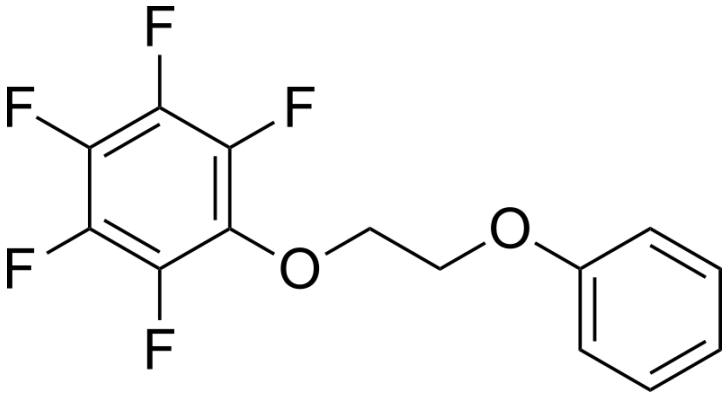
Gas Electron Diffraction
for
Large and Flexible Molecules

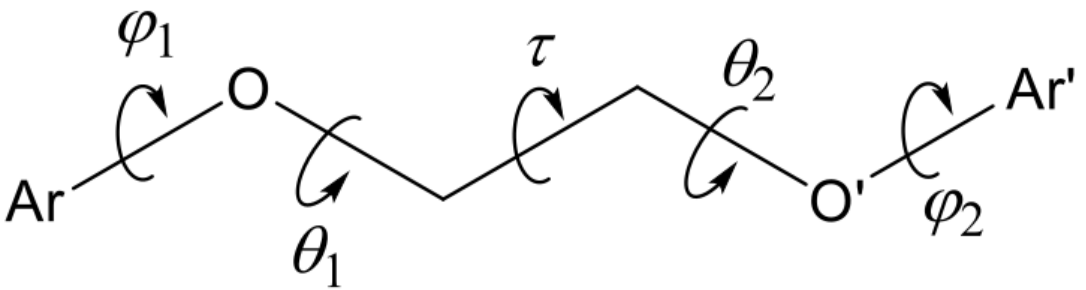
Yury V. Vishnevskiy

Uni-Bielefeld

August 2020

An example



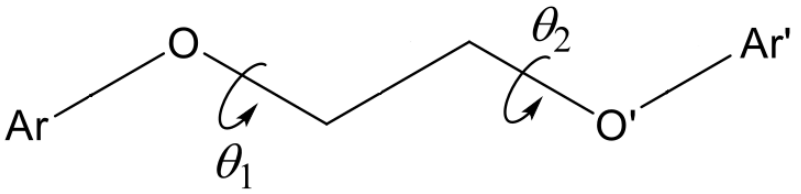


Up to 14 Å

J.-H. Weddeling, Yu. V. Vishnevskiy, B. Neumann, H.-G. Stammler, N. W. Mitzel,
Inter- and Intramolecular Aryl-Aryl-Interactions in Partially Fluorinated Ethylenedioxy-Bridged Bisarenes,
ChemRxiv, Preprint, 2020, DOI: 10.26434/chemrxiv.12643154

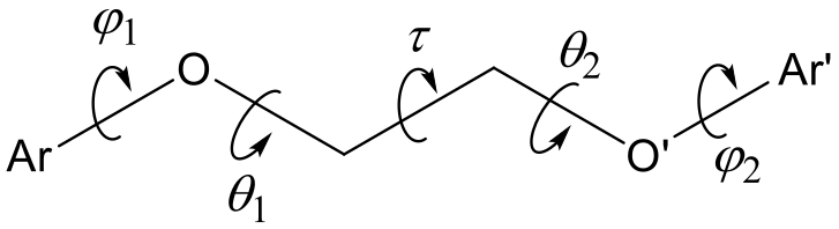
Calculations

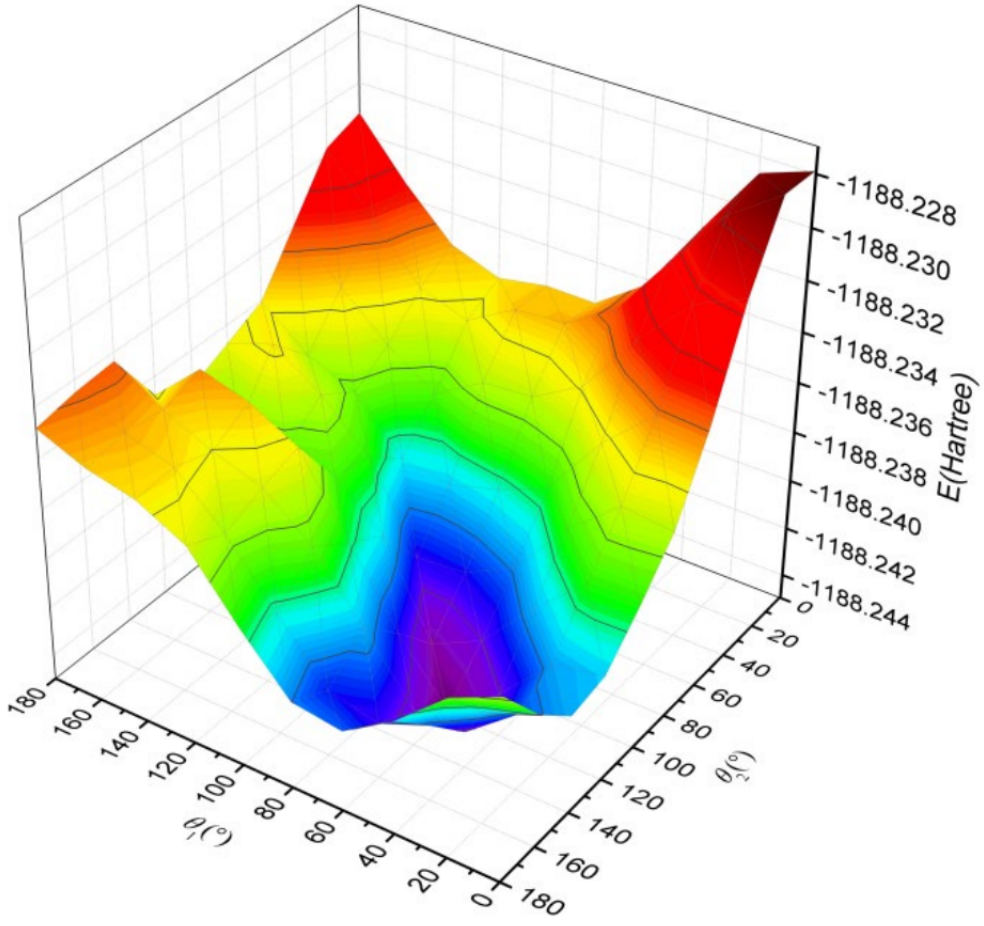
Systematic PES scan?



Single conformation?

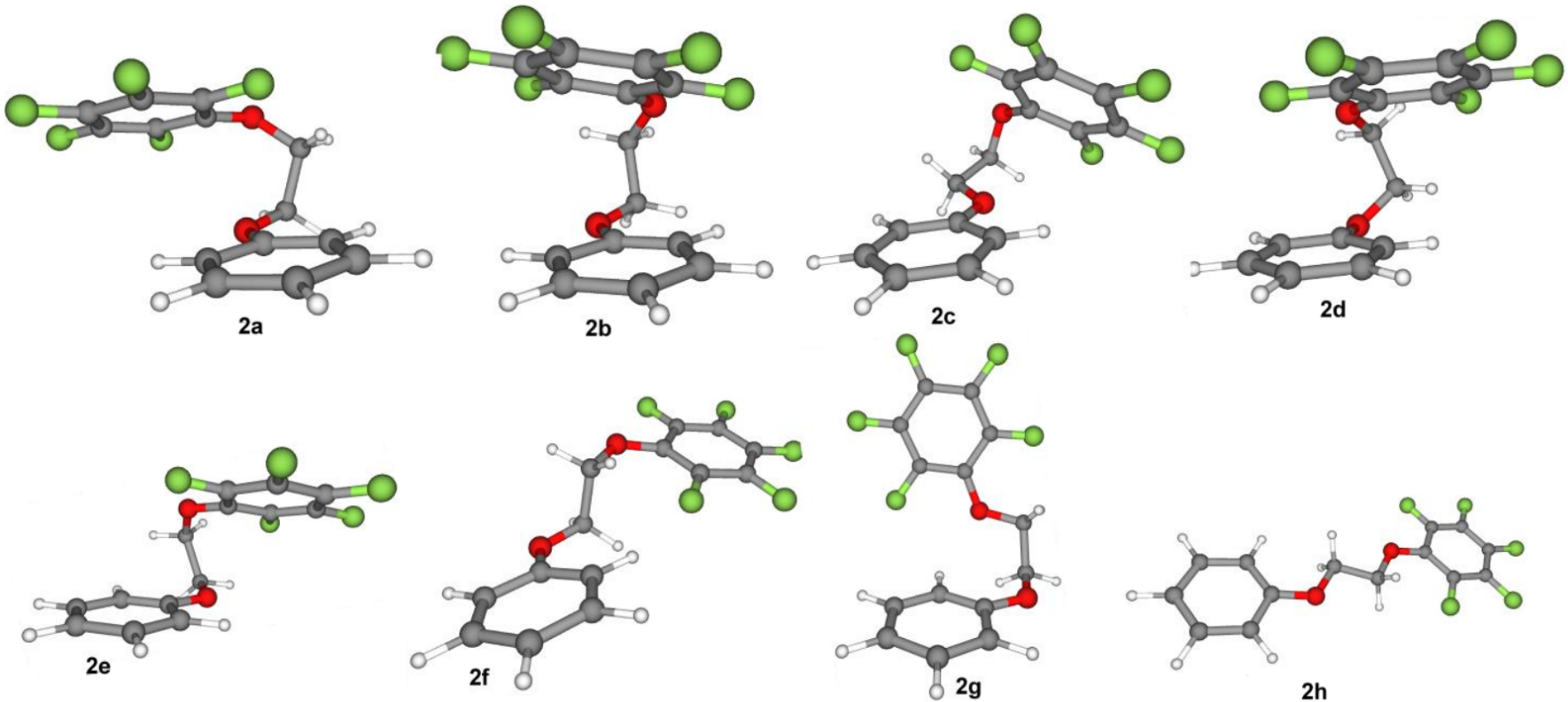
Too expensive for a 5D scan!





Need more advanced methods for conformational search!

Heuristic search



GFN2-xTB, CREST 2.10 (P. Pracht, F. Bohle, S. Grimme)

Check for possible duplicates!

Calculation of equilibrium structures

DFT in most cases.

Include D3(BJ), D4.

Recommended: PBEh-3c, B97-3c, TPSSh, PBE0, B3PW91, ...

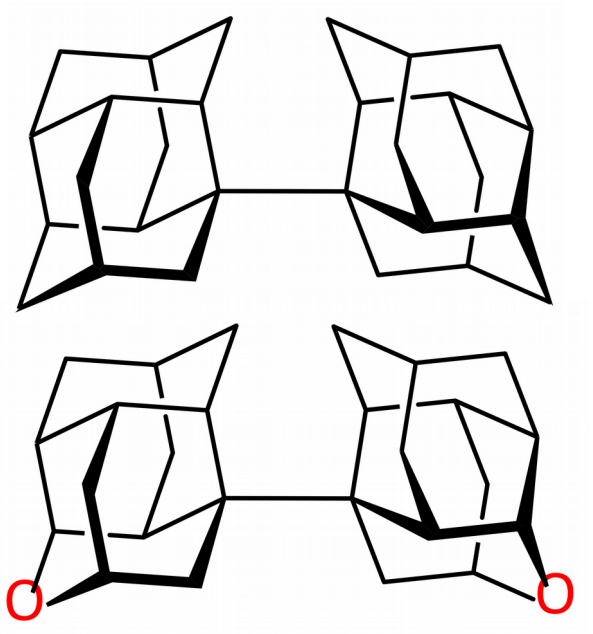
Turbomole, Orca, Gaussian, ...

MP2 – unclear accuracy together with high computational requirements (even with RI-approx.)

PNO-CCSD, PNO-CCSD(T) need benchmarking.

Benchmark: central C—C bond

With dispersion correction is better.



Method	Oxa-dimer	Dimer
XRD	1.643(1)	1.647(4)
GED	1.632(9)	**1.630(5)**
GED+MW	**1.632(5)**	-
B3LYP/cc-pVTZ	1.662	1.674
TPSS/cc-pVTZ	1.658	1.668
HF/cc-pVTZ	1.652	1.664
B97-D3/cc-pVTZ	1.651	1.662
B3PW91/cc-pVTZ	1.646	1.657
TPSS-D3/cc-pVTZ	1.642	1.652
B3LYP-D3/cc-pVTZ	1.642	1.653
ωB97XD/cc-pVTZ	1.638	1.648
PBE0/cc-pVTZ	1.637	1.648
M06-2X/cc-pVTZ	1.636	1.647
PBEh-3c	**1.632**	1.642
SCS(1.2;2/3)-MP2/def2-QZVP	1.629	1.640
PBE0-D3/cc-pVTZ	1.628	1.638
B3PW91-D3/cc-pVTZ	1.627	1.636
PW6B95-D3/def2-QZVP	1.626	1.636
ae-MP2/cc-pwCVTZ	1.622	**1.633**

Fokin, A. A.; Zhuk, T. S.; Blomeyer, S.; Pérez, C.; Chernish, L. V.; Pashenko, A. E.; Antony, J.; Vishnevskiy, Y. V.; Berger, R. J. F.; Grimme, S.; Logemann, C.; Schnell, M.; Mitzel, N. W.; Schreiner, P. R. *Intramolecular London Dispersion Interaction Effects on Gas-Phase and Solid-State Structures of Diamondoid Dimers.* J. Am. Chem. Soc. 2017, 139 (46), 16696–16707. DOI: 10.1021/jacs.7b07884.

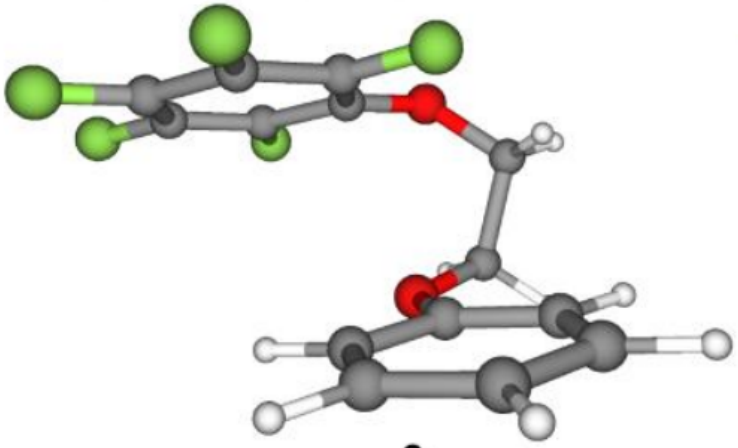
Frequencies

Mostly very complicated problem.

PBE0/6-31G(d,p):

Frequencies, cm^{-1}:

Harmonic	Anharmonic VPT2
12	-95
15	-30
34	-21
51	15
75	55



How to calculate amplitudes and corrections?
Small amplitude approximation? Shrink? VibModule? ElDiff?
Dynamic model?

Alternative: molecular dynamics simulations

In application to GED:

$$r_a = \left\langle \frac{1}{r} \right\rangle^{-1},$$

$$l^2 = \langle r^2 \rangle - \langle r \rangle^2,$$

$$\kappa = \frac{1}{6}\left(\langle r^3 \rangle - 3\langle r \rangle \langle r^2 \rangle + 2\langle r \rangle^3\right),$$

Wann, D. A.; Less, R. J.; Rataboul, F.; McCaffrey, P. D.; Reilly, A. M.; Robertson, H. E.; Lickiss, P. D.; Rankin, D. W. H. *Accurate Gas-Phase Experimental Structures of Octasilsesquioxanes (Si8O12X8; X = H, Me).* Organometallics 2008, 27 (16), 4183–4187.

Classical MD in GED

Problem: quantum effects are poorly recovered!

