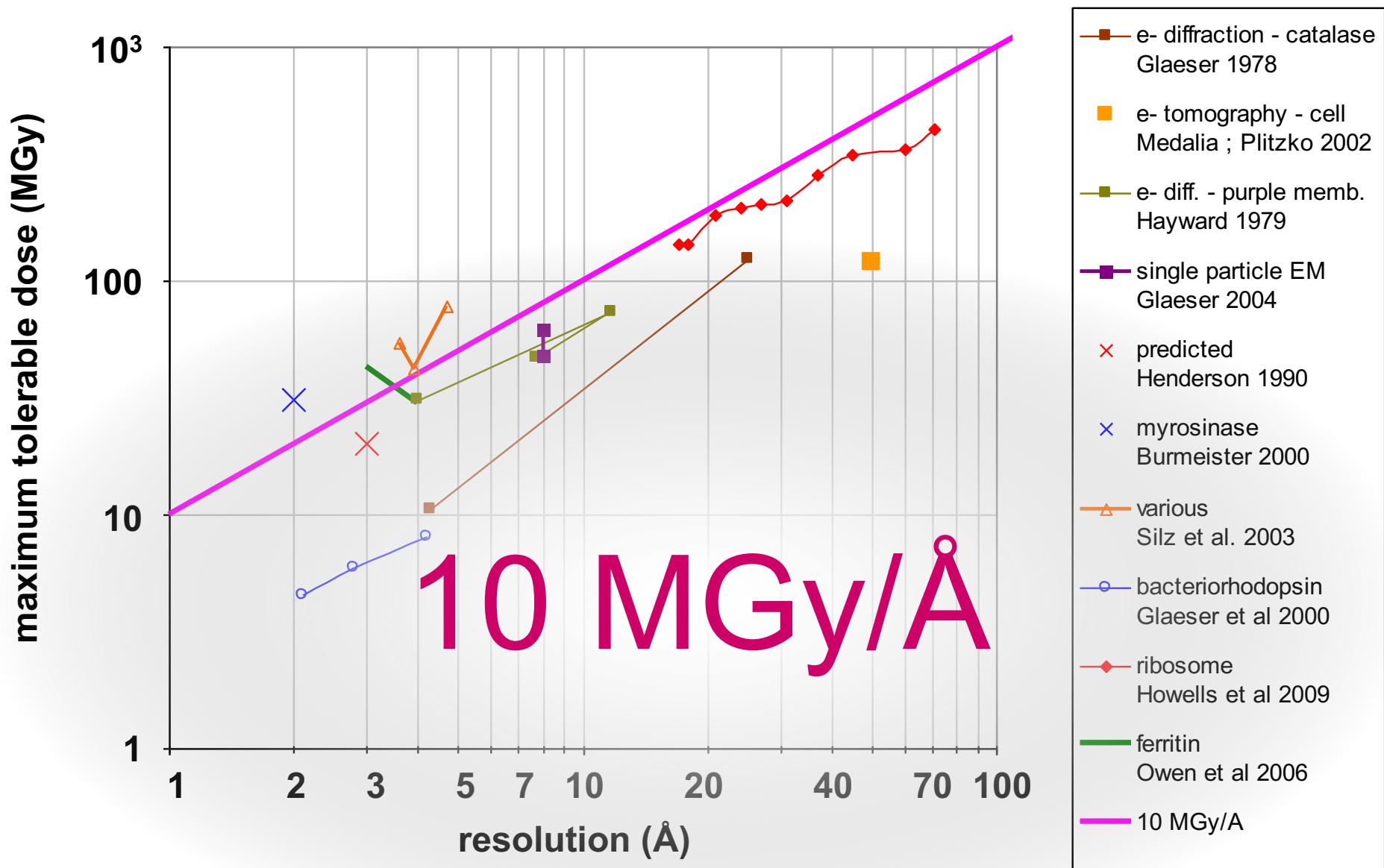
0.0e+00 photons/ μm^2



resolution dependence of global damage



resolution dependence of global damage



what the is a MGy?

$$1 \text{ MGy} = 10^6 \text{ J/kg}$$

[http://bl831.als.lbl.gov/
damage_rates.pdf](http://bl831.als.lbl.gov/damage_rates.pdf)

How long will my crystal last?

