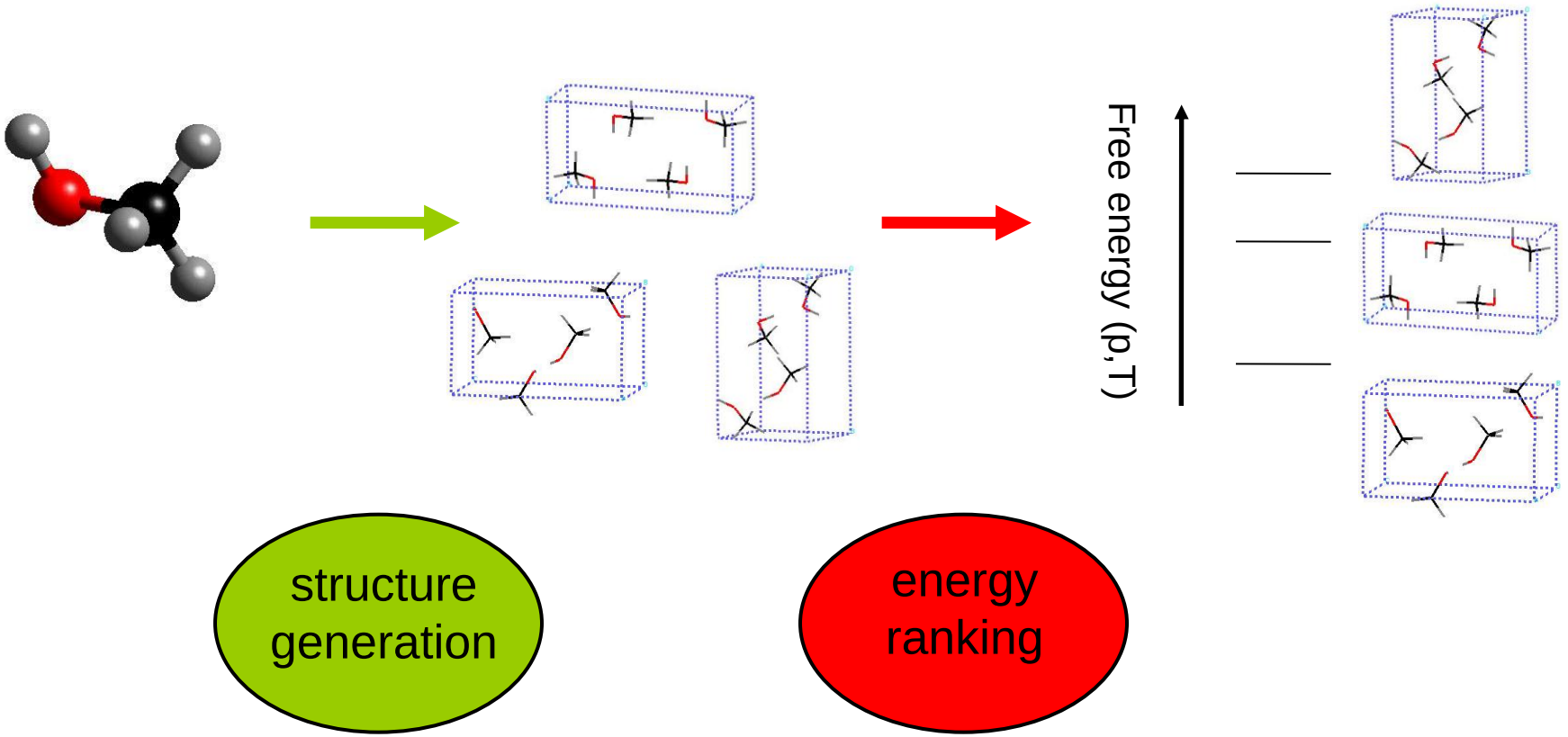
Theoretical Calculations to Assist Experimental Crystal Form Screening