Table S13 Refined and calculated RMS amplitudes of vibration (u), associated r_a distances and corresponding perpendicular correction values (k) for $Si_8O_{12}Me_8$, **2**.[a]

	Atom pair	r_a	$u_{(GED)}$	k	$u_{MD(calc.)}$
u_1	C(21) –H(22)	110.6(7)	8.3(6)	0.5	7.5[b]
u_2	Si(1)–O(3)	162.5(1)	4.9(1)	0.8	4.2
u_3	H(22)···H(23)	178.3(17)	8.1(fixed)	0.1	8.1
u_4	Si(1)–C(29)	183.9(3)	5.9(4)	1.0	5.3
u_5	Si(1)···H(30)	245.1(11)	10.3(tied to u_6)	0.4	10.8
...					
u_{50}	H(22)···H(36)	995.4(24)	24.4(fixed)	–9.8	24.4

[a] Distances in pm. Values in parentheses are the standard deviations on the last digits. See Figure 1(b) for atom numbering. [b] Assumed value.

Wann, D. A.; Less, R. J.; Rataboul, F.; McCaffrey, P. D.; Reilly, A. M.; Robertson, H. E.; Lickiss, P. D.; Rankin, D. W. H. *Accurate Gas-Phase Experimental Structures of Octasilsesquioxanes (Si8O12X8; X = H, Me).* Organometallics 2008, 27 (16), 4183–4187.

Corrected MD: Qassandra

Vishnevskiy, Y. V.; Tikhonov, D. *Quantum Corrections to Parameters of Interatomic Distance Distributions in Molecular Dynamics Simulations.* Theoretical Chemistry Accounts 2016, 135, 88. DOI: 10.1007/s00214-016-1848-2

Benchmarking:

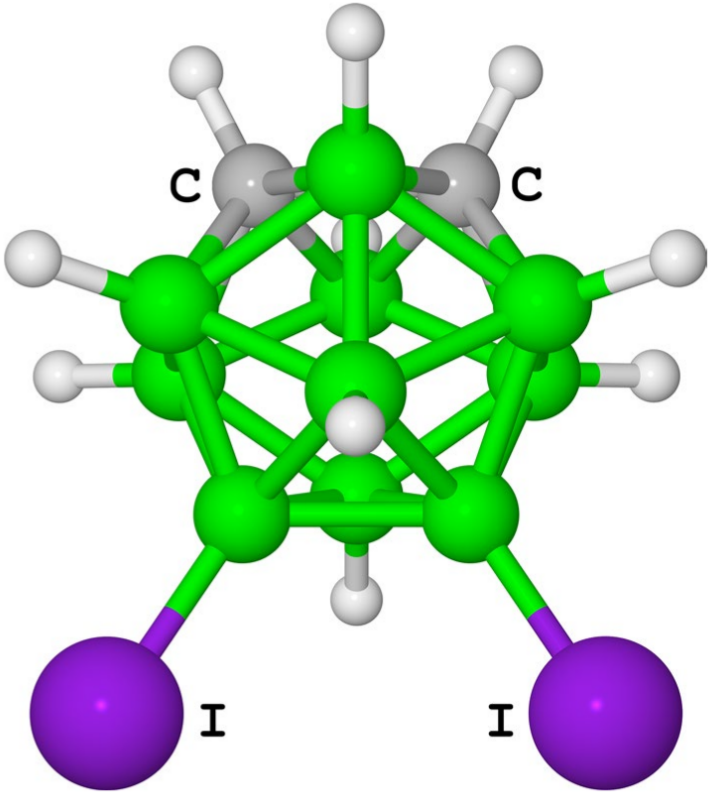
Tikhonov, D. S.; Sharapa, D. I.; Schwabedissen, J.; Rybkin, V. V. *Application of Classical Simulations for the Computation of Vibrational Properties of Free Molecules.* Phys. Chem. Chem. Phys. 2016, 18 (40), 28325–28338. DOI: 10.1039/C6CP05849C

Tikhonov, D. S.; Otlyotov, A. A.; Rybkin, V. V. *The Effect of Molecular Dynamics Sampling on the Calculated Observable Gas-Phase Structures.* Phys. Chem. Chem. Phys. 2016, 18 (27), 18237–18245. DOI: 10.1039/C6CP02973F

MD: ethane

Term	r_e (Å)	Time step (fs)	l (Å)				$r_e - r_a$ (Å)				$\kappa \cdot 10^6$ (Å^3)		
			MD	**MDC**	**QFF**	GED[a]	**MD**	**MDC**	**QFF**	GED[b]	**MD**	**MDC**	**QFF**
C–H	1.114	0.125	0.039	0.077	0.076	0.076 (1)	−0.004	−0.008	−0.017	−0.018 (1)	1.4	7.4	12.1
		0.250	0.039	0.078			−0.004	−0.007			1.2	6.1	
		0.500	0.040	0.078			−0.005	−0.007			1.4	6.4	
		1.000	0.040	0.078			−0.005	−0.008			1.6	7.3	
C–C	1.533	0.125	0.022	0.049	0.051	0.050 (1)	−0.003	−0.004	−0.011	−0.009 (1)	0.1	0.8	2.1
		0.250	0.022	0.049			−0.003	−0.005			0.1	0.9	
		0.500	0.022	0.050			−0.003	−0.004			0.1	0.7	
		1.000	0.020	0.049			−0.003	−0.004			0.1	0.7	
H···H	1.794	0.125	0.057	0.125	0.125		−0.007	−0.010	−0.017		3.1	21.9	3.3
		0.250	0.054	0.124			−0.007	−0.012			3.0	25.3	
		0.500	0.053	0.124			−0.007	−0.009			2.2	19.2	
		1.000	0.059	0.126			−0.007	−0.005			1.9	11.7	
C···H	2.202	0.125	0.040	0.091	0.108	0.107 (2)	−0.005	−0.012	−0.019	−0.019 (2)	1.4	11.6	1.1
		0.250	0.042	0.092			−0.005	−0.010			1.3	9.3	
		0.500	0.043	0.092			−0.006	−0.011			1.5	10.3	
		1.000	0.045	0.094			−0.006	−0.010			1.5	9.2	
H···H	2.570	0.125	0.077	0.154	0.183		−0.005	−0.004	−0.021		3.9	20.5	−28.9
		0.250	0.079	0.157			−0.005	−0.001			1.9	9.2	
		0.500	0.072	0.156			−0.006	−0.005			3.2	22.2	
		1.000	0.078	0.156			−0.007	−0.010			7.5	39.9	
H···H	3.135	0.125	0.056	0.115	0.126		−0.006	−0.013	−0.021		3.1	18.7	13.2
		0.250	0.054	0.114			−0.006	−0.010			1.9	11.8	
		0.500	0.057	0.113			−0.007	−0.010			2.6	12.0	
		1.000	0.061	0.117			−0.006	−0.010			3.1	13.7	

MD: 9,12-I_2-*closo*-1,2-$C_2B_{10}H_{10}$



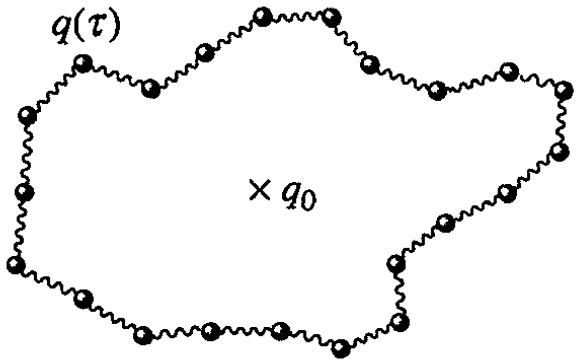
MD: 9,12-I_2-*closo*-1,2-$C_2B_{10}H_{10}$

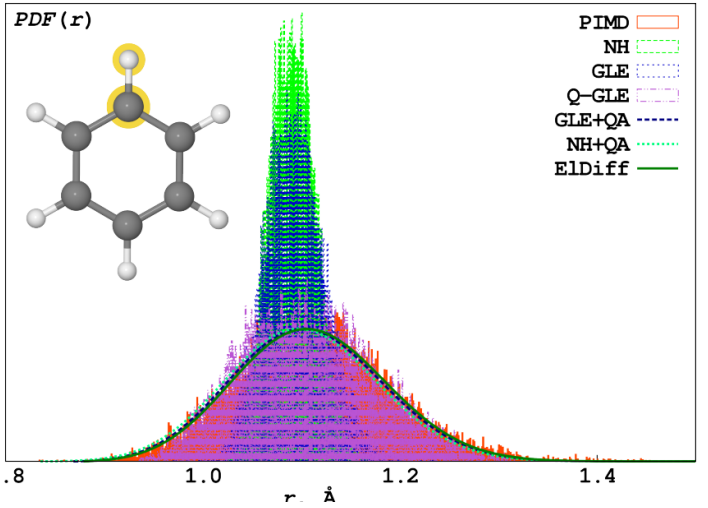
Term	r_e (Å)	l (Å)				$r_e - r_a$ (Å)				$\kappa \cdot 10^6$ (Å^3)			
		MD	**MDC(3)**	**MDC(5)**	**QFF**	**MD**	**MDC(3)**	**MDC(5)**	**QFF**	**MD**	**MDC(3)**	**MDC(5)**	**QFF**
C–H	1.091	0.046	0.074	0.063	0.073	−0.005	−0.011	−0.038	−0.016	3.8	8.9	36.0	10.1
B–H	1.202	0.036	0.082	0.062	0.080	−0.004	−0.002	−0.049	−0.016	0.3	2.0	42.5	12.2
C–C	1.612	0.037	0.060	0.060	0.066	−0.003	−0.008	−0.011	−0.013	1.8	6.8	9.1	10.8
B–B	1.790	0.043	0.074	0.073	0.069	−0.016	−0.011	−0.016	−0.013	0.4	−5.6	3.9	8.7
B–I	2.165	0.057	0.082	0.075	0.061	−0.021	−0.023	−0.038	−0.011	8.3	13.8	72.6	6.3
H…H	2.738	0.120	0.191	0.189	0.194	−0.009	−0.002	−0.010	−0.020	4.7	10.9	81.3	−22.1
B…I	3.490	0.151	0.167	0.162	0.108	−0.042	−0.038	−0.043	−0.019	−56.4	−127.7	111.9	12.9
H…I	3.697	0.271	0.293	0.288	0.217	−0.018	−0.015	−0.023	−0.013	87.6	78.9	1140.7	−34.6
I…I	4.188	0.338	0.181	0.208	0.215	−0.090	−0.134	−0.054	−0.029	−1268.7	−304.1	581.5	313.8
H…H	4.681	0.110	0.179	0.178	0.172	−0.013	−0.002	−0.006	−0.037	−15.6	−48.3	−19.1	−28.7
H…I	5.572	0.117	0.134	0.133	0.142	0.012	0.012	0.009	−0.017	9.3	7.4	36.6	−28.1

In general:

- For amplitudes MDC is more accurate than MD.
- For corrections there is some inconsistency.
- Asymmetry constants are very unstable both in MD and MDC.

Path-Integral MD





***

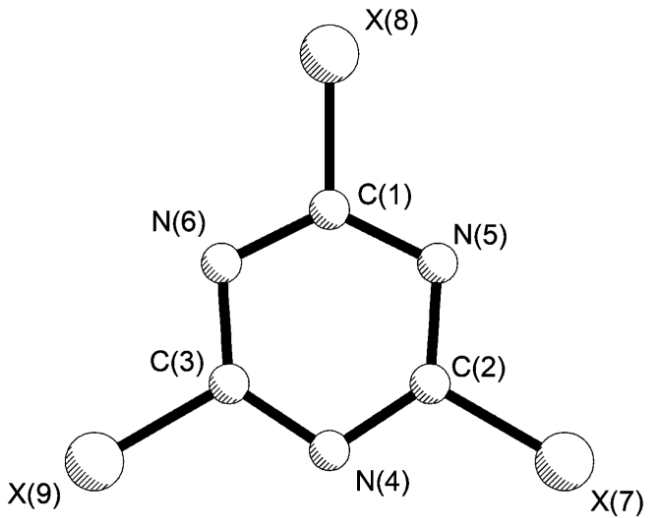
Cao, J.; Voth, G. A. *The Formulation of Quantum Statistical Mechanics Based on the Feynman Path Centroid Density. I. Equilibrium Properties.* J. Chem. Phys. 1994, 100 (7), 5093–5105. DOI: 10.1063/1.467175

*** Tikhonov, D. S.; Otlyotov, A. A.; Rybkin, V. V. *The Effect of Molecular Dynamics Sampling on the Calculated Observable Gas-Phase Structures.* Phys. Chem. Chem. Phys. 2016, 18 (27), 18237–18245. DOI: 10.1039/C6CP02973F.

Path-Integral MD

First application to GED:

Wann, D. A.; Zakharov, A. V.; Reilly, A. M.; McCaffrey, P. D.; Rankin, D. W. H. *Experimental Equilibrium Structures: Application of Molecular Dynamics Simulations to Vibrational Corrections for Gas Electron Diffraction.* J. Phys. Chem. A 2009, 113 (34), 9511–9520. DOI: 10.1021/jp904185g



PIMD in GED

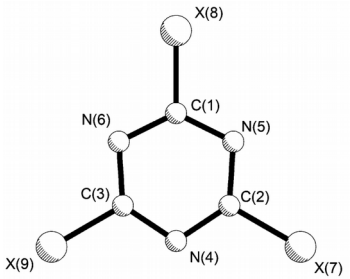


TABLE 3: PIMD-Calculated rms Amplitudes of Vibration (u), Distance Corrections ($r_a - r_e$), and Morse Constants (a), for $C_3N_3H_3$ for Various Numbers of Beads, P, Together with Experimental rms Amplitudes of Vibration[a]

	$P = 1$			$P = 12$			$P = 32$			exp[b]
atom pair	u	$r_a - r_e$	a	u	$r_a - r_e$	a	u	$r_a - r_e$	a	u
C(1)−N(5)	3.4(1)	0.4	22.3	4.6(1)	0.7	17.2	4.7(1)	0.8	13.1	5.1(2)
C(1)−H(8)	3.3(1)	0.5	23.2	7.7(1)	1.6	10.1	7.8(1)	1.8	10.7	5.3(11)
C(1)···C(2)	4.1(1)	0.3	4.5	5.2(1)	0.8	6.6	5.2(1)	0.9	5.5	6.1(5)
N(4)···N(5)	4.5(1)	0.4	9.8	5.5(1)	0.8	4.8	5.6(1)	0.9	7.6	6.2(4)
N(4)···H(7)	5.9(1)	0.5	1.9	9.7(1)	1.5	2.7	9.9(1)	1.7	3.9	12.4(12)
C(1)···N(4)	5.2(1)	0.2	2.1	6.0(1)	0.8	2.5	6.2(1)	0.9	4.8	6.5(4)
C(1)···H(7)	5.6(1)	0.1	2.8	9.3(1)	1.5	4.1	9.4(1)	1.7	4.9	13.1(16)
N(4)···H(8)	6.2(1)	−0.3	0.3	9.4(1)	1.3	3.6	9.6(1)	1.5	3.6	12.0(29)
H(7)···H(8)	7.6(1)	−0.2	0.0	12.7(1)	1.9	2.4	12.9(1)	2.2	2.1	12.6(fixed)[c]

[a] Distances and amplitudes are in pm, while Morse constants are in nm^{-1}. Values in parentheses are the standard deviations on the last digits. See Figure 1c for atom numbering. [b] Values from the combined GED/IR/Raman/MW/LCNMR refinement from ref 11. [c] Fixed at the value determined by an experimental force field.

PIMD in GED

PIMD works very well, but is expensive with ab initio.
However, we can use semi-empirical methods.

