Electronic Supplementary Information for:

Modelling Organic Crystal Structures using Distributed Multipole and Polarizability-Based Model Intermolecular Potentials.

Sarah L. Price, Maurice Leslie Gareth W. A. Welch, Matthew Habgood, Louise S. Price, Panagiotis G. Karamertzanis and Graeme M. Day

An example of distributed polarizabilities and induced moments

The dipole moments induced by the field within the crystal structure of trifluoromethyl benzaldehyde oxime (Fig. 1a) are shown in Table ESI1 alongside the corresponding atomic polarizability tensors. The largest dipole polarizabilities are the $(11c,11c)$, $(11s,11s)$ and $(11c,11s)$ components for the carbon, oxygen and nitrogen atoms, which as they lie approximately in the $\mathbf{x_M,y_M}$ plane (Fig. 1b) can be rationalized by charge density moving along the direction of covalent bonds. The off-diagonal terms $\alpha(10,11c)$ and $\alpha(10,11s)$, which are those that change sign for the inverted molecule, are small. The induced dipoles are generally small, with the largest being on the nitrogen atom which is the hydrogen bond acceptor. There is also some polarization of the carbon atoms closest to the virtually coplanar $R_2^2(6)$ hydrogen bonding motif. This change in the molecular charge distribution within the crystal results in an undamped induction energy of 7 kJ mol^{-1}, with about 1 kJ mol^{-1} coming from iterating the induced dipoles to consistency. If damping is applied (β=1.6904 au), then the induction energy is reduced by about 1 kJ mol^{-1}, mainly due to the induced moments on the hydrogen-bonding proton (H1) being reduced to 0.033 ea$_0$ and on the hydrogen-bond acceptor (N1) reduced to 0.105 ea$_0$.

Table ESI1 Independent atomic polarizability tensor components for (E)-4-(trifluoromethyl)benzaldehyde oxime, and the undamped induced dipole moment components, iterated to convergence, and the magnitude in the crystal structure.[a]

	Polarizability tensor components / α_0^3						Induced dipole components and magnitude / ea$_0$			
	$\alpha(10,10)$	$\alpha(11c,11c)$	$\alpha(11s,11s)$	$\alpha(10,11c)$	$\alpha(10,11s)$	$\alpha(11c,11s)$	ΔQ_{10}	ΔQ_{11c}	ΔQ_{11s}	$|\mu|$
C1	3.12	17.27	12.99	0.52	-0.34	-6.55	-0.004	-0.062	0.073	0.096
C2	4.75	16.68	7.76	0.08	0.24	9.50	-0.006	-0.062	-0.032	0.070
C3	4.44	19.36	7.89	-0.35	-0.11	-7.55	0.001	0.007	-0.012	0.014
C4	2.77	18.83	9.39	-0.02	-0.65	0.35	0.005	0.044	0.006	0.045
C5	4.68	10.13	9.28	-0.08	0.11	5.40	0.007	0.034	0.034	0.049
C6	5.03	17.55	8.04	-0.09	-0.02	-3.25	-0.001	-0.002	0.027	0.027
C7	2.34	22.52	7.00	0.38	-0.17	-4.13	-0.001	-0.046	0.055	0.072
C8	3.94	2.94	4.56	-0.01	-0.04	0.41	0.003	0.013	0.007	0.015
O1	4.60	7.86	6.06	0.16	-0.06	-1.17	0.013	0.047	-0.048	0.068
N1	4.93	18.55	7.04	0.72	-0.47	-3.57	-0.016	0.080	0.079	0.114
F1	3.06	3.77	4.15	-0.05	0.08	-0.59	-0.001	0.010	0.000	0.010
F2	4.64	3.91	3.53	0.72	0.57	0.14	0.001	0.013	0.002	0.013
F3	4.11	3.91	3.92	-0.37	-0.49	0.17	0.001	0.027	0.009	0.029
H1	1.06	1.12	1.02	-0.01	-0.02	0.04	0.012	0.003	-0.037	0.039
H2	2.57	1.67	2.66	-0.02	-0.13	-1.27	-0.011	-0.022	0.019	0.031
H3	1.84	-1.00	1.81	0.06	0.28	1.00	-0.006	-0.009	-0.015	0.018
H5	1.85	0.93	2.23	0.05	0.15	-1.68	0.005	0.000	-0.002	0.006
H6	2.17	-0.29	2.14	0.01	0.01	0.72	-0.003	0.003	0.013	0.014
H7	2.13	-0.65	3.34	-0.32	-0.06	0.40	0.001	0.003	0.022	0.022

[a] The quantities which have the opposite sign for molecules related by inversion are in italics.