Jacco van de Streek

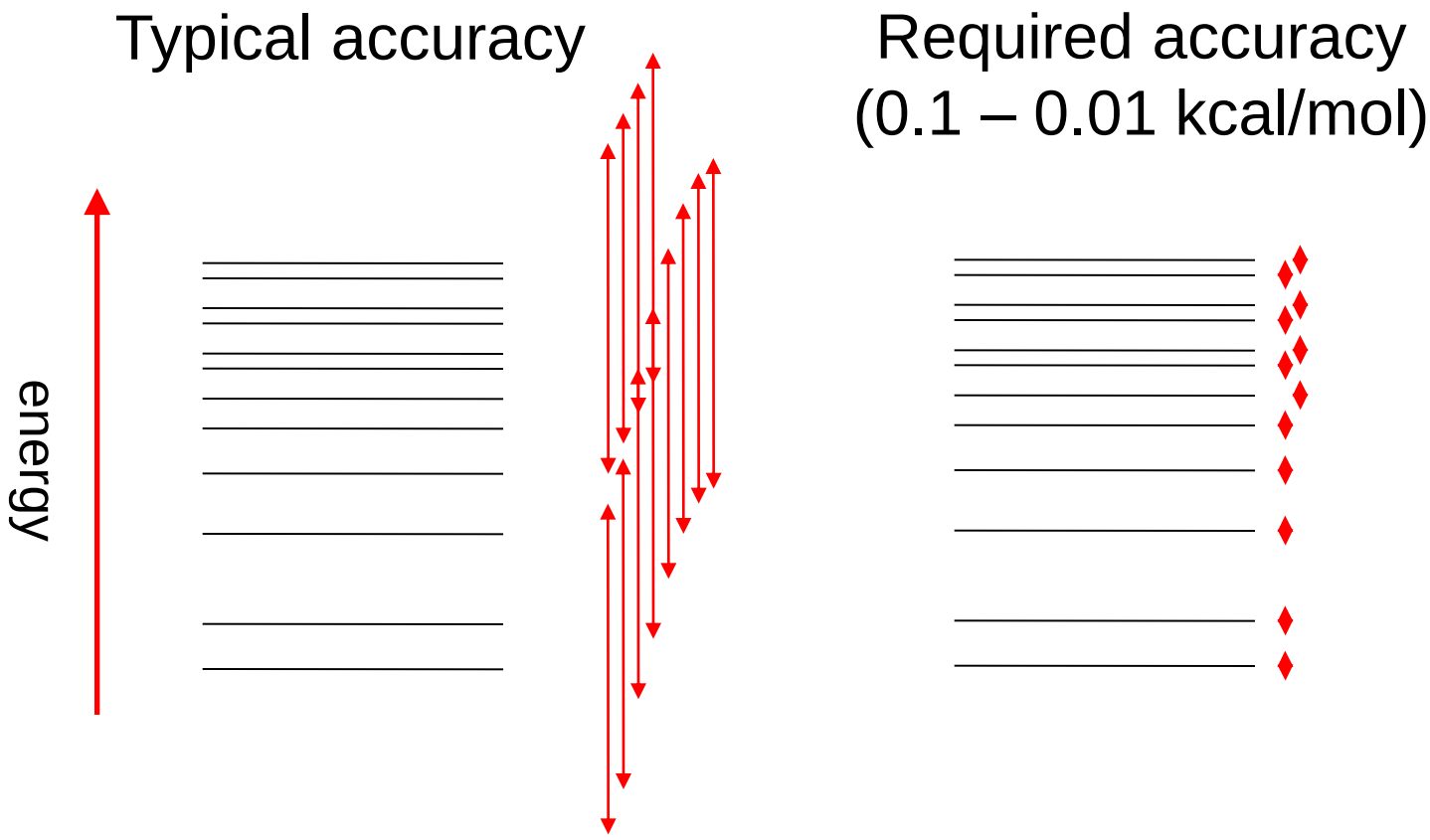
University of Copenhagen

Department of Pharmaceutics and Analytical Chemistry

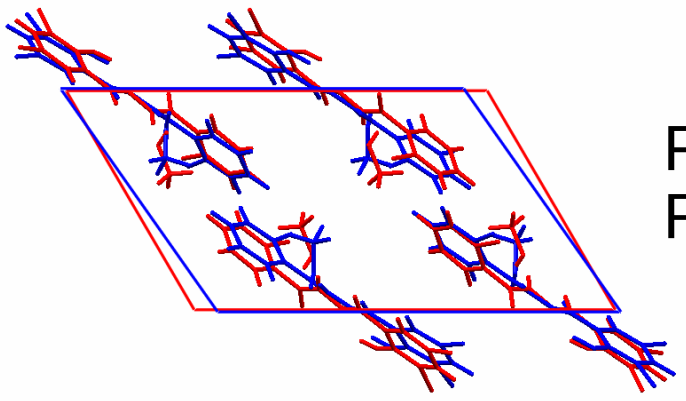
Crystal Structure Prediction: Basics



Why Force Fields Do not Work



Dispersion-corrected DFT (DFT-D)

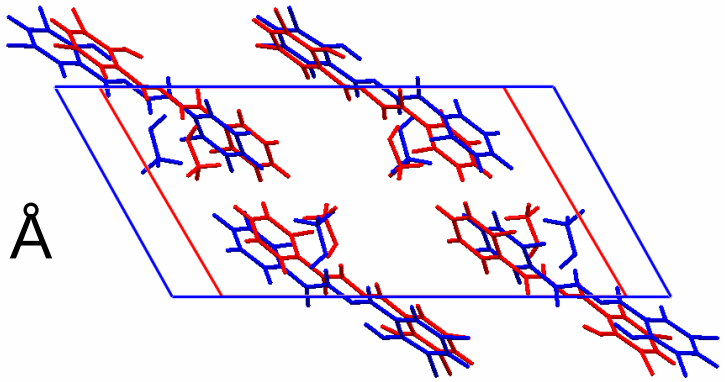


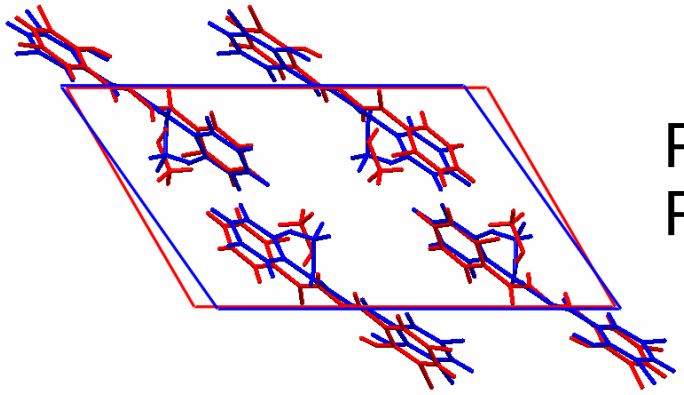
Force fields...
RMS = 0.497 Å

Predicted
Experiment

Dispersion-corrected DFT (DFT-D)

Pure DFT...
RMS = 0.833 Å



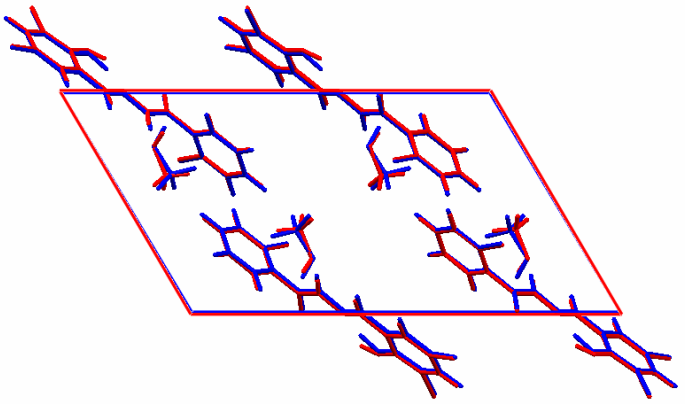


Force fields...
RMS = 0.497 Å

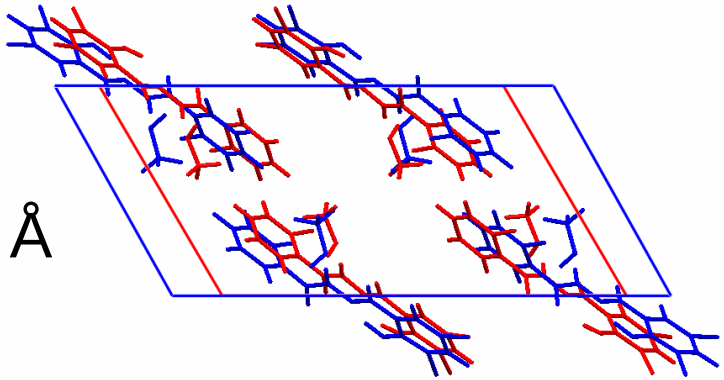
Predicted
Experiment

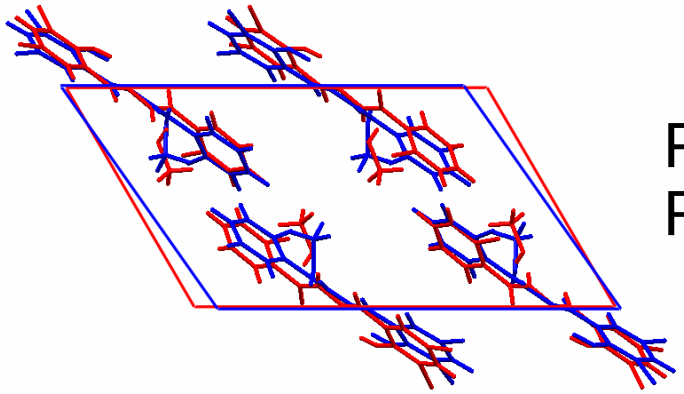
Dispersion-corrected DFT (DFT-D)

