Acta Cryst. (1997). B53, 1006–1016

Halogen···O(Nitro) Supramolecular Synthon in Crystal Engineering: A Combined Crystallographic Database and *Ab Initio* Molecular Orbital Study

FRANK H. ALLEN,[a]* JOS P. M. LOMMERSE,[a] VANESSA J. HOY,[b] JUDITH A. K. HOWARD[b] AND GAUTAM R. DESIRAJU[c]

[a]*Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England,* [b]*Department of Chemistry, University of Durham, South Road, Durham DH1 3LE, England, and* [c]*School of Chemistry, University of Hyderabad, Central University PO, Hyderabad 500 046, India. E-mail: allen@ccdc.cam.ac.uk*

(*Received 2 June 1997; accepted 15 July 1997*)

Abstract

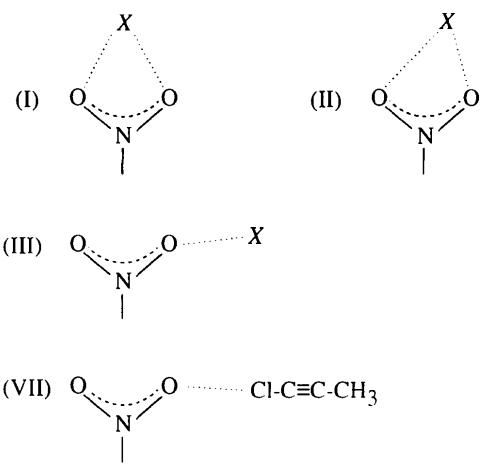
Crystallographic data for 309 C—*X*···O(nitro)—C, *X* = Cl, Br, I, interactions, involving 245 symmetry-independent *X* atoms, have been investigated out to 0.2 Å above van der Waals radii limits [*v*(O) + *v*(*X*)]. A total of 138 (45%) of these interactions are shorter than *v*(O) + *v*(*X*), with the degree of interpenetration of the *X* and O atoms increasing in the order Cl < Br < I. The analysis also shows that: (*a*) the C—*X*···O angle tends to linearity as the *X*···O distance shortens, (*b*) the angle of approach of *X* to the C—NO_2 plane is preferentially less than 45° for Cl, with an increasing tendency to in-plane *X* approach in the order Cl < Br < I, and (*c*) the halogen (*X*) forms either (i) mono-coordinate interactions with one nitro O atom, with *X*···O in a *cis* relationship to the nitro C substituent about the N—O bond, or (ii) approaches both nitro O atoms in a bifurcated manner; the tendency to form such bifurcated motifs increases in the order Cl < Br < I. Only iodine consistently forms short interactions with both nitro O atoms. *Ab-initio*-based molecular orbital calculations, using intermolecular perturbation theory (IMPT) applied to a nitromethane–1-chloro-2-methylacetylene model dimer, agree with the analysis of experimental crystal structure geometries. The IMPT calculations yield an attractive interaction energy of *ca* −6 kJ mol^{-1} for Cl···O at the 6-31G* basis set level. Calculations for Br···O at the (only available) [6*s*4*p*1*d*] basis set level indicate that this interaction is more attractive than Cl···O.

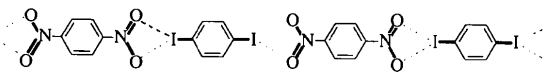
1. Introduction

The presence of certain covalently bonded atoms or functional groups in the same or different molecules can often direct the formation of supramolecular assemblies through the mediation of non-covalent interactions involving those atoms or groups. Desiraju (1995) has defined these non-covalent substructural units as supramolecular synthons and, in recent years, we have identified the importance of the halogen (*X* = Cl, Br, I)···O(nitro) synthon in molecular recognition processes, and have established its use in retrosynthetically guided supramolecular synthesis.