synch	line	type	flux ph/s	beamsize μm	flux density ph/μm ² /s	dose rate	max xtal lifetime	min site lifetime
ALS	4.2.2	MAD	2.2e11	75x80	3.7e+07	27.3 kGy/s	18 m	73 s
ALS	5.0.1	mono	1.6e11	100	2.0e+07	10.4 kGy/s	48 m	3.2 m
ALS	5.0.2	MAD	8e11	100	1.0e+08	51.8 kGy/s	9.6 m	39 s
ALS	5.0.3	mono	1.7e11	100	2.2e+07	11 kGy/s	45 m	3 m
ALS	8.2.1	MAD	1.8e11	100	2.3e+07	11.7 kGy/s	43 m	2.9 m
ALS	8.2.2	MAD	2.3e11	100	2.9e+07	14.9 kGy/s	34 m	2.2 m
ALS	8.3.1	MAD	9e11	70	2.3e+08	119 kGy/s	4.2 m	17 s
ALS	8.3.1	typical	6e11	70	1.6e+08	115 kGy/s	4.3 m	17 s
ALS	12.3.1	MAD	1.8e11	100	2.3e+07	11.7 kGy/s	43 m	2.9 m
ALS	12.3.1	ML	4.0e12	100	5.1e+08	513 kGy/s	58 s	3.9 s
APS	8-BM	MAD	1e11	200	2.5e+06	1.27 kGy/s	6.6 h	26 m
APS	14-BM-C	mono	5.8e10	200	1.4e+06	738 Gy/s	11 h	45 m
APS	14-BM-D	MAD	3.3e9	200	8.2e+04	42 Gy/s	8.3 d	13 h
APS	14-ID-B	MAD	6.0e10	200	1.5e+06	763 Gy/s	11 h	44 m
APS	17-BM	MAD	1.1e11	200	2.8e+06	1.4 kGy/s	6 h	24 m
APS	17-ID	MAD	2.3e11	200	5.8e+06	2.93 kGy/s	2.8 h	11 m
APS	19-BM	MAD	2.0e11	70x60	4.8e+07	24.2 kGy/s	21 m	83 s
APS	19-ID	MAD	1.3e13	80x40	4.1e+09	2.07 MGy/s	15 s	0.97 s
APS	19-ID	typical	5.5e11	100x100	5.5e+07	28 kGy/s	18 m	71 s
APS	22-BM	MAD	7e12	80x40	2.2e+09	1.23 MGy/s	24 s	1.6 s
APS	22-ID	MAD	7e12	80x40	2.2e+09	1.23 MGy/s	24 s	1.6 s
APS	22-ID	typical	1.5e12	80	2.3e+08	119 kGy/s	4.2 m	17 s
APS	23-ID-B	MAD	1e13	75x25	5.3e+09	3.01 MGy/s	10 s	0.66 s
APS	23-ID	typical	1.5e12	80	2.3e+08	119 kGy/s	4.2 m	17 s
APS	24-ID-C	MAD	1.3e13	20x60	1.1e+10	5.23 MGy/s	5.7 s	0.38 s
APS	24-ID-E	MAD	0.5e13	20x100	2.5e+09	1.19 MGy/s	25 s	1.7 s
APS	31-ID	MAD	2e12	70	4.1e+08	194 kGy/s	2.6 m	10 s

expected crystal lifetime calculator

source =

full flux =

7.0e+12 photons/s

attenuation =

transmittance = beam size_{horiz} =

40.0 microns

beam size_{vert} = 80.0 microns

wavelength =

1 Ang

k_{dose} = 2000 photons/micron²/Gy

dose rate =

1.1e+6 Gy/s

experiment goal =

high resolution (cryo)

resolution =

3 Ang

dose limit =

30 MGy

exposure time =

1 seconds/image

xtal size_{horiz} =

50 microns

xtal size_{vert} = 50 microns

translation during dataset =

0 microns

rotisserie factor disable warnings

max images =

28 at damage limit

inverse beam =

no

number of wavelengths =

1

images/wedge =

28

required number of crystals calculator - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://bl831.als.lbl.gov/xtalsize.html Wikipedia (en)

required number of crystals calculator

Required crystal number or size calculator

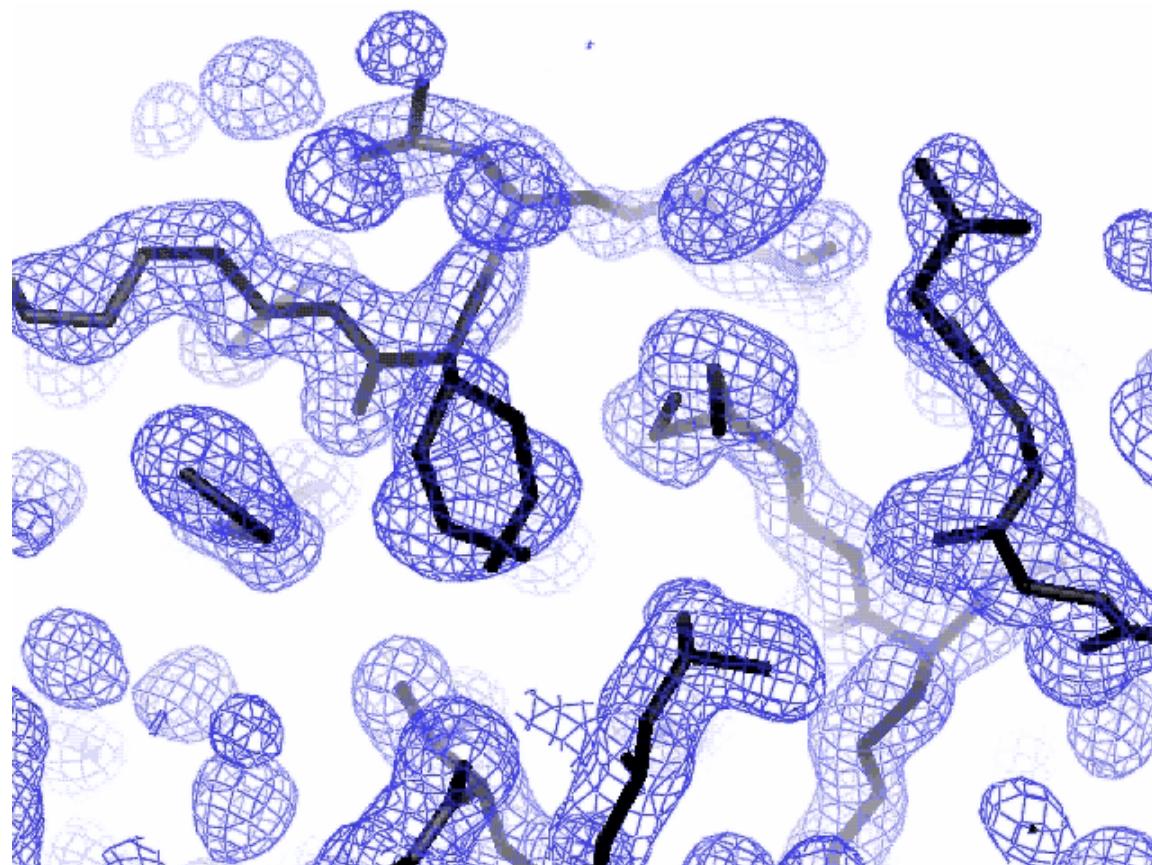
$n_{\text{xtals}} = \langle I_{\text{DL}} \rangle / 20 * f_{\text{NH}} * \text{MW} * V_M^2 / \exp(-0.5 * B/\text{reso}^2) / \text{xtalsize}^3 / (\text{reso}^3 - 1.53)$