Dispersion-corrected DFT...
RMS = 0.084 Å



Pure DFT...
RMS = 0.833 Å





Force fields...
RMS = 0.497 Å

Predicted
Experiment

Dispersion-corrected DFT (DFT-D)

$$E_{\text{lattice}} = E_{\text{DFT}} + E_{\text{vdW}}$$

VASP

- PAW potentials
- Plane-wave basis set
- GGA – PW91 / GGA – PBE
- 520 eV enery cut-off
- 0.07 Å^{-1} k-point spacing

- Pair potentials
- Element dependent
- Hybridisation dependent
- $-C_6 \cdot r^{-6}$

"0 K" calculations, no free energies

M. A. Neumann & M.-A. Perrin (2005) *J. Phys. Chem. B* **109**, 15531-15541
G. Kresse & J. Hafner (1993) *Phys. Rev. B* **47**, 558-561

Crystal Structure Generation

Parallel tempering Monte-Carlo algorithm

Tailor-made force field: accurate force field, fitted to artificial DFT-D reference data for each compound

One and two independent molecules in all 230 space groups

M. A. Neumann (2008) *J. Chem. Phys. B* **112**, 9810-9829

Examples: Blind Tests

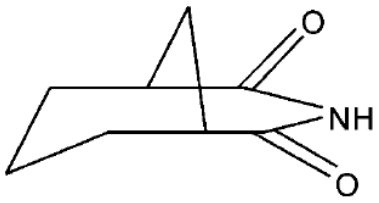
Crystal Structure Prediction Blind Tests:

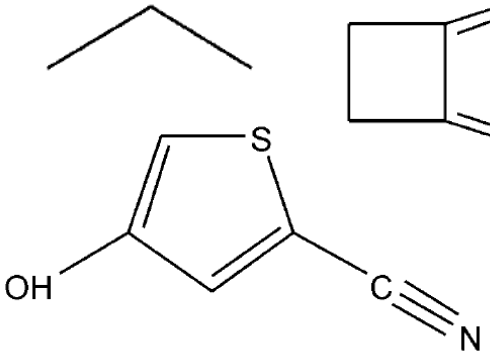
Blind Tests in 1999, 2001, 2004, 2007 and 2010

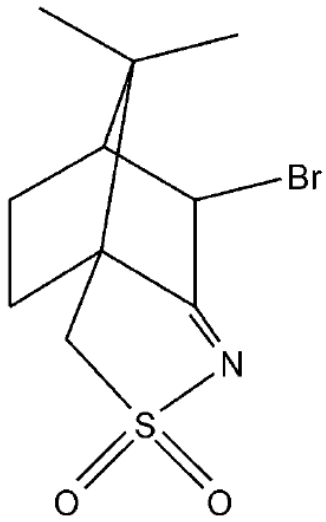
Good compounds for validation

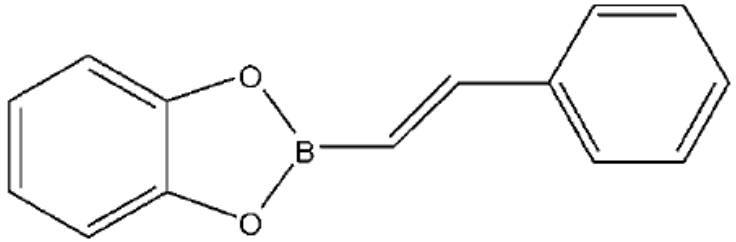
G. M. Day *et al*. (2009) *Acta Cryst.* B**65**, 107-125

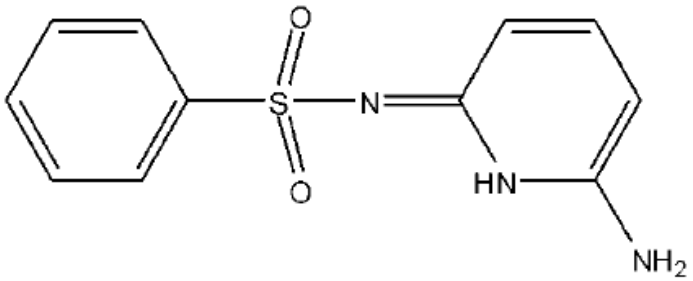
Previous Blind Test Compounds

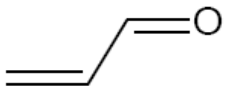


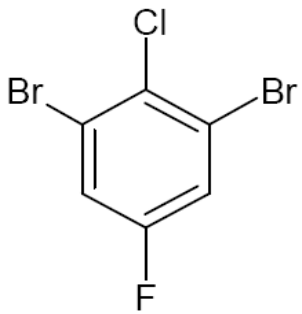


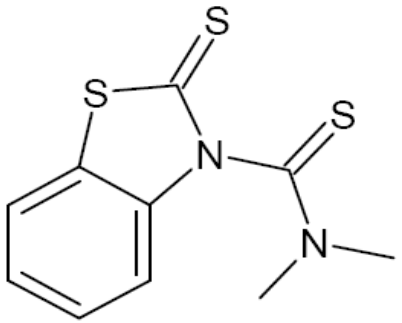


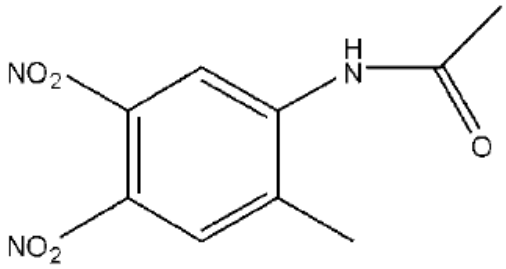


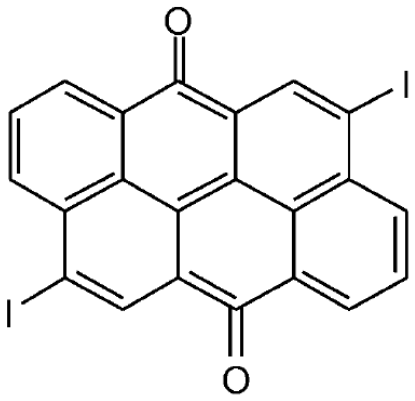


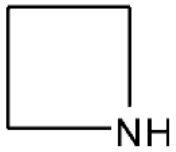


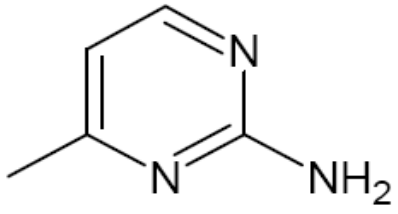




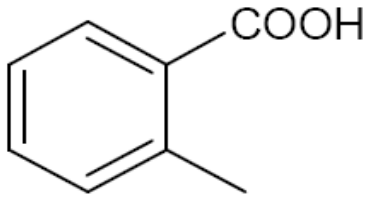




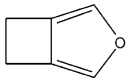




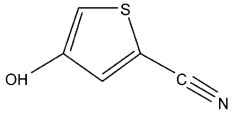
:



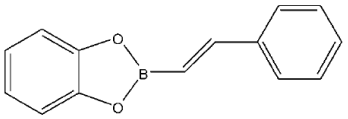
Ranking Results (*NOT* Full Studies)



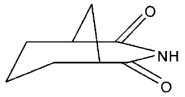
1 & 2



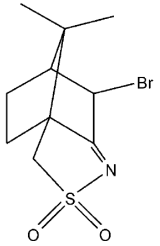
2



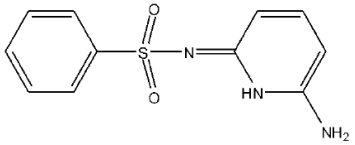
1



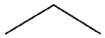
1 & 2



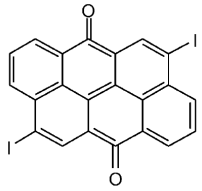
4



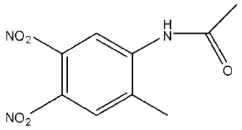
1 & 2



1

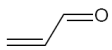


2

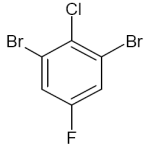


1

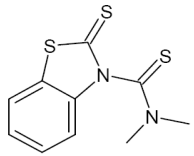
1



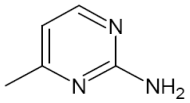
1

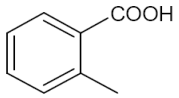


1



1





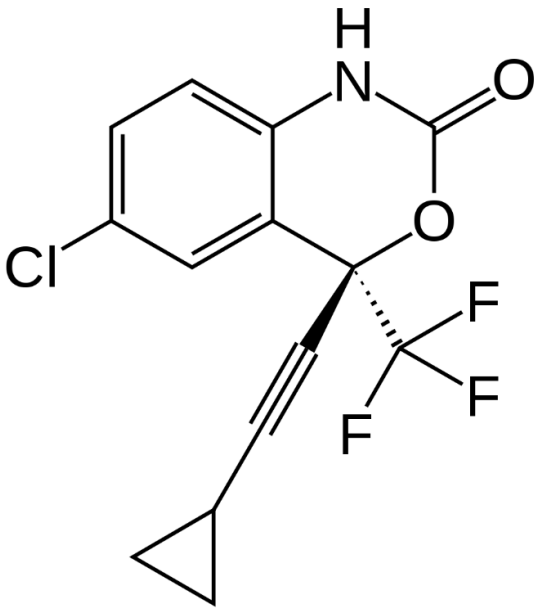
1

A. Asmadi, M. A. Neumann, J. Kendrick, P. Girard, M.-A. Perrin & F. J. J. Leusen (2009) *J. Phys. Chem. B* **113**, 16303-16313

Previous Blind Test Compounds

Conclusion

Dispersion-corrected DFT appears to work well for energy-ranking crystal structures (80% success rate)



Efavirenz

Number of patents: 5
Number of claimed forms: 24
Number of crystal structures known: 0

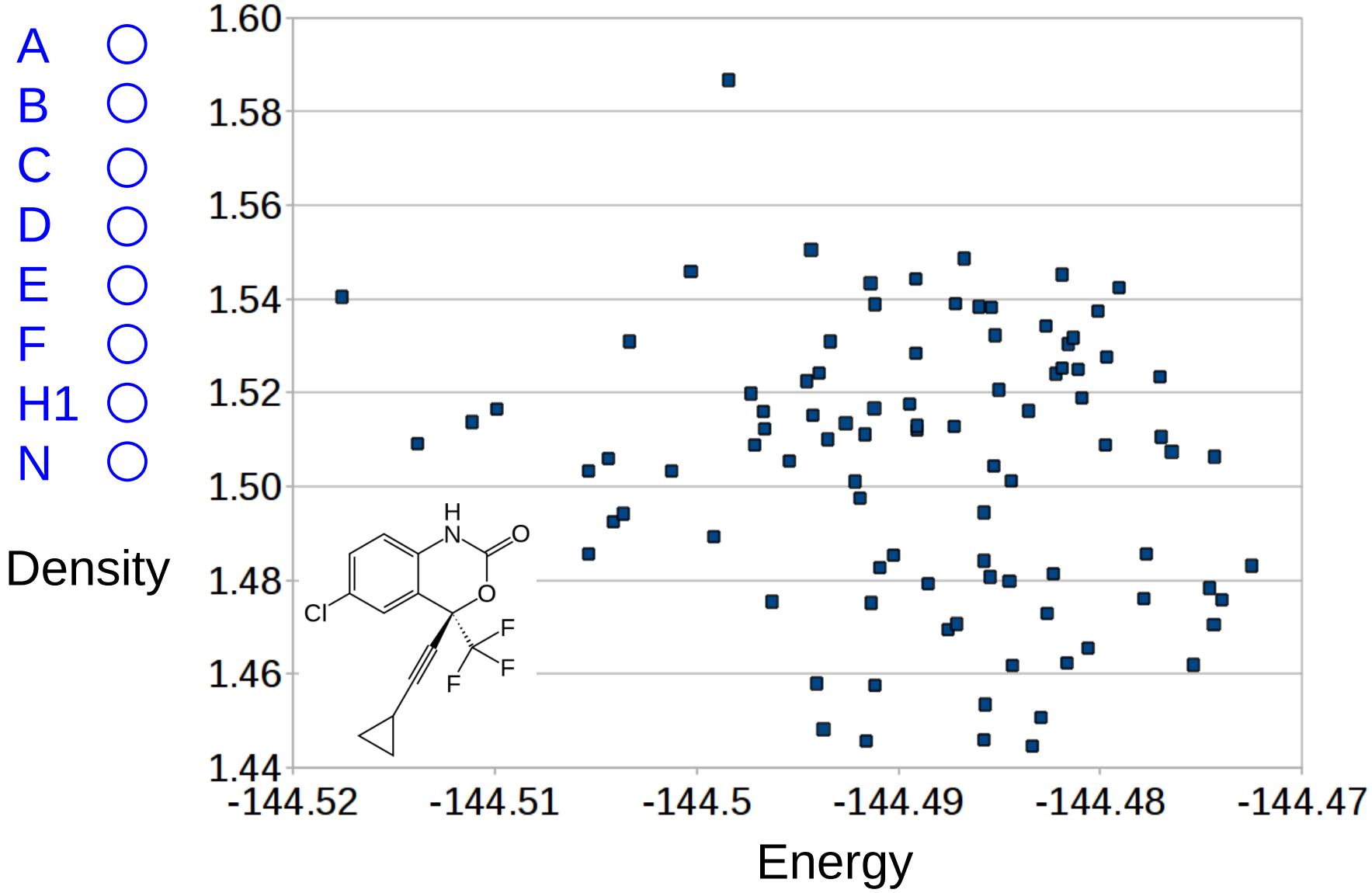
Efavirenz

Experimental polymorph screen by Dr Eva Dova (Avantium)
Scan of patent literature by Dr Menno Deij (Avantium)

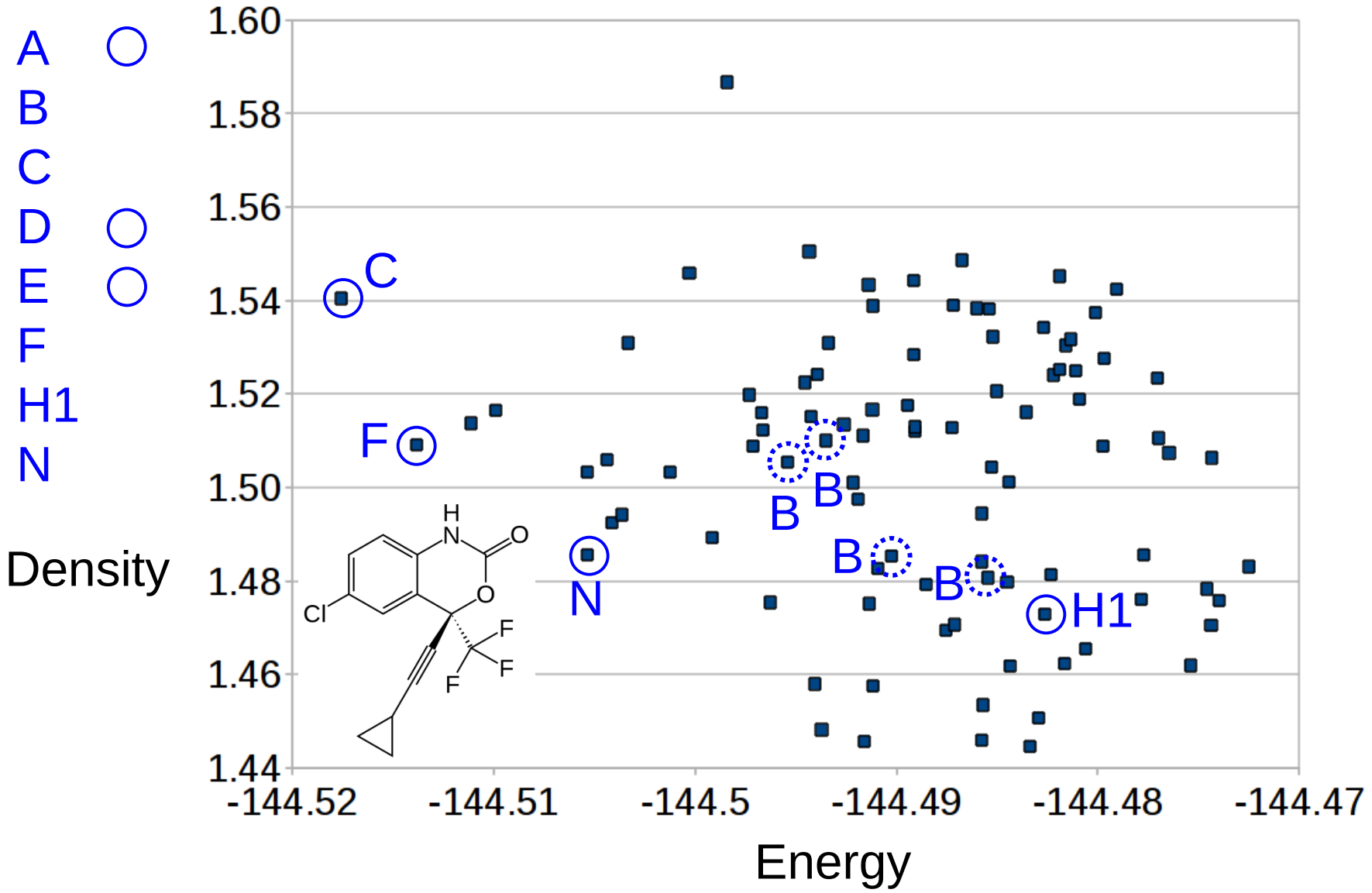
The 24 forms boil down to eight distinct forms:
A, B, C, D, E, F, H1, N

Efavirenz (*Z*'=1-2)