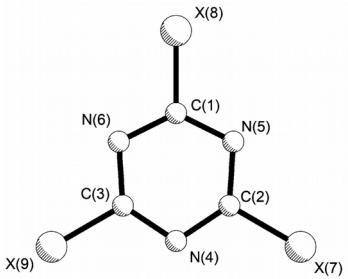


TABLE 4: Calculated and Experimental rms Amplitudes of Vibration (u), Distance Corrections ($r_a - r_x$), and Morse Constants for $C_3N_3Cl_3$[a]

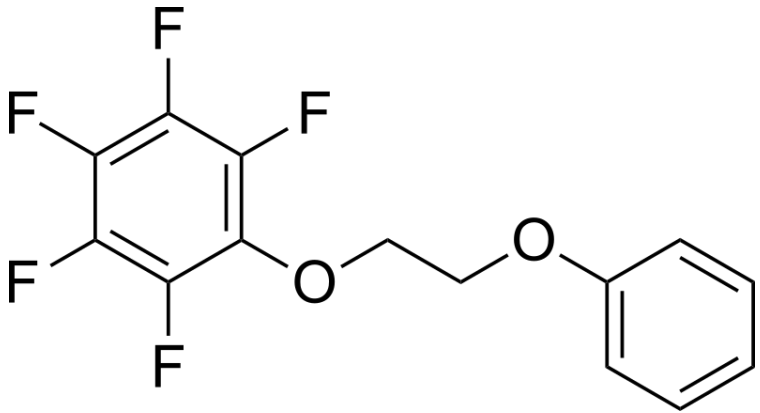
	rms amplitudes of vibration			distance corrections			Morse constants	
atom pair	u_{h1}	$u_{DFT\text{-}PIMD}$	u_{GED}	$r_a - r_{h1}$	$r_a - r_{a3,1}$	$r_a - r_{e,DFT\text{-}PIMD}$	SHRINK[b]	MD
C(1)−N(5)	4.5	4.6(1)	4.7(1)	−0.1	0.6	0.6	22.7	21.7
C(1)−Cl(8)	4.8	5.2(1)	4.7(1)	0.0	0.8	0.9	18.3	21.5
C(1)···C(2)	5.2	5.3(1)	5.4 [tied to N(4)N(5)]	−0.6	0.6	0.7	0.0	0.3
N(4)···N(5)	5.3	5.5(1)	5.7(3)	−0.3	0.8	0.8	0.0	10.3
N(4)···Cl(7)	6.2	6.6(1)	6.7(1)	−0.2	1.0	1.1	0.0	9.5
C(1)···N(4)	5.9	6.2(1)	6.3 [tied to N(4)Cl(7)]	−0.7	0.7	0.7	0.0	2.1
C(1)···Cl(7)	6.2	6.6(1)	7.2(1)	−1.3	0.8	0.8	0.0	2.8
N(4)···Cl(8)	6.3	6.8(1)	7.7(2)	−1.6	0.7	0.6	0.0	1.3
Cl(7)···Cl(8)	8.4	8.5(1)	10.0(1)	−2.0	0.8	0.8	0.0	0.0

[a] Distances and amplitudes are in pm, Morse parameters are in nm^{-1}. Values in parentheses are the standard deviations on the last digits. See Figure 1c for atom numbering. [b] Tabulated values given in the SHRINK output, which are taken from ref 44. For nonbonded atom pairs values were set to zero.

Parameters of MD trajectory

Lowest harmonic frequencies (GFN1-xTB, cm^{-1}):

2a:	2b:	2c:	2d:	….
16	14	11	11	
27	18	19	29	
37	28	31	36	
47	36	41	52	
73	48	79	67	
….	….	….	….	



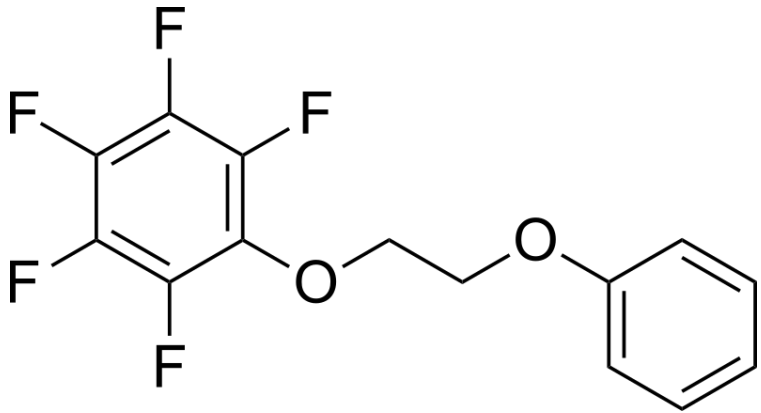
Period for $\nu = 10\ cm^{-1}$ is 3.35 ps.
We need at least several periods.

Δt = 0.5 fs is enough for highest frequency $\nu = 3000\ cm^{-1}$ (even 1.0 fs can be good!)

NVT type of thermostat, Nose-Hoover (chains), (there are also more promising possibilities!)

Use path-integral MD whenever possible! (CP2K)

Qassandra (PIMD) vs. VibModule



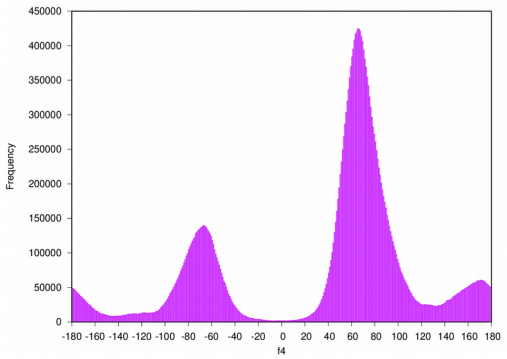
Selected terms in **2a** (Å):

Term	r_e	l		r_e-r_a	
		PIMD	VibMod	PIMD	VibMod
C—H	1.1	0.074	0.075	-0.015	-0.015
C—F	1.3	0.043	0.042	-0.006	-0.006
C...C	6.0	0.43	1.00	-0.27	-0.76
C...F	7.2	0.49	0.94	-0.015	-0.62

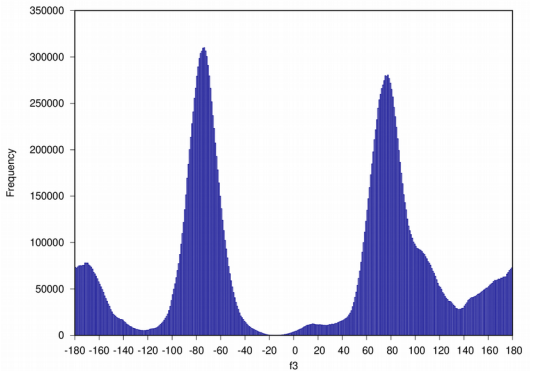
As expected:

- Agreement for bonded terms.
- Disagreement for others.

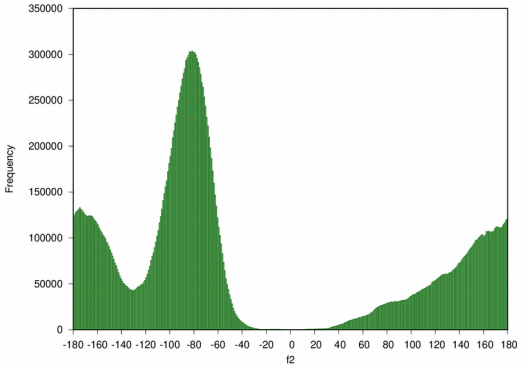
MD: distributions of torsion angles



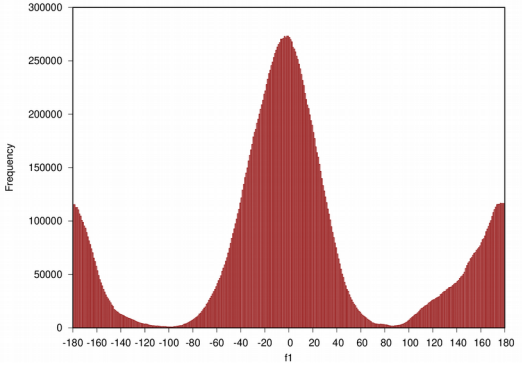
θ_2



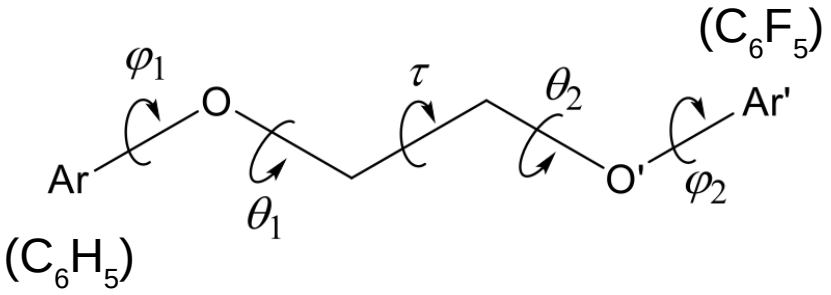
τ

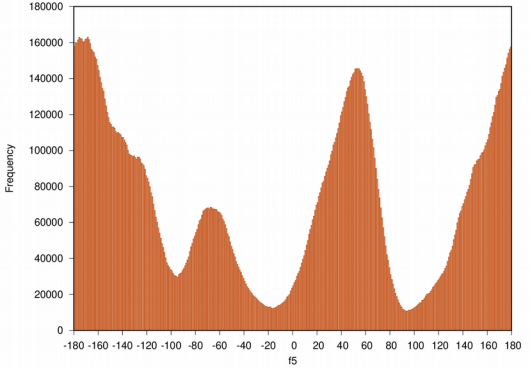


θ_1



φ_1





φ_2

16 beads, total t = 800 ps (Δt = 0.5 fs)

Refinements

Available methods

Least squares:
Jones, M. E.; Hedberg, K.; Schomaker, V. *J. Am. Chem. Soc.* 1955, 77 (20), 5278–5280.

MOCED:
Chiu, N. S.; Ewbank, J. D.; Askari, M.; Schäfer, L. *J. Mol. Struct.* 1979, 54, 185–195.

Predicate observations:
Bartell, L. S.; Romenesko, D. J.; Wong, T. C. *Augmented Analyses: Method of Predicate Observations.* In Molecular Structure by Diffraction Methods; Sim, G. A., Sutton, L. E., Eds.; The Chemical Society, Burlington House, London, 1975; Vol. 3, pp 72–79.

Regularization together with vibrationally-consistent solution:
Kochikov, I. V.; Tarasov, Yu. I.; Kuramshina, G. M.; Spiridonov, V. P.; Yagola, A. G.; Strand, T. G. *J. Mol. Struct.* 1998, 445 (1–3), 243–258.

SARACEN:
Mitzel, N. W.; Rankin, D. W. H. *Dalton Trans.* 2003, No. 19, 3650–3662.

In UNEX:
Vishnevskiy, Y. V.; Abaev, M. A.; Rykov, A. N.; Gurskii, M. E.; Belyakov, P. A.; Erdyakov, S. Yu.; Bubnov, Y. N.; Mitzel, N. W. *Chem. Eur. J.* 2012, 18 (34), 10585–10594.

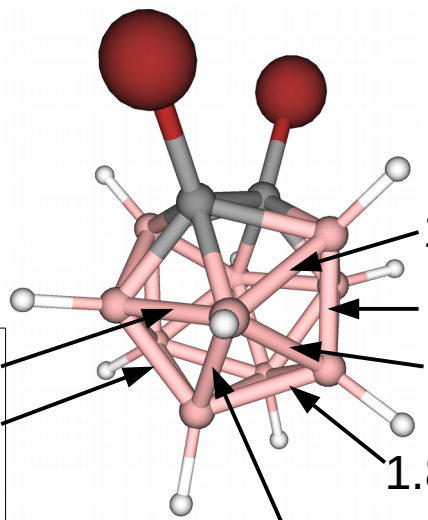
MOCED: Problems

How to define geometry?
Independent natural (internal) parameters (Z-matrices) ?

Problems with Z-matrices:
1. Difficult to construct for large mols.
2. Error prone.
3. Not unique.

Rigid constraints on independent internal parameters lead to biased results!

Example:
1,2-Br_2-1,2-dicarba-*closo*-dodecaborane

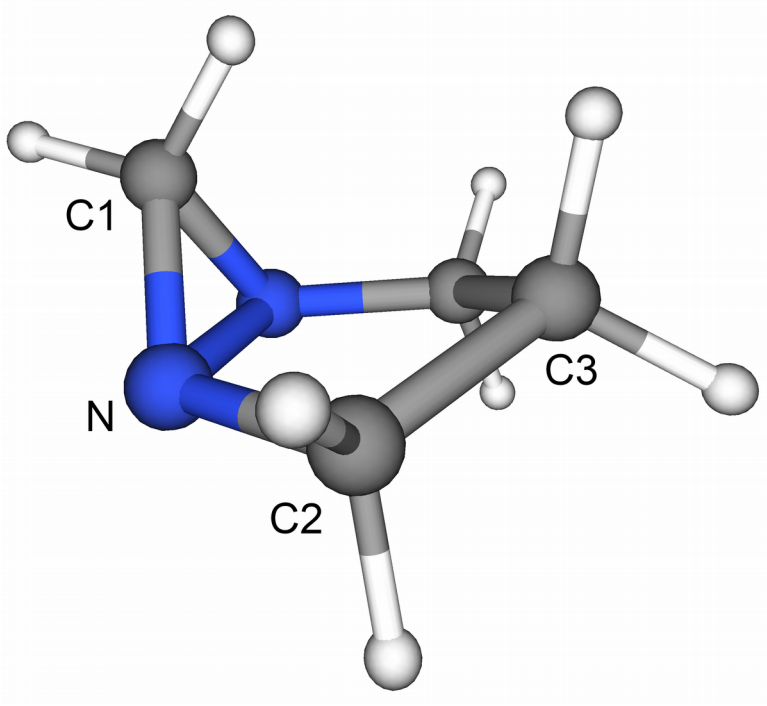


dependent 1.728(18)
dependent 1.744(16)
Δr(MP2) = -0.013 vs. Δr(GED) = 0.016

No control over dependent parameters!
No simple way to fix differences between dependent parameters!

Yu. Vishnevskiy, D. Tikhonov, E. Altova, et al., in preparation.

MOCED: Misleading standard deviations



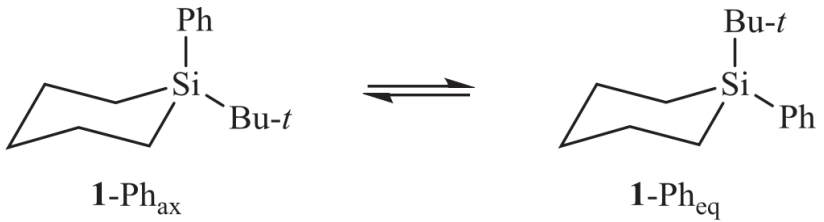
Combined in one group:

Parameter	r_g, Å	$3\sigma_{LSQ}$	Est. total error	
C1-N	1.453	0.001	0.002	
N-N	1.518	0.001	0.013	!!!
N-C2	1.479	0.001	0.004	
C2-C3	1.535	0.001	0.005	

Vishnevskiy, Y. V.; Vogt, N.; Vogt, J.; Rykov, A. N.; Kuznetsov, V. V.; Makhova, N. N.; Vilkov, L. V. *Molecular Structure of 1,5-Diazabicyclo[3.1.0]Hexane as Determined by Gas Electron Diffraction and Quantum-Chemical Calculations.* J. Phys. Chem. A 2008, 112 (23), 5243–5250. DOI: 10.1021/jp801346v

MOCED: Misleading standard deviations

Difficulties with interpretation of results for models of mixtures!



1-Ph_{ax} **1**-Ph_{eq}

	Conformer	Ph_{ax}	Ph_{eq_90}	
	Method	GED	GED	
Group 1	Si–C2	1.883(4)	1.889(4)	!!!
	Si–C7	1.888(4)	1.884(4)	
	Si–C13	1.900(4)	1.907(4)	
Group 2	C2–C3	1.544(3)	1.543(3)	
	C7–C8	1.403(3)	1.403(3)	
	C13–C14	1.537(3)	1.537(3)	
Group 3	C7SiC13	108.6(17)	106.9(17)	
Group 4	SiC13C14	110.7(5)	111.4(05)	
		92(7) %	8(7) %	

Shainyan, B. A.; Suslova, E. N.; Phien, T. D.; Shlykov, S. A.; Oznobikhina, L. P. *1-t-Butyl-1-Phenyl-1-Silacyclohexane: Synthesis, Conformational Analysis in Gas and Solution by GED, FT-IR and Theoretical Calculations.* J. Organomet. Chem. 2020, 923, 121433. DOI: 10.1016/j.jorganchem.2020.121433.

Solution of some problems

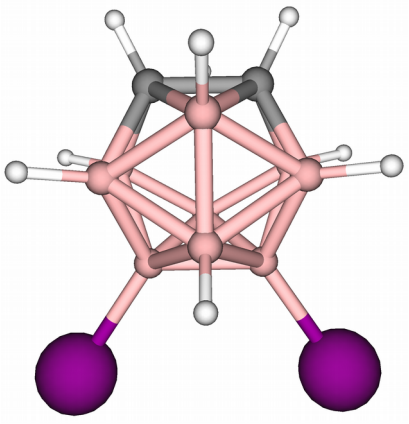
- Definition and refinement of the structure in Cartesian coordinates.
- Regularization of Cartesian coordinates.

$$Q = \sum_i w_i \left[sM(s)_{\exp} - sM(s)_{model} \right]^2 + \alpha \sum_j^{3N} w_j \left(x_{j,0} - x_{j,model} \right)^2 \rightarrow min$$

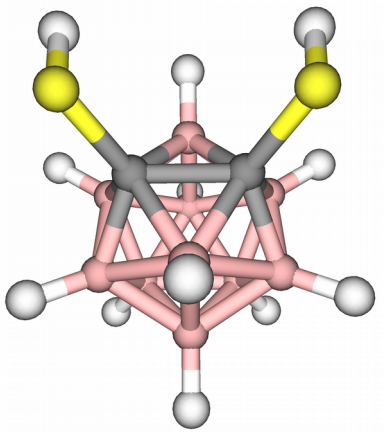
Advantages over Z-matrix method with rigid constraints:

- Simplicity (important for large molecules).
- Equal status of all internal natural parameters (with respective consequences).
- Better control of the model flexibility.

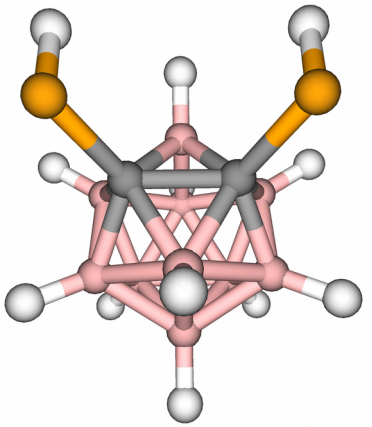
Examples: *closo*–1,2–$C_2B_{10}H_{10}$



9,12–I_2–



1,2–$(SH)_2$–



1,2–$(SeH)_2$–

Baše, T.; Holub, J.; Fanfrlík, J.; Hnyk, D.; Lane, P. D.; Wann, D. A.; Vishnevskiy, Y. V.; Tikhonov, D.; Reuter, C. G.; Mitzel, N. W. *Icosahedral Carbaboranes with Peripheral Hydrogen–Chalcogenide Groups: Structures from Gas Electron Diffraction and Chemical Shielding in Solution.* Chem. Eur. J. 2019, 25 (9), 2313–2321. DOI: 10.1002/chem.201805145

Vishnevskiy, Y. V.; Tikhonov, D. S.; Reuter, C. G.; Mitzel, N. W.; Hnyk, D.; Holub, J.; Wann, D. A.; Lane, P. D.; Berger, R. J. F.; Hayes, S. A. *Influence of Antipodally Coupled Iodine and Carbon Atoms on the Cage Structure of 9,12-I_2-closo-1,2-$C_2B_{10}H_{10}$: An Electron Diffraction and Computational Study.* Inorg. Chem. 2015, 54 (24), 11868–11874. DOI: 10.1021/acs.inorgchem.5b02102

In case of conformational flexibility

What if
- model parameters change significantly?
- we want to restrain particular natural geometrical parameters?

$$Q = \sum_{i}^{N} w_i \left[sM(s)_{\exp} - sM(s)_{model} \right]^2 + \alpha \sum_{j}^{M} w_j \left(p_{j,0} - p_{j,model} \right)^2 \rightarrow min$$

Combination of advantages:
- Cartesian coordinates as independent refined parameters. (Simplicity!)
- Set of internal geometrical parameters for regularization. (Control!)

J.-H. Weddeling, Yu. V. Vishnevskiy, B. Neumann, H.-G. Stammler, N. W. Mitzel,
Inter- and Intramolecular Aryl-Aryl-Interactions in Partially Fluorinated Ethylenedioxy-Bridged Bisarenes,
ChemRxiv, Preprint, 2020, DOI: 10.26434/chemrxiv.12643154

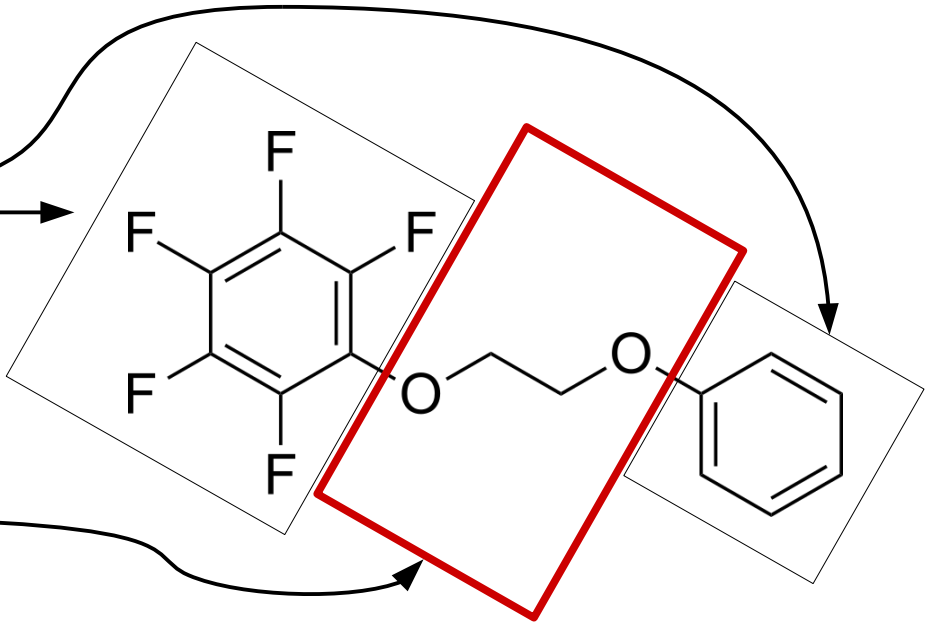
An example

$$Q = \sum_{i}^{N} w_i [\, sM(s)_{\exp} - sM(s)_{model} \,]^2 + \alpha \sum_{j}^{M} w_j (p_{j,0} - p_{j,model})^2 \rightarrow min$$

Full set p_0 generated automatically.

Accurately predicted angles, w_j can be increased.

Poorly predicted angles, w_j can be decreased.



J.-H. Weddeling, Yu. V. Vishnevskiy, B. Neumann, H.-G. Stammler, N. W. Mitzel,
Inter- and Intramolecular Aryl-Aryl-Interactions in Partially Fluorinated Ethylenedioxy-Bridged Bisarenes,
ChemRxiv, Preprint, 2020, DOI: 10.26434/chemrxiv.12643154

Amplitudes

- Approximation of point-like scattering volume leads to biased and not consistent amplitudes!

<u>Effect of finite sample size:</u>

Kuchitsu, K. *Electron Diffraction Investigation on the Molecular Structure of N-Butane.* Bull. Chem. Soc. Jpn. 1959, 32 (7), 748–769. DOI: 10.1246/bcsj.32.748.

Karle, I. L.; Karle, J. *Internal Motion and Molecular Structure Studies by Electron Diffraction. III. Structure of CH2CF2 and CF2CF2.* J. Chem. Phys. 1950, 18 (7), 963--971. DOI: 10.1063/1.1747820.

Harvey, R. B.; Keidel, F. A.; Bauer, S. H. *Some Effects of Nozzle Design on the Diffraction of Electrons by Gases.* J. Appl. Phys. 1950, 21 (9), 860–874. DOI: 10.1063/1.1699775.

<u>Benchmark:</u>

Vishnevskiy, Y. V.; Blomeyer, S.; Reuter, C. G. *Gas Standards in Gas Electron Diffraction: Accurate Molecular Structures of CO2 and CCl4.* Struct. Chem. 2020, 31 (2), 667–677. DOI: 10.1007/s11224-019-01443-5.

Amplitudes

- No simple way to create a dynamic model (see above).
- In semi-rigid GED models for non-rigid molecules many amplitudes are essentially effective!
- They can be calculated using (PI)MD.
- Still, their accuracy can be questionable (we need benchmarks)…
- … and they can influence conformations.
- Thus, we want to refine them!

How to refine?

- In groups, for minimizing correlations.
- Fix ratios, not differences!
- We may regularize amplitudes or their scale factors.

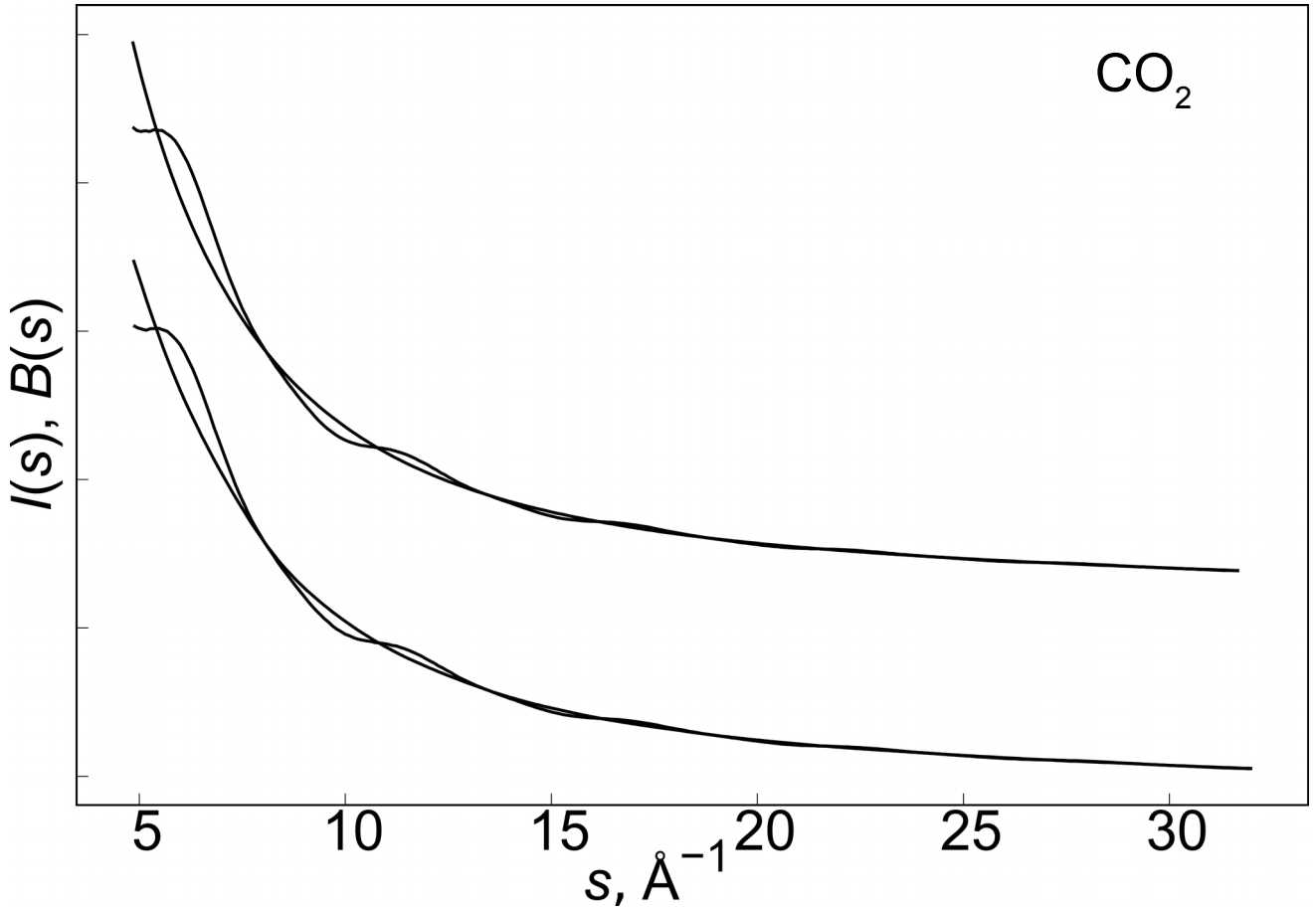
Background

Background

Multiplicative background *B*:

$$sM(s) = \frac{I(s) - B(s)}{B(s)} s$$

Background smoothness?
2nd derivatives?
Inflection points?
Maximal degree of polynomial?



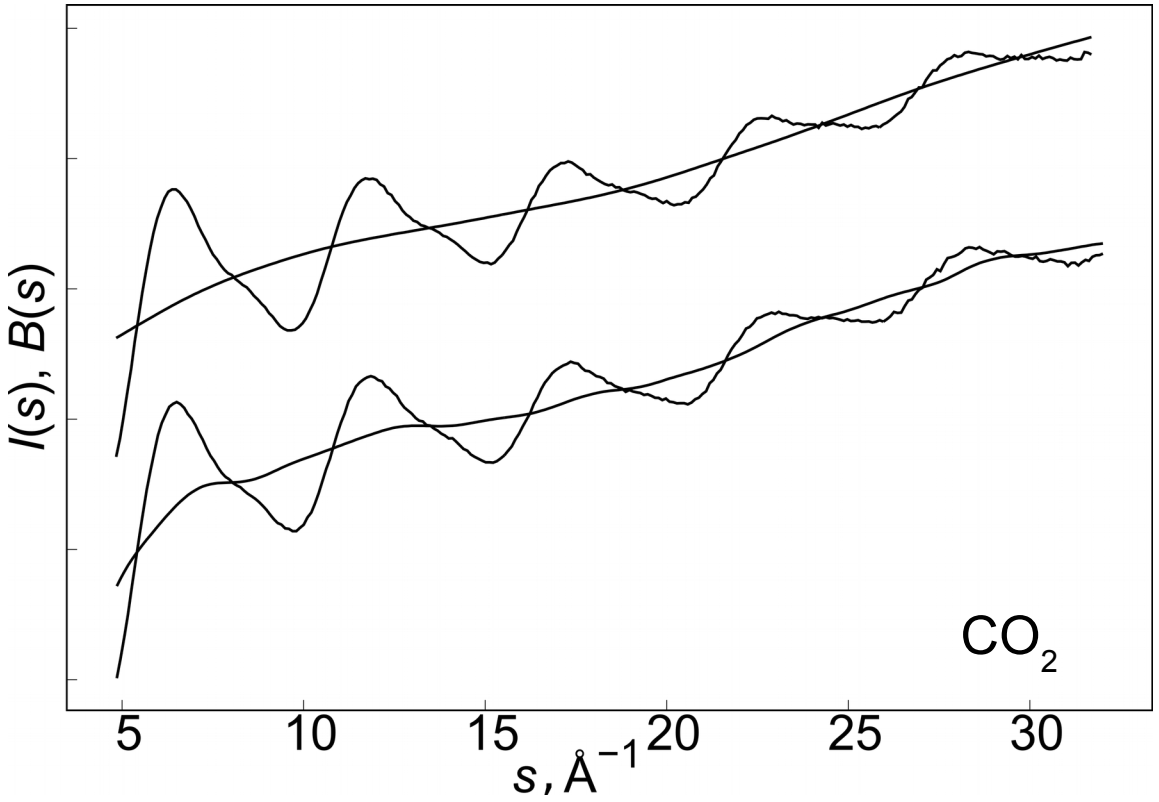
Reduced intensity and background

$$sM(s) = \frac{I(s) - B(s)}{B(s)} s$$

Reduced $I(s)$, $B(s)$:

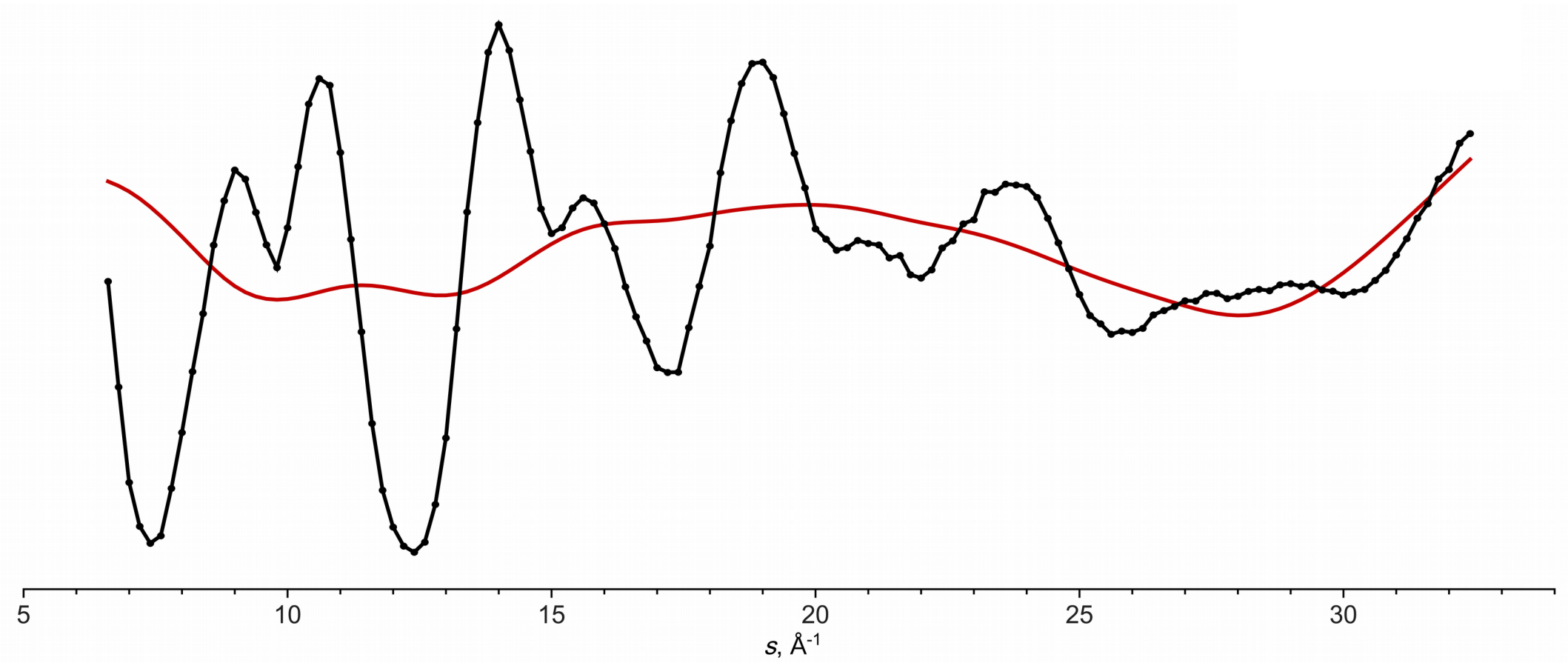
$$\frac{I(s)}{I_{at}(s)K} \qquad \frac{B(s)}{I_{at}(s)K}$$

Visual control!

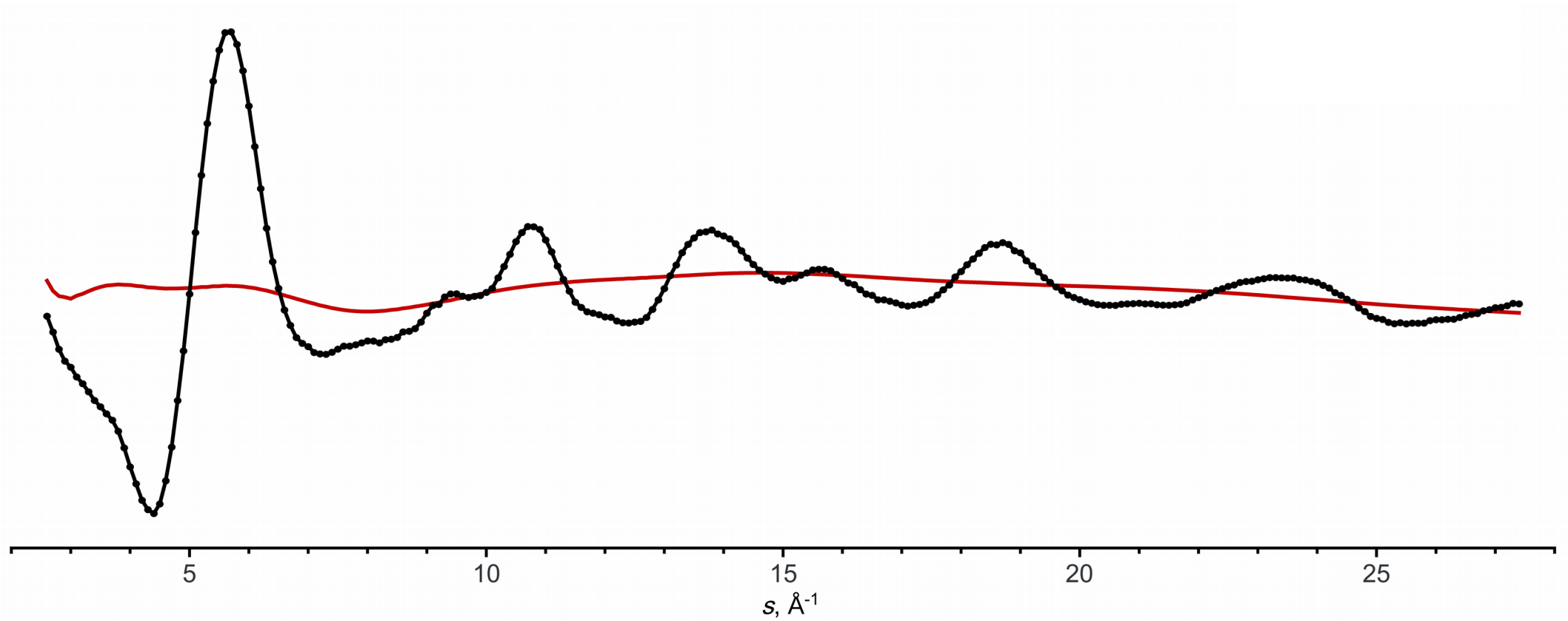


Vishnevskiy, Y. V.; Blomeyer, S.; Reuter, C. G. *Gas Standards in Gas Electron Diffraction: Accurate Molecular Structures of CO_2 and CCl_4.* Struct. Chem. 2020, 31 (2), 667–677. DOI: 10.1007/s11224-019-01443-5.

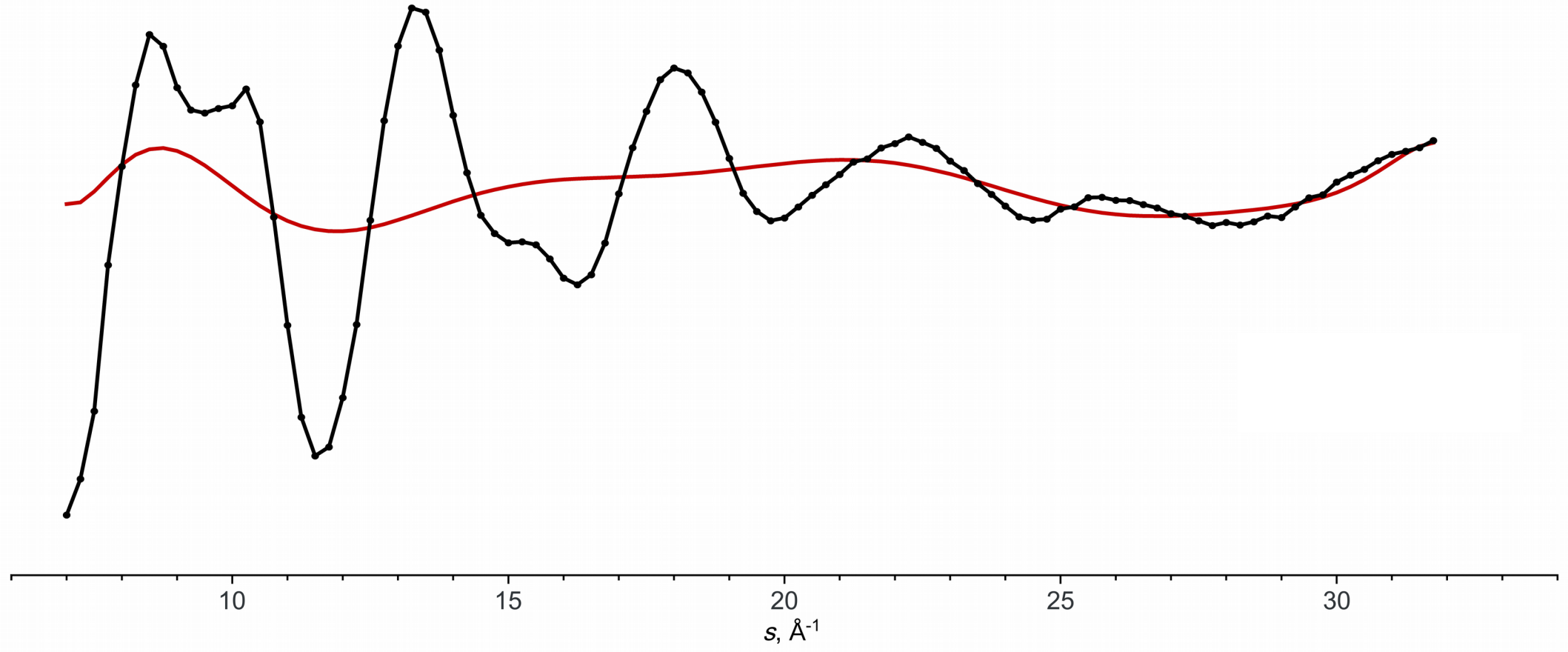
Example # 1 from literature



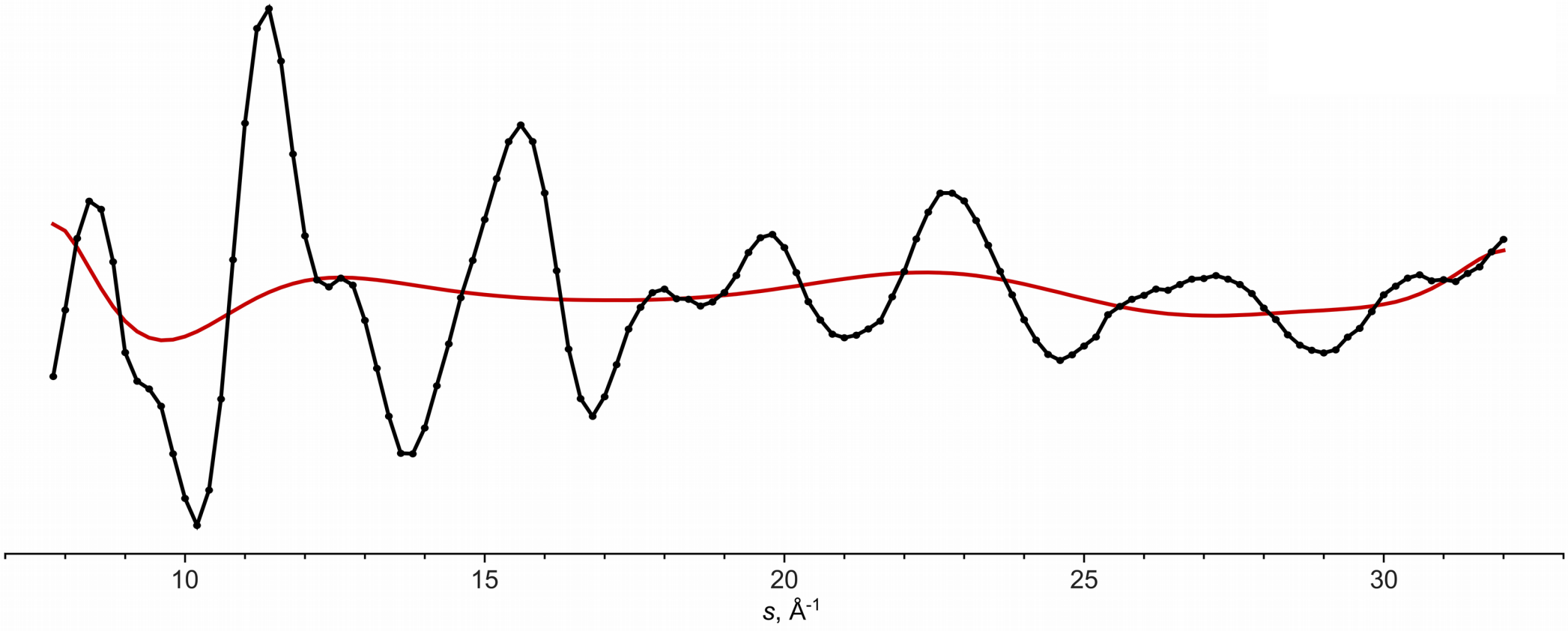
Example # 2



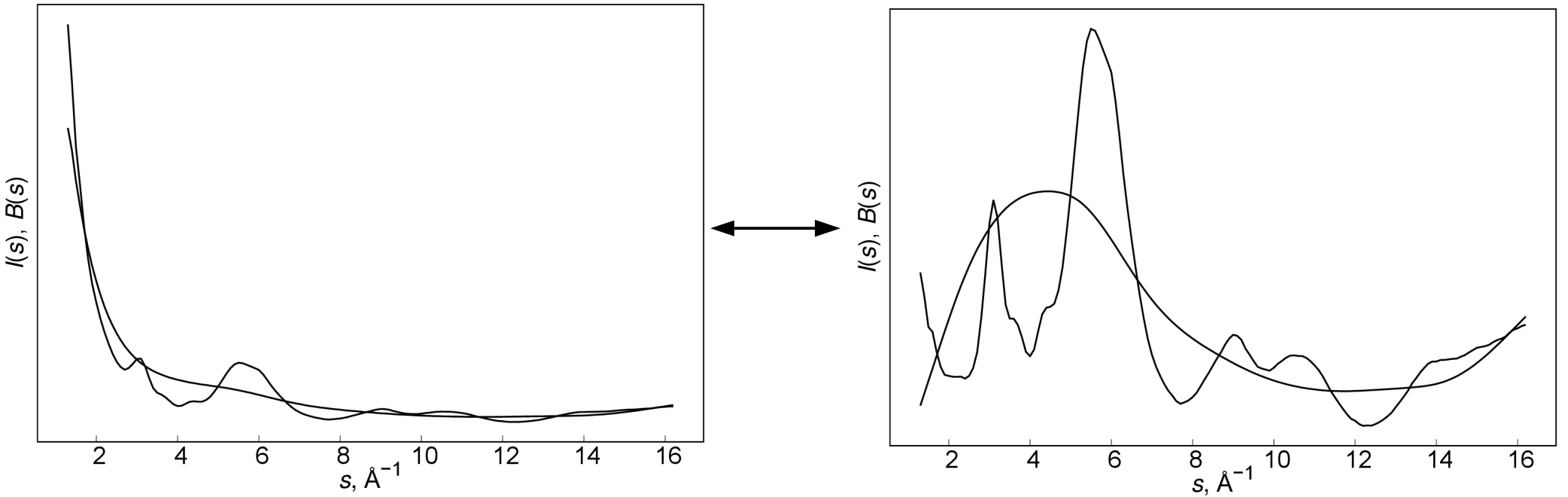
Example # 3



Example # 4

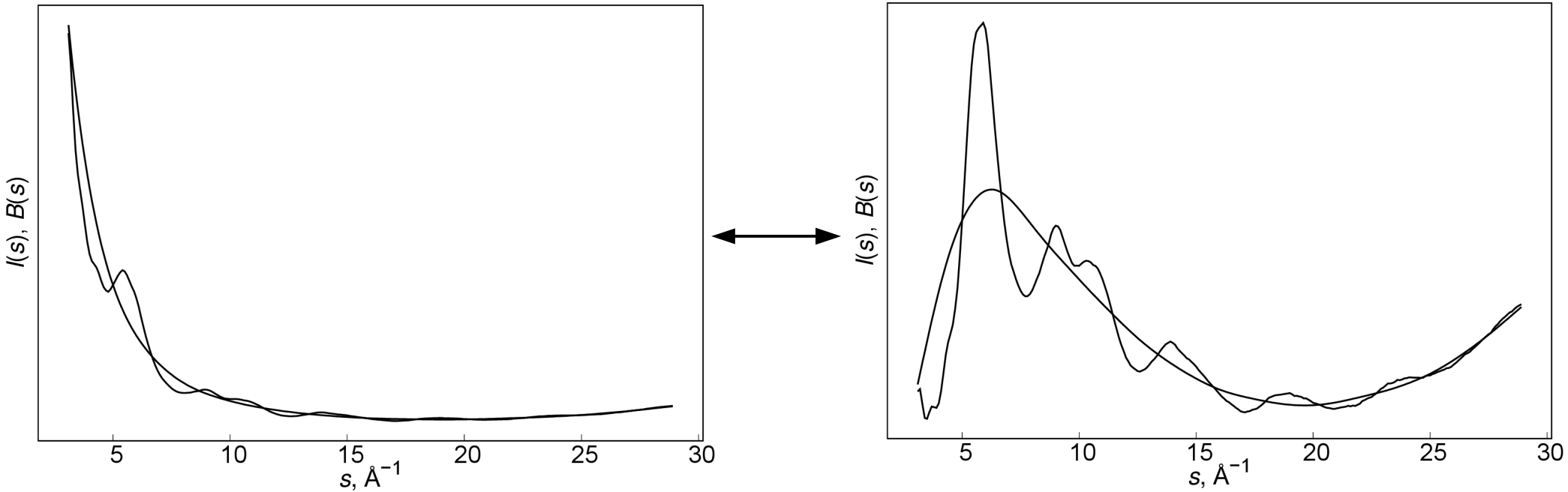


Example # 5



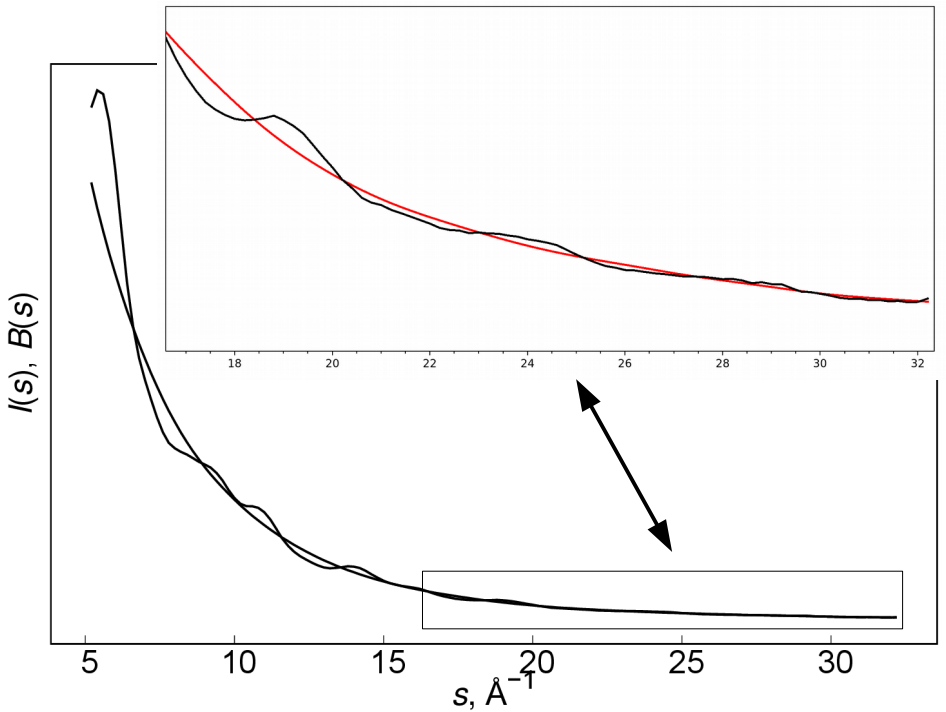
Kurochkin, I. Y.; Pogonin, A. E.; Otlyotov, A. A.; Kiselev, A. N.; Krasnov, A. V.; Shlykov, S. A.; Girichev, G. V. *Molecular Structure of 5,10,15,20-Tetrakis(4'-Fluorophenyl)Porphyrin by Combined Gas-Phase Electron Diffraction/Mass Spectrometry Experiment and DFT Calculations.* J. Mol. Struct. 2020, 1221, 128662. DOI: 10.1016/j.molstruc.2020.128662

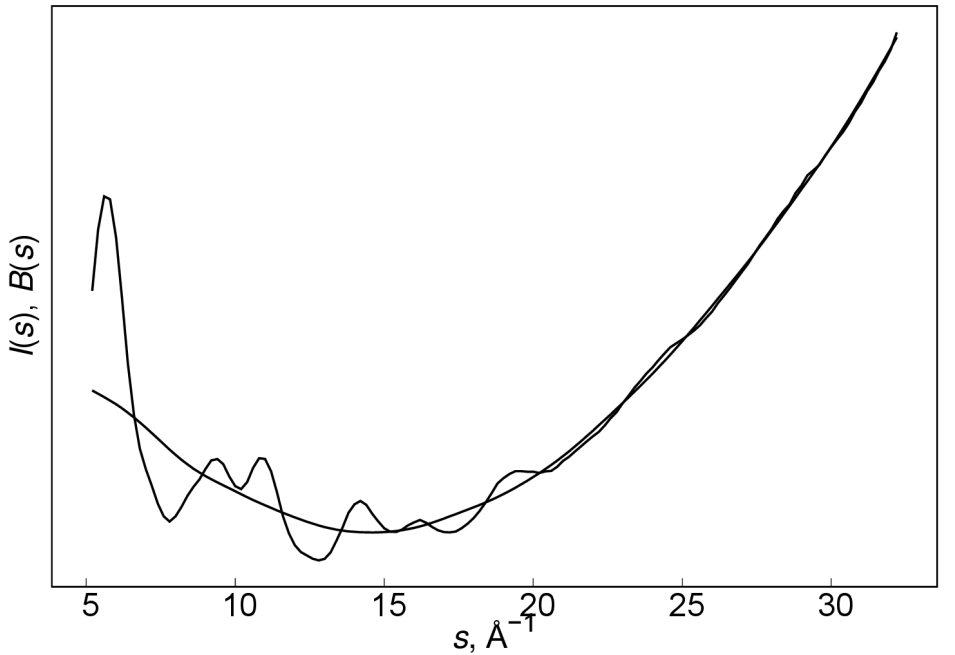
Example # 6



Kurochkin, I. Y.; Pogonin, A. E.; Otlyotov, A. A.; Kiselev, A. N.; Krasnov, A. V.; Shlykov, S. A.; Girichev, G. V. *Molecular Structure of 5,10,15,20-Tetrakis(4'-Fluorophenyl)Porphyrin by Combined Gas-Phase Electron Diffraction/Mass Spectrometry Experiment and DFT Calculations.* J. Mol. Struct. 2020, 1221, 128662. DOI: 10.1016/j.molstruc.2020.128662

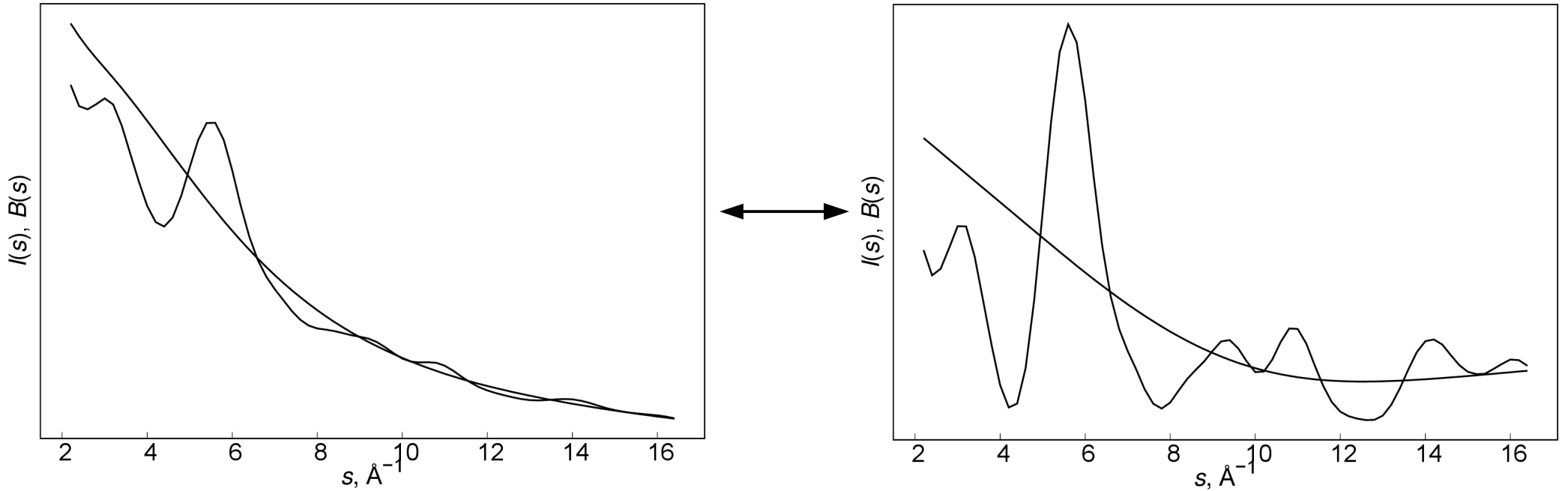
Example # 7





J.-H. Weddeling, Yu. V. Vishnevskiy, B. Neumann, H.-G. Stammler, N. W. Mitzel,
Inter- and Intramolecular Aryl-Aryl-Interactions in Partially Fluorinated Ethylenedioxy-Bridged Bisarenes,
ChemRxiv, Preprint, 2020, DOI: 10.26434/chemrxiv.12643154

Example # 8



J.-H. Weddeling, Yu. V. Vishnevskiy, B. Neumann, H.-G. Stammler, N. W. Mitzel,
Inter- and Intramolecular Aryl-Aryl-Interactions in Partially Fluorinated Ethylenedioxy-Bridged Bisarenes,
ChemRxiv, Preprint, 2020, DOI: 10.26434/chemrxiv.12643154

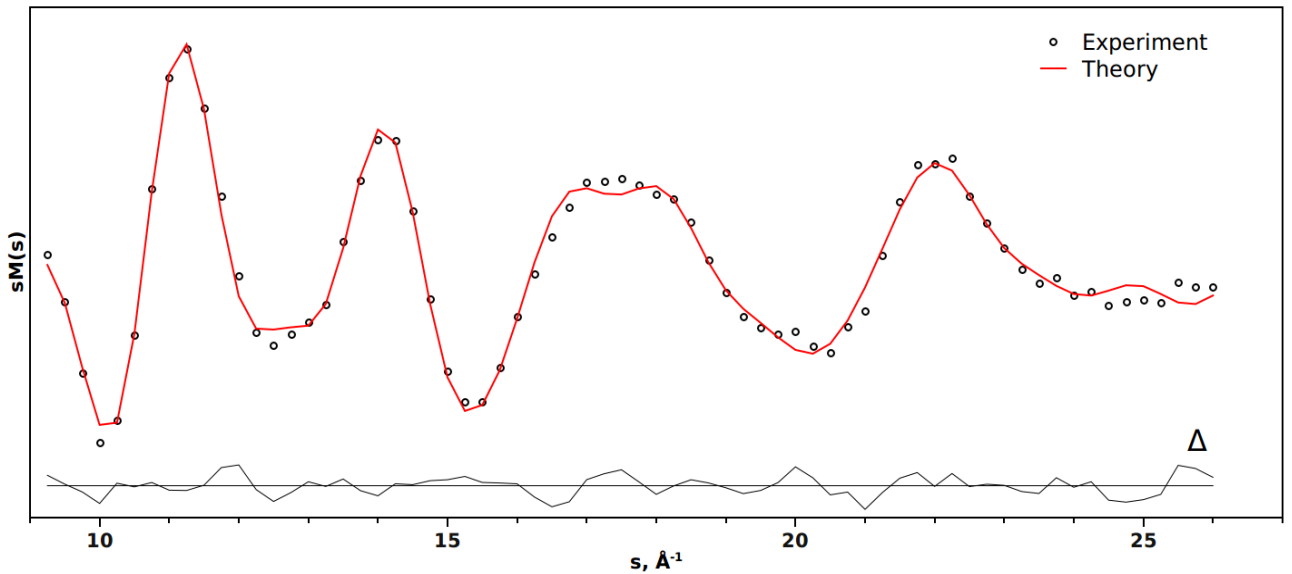
Quality of fit

Structural *R*-factor

$$wR_{str} = \sqrt{\frac{\sum_i w_i \left(sM(s_i)_{\exp} - sM(s_i)_{model}\right)^2}{\sum_i w_i \left(sM(s_i)_{\exp}\right)^2}} \times 100\,\%$$

wR_{str} = 5.0%

How much is this?

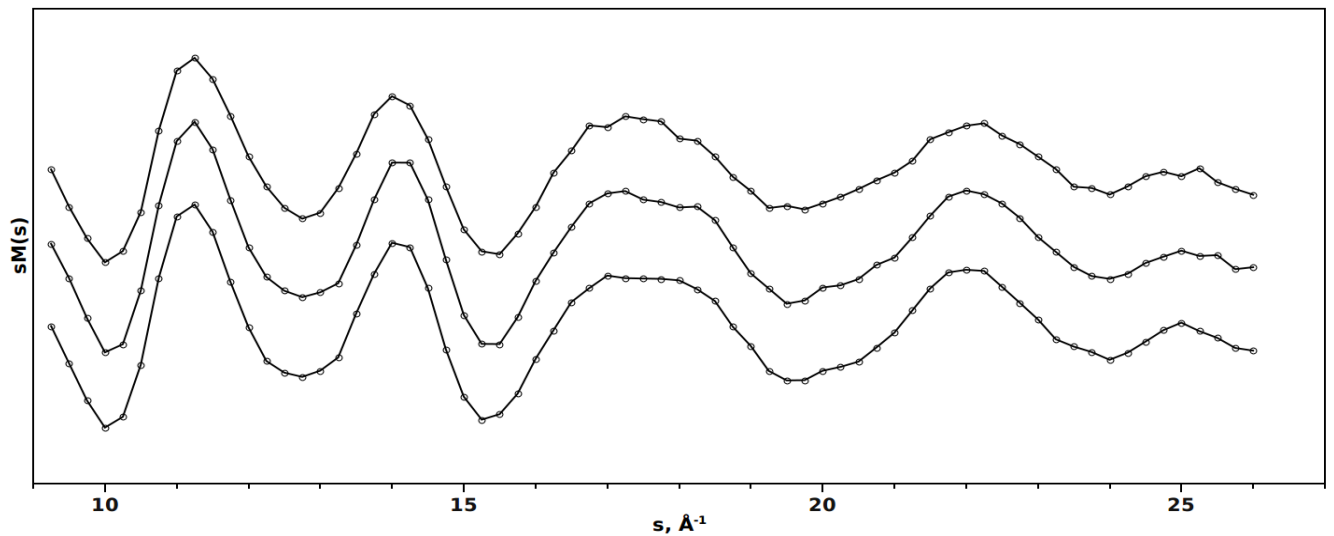


Experimental *R*-factor

$$wR_{\text{exp}} = \sqrt{\frac{\sum_{i}^{N}\sum_{j}^{M_i} w_{i,j}\left(sM(s_j)_{i,\text{exp}} - sM(s_{i,j})_{\text{av,exp}}\right)^2}{\sum_{i}\sum_{j} w_{i,j}\left(sM(s_{i,j})_{\text{av,exp}}\right)^2}} \times 100\,\%$$

wR_{exp} = 5.0%

Compare with wR_{str}



Vishnevskii, Yu. V.; Shishkov, I. F.; Khristenko, L. V.; Rykov, A. N.; Vilkov, L. V.; Oberhammer, H. *Molecular Structures of o- and m-Fluoro(Trifluoromethoxy)Benzenes According to Gas Electron Diffraction and Quantum-Chemical Studies: Comparison of the Structures of Trifluoromethoxybenzene and Its Fluorinated Derivatives.* Russ. J. Phys. Chem. 2005, 79 (10), 1537–1547.

wR_{exp} *vs.* wR_{str}

Possible cases:

- $wR_{str} >> wR_{exp}$: underfitting, data are reproducible but model cannot describe them; the model should be improved.
- $wR_{str} << wR_{exp}$: overfitting, model describes data too well, probably not reproducible data features are fitted; better data are needed.
- $wR_{str} \approx wR_{exp}$: optimal solution if both values are small.

Note: we need high-quality backgrounds!

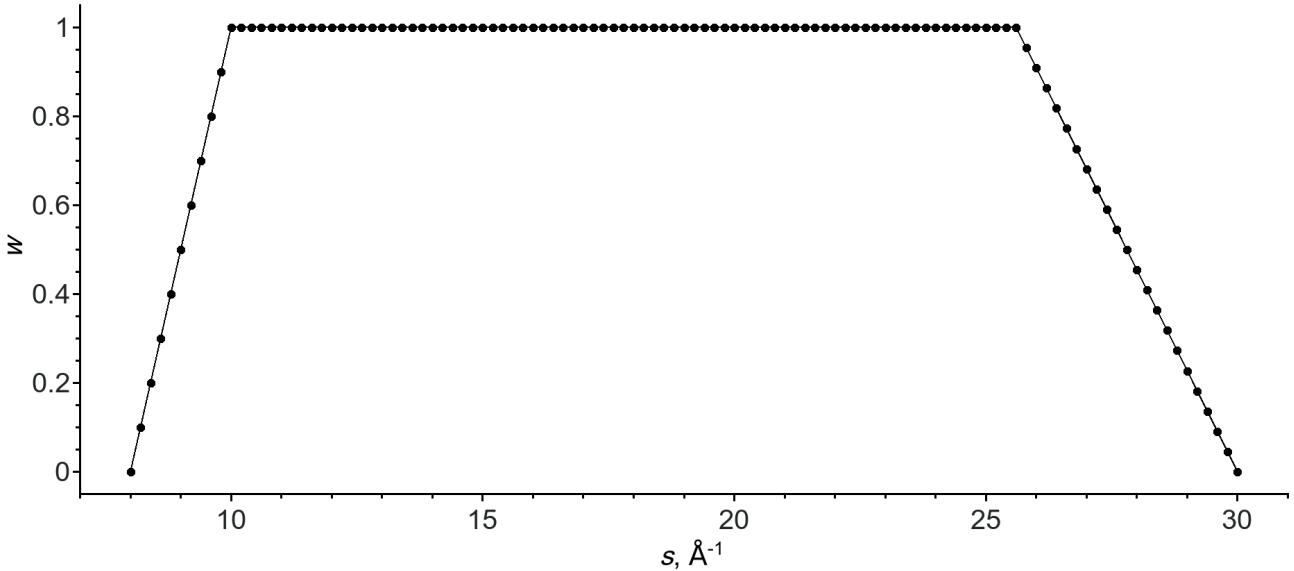
Best wR_{exp} @ Uni-Bielefeld: 1.6 %

For typical values see:
Otlyotov, A. A.; Girichev, G. V.; Rykov, A. N.; Glodde, T.; Vishnevskiy, Y. V. *Molecular Structure of Pyrazinamide: A Critical Assessment of Modern Gas Electron Diffraction Data from Three Laboratories.* J. Phys. Chem. A 2020, 124 (25), 5204–5211. DOI: 10.1021/acs.jpca.0c03531

Data weighting

Typically:
- No weighting
- Empirical weighting

$$
\begin{aligned}
w_{ii} &= (s_i - s_{\min})/(sw_1 - s_{\min}) && s_{\min} \le s_i \le sw_1 \\
w_{ii} &= 1 && sw_1 \le s_i \le sw_2 \\
w_{ii} &= (s_{\max} - s_i)/(s_{\max} - sw_2) && sw_2 \le s_i \le sw_{\max} \\
w_{ij} &= 0 && i \ne j \pm 1 \\
w_{ij} &= -0.5(w_{ii} + w_{ij})q && i = j \pm 1
\end{aligned}
$$



Often do not correspond to real data precision!

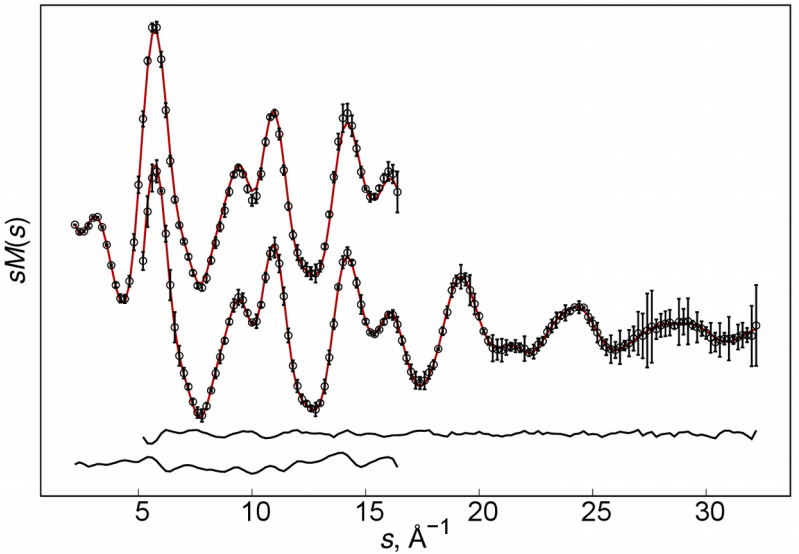
Hinchley, S. L.; Robertson, H. E.; Borisenko, K. B.; Turner, A. R.; Johnston, B. F.; Rankin, D. W. H.; Ahmadian, M.; Jones, J. N.; Cowley, A. H. *The Molecular Structure of Tetra-Tert-Butyldiphosphine: An Extremely Distorted, Sterically Crowded Molecule.* Dalton Trans. 2004, No. 16, 2469–2476. DOI: 10.1039/B407908F.

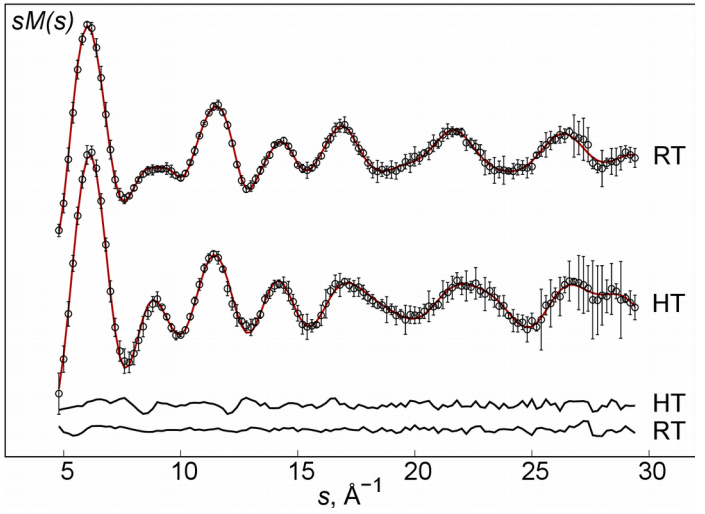
Data weighting: using standard deviations

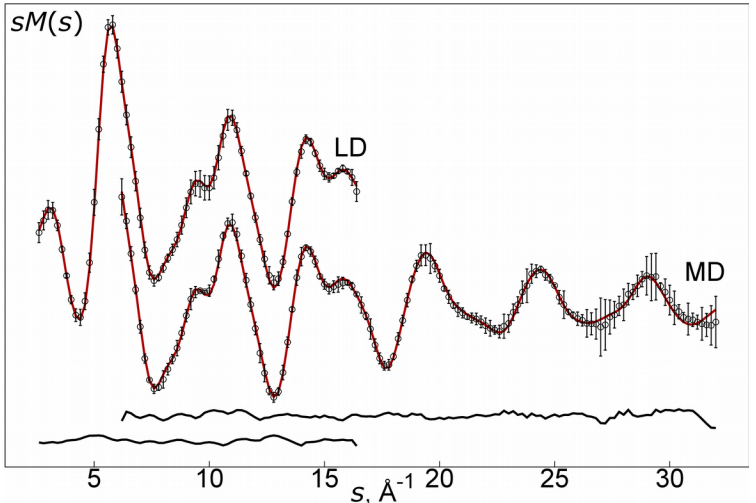
Standard deviations for $sM(s)$

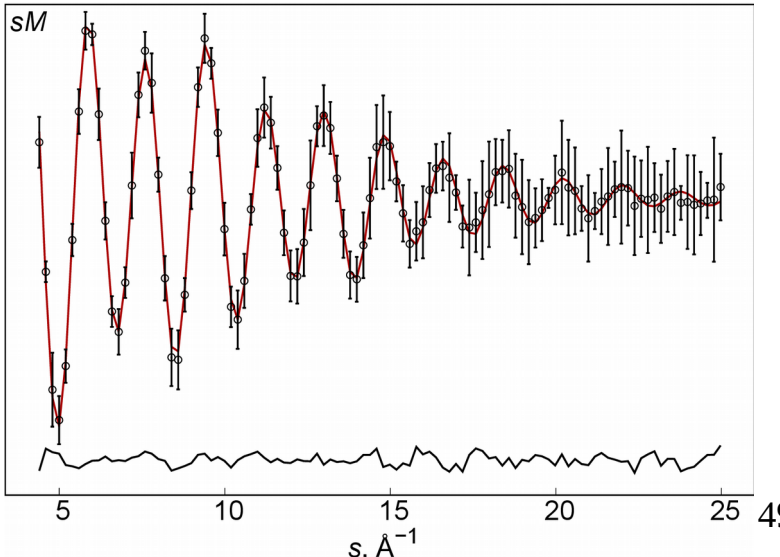
Calculated along with wR_{exp}

For LSQ:
$w_i = \sigma_i^{-2}$





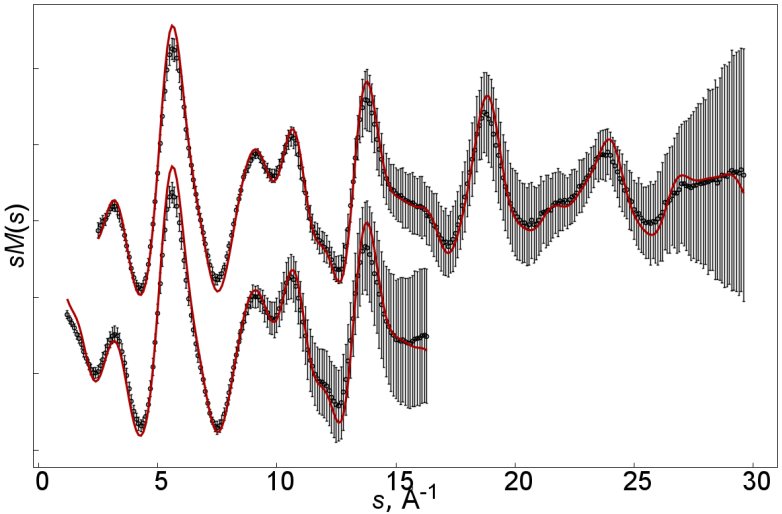




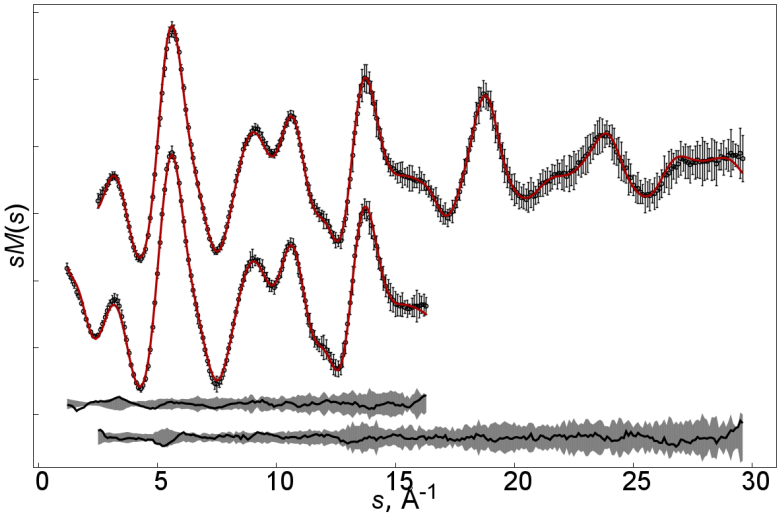
What to average

(A. Otletov et al., Acenaphtene, in preparation):

Averaged $I(s)$, $\sigma_{I(s)}$ propagated to $\sigma_{sM(s)}$



Averaged $sM(s)$, $\sigma_{sM(s)}$ calculated directly



Variant for $M(s)$ $\sigma_{I(s)} \rightarrow \sigma_{M(s)}$:

Tarasov, Y. I.; Kochikov, I. V.; Kovtun, D. M.; Polenov, E. A.; Ivanov, A. A. *Internal Rotation and Equilibrium Structure of the 2-Methyl-2-Nitropropane Molecule from Joint Processing of Gas Phase Electron Diffraction Data, Vibrational and Microwave Spectroscopy Data, and Quantum Chemical Calculation Results.* J Struct Chem 2017, 58 (3), 498–507.

But this is not a solution because signal-to-noise ratio remains the same!

Status of refined parameters

Table S42. Refined (r_{h1}) and theoretical[a] (r_e) parameter values[b] and SARACEN restraints[c] applied in the least-squares refinement of **1** and **2**.

	r_{h1}	r_e	Restraint
p_1	215.8(7)	217.6	–
p_2	−13.8(10)	−13.0	−13.0(7)
p_3	33.6(11)	30.1	30.1(18)
p_4	22.1(21)	21.5	21.5(18)
p_5	260.0(7)	257.4	–
p_6	317.2(7)	318.5	–
p_7	125.8(2)	124.8	–
p_8	−0.5(1)	−0.6	−0.6(1) ← ?
p_9	3.4(4)	3.4	3.4.(4) ← ???
p_{10}	153.1(6)	154.9	–
p_{11}	133.7(1)	132.7	–
p_{12}	110.7(1)	110.4	110.4(1)
p_{13}	120.8(3)	120.8	120.8(3) ← ???
p_{14}	97.5(3)	97.5	97.5(3) ← ???

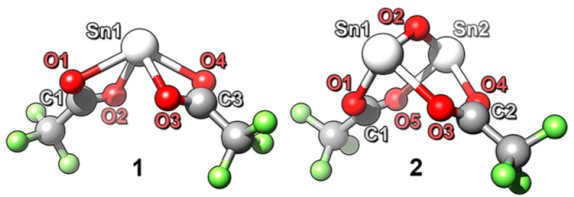
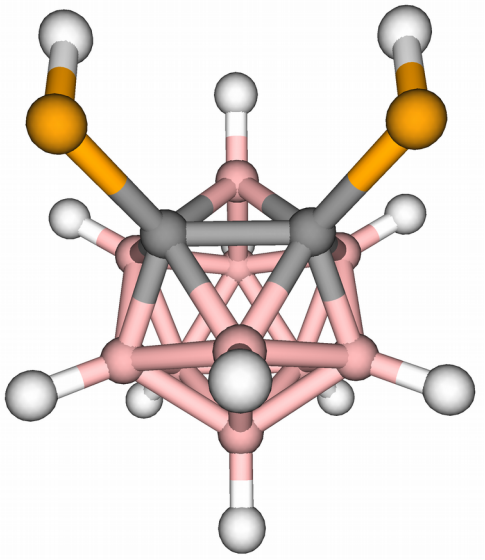


Figure 1. Gas phase products of the thermal depolymerization of **3** determined by GED: tin(II) bis(trifluoroacetate) (**1**, left) and ditin(II) μ-oxybis(μ-trifluoroacetate) (**2**, right).

Bačić, G.; Rankine, C. D.; Masuda, J. D.; Wann, D. A.; Barry, S. T. *Volatile and Thermally Stable Polymeric Tin Trifluoroacetates.* Inorg. Chem. 2020, 59 (2), 996–1005. DOI: 10.1021/acs.inorgchem.9b01308.

Contribution of experimental data: W2 method

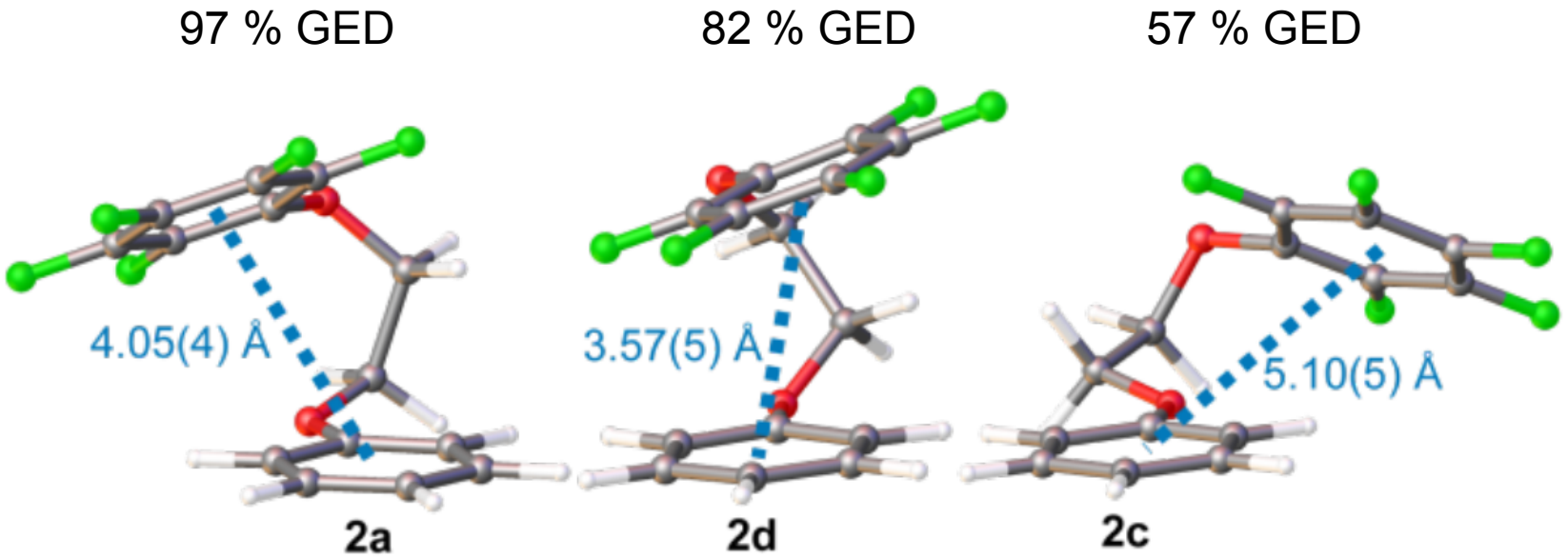
1,2–$(SeH)_2$–*closo*–1,2–$C_2B_{10}H_{10}$



Parameter	MP2/ cc-pVTZ	GED			
		α = 0.7		α = 43	
	r_e	r_e	w_{GED}	r_e	w_{GED}
*r*Se–H	1.454	1.508(17)	**0.388**	1.457(4)	**0.024**
*r*B–B	1.782	1.777(19)	**0.943**	1.775(3)	**0.236**
*r*C–Se	1.904	1.902(10)	**0.987**	1.904(3)	**0.521**
∠B–B–B	60.0	60.0(7)	**0.764**	60.0(1)	**0.055**
∠Se–C–B	118.9	119.1(9)	**0.908**	118.9(2)	**0.172**
∠B–B–H	123.1	123.0(12)	**0.506**	123.1(2)	**0.016**
|∠CCSeH|	92.1	93.9(7)	**0.243**	92.7(1)	**0.002**
R_f, %	10.8	4.0		5.1	

Baše, T.; Holub, J.; Fanfrlík, J.; Hnyk, D.; Lane, P. D.; Wann, D. A.; Vishnevskiy, Y. V.; Tikhonov, D.; Reuter, C. G.; Mitzel, N. W. *Icosahedral Carbaboranes with Peripheral Hydrogen–Chalcogenide Groups: Structures from Gas Electron Diffraction and Chemical Shielding in Solution.* Chem. Eur. J. 2019, 25 (9), 2313–2321. DOI: 10.1002/chem.201805145

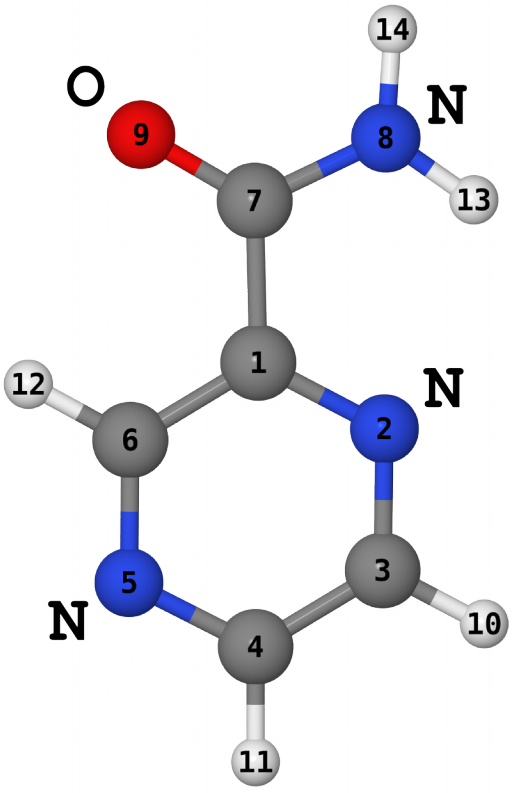
W2 for centroid-centroid distances



Essentially “experimental” results!

J.-H. Weddeling, Yu. V. Vishnevskiy, B. Neumann, H.-G. Stammler, N. W. Mitzel,
Inter- and Intramolecular Aryl-Aryl-Interactions in Partially Fluorinated Ethylenedioxy-Bridged Bisarenes,
ChemRxiv, Preprint, 2020, DOI: 10.26434/chemrxiv.12643154

W12 Method



	MP2(full)/ cc-pwCVTZ	GED		
	r_e	r_e	*w*12	*w*2
*r*C1–N2	1.336	1.341(2)	1.00	0.96
*r*C1–C6	1.391	1.404(2)	1.00	0.96
*r*C3–H10	1.081	1.082(4)	0.02	0.21
∠N2–C3–C4	122.0	122.1(2)	0.90	0.68
∠N8–C7–O9	125.2	124.9(2)	1.00	0.88
$\angle(X–C3–H10)_{av}$	119.1	119.0(3)	0.26	0.39
∠H13–N8–H14	121.8	121.8(3)	0.00	0.03

Tikhonov, D. S.; Vishnevskiy, Y. V.; Rykov, A. N.; Grikina, O. E.; Khaikin, L. S. *Semi-Experimental Equilibrium Structure of Pyrazinamide from Gas-Phase Electron Diffraction. How Much Experimental Is It?* J. Mol. Struct. 2017, 1132, 20–27. DOI: 10.1016/j.molstruc.2016.05.090

Pure experimental errors

LSQ functional:

$$Q = \sum_i w_i \left[sM(s)_{\exp} - sM(s)_{model} \right]^2 + \alpha \sum_j^{3N} w_j \left(x_{j,0} - x_{j,model} \right)^2 \rightarrow min$$

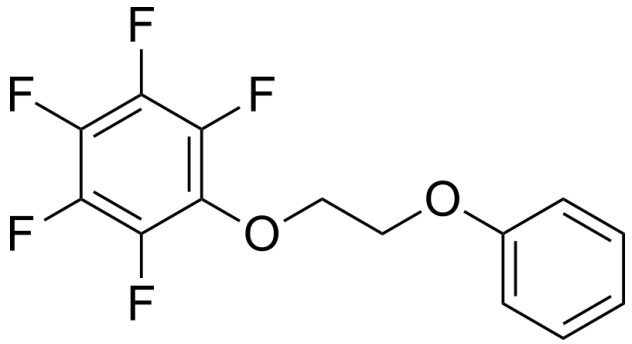
Experimental data,
contribute to the precision of
refined parameters.

Theoretical data,
artificially improves the precision of
refined parameters;
we want to avoid this!

A method for calculation of pure experimental errors of refined parameters:

Tikhonov, D. S.; Vishnevskiy, Y. V.; Rykov, A. N.; Grikina, O. E.; Khaikin, L. S. *Semi-Experimental Equilibrium Structure of Pyrazinamide from Gas-Phase Electron Diffraction. How Much Experimental Is It?* J. Mol. Struct. 2017, 1132, 20–27. DOI: 10.1016/j.molstruc.2016.05.090

An example: LSQ standard deviations



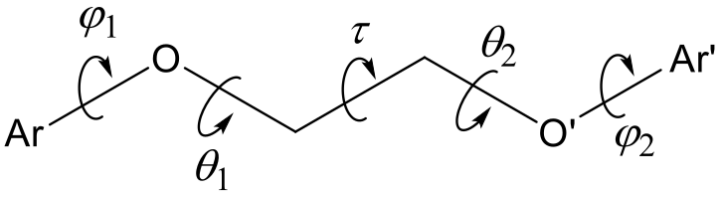


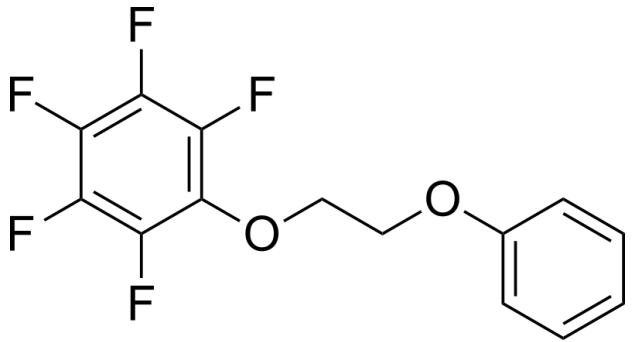
Table S39: Refined torsion angles from GED data in models of **2**.[a]

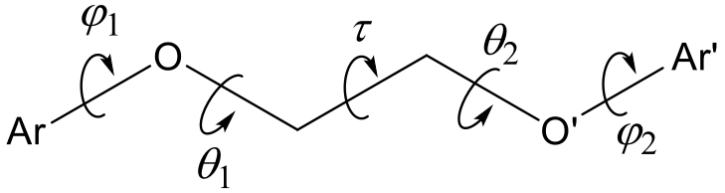
	φ_1	θ_1	τ	θ_2	φ_2	*wR*, %
2a	−17.8(32)	−65.9(9)	−59.8(5)	75.9(9)	72.8(18)	3.62
2b	−8.1(41)	−120.7(11)	82.1(6)	−54.3(11)	−76.4(14)	4.78
2c	14.6(42)	166.7(9)	−71.0(5)	71.6(9)	−105.0(14)	3.59
2d	−47.1(32)	−104.8(9)	58.9(5)	36.7(9)	−129.6(13)	3.60
2e	23.3(45)	70.9(10)	50.2(5)	43.0(10)	71.4(14)	5.15
2f	−20.0(70)	−80.5(15)	−179.7(8)	−78.5(15)	−57.2(20)	5.24
2g	157.3(37)	109.7(8)	−70.9(5)	145.3(8)	63.4(15)	4.14
2h	0.0	180.0	180.0	180.0	180.0	11.10

[a] In parentheses are 1σ standard deviations from least squares refinements.

J.-H. Weddeling, Yu. V. Vishnevskiy, B. Neumann, H.-G. Stammler, N. W. Mitzel,
Inter- and Intramolecular Aryl-Aryl-Interactions in Partially Fluorinated Ethylenedioxy-Bridged Bisarenes,
ChemRxiv, Preprint, 2020, DOI: 10.26434/chemrxiv.12643154

An example: pure experimental errors

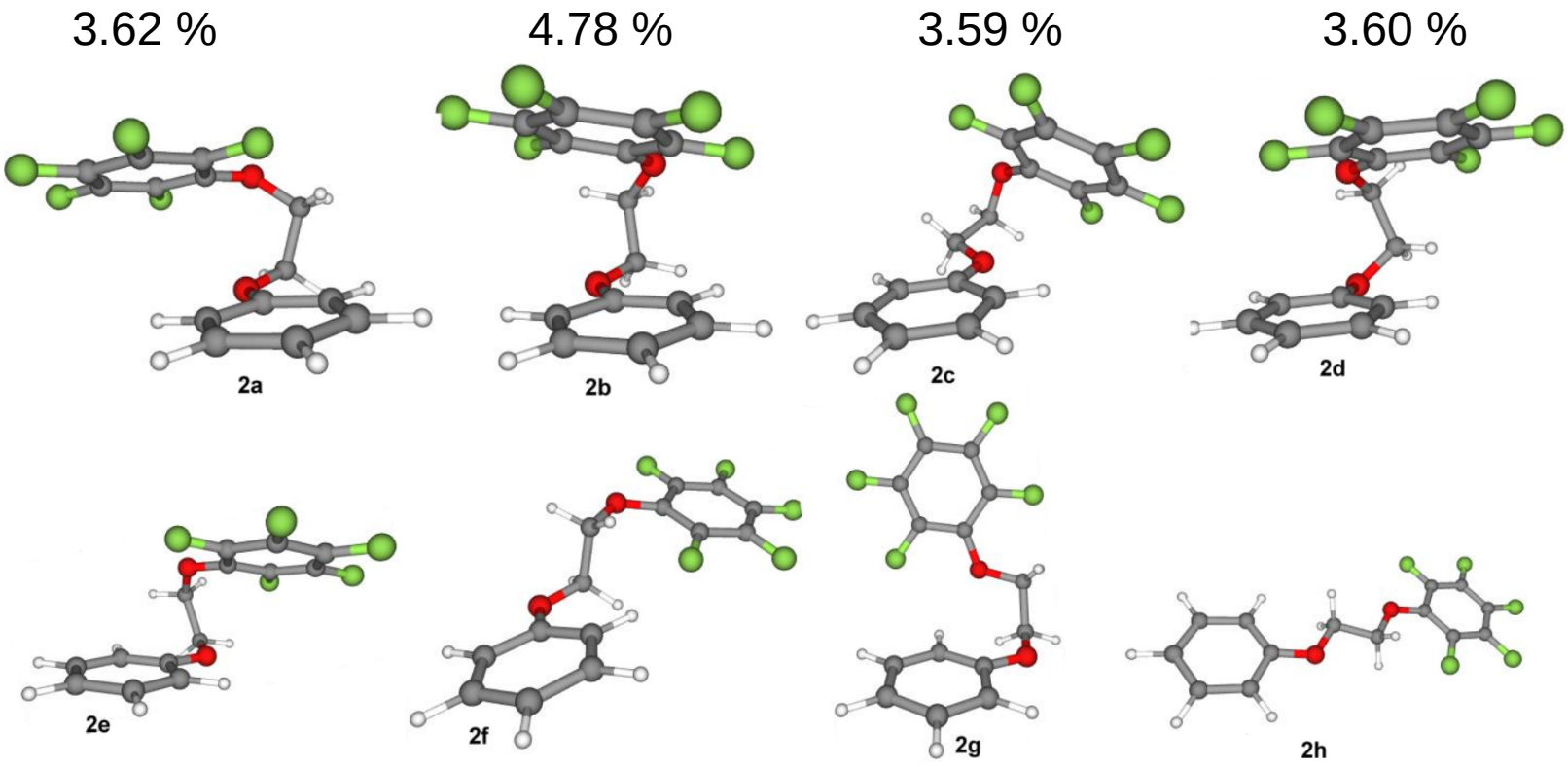




Parameter	w_{GED}, %	Least squares standard deviation, °	Pure experimental st. deviation, °
2a, φ_1	98	3.2	3.2
2a, θ_1	7	0.9	3.4
2d, φ_1	50	3.2	4.5
2c, τ	2	0.5	3.6

J.-H. Weddeling, Yu. V. Vishnevskiy, B. Neumann, H.-G. Stammler, N. W. Mitzel,
Inter- and Intramolecular Aryl-Aryl-Interactions in Partially Fluorinated Ethylenedioxy-Bridged Bisarenes,
ChemRxiv, Preprint, 2020, DOI: 10.26434/chemrxiv.12643154

Results, *wR*-factors



Equal conditions for all models and high-quality backgrounds!

Results, *wR* vs. *ΔE*

kJ mol^{-1}

	φ_1	θ_1	τ	θ_2	φ_2	wR[b]	ΔE (GFN1-xTB)	ΔE (PBEh-3c)	
2a	−17.8(32)	−65.9(34)	−59.8(21)	75.9(14)	72.8(22)	3.62	0.00	0.00	!
2b	−8.1(71)	−120.7(31)	82.1(25)	−54.3(47)	−76.4(16)	4.78	3.89	0.46	?
2c	14.6(59)	166.7(51)	−71.0(36)	71.6(36)	−105.0(18)	3.59	3.10	0.21	
2d	−47.1(45)	−104.8(28)	58.9(19)	36.7(24)	−129.6(14)	3.60	6.44	6.02	?
2e	23.3(150)	70.9(43)	50.2(10)	43.0(18)	71.4(17)	5.15	7.78	9.92	
2f	−20.0(300)	−80.5(130)	−179.7(25)	−78.5(88)	−57.2(20)	5.24	6.23	12.38	
2g	157.3(74)	109.7(21)	−70.9(33)	145.3(25)	63.4(20)	4.14	5.69	13.05	?
2h	0.0	180.0	180.0	180.0	180.0	11.10	8.83	18.66	!

[a] In parentheses are 1σ pure experimental errors, see text for details.
[b] Calculated as $wR = [\sum w_i\{sM_{(i),\mathrm{exp}} - sM_{(i),\mathrm{model}}\}^2 / \sum w_i\{sM_{(i),\mathrm{exp}}\}^2]^{½} \times 100\ \%$

Interpretation of results

- Bond lengths – typical values, not relevant to the problem.
- Main conformation is not stretched.
- Can we exclude existence of small fractions of stretched conformer(s)? No! (stats., temp., …)
- Can we say, which is the most stable conformation? No!
- Could be overlook other conformation(s)? Yes!
- Can we determine, whether we have a mixture? No!
- How accurate are the refined torsion angles? Probably highly affected by uncertainty in model!

Summary

- Use DFT with dispersion corrections for supplementary calculations.
- Calculate amplitudes and corrections using MD or even better PIMD whenever possible.
- Define geometry in Cartesian coordinates, unless you do something special.
- Apply regularization to natural geometrical parameters.
- Fix ratios of amplitudes.
- Use data weighting based on experimental standard deviations.
- Control levelled (reduced) background.
- Calculate contributions W2 or W12.
- Calculate pure experimental standard deviations.
- Check and publish experimental *wR*-factors.
- Please, provide all types of intensities in SI!

Topics not covered:

- Data reduction.
- Calibration.
- Refinement of more complicated models, with mixtures, ...
- Calculation of total errors, correlations.
- Technical details (usage of programs, etc...)

Thank you!