Enter values:

experiment goal =	<input type="button" value="faint spots (MR)"/>	at this resolution	
molecular weight =	<input type="text" value="30"/> kDa in asymmetric unit	<input type="text" value="2"/>	
resolution =	<input type="text" value="2.4"/> Å	signal to noise =	<input type="text" value="2"/>
reso on snapshot =	<input type="text" value="2.4"/> Å	→ Wilson B =	<input type="text" value="35"/> Å ²
background level =	<input type="text" value="260"/> ADU/pixel	multiplicity =	<input type="text" value="2"/>
spot size =	<input type="text" value="5"/> pixels		
detector type =	<input type="button" value="ADSC Q210/315r (hwbin)"/>		
solvent content =	<input type="text" value="50"/> %		
xtal size _{beam} =	<input type="text" value="20"/> microns	beam size _{vert} =	<input type="text" value="100"/> microns
xtal size _{vert} =	<input type="text" value="20"/> microns	beam size _{spindle} =	<input type="text" value="100"/> microns
xtal size _{spindle} =	<input type="text" value="20"/> microns		
<input type="button" value="Calculate n_xtals"/>	<input type="button" value="Calculate size ↑"/>		
$n_{\text{xtals}} =$	<input type="text" value="0.29"/> xtals you will need to merge	$\leftarrow \langle I_{\text{DL}} \rangle$	<input type="text" value="160"/> photons/hkl

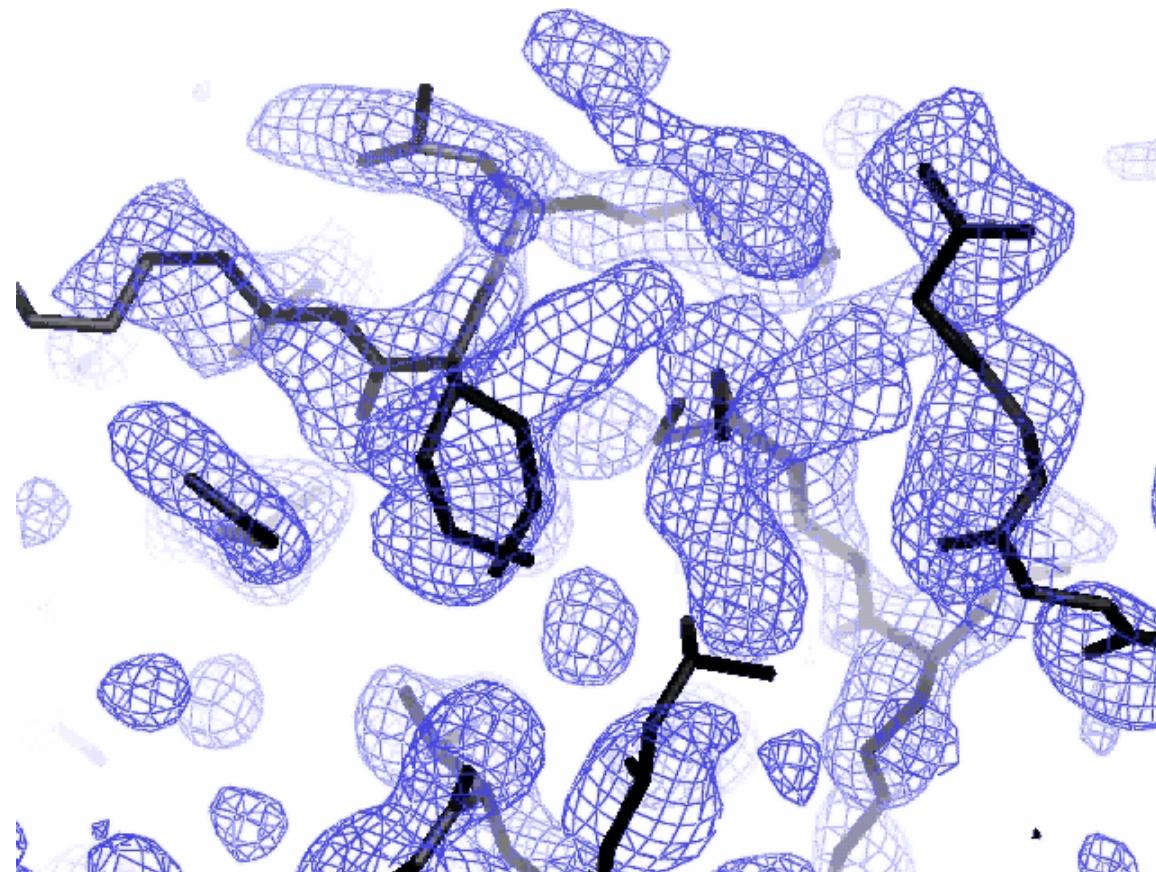
Done

Specific damage



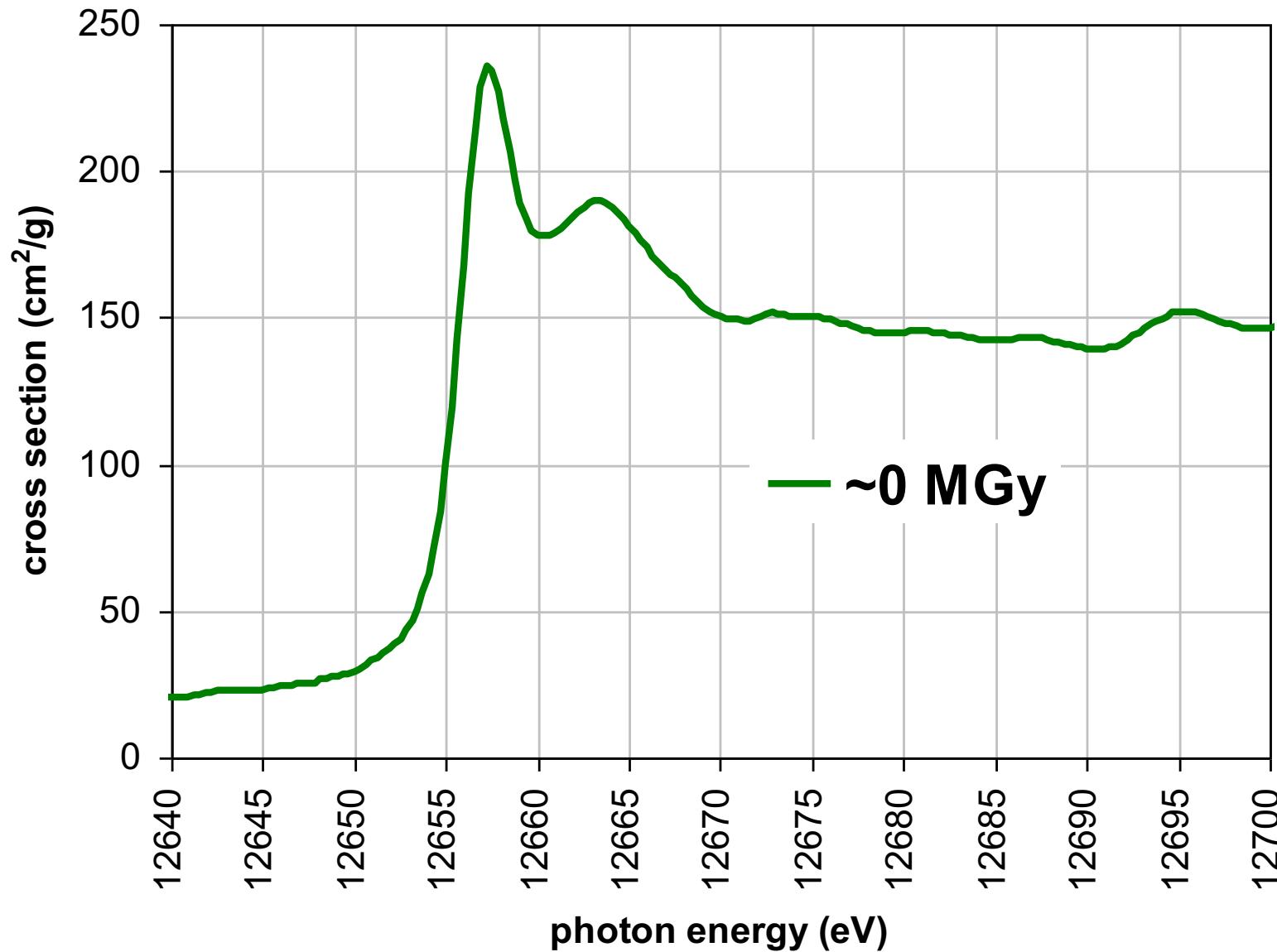
final, refined phases

Specific damage

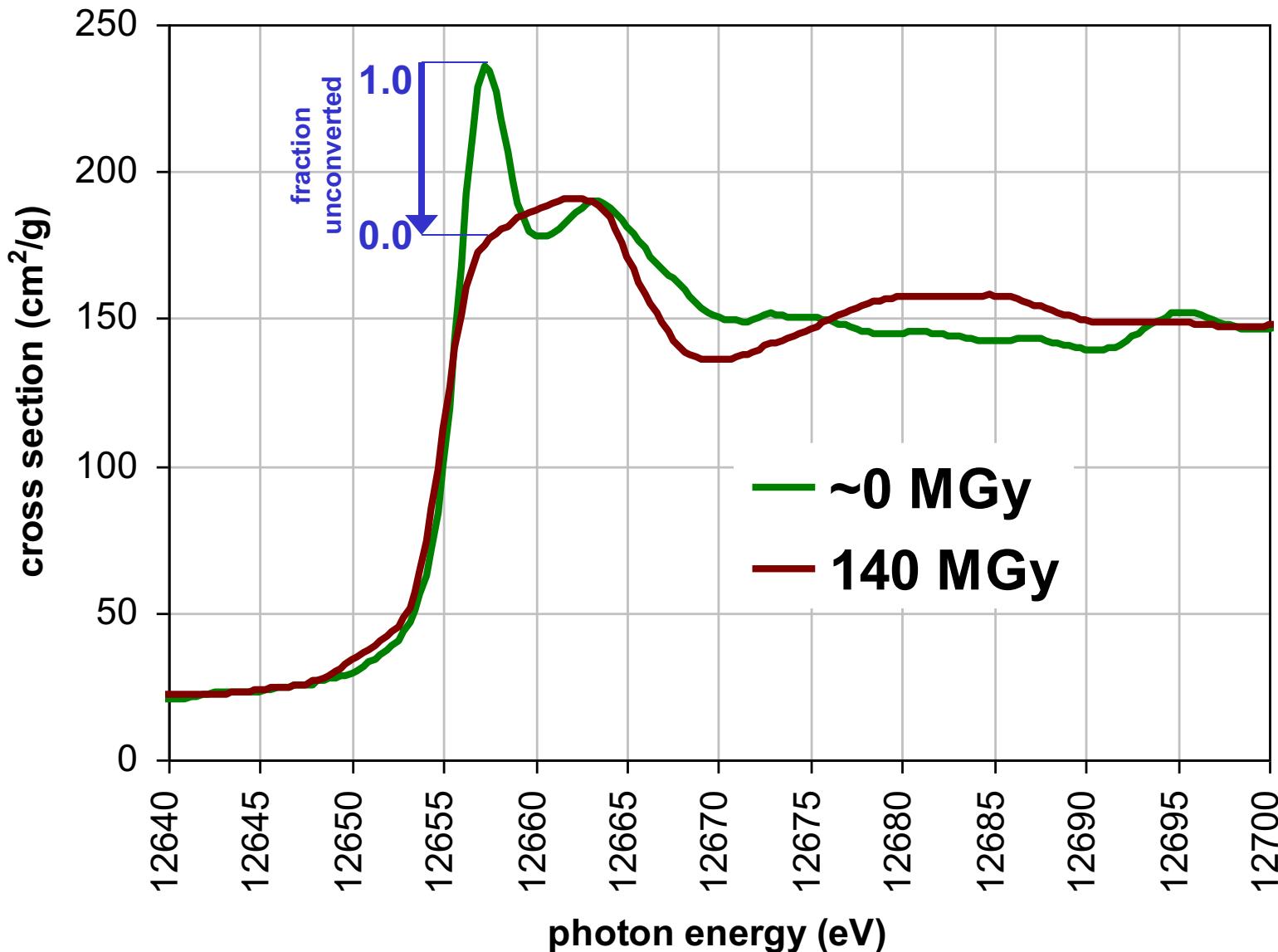


experimentally-obtained phases

damage changes absorbance



damage changes absorbance



Radiation Damage World Records

MGy	reaction	reference
~45	global damage	Owen <i>et al.</i> (2006)
10/Å	global damage	Howells <i>et al.</i> (2009)
5	Se-Met	Holton (2007)
4	Hg-S	Ramagopal <i>et al.</i> (2004)
4	R-C-COOH	Garman <i>et al.</i> (2015)
3	S-S	Murray <i>et al.</i> (2002)
1	Br-RNA	Olieric <i>et al.</i> (2007)
~1?	Cl-C	???
0.5	Mn in PS II	Yano <i>et al.</i> (2005)
0.06	putidaredoxin	Corbett <i>et al.</i> (2007)
0.02	Fe in myoglobin	Denisov <i>et al.</i> (2007)

Life-halving concentration

molar, at the Se edge based on RADDOSE

Na	44	As	0.44
Mg	27	Se	0.42
P	9	Br	2.9
S	7	I	0.47
Cl	5	Gd	0.21
K	3.3	Ta	0.13
Ca	2.7	Pt	0.17
Fe	1.0	Au	0.16
Cu	0.67	Hg	0.16
Zn	0.58	U	0.23

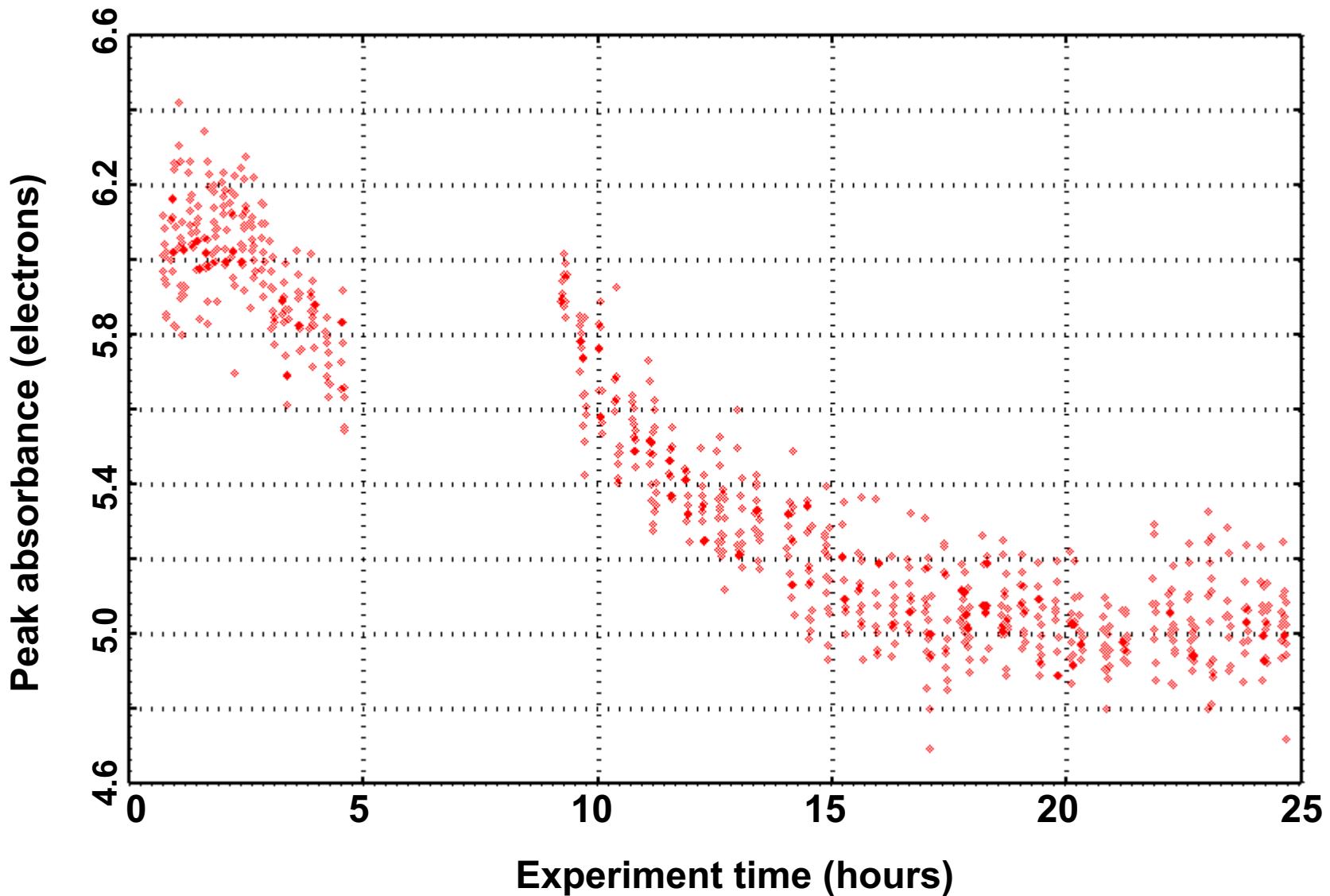
Life-halving concentration

molar, **at the Se edge** based on RADDOSE

Na	44
Mg	27
P	9
S	7
Cl	5
K	3.3
Ca	2.7
Fe	1.0
Cu	0.67
Zn	0.58

As	0.44
Se	0.42
Br	2.9
I	0.47
Gd	0.21
Ta	0.13
Pt	0.17
Au	0.16
Hg	0.16
U	0.23

No Damage in the Dark!



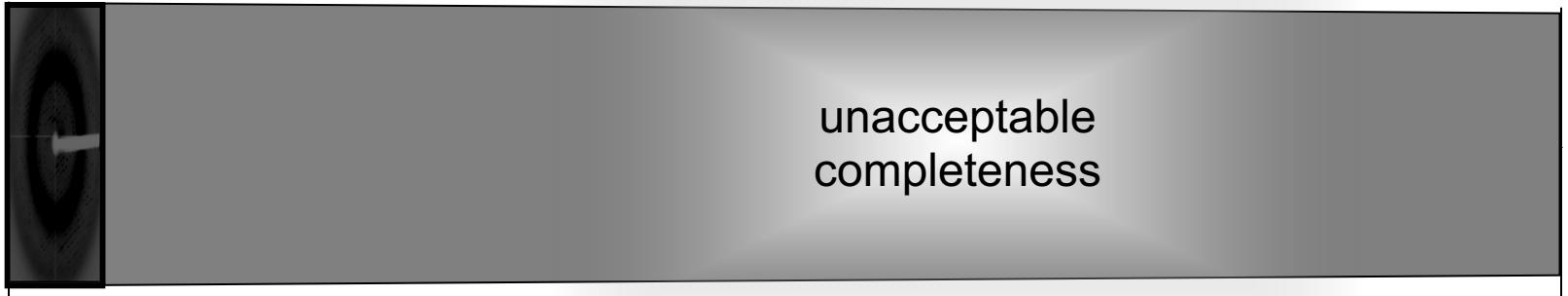
Rad dam take-home lesson:
The number of **photons** scattered
before crystal is dead
is **independent**
of flux & time

1 $\text{um}^3 = 10^5$ photons (roomT)
1 $\text{um}^3 = 10^6$ photons (synch)
1 $\text{um}^3 = 10^8$ photons (XFEL)

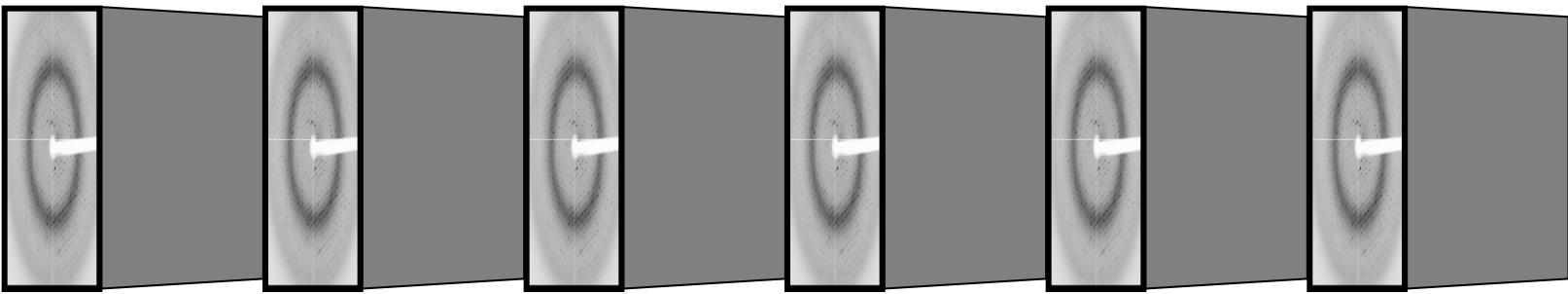
Henderson, 1990; Gonzalez & Nave, 1994; Glaeser *et al.*, 2000; Sliz *et al.*, 2003; Leiros *et al.*, 2006; Owen *et al.*, 2006; Garman & McSweeney, 2006; Garman & Nave, 2009; Holton, 2009

Dose slicing

N
photons



N
photons



N
photons



Zero-dose extrapolation (ZDE)

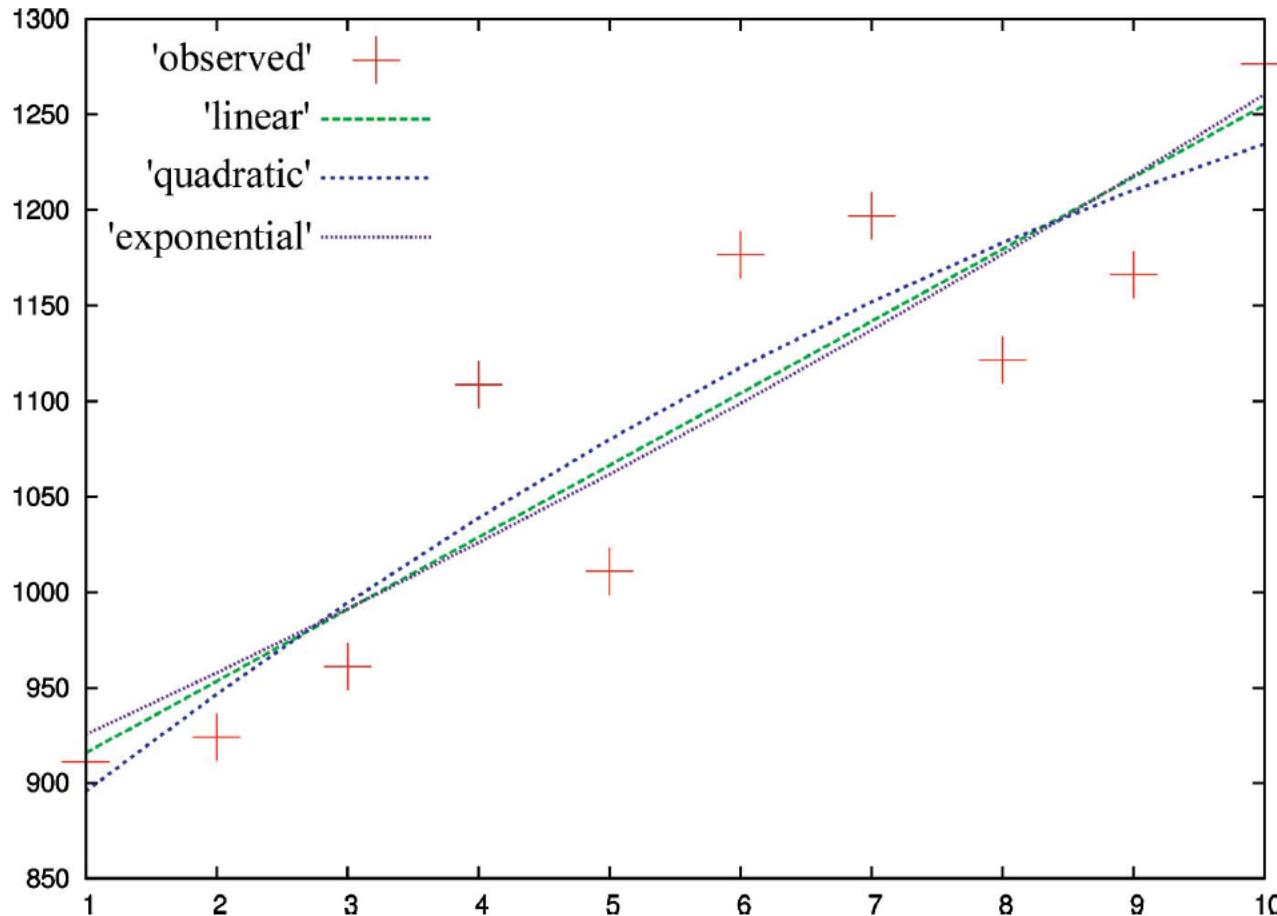


Figure 2

Example of noisy data points (marked by '+') fitted by a linear (green), a quadratic (blue) and an exponential function (purple). Near 1/4 and near 3/4 of the range, the values of the functions used for fitting coincide.

Diederichs, K. (2006). "Some aspects of quantitative analysis and correction of radiation damage." *Acta Cryst. D* **62**(1): 96-101.

Basic Principles

“Hell, there are **NO RULES** here - we're trying to accomplish something.”

Thomas A. Edison – inventor

“You've got to have an **ASSAY**.”

Arthur Kornberg – Nobel Laureate

“Control, control, you must learn **CONTROL!**”

Yoda – Jedi Master

Data quality vs Model Interpretation

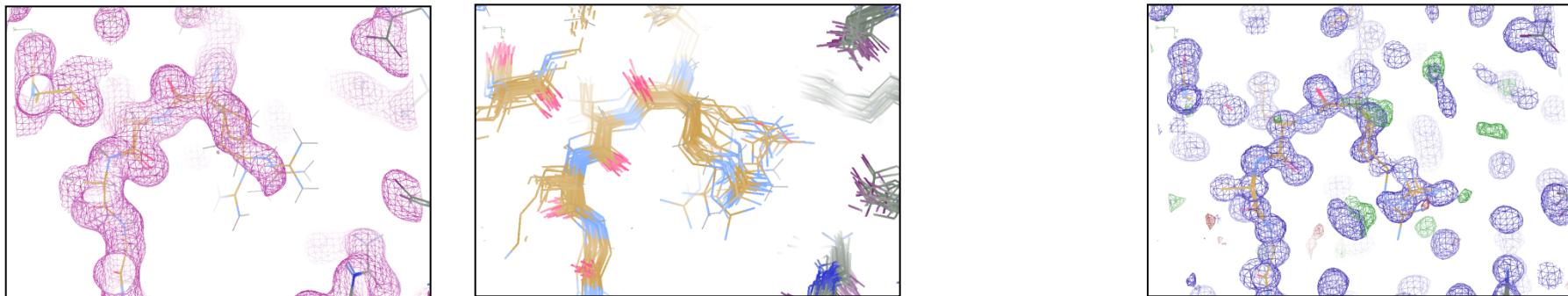
- Big difference features first
 - Don't build anything you don't have to
- Ligand bound or not?
 - Omit maps, “polder” maps, chemistry
- R factor good enough?
 - Rules of thumb
 - Nothing beats a “control”
- Radiation damage?
 - Only control: early vs late data

PowerPoint File available:

[http://bl831.als.lbl.gov/~jamesh/powerpoint/
UCSF_DQvMI_2017.pptx](http://bl831.als.lbl.gov/~jamesh/powerpoint/UCSF_DQvMI_2017.pptx)

Molecular Dynamics vs Observation

1aho 64-residue scorpion toxin in water to 1.0 Å resolution



refined_vs_Fsim.pdb

F_{sim}

1aho.cif

1aho.pdb

F_{calc}

F_{obs}

F_{calc}

$R_{work} = 0.137$

$R_{isolf} = 0.69$

$R_{vault} = 0.48 \text{ to } 4 \text{ \AA}$

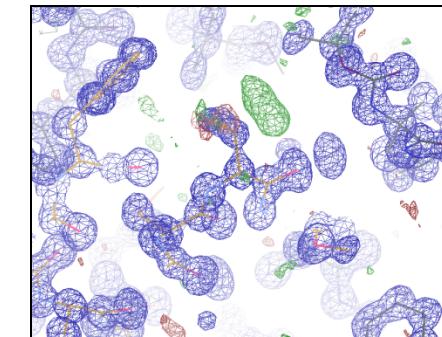
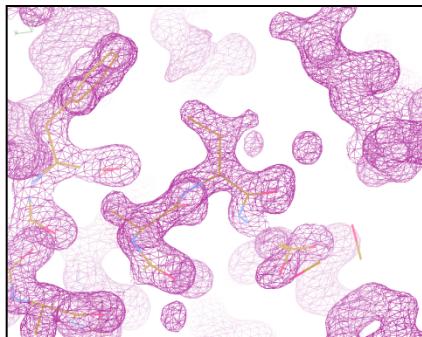
$R_{work} = 0.116$

$\text{rmsd} = 1.05 \text{ \AA}$

$\text{LSQ rmsd} = 0.43 \text{\AA}$

Molecular Dynamics vs Observation

“fav8” 8-residue aromatic peptide with 4 waters to 1.0 Å resolution



refined_vs_Fsim.pdb

F_{calc}

F_{sim}

fav8.fcf

fav8.cif

F_{obs}

F_{calc}

$R_{\text{work}} = 0.13$

$R_{\text{vault}} = 0.23$

$R_{\text{vault}} = 0.20 \text{ to } 2 \text{ Å}$

$R_{\text{work}} = 0.16$

$R1 = 0.041$
With 4σ cutoff!

$\text{rmsd} = 0.20 \text{ Å}$

LSQ rmsd = 0.091 Å & 0.15 Å

Is the structure any good?

“do the data presented justify the conclusions drawn?”