The FIT and W99 empirical "repulsion-dispersion" potentials

Most of the crystal structure modelling using DMACRYS that has not derived a specific potential for the molecule[1,2] has supplemented the Distributed Multipole electrostatic model with an empirically fitted isotropic atom-atom potential of the form:

$$U = \sum_{i \in M, k \in N} \sqrt{A_{\iota\iota} A_{\kappa\kappa}} \exp\left(-\left(\frac{B_{\iota\iota} + B_{\kappa\kappa}}{2}\right) R_{ik}\right) - \frac{\sqrt{C_{\iota\iota} C_{\kappa\kappa}}}{R_{ik}^6}$$

where atom i of type ι in molecule M is separated by R_{ik} from atom k of type κ in molecule N. Two widely used sets of parameters are tabulated below, with some notes and comments on their use and references to the definitive literature.

W99 was derived[3] by fitting to a wide range of crystal structures and validated against peptide and nucleoside structures. A key feature is that the hydrogen interaction sites are moved by 0.1 Å into the H-X bond from their neutron or ab initio optimized positions. DMACRYS can perform this foreshortening, which has to also be used in the DMA analysis.

Potential	Atom pair	Description	A_{ik} /kJ mol^{-1}	B_{ik}/Å^{-1}	C_{ik}/kJ mol^{-1} Å^6
W99	C(2)···C(2)	C bonded to 2 atoms	103235	3.60	1435.09
W99	C(3)···C(3)	C bonded to 3 atoms	270363	3.60	1701.73
W99	C(4)···C(4)	C bonded to 4 atoms	131571	3.60	978.36
W99	H(1)···H(1)	H bonded to C	12680	3.56	278.37
W99	H(2)···H(2)	H in alcoholic group	361.30	3.56	0
W99	H(3)···H(3)	H in carboxyl group	115.70	3.56	0
W99	H(4)···H(4)	H in N-H group	764.90	3.56	0
W99	N(1)···N(1)	N in triple bond	96349	3.48	1407.57
W99	N(2)···N(2)	other N with no bonded H	102369	3.48	1398.15
W99	N(3)···N(3)	N bonded to 1 H atom	191935	3.48	2376.55
W99	N(4)···N(4)	N with 2 bonded H atoms	405341	3.48	5629.82
W99	O(1)···O(1)	O bonded to 1 other atom	241042	3.96	1260.73
W99	O(2)···O(2)	O bonded to 2 other atoms	284623	3.96	1285.87

This potential was successfully used in a large survey of crystal structure predictions in conjunction with both a point charge[4] and Distributed Multipole model.[5] However problems have been observed in underestimating hydrogen bond distances when combined with a Distributed Multipole Model, presumably because the electrostatic forces can be stronger than with the point charge model used in Williams' parameterization. In certain cases, of (carboxylate)O···H-O and N···H-O hydrogen bonds the underestimate has led to the hydrogen bonds becoming within the covalent bond range. (An ad hoc fix for the latter[6] is to replace the purely repulsive, Williams potential between pyridine nitrogen and carboxylic proton (N(2).H(3)) to A=75719.47 kJ mol^{-1} and B=5.1765 Å^{-1} used to model the cocrystal of 4-aminobenzoic acid with 4-nitrophenylacetic acid. This potential is much steeper at unphysically short distances without penalizing the lattice energy at typical contacts compared with the original Williams parameterization.) Graeme Day's group noted that O-H(3)···N(2) distances

were unreasonably short for the cocrystal in the 2007 blind test,[7] and found substituting the alcohol H(2) parameters for H(3) in this carboxylic acid produced more reasonable results.