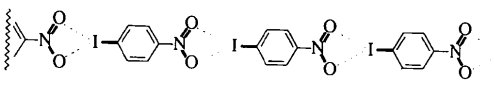
Initial interest in NO_2···*X* interactions arose from the different packings of 4-chloro- and 4-bromo-β-nitrostyrene (Desiraju & Pedireddi, 1989; Pedireddi, Sarma & Desiraju, 1992). The chloro compound is characterized by C—H···O(nitro) and C—H···Cl interactions, but the bromo compound, although exhibiting the C—H···O(nitro) bonds, had the Br atom directed towards, and almost equidistant from, the O atoms of the nitro group, as depicted in motif (I). Later, Desiraju, Pedireddi, Sarma & Zacharias (1993) proposed a model for the 4-iodo compound, which had proved intractable to crystallize. However, these workers also carried out a preliminary analysis of existing short NO_2···*X* contacts using data retrieved from the Cambridge Structural Database (CSD: Allen *et al.*, 1991). They identified (*a*) the symmetrical bifurcated motif (I: *D*1 = *D*2) and (*b*) the asymmetric motif (II, *D*1 < *D*2), in which *X* approaches between the two nitro O atoms (*X*···O *trans* to C—N), and (*c*) motif (III), in which *X*···O is *cis* to C—N. In motif (III) *X* can only interact with one of the nitro O atoms and the motif is termed 'mono-coordinate'. Motifs (I) and (II) are both described as 'bifurcated' in this paper, although the strength of the second *X*···O interaction may be rather weak in many examples of the asymmetric motif (II).

Recently, we have been able to crystallize the 1:1 complex of 1,4-diiodobenzene and 1,4-dinitrobenzene

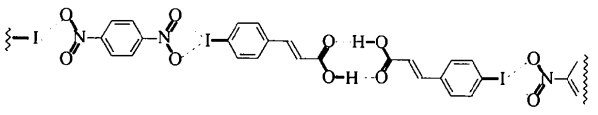




(IV)



(V)



(VI)

(IV), 4-iodonitrobenzene (V) and the 1:2 complex between 1,4-dinitrobenzene and 4-iodocinnamic acid (VI). Their crystal structures [(IV): Allen, Goud, Hoy, Howard & Desiraju, 1994; (V), (VI): Thalladi *et al.*, 1996] all show the perfectly symmetric [(IV): $D1 = D2 = 3.45\,(1)$ Å; (V): $D1 = D2 = 3.466\,(5)$ Å] or near symmetric [(VI): $D1 = 3.306\,(5)$, $D2 = 3.655\,(5)$ Å] iodo···O(nitro) interactions of motif (I). Whilst structures (IV) and (V) may have few other crystallization options, the iodo···O(nitro) synthon maintains its integrity in structure (VI) even in the presence of a much more robust synthon, *e.g.* the hydrogen-bonded carboxylic acid dimer.

These developments now make it appropriate to perform a more complete analysis of available crystallographic results for C—X···O_2N—C interactions and, additionally, to quantify the approximate strength of these interactions using *ab-initio*-based molecular orbital calculations. This combination of database analysis and molecular orbital calculations has already proved successful in studying the nature of short non-covalent interactions between carbon-bound Cl atoms and the O atoms of, particularly, carbonyl groups (Lommerse, Stone, Taylor & Allen, 1996; hereinafter referred to as LSTA96).

2. Methodology

2.1. *Database analysis*

Version 5.10 of the Cambridge Structural Database (October 1995; 146,272 entries) was used in this study. Searches for covalently bonded substructures and for synthons involving non-covalent X···O(nitro) interactions were carried out using the program *QUEST3D* (CSD, 1994). We note that the synthon structure may be supplied to the search program directly with all appropriate geometrical constraints specified. Statistical analyses and data visualizations were performed using *VISTA*2.0 (CSD, 1995). All substructure and synthon searches were restricted to those CSD entries that satisfied the following secondary search criteria: (*a*) error-free after CSD evaluation procedures, (*b*) no reported structural disorder, (*c*) organic compounds within CSD chemical class definitions, (*d*) perfect matching of their chemical and crystallographic connectivity representations and (*e*) $R \leq 0.075$. Results for the searches described below are collected in Table 1.

Initial searches (denoted *P* in Table 1) located those CSD entries that contained both the C—NO_2 and either C—Cl, C—Br or C—I fragments, *i.e.* structures which have the potential for containing the supramolecular synthons (I), (II) or (III). The total numbers of symmetry-independent C—NO_2 and C—X fragments were established in each case. These searches were further subdivided according to the dominant carbon types [Csp^3 or aromatic C_{ar}] observed in the general search for the C—NO_2 fragment.

Examples of the C—X···O_2N—C supramolecular synthon were located using the generalized non-covalent substructure depicted in Fig. 1. Here $D1$, the distance between the halogen and one of the nitro O atoms, was used as the primary constraint in the *QUEST3D* (CSD, 1994) intermolecular non-bonded search routine. Two sets of searches were performed for each halogen: set *A* (Table 1) located $D1 \leq v(O) + v(X) + 0.2$ Å, set *B* located $D1 \leq v(O) + v(X)$ Å, where the van der Waals radii v(O, Cl, Br, I) were taken as 1.52, 1.75, 1.85 and 1.98 Å, respectively (Bondi, 1964).

For each occurrence of the non-covalent fragment of Fig. 1, the O···X distances $D1$ and $D2$, the C—X···O1 angle (α) and the X···O1—N angle (β) were calculated. We also calculated the angle of elevation (χ) and the angle of rotation (φ) that characterize the approach of the halogen (X) to the C—N—O1 plane. In the construction of Fig. 1, φ is the valence angle N—O···X', where X' represents the projection of the halogen (X) atom onto the C—N—O1 plane. Positive values of φ were chosen to correspond to motifs (I) and (II) ($\varphi = 0$–$180°$), thus negative values correspond to motif (III) ($\varphi = 0$ to $-180°$). The general non-bonded searches considered any carbon type

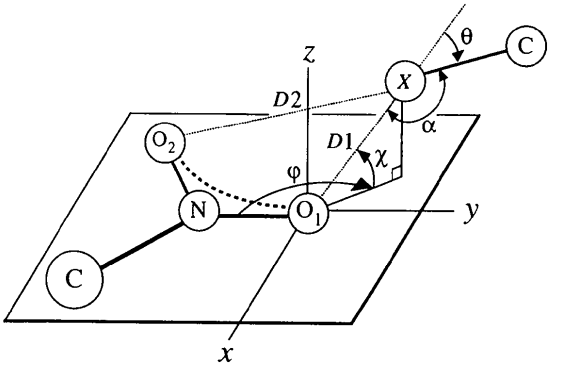


Fig. 1. Geometrical parameters used to describe C—NO_2···X—C interactions.

Table 1. *Summary of search results for halogen (X)···O(nitro) interactions*

(*a*) Searches (*P*) for potential interactions, *i.e.* C—*X* and C—NO_2 occur in the same crystal structure. C type is the carbon type in C—NO_2, N_e is the number of crystal structures located, N_X is the number of symmetry-independent halogen atoms in those structures and N_{nitro} is the number of symmetry-independent nitro groups.

		C—Cl and C—NO_2			C—Br and C—NO_2			C—I and C—NO_2		
Search	C type	N_e	N_{Cl}	N_{nitro}	N_e	N_{Br}	N_{nitro}	N_e	N_I	N_{nitro}
*P*1	Any	216	410	357	85	107	133	10	14	15
*P*2	Csp^3	37	59	75	31	46	51	4	5	4
*P*3	C_{ar}	157	317	251	45	50	69	5	5	7

(*b*) Non-bonded searches (*A*) and (*B*) for *X*···O(nitro) contacts within $v(O) + v(X) + 0.2$ Å (series *A* searches), or within $v(O) + v(X)$ Å (series *B* searches). C-type is the carbon type in C—NO_2, N_e is the number of crystal structures, N_f is the number of non-bonded fragments located in those structures, and N_X is the number of symmetry-independent halogen atoms involved in those fragments (see *Methodology*). The percentages are $100N_X$ (search *A* or *B*)/N_X (search *P*) and represent the frequency of formation of the *X*···O interaction under search criteria *A* or *B*.

		C—Cl···O(nitro)				C—Br···O(nitro)				C—I···O(nitro)			
Search	C type	N_e	N_f	N_{Cl}	%	N_e	N_f	N_{Br}	%	N_e	N_f	N_I	%
*A*1	Any	119	228	180	44	46	67	56	52	6	14	9	64
*A*2	Csp^3	26	48	37	63	18	35	26	56	1	1	1	20
*A*3	C_{ar}	83	163	130	41	21	24	21	42	3	3	3	60
*B*1	Any	75	96	93	23	30	34	34	32	4	8	7	50
*B*2	Csp^3	17	23	22	37	11	14	14	30	0	0	0	
*B*3	C_{ar}	51	65	63	20	12	12	12	24	2	2	2	40

in the C—NO_2 segment and, again, were repeated for the Csp^3— and C_{ar}—NO_2 subsets. In the latter case we also calculated the torsion angle (τ) that describes the conformation of the —NO_2 group with respect to the aromatic ring: $\tau = 0°$ for an in-plane nitro group, $\tau = 90°$ for perpendicular examples.

The exact or near-exact C_{2v} symmetry of interaction motif (I) causes problems in parts of the data analysis. We need to count the unique number of halogen (*X*) atoms that interact with nitro O atoms, but for motif (I) both *X*···O1 and *X*···O2 may satisfy the primary distance constraint. Thus, within CSD software, each occurrence of I would then generate two sets of the geometrical parameters described above and imply the involvement of two *X* atoms rather than one. True halogen counts were obtained through visual inspection of search hits, followed by manual editing of the geometrical outputs: only those hits for the shortest of *D*1 and *D*2 being retained.

2.2. *Molecular orbital calculations*

The intermolecular perturbation theory (IMPT) of Hayes & Stone (1984), as implemented within the *CADPAC*6.0 program package (Amos, 1996), was used to quantify the *X*(= Cl, Br)···O(nitro) interaction energies. For systems containing chlorine, 6-31G* basis set functions were used that contained polarization functions for the heavy atoms (C, N, O, Cl). For Br, only the [6*s*4*p*1*d*] basis set functions of Binning & Curtiss (1990) were available;† this basis set is a little larger than, but should be well balanced with, the 6-31G basis set descriptions used for other atoms in a particular molecule. The lack of polarization functions in these basis sets evidently will lead to an underestimation of the intermolecular polarization or induction energy. Simple model monomer molecules, whose choice is discussed later, were constructed and their geometries were optimized. Interacting dimers were then constructed having various orientations of the C—*X* vector with respect to the nitro group; intermolecular distances involving non-interacting atoms of the dimer were kept as long as possible. Application of the IMPT method within *CADPAC* (Amos, 1996) is fully described in LSTA96 and only a brief summary is given here.

An important feature of the IMPT method is that it calculates separate interaction energy terms which have distinct physical significance. Further, the sum of the significant interaction energy terms yields a total IMPT energy which is free of basis set superposition error (Stone, 1993). At first order, these separate terms are (*a*) E_{es}, the (attractive or repulsive) electrostatic energy that describes the classical Coulombic interaction, and (*b*) E_{er}: the exchange–repulsion term, the sum of an attractive part, due to the exchange of electrons of parallel spin, and a repulsive part, as a result of the Pauli exclusion principle which prevents electrons with parallel spins occupying the same region in space. At second order the IMPT gives (*c*) E_{pol}, the polarization or induction energy, (*d*) E_{ct}, the charge-transfer energy, and (*e*) E_{disp}, the dispersion energy term. We note that the dispersion energy is largely dependent on the size of the basis set used. Thus, even the 6-31G* basis set is too small to describe accurately the dynamic polarizability that underlies the dispersion energy and this results in an underestimation of E_{disp}. The lack of extra polarization functions in the 6-31G and [6*s*4*p*1*d*] basis sets used in our Br calculations worsens this situation. However

† The coefficient of the third primitive in the first *s* function should read 0.0087457 rather than 0.087457 (Stone & Amos, 1996).

(LSTA96; Stone, 1993), dispersion effects are only weakly orientation dependent and their underestimation does not affect the validity of the general conclusions derived in this work.

All database searches, data analyses and molecular orbital calculations were carried out on SUN or Silicon Graphics workstations on the CCDC Unix network.

3. Analysis of crystal structure results

3.1. *Frequency of formation of halogen(X)···O(nitro) interactions*

Table 1 shows that 531 carbon-bound halogen atoms (X) occur with C—NO_2 groups in crystal structures that fall within our secondary search criteria. Of these halogen atoms, 46% form $X\cdots$O(nitro) contacts within distance criterion A [$D1 \leq v(O) + v(X) + 0.2$ Å] and 25% form contacts within distance criterion B [$D1 \leq v(O) + v(X)$ Å]. This latter percentage represents a high frequency of formation for an interaction that is assumed to be energetically quite weak. Lines $A1$ and $B1$ of Table 1 show that the frequency of formation of $X\cdots$O interactions for individual halogens increases from Cl to Br to I at both distance limits A and B. This ordering was also observed by LSTA96 and parallels the order of polarizabilities of these elements.

Table 1 also shows that C_{ar}—NO_2 is the dominant carbon type and that the frequencies of formation of $X\cdots$O interactions within this subgroup (lines $A3$ and $B3$) closely mirror the overall results. Further, it appears that O atoms from Csp^3—NO_2 substructures are equally effective in forming $X\cdots$O interactions. Indeed, the data in lines $A2$ and $B2$ of Table 1, especially for X = Cl, might indicate that they are more effective, but the small numbers of observed interactions limit the statistical validity of this indication. This undiminished ability of O atoms from Csp^3—NO_2 substructures to form short non-covalent contacts to halogen atoms, by comparison with those from C_{ar}—NO_2 substructures, has important implications for the choice of model compounds for the IMPT calculations.

We also checked for any possible conformational dependency on the frequency of formation of $X\cdots$O interactions involving C_{ar}—NO_2 substructures. Of the 163 Cl···O interactions of $A3$ in Table 1, 101 have the NO_2 group coplanar with the aromatic ring, $0 < \tau < 20°$, 22 interactions have $20 < \tau < 70°$, while a further 40 form to perpendicular NO_2 groups, $70 < \tau < 90°$. The percentage of these NO_2-orientation classes that form the shortest Cl···O contacts (search B3, of Table 1) are 41, 41 and 33%, indicating that the formation of the strongest interactions is independent of the nitro-group conformation.

3.2. *Directionality of nitro O approach to X—C*

Initial histograms (not shown) of the C—Cl···O angle (α) from searches A and B indicated that nearly all values of $\alpha < 135°$ arise from Cl···O > [v(O) + v(Cl)]. The mean values of α were 133 (2)° for search A and 154 (2)° for search B. Using the angular definitions in Fig. 1, in which $\theta = 180 - \alpha$, we may generate a standard orthogonal plot of the distribution of O atoms around C—Cl, as shown in Fig. 2. However, in order to compare directly and visually the dispositions of O atoms relative to the C—X bond vector for all three halogens X = Cl, Br, I, we have used the geometrical constructs of LSTA96. A rapid and effective comparison of O···X—C geometries is then provided by scatterplots (Figs. 3*a*, *b* and *c*) of R^3 *versus* $(1 - \cos\theta)$, where R is the O···X distance normalized to the appropriate sum of van der Waals radii: $R = D1/[v(O) + v(X)]$. This normalization means that data points in all three plots that have $R^3 \leq 1.0$ correspond to values of $D1 \leq v(O) + v(X)$. However, the most important feature of the axial transformations is that equal areas of the plots of Figs. 3(*a*), (*b*) and (*c*) correspond to equal volumes in space: a uniform distribution of data points in 3D (three-dimensional) space is transformed into a uniform distribution of points in the two-dimensional scattergrams. All of the plots in Fig. 3 utilize the full number of $X\cdots$O interactions from search series $A1$ of Table 1.

Fig. 3 shows quite clearly that interpenetration of O and X atoms, $R^3 \leq 1.0$, occurs preferentially as $1 - \cos\theta$ approaches zero, *i.e.* $\theta \rightarrow 0°$ and $\alpha \rightarrow 180°$, indicating a strong preference for linear O···X—C geometries for all three halogens. The numbers of data points having $R^3 \leq 1.0$ and $1 - \cos\theta \leq 0.18$ ($\alpha \geq 145\alpha$) are 78, 30 and 7 for Cl, Br and I, respectively, or 81, 88 and 88% of the short O···X—C contacts located in search $B1$ of Table 1. While it is impossible for linear geometries to be maintained by either O···X—C interaction in the symmetrical motif (I), simple calculations using N—O = 1.20 Å, O—N—O = 120° and $D1 = D2 = v(O) + v(X)$ Å yield α values of 161, 162 and 163° for Cl, Br and I in this situation, where the C—X axis bisects the O—N—O angle. Experimental values in (IV) (Allen *et al.*, 1994)

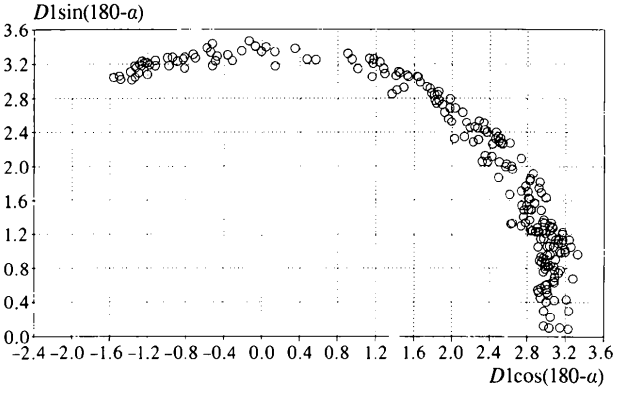


Fig. 2. Radial distribution of nitro O atoms around carbon-bound chlorine. The C—Cl bond vector is along the x axis of the plot, with Cl at the origin.

and (V) (Thalladi *et al.*, 1996) are 161.3 and 163.8°. Thus, the preferred O···*X*—C linearity is relatively little distorted in motif (I).

Fig. 3 gives the visual impression that the degree of interpenetration of O and *X* increases in the order Cl < Br < I. This impression is quantified in Table 2, which shows that the percentage of contacts having $R^3 \leq 0.85$ indeed increases in the above order, while the mean value of R^3 decreases in moving from Cl through Br to I. This ordering is identical to that observed in C=O···*X*—C contacts (LSTA96), where the numbers of short contacts were much higher for each halogen. There it was observed that the O···*X*—C angle (α) moved closer to 180° with increasing interpenetration of the O and *X*

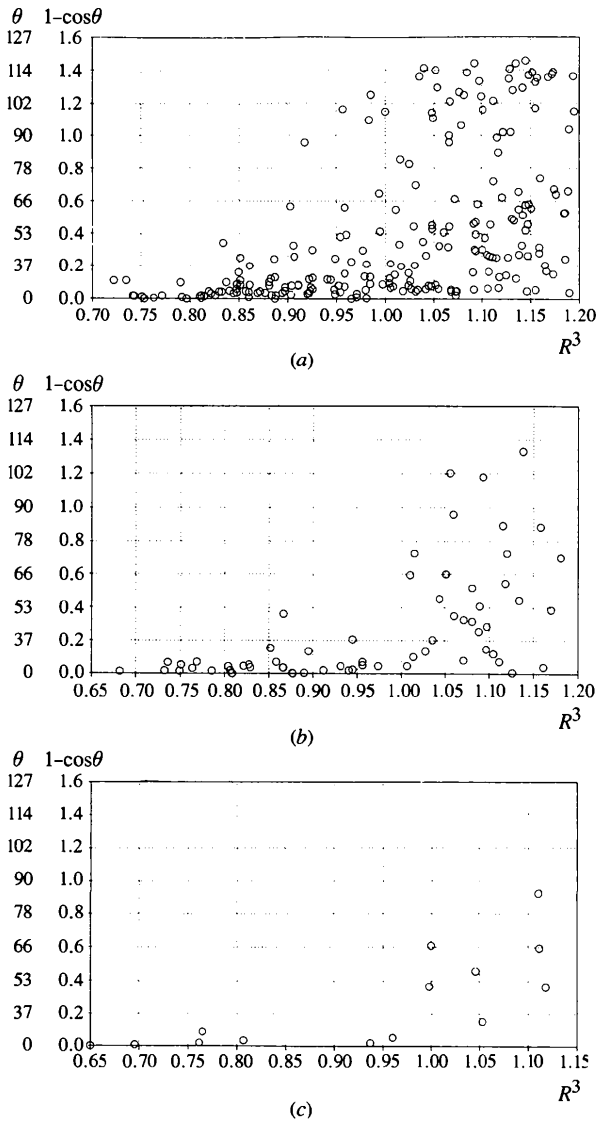


Fig. 3. Plots of $1 - \cos\theta$ *versus* R^3 (see text) for (*a*) C—Cl···O, (*b*) C—Br···O and (*c*) C—I···O interactions.

Table 2. *Statistics for X···O(nitro) interactions having* $R^3 \leq 0.85$

R is the value of *D*1 (Fig. 1*a*) normalized to the sum of van der Waals radii: $R = D1/[v(O) + v(X)]$, *N* is the number of contacts having $R^3 \leq 0.85$, with (%) expressing *N* as a percentage of N_t from search *B*1 of Table 1, *i.e.* as a percentage of contacts having $R^3 \leq 1.0$. Mean values of *D*1 and α are in Å and ° respectively, with e.s.d.'s of the mean value in parentheses.

	X = Cl	*X* = Br	*X* = I
N (%)	29 (30%)	17 (50%)	5 (63%)
$<R^3>$	0.803 (7)	0.770 (11)	0.736 (25)
$<R>$	0.929 (3)	0.916 (5)	0.902 (11)
$<D1>$	3.039 (9)	3.088 (15)	3.158 (40)
$<\alpha>$	164 (2)	167 (2)	168 (4)

atoms. Despite the small numbers of observations, the mean α values of Table 2 also reflect this effect.

The preference for carbon-bound halogens (*X*) to form shorter contacts to oxygen and other electronegative elements in the forward direction of the C—*X* vector (α = 180°), by comparison with the longer contacts formed perpendicular to this vector, is well known (see *e.g.* Nyburg & Faerman, 1985; Pedireddi *et al.*, 1994; Price, Stone, Lucas, Rowland & Thornley, 1994; Tsirelson, Zou, Tang & Bader, 1995). Debate has centred on whether the shorter contacts at α = 180° result from close packing of anisotropic atoms (the polar flattening effect) or from some specific attractive force. The recent IMPT calculations on C—Cl···O=C systems (LSTA96) indicate that the anisotropic electron distribution around Cl causes a decreased repulsive wall and an increased electrostatic attraction for electronegative atoms, in the forward direction. The crystallographic data for C—*X*···O(nitro) interactions are clearly in accord with this hypothesis.

3.3. *Directionality of halogen approach to O(nitro)*

The spherical polar angles χ and φ (Fig. 1) describe the orientation of the *X*···O vector with respect to the C—NO_2 plane. The sign of χ is not discriminatory in this study, hence absolute values $|\chi|$ are used throughout. However, the sign of φ is important: φ is measured in-plane and relative to the unique N—O1 vector, as established by the search criteria. Thus (Fig. 1), positive φ values place the incoming halogen *trans* to the C atom about the N—O1 bond, corresponding to motifs (I) and (II), while negative φ values place the halogen in a *cis*-relationship to C as in motif (III).

Circular scatterplots of φ *versus* $(1 - \sin|\chi|)^{1/2}$ are presented in Figs. 4(*a*), (*b*) and (*c*) for *X* = Cl, Br, I, respectively. Only the symmetry-independent *X*···O interactions from search series *B*1 of Table 1 are included in Fig. 4, as discussed in the *Methodology*. The use of $(1 - \sin|\chi|)^{1/2}$ generates an equal-area projection: equal areas on the sphere of *X* approach to the nitro group map to equal areas on the plane of projection depicted in Fig. 4. Our derivation is based on the Lambert, or Schmidt,

projection of spherical data (see Fisher, Lewis & Embleton, 1987). Consider the surface of a hemisphere, radius R_s, that has the nitro group in the equatorial plane, is centred on the unique oxygen O1 of Fig. 1 and in which χ has positive values. A latitudinal strip of width $d\chi$ then has a surface area of $2\pi R_s 2\cos\chi d\chi$, which is the $\cos\chi d\chi$ part of the surface area of the full hemisphere. Integrating for this part from $\chi = 0$ to $\chi = \chi$ gives $\sin\chi$. If we now take a circle of radius R_c, then a concentric circular strip at distance $r = xR_c$ from the centre and having a width $d(xR_c)$ has a surface area of of $2\pi r dr = 2\pi R_c^2 x\, dx$, or the $2xdx$ part of the area of the full circle. Integrating from $x = 0$ to $x = x$ gives x^2. Therefore, if we use $x \propto (\sin\chi)^{1/2}$, then normalized equal surface areas on the hemisphere ($\chi = 0 \rightarrow \chi$) correspond to equal areas on the circle ($x = 0 \rightarrow x$). Similarly, to project the observations inwards from the circumference of the circle, *i.e.* from $x = 1$ to 0, we can integrate from $x = x$ to $x = 1$ to give $1 - x^2$. Thus, $x \propto (1 - \sin\chi)^{1/2}$ and, in Fig. 4, data points having $|\chi|$ close to zero will project onto the circumference of the circular scatterplot, while those with $|\chi|$ close to 90° project close to its centre. Hence, Fig. 4 provides an equal-area projection, onto the equatorial plane, of the *X*-approaches to nitro O as viewed along the *z* axis of Fig. 1.

Fig. 4(*a*) shows a clear preference for Cl to approach O1 in the mono-coordinate mode of motif (III), with a much smaller peak at $\varphi \simeq 110$–120°, which might be considered to arise from symmetric or near-symmetric interactions of Cl with both nitro-O atoms. However, examination of the results of search A1 (Table 1) shows that only 11 interactions have the second Cl···O contact, *D*2, in Fig. 1, less than the search limit of 3.47 Å, and that these contacts all have $\varphi \leq 103°$. Such an φ-limit is a

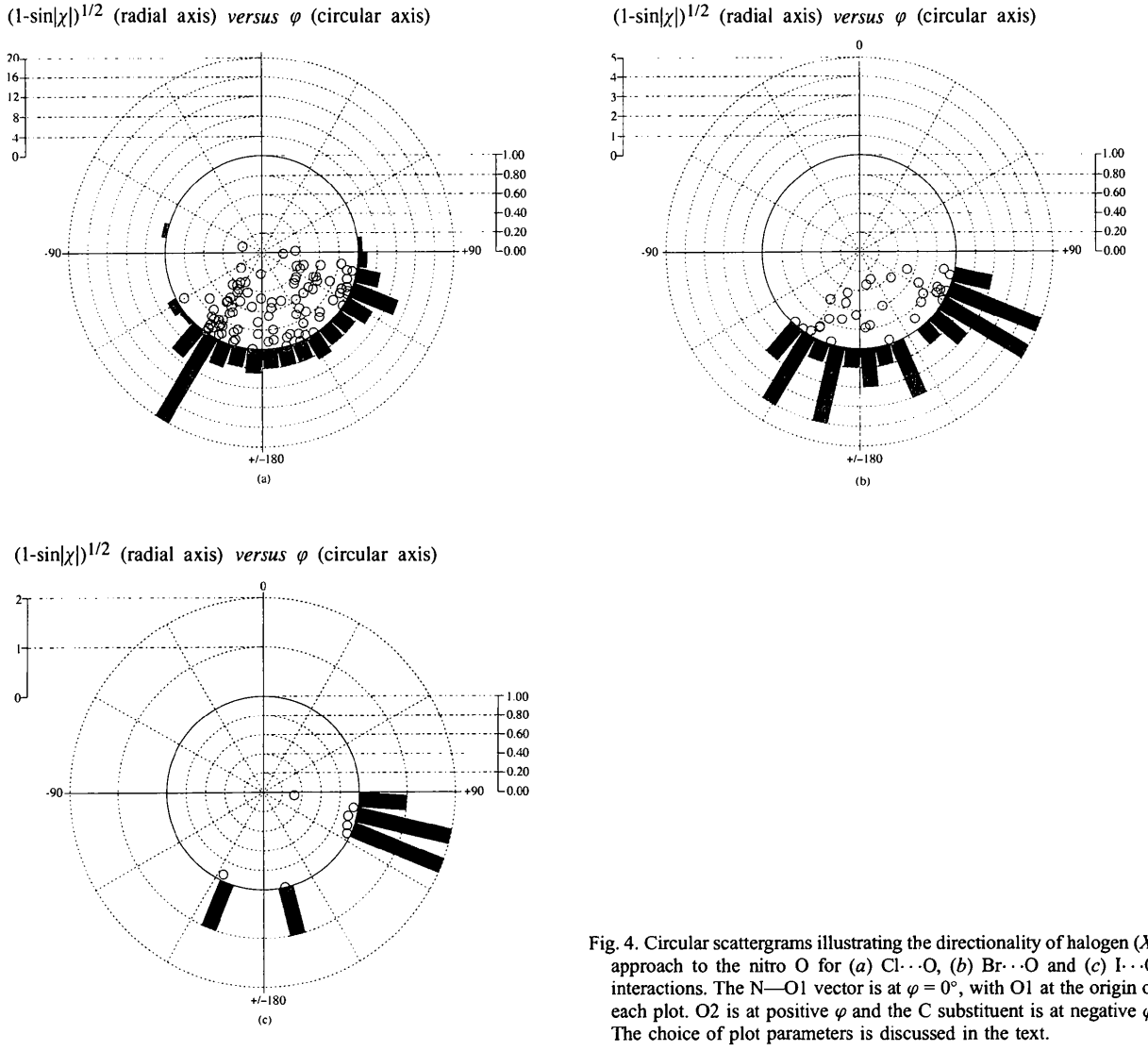


Fig. 4. Circular scattergrams illustrating the directionality of halogen (*X*) approach to the nitro O for (*a*) Cl···O, (*b*) Br···O and (*c*) I···O interactions. The N—O1 vector is at $\varphi = 0°$, with O1 at the origin of each plot. O2 is at positive φ and the C substituent is at negative φ. The choice of plot parameters is discussed in the text.

natural geometrical constraint: simple calculations (based on N—O = 1.20 Å, O—N—O = 120° and van der Waals radii sums for O and *X*) yield N—O···*X* angles of 101–104° for symmetrical in-plane approach of *X* to the two nitro O atoms. The N—O···*X* angle is, of course, equivalent to φ when *X* is coplanar with the C—NO_2 system. Thus, the majority of