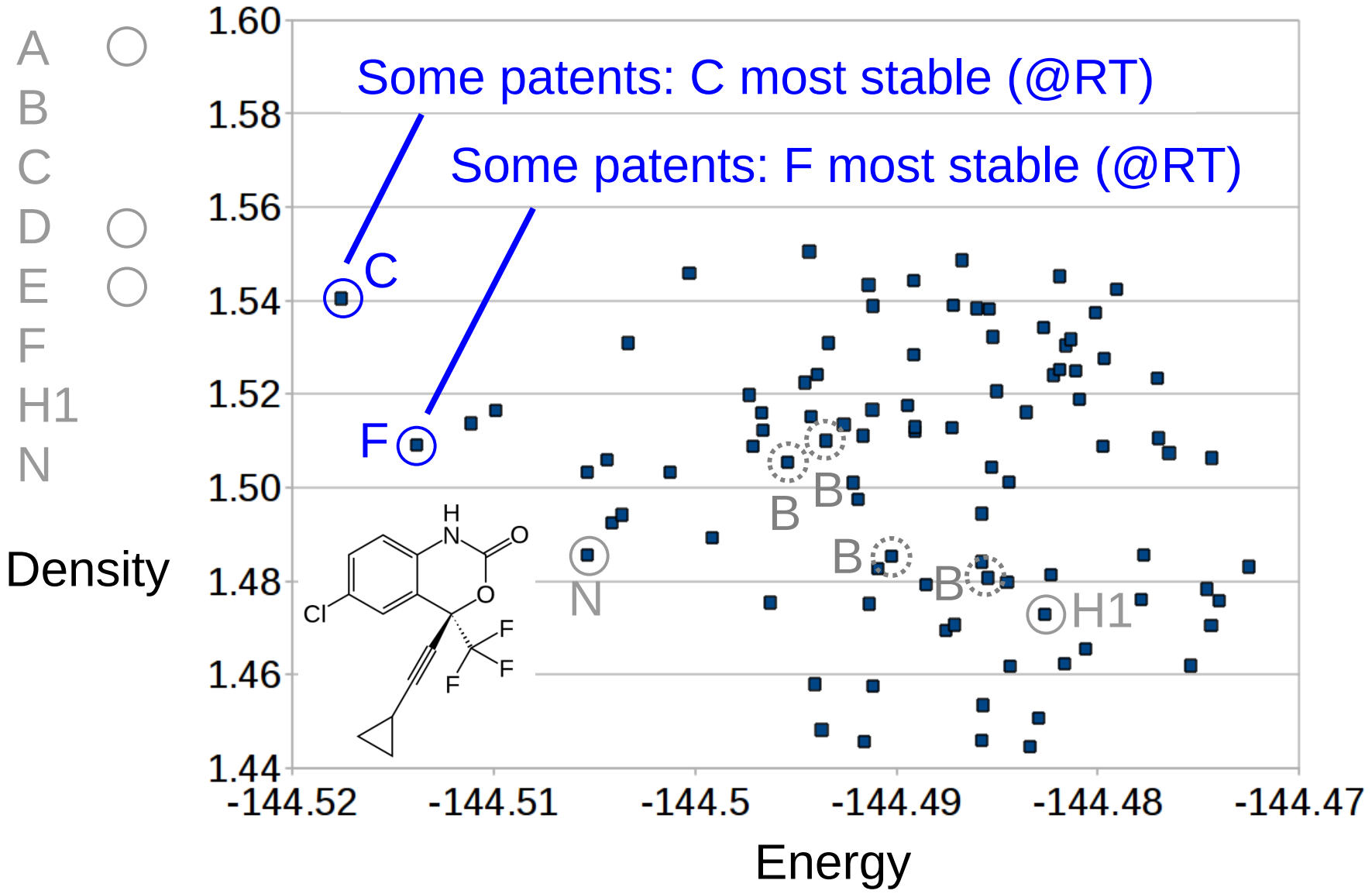
A ○
B ○
C ○
D ○
E ○
F ○
H1 ○
N ○



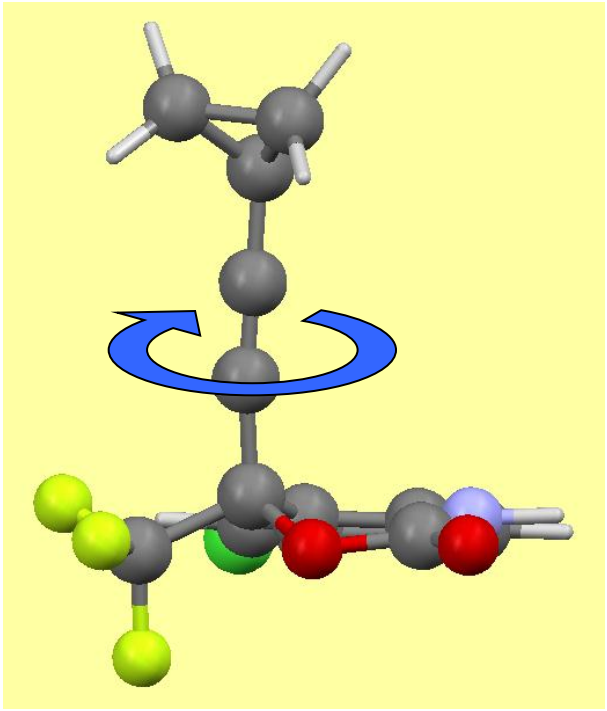
Efavirenz (*Z*'=1-2)



Efavirenz (*Z*'=1-2)



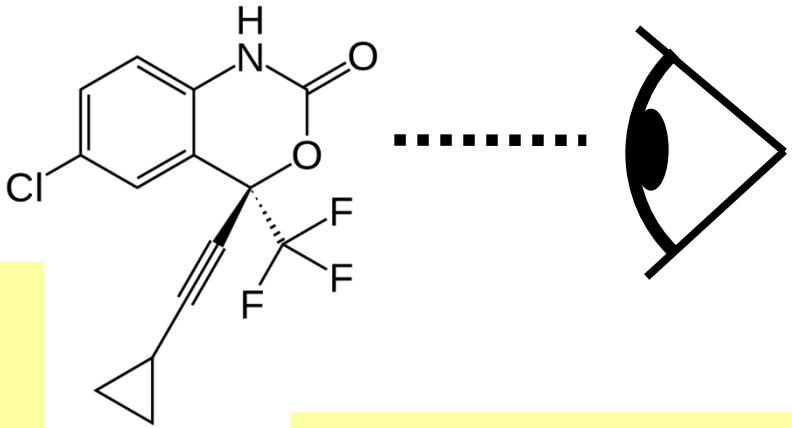
Efavirenz

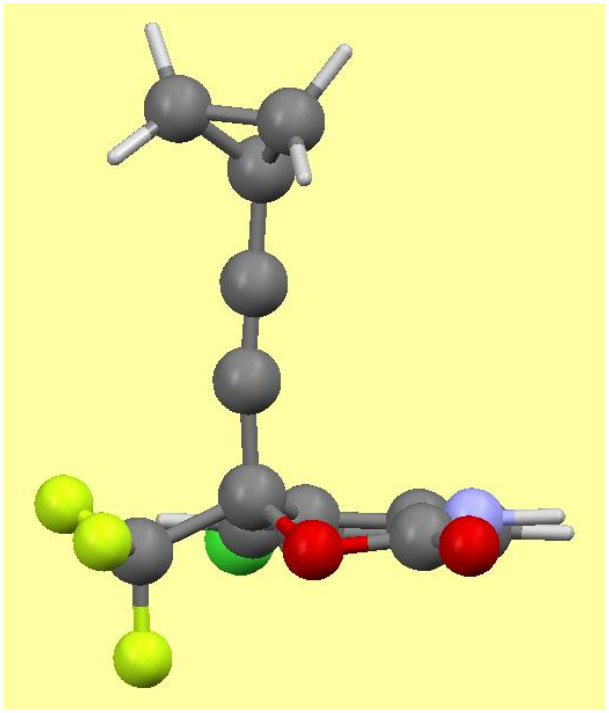


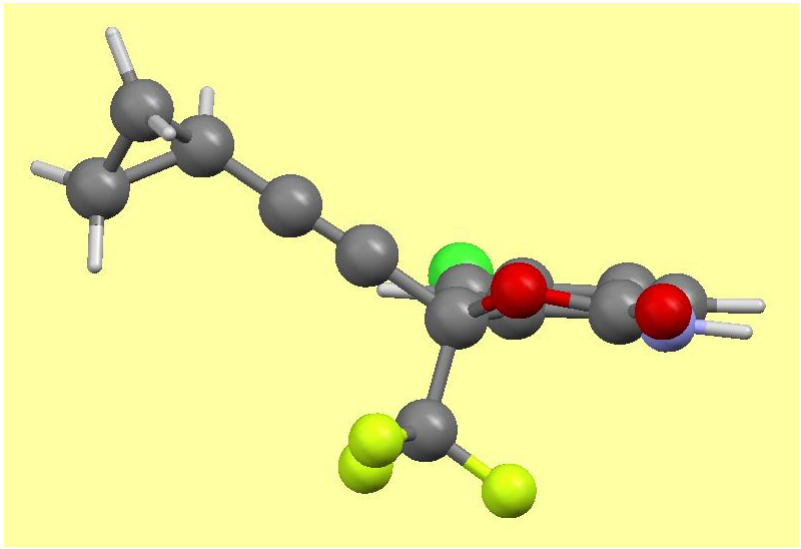
Form B is disordered, four orientations of the cyclopropane group are found in the search (ranks 20, 27, 40, 57)

Disorder means that our 0 K energies are not reliable

S. Cuffini, R. E. Howie, E. R. T. Tiekink, J. L. Wardell &
S. M. S. V. Wardell (2009) *Acta Cryst.* E**65**, o3170-o3171



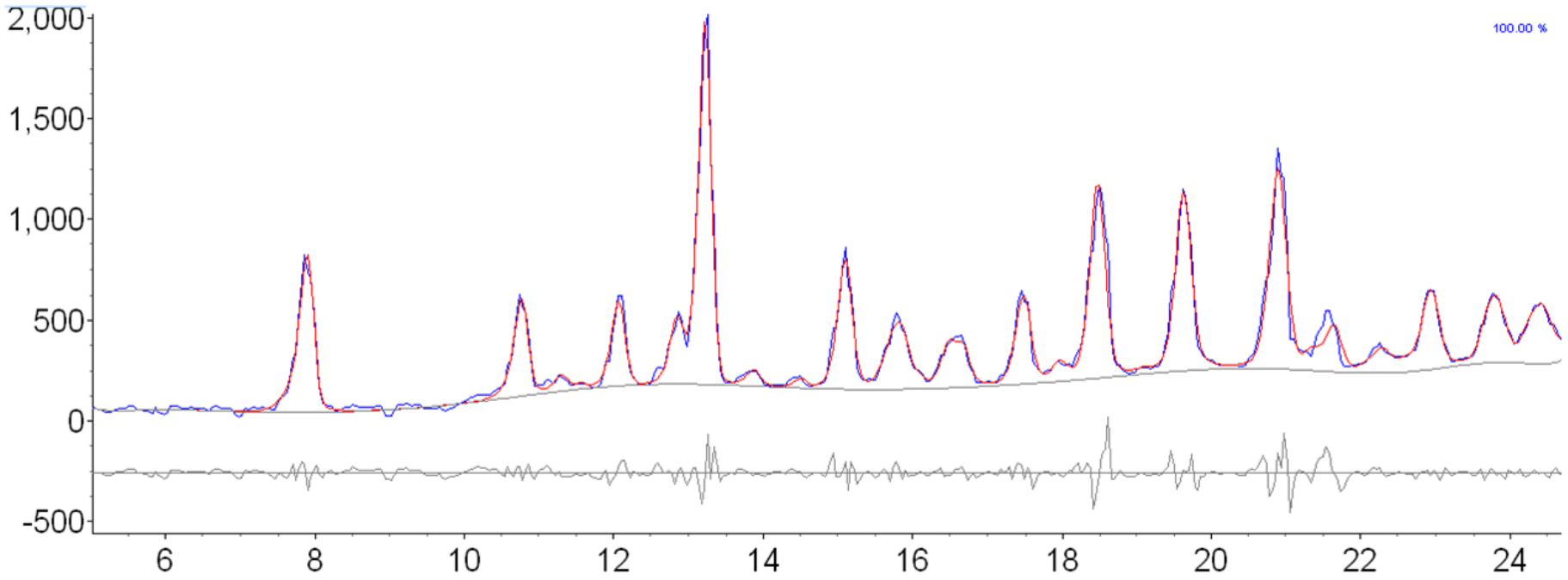




- Form N is *Z*'=2
- One axial -CF_3, one equatorial -CF_3
- Requires fully flexible search with *Z*'=2

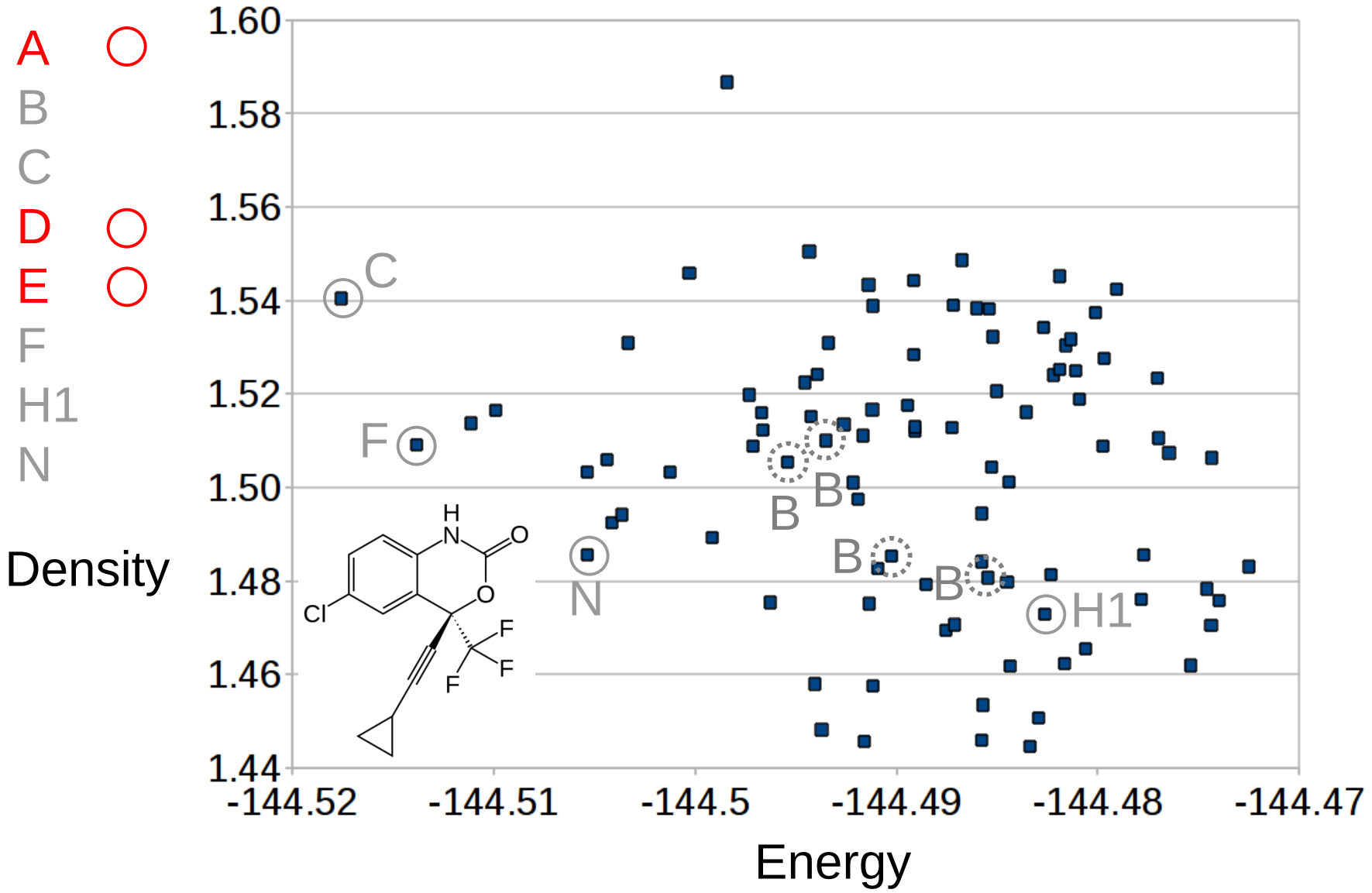
Efavirenz

Rietveld refinement with *TOPAS* for form N (*Z*'=2)



We can solve structures from poor quality laboratory powder patterns, scanned from a patent: low resolution, preferred orientation.

What about Forms A, D & E?



Efavirenz

A	*Z*'=3 / *Z*'=6 [1]
B	Found, Disordered [2]
C	Found
D	Solvate (from TGA)
E	?
F	Found [3]
H1	Found
N	Found

[1] S. Mahapatra, T. S. Thakur, S. Joseph, S. Varughese & G. R. Desiraju (2010) *Cryst. Growth Des*. **10**, 3191-3202

[2] S. Cuffini, R. E. Howie, E. R. T. Tiekink, J. L. Wardell & S. M. S. V. Wardell (2009) *Acta Cryst.* E**65**, o3170-o3171

[3] K. Ravikumar & B. Sridhar (2009) *Mol. Cryst. Liq. Cryst*. **515**, 190-198

Efavirenz

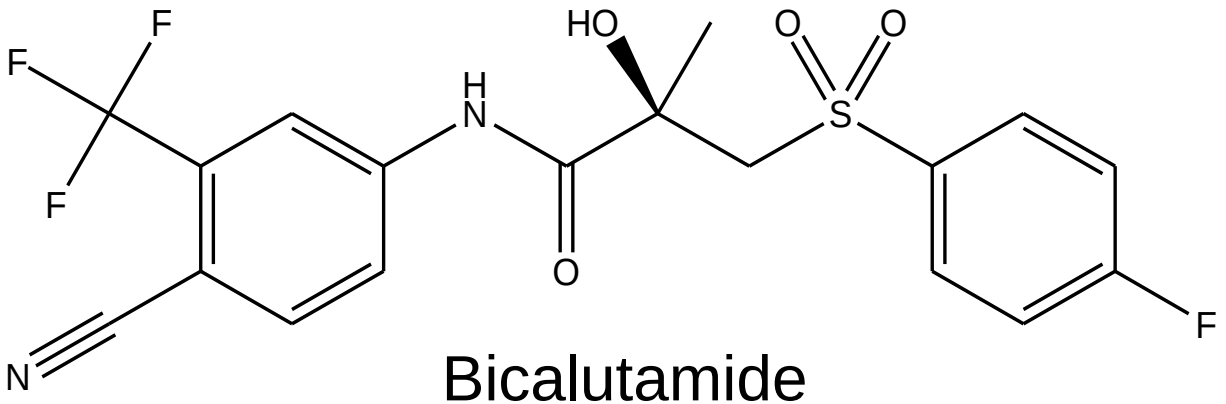
Conclusions

- Real life very complicated:
- *Z*'>2
- Disorder: entropy contribution
- Solvates
- Form E ambiguous

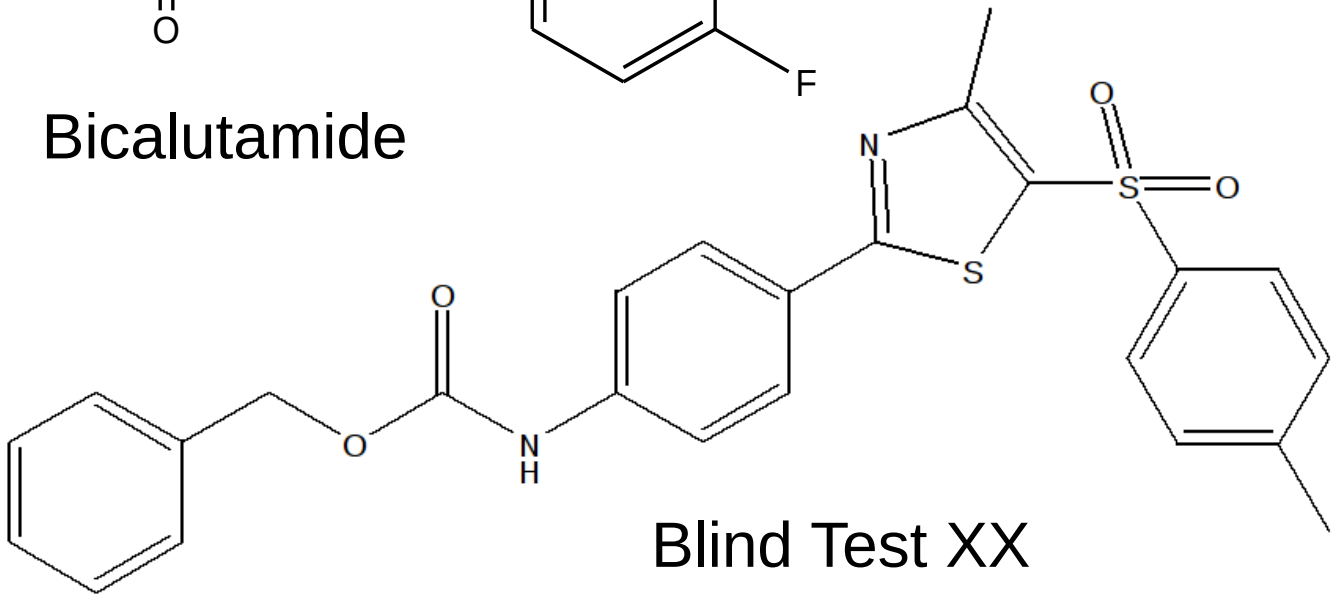
- Crystal structures from very poor quality powder patterns: use crystal-structure prediction

- No more stable form found: it is unlikely that one turns up in the future

Current complexity limit



Bicalutamide



Blind Test XX

- Six months on a 64 CPU quad-core Xeon or Opteron cluster
- Quasicomplete screen for *Z*'=1 in 230 space groups
- 50 – 90 % complete screen for *Z*'=2 in 230 space groups

Acknowledgements

Marcus Neumann - Avant-garde Materials Simulation

Eva Dova - Avantium

Menno Deij - Avantium