FIT has evolved through using Williams older parameterizations, which were fitted to crystals which had each element and C and H only. The H nuclei are used as the interaction sites, although whether Williams corrected for the X-ray foreshortening in the parameterisation is unclear. Note that NEIGHBOURS does not distinguish between H-O and H-N and so many papers have use the H_N (=H_p) parameters for any polar hydrogen, including in hydrates and ice.[8]

Reference	Atom pair	Description	A_{ik} /kJ mol^{-1}	B_{ik}/Å^{-1}	C_{ik}/kJ mol^{-1} Å^6
9	C$\cdots$C	Any C atom	369743	3.60	2439.8
9	H$\cdots$H	H bonded to C	11971	3.74	136.4
10	$H_N\cdots H_N$	H bonded to N	5029.68	4.66	21.50
11	$H_O\cdots H_O$	H bonded to O	2263.3	4.66	21.50
9	N$\cdots$N	Any N atom	254529	3.78	1378.4
12	O$\cdots$O	Any O atom	230064.1	3.96	1123.59
13	F$\cdots$F	Any F atom	363725	4.16	844
14	Cl$\cdots$Cl	Any Cl Atom	924675	3.51	7740.48
Other parameters that have been used in conjunction with FIT for a few applications					
15	S$\cdots$S	S in SO_2	236501.5	2.9	8397.115
15	O$\cdots$O	O in SO_2	202983.6	3.96	640.8628
7	S$\cdots$S	S in blind test (C=S and thioether)	401034	3.30	5791

NEIGHCRYS will automatically provide the FIT or W99 typing and can allow a user specified typing, allowing the ability of DMACRYS to use more atomic types for a given atom to be applied without manual editing of the input file. The extent to which the parameters in combination have been tested beyond[10] is limited, though there has been some validation for the F parameters in Ashley Hulme's thesis and for the isotropic chlorine in situations without close Cl$\cdots$Cl contacts.[16] There are various failures: problems in stacking of some rigid aromatics differences led Tom Lewis to reduce the C parameters by 25% in his studies.[17]

Both these empirically fitted potentials are effectively modelling the total intermolecular potential with the electrostatic component removed, as well as it is sampled in the crystal structures used for fitting and validation. It is not surprising that the results can be sensitive to the quality of the electrostatic model used, and may be very poor for atypical short contacts. Since they are empirically based, the choice of which to use can only be made by empirical testing on related crystal structures. If neither are satisfactory, or the study warrants greater accuracy and confidence in the model potential, then a molecule specific, anisotropic atom-atom potential should be derived.

Reference List

1. S. L. Price and L. S. Price, in *Intermolecular Forces and Clusters I*, ed. D. J. Wales, Springer-Verlag, Berlin, Heidelberg, Germany, 2005 Chapter 3, pp 81-123.

2. S. L. Price, *CrystEngComm*, 2004, **6**, 344-353.

3. D. E. Williams, *J. Comput. Chem.*, 2001, **22**, 1154-1166.

4. G. M. Day, J. Chisholm, N. Shan, W. D. S. Motherwell and W. Jones, *Cryst. Growth Des.*, 2004, **4**, 1327-1340.

5. G. M. Day, W. D. S. Motherwell and W. Jones, *Cryst. Growth Des.*, 2005, **5**, 1023-1033.

6. P. G. Karamertzanis, A. V. Kazantsev, N. Issa, G. W. A. Welch, C. S. Adjiman, C. C. Pantelides and S. L. Price, *J. Chem. Theory Comput.*, 2009, **5**, 1432-1448.

7. G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. Wolf and B. Schweizer, *Acta Crystallogr. , Sect. B*, 2009, **65**, 107-125.

8. A. T. Hulme and S. L. Price, *J. Chem. Theory Comput.*, 2007, **3**, 1597-1608.

9. D. E. Williams and S. R. Cox, *Acta Crystallogr. , Sect. B*, 1984, **40**, 404-417.

10. D. S. Coombes, S. L. Price, D. J. Willock and M. Leslie, *J. Phys. Chem.*, 1996, **100**, 7352-7360.

11. T. Beyer and S. L. Price, *J. Phys. Chem. B*, 2000, **104**, 2647-2655.

12. S. R. Cox, L. Y. Hsu and D. E. Williams, *Acta Crystallogr. , Sect A.*, 1981, **37**, 293-301.

13. D. E. Williams and D. J. Houpt, *Acta Crystallogr. , Sect. B*, 1986, **42**, 286-295.

14. L. Y. Hsu and D. E. Williams, *Acta Crystallogr. , Sect A.*, 1980, **36**, 277-281.

15. W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallogr. , Sect. B*, 2002, **58**, 647-661.

16. S. A. Barnett, A. Johnson, A. J. Florence, S. L. Price and D. A. Tocher, *Cryst. Growth Des.*, 2008, **8**, 24-36.

17. T. C. Lewis, D. A. Tocher and S. L. Price, *Cryst. Growth Des.*, 2005, **5**, 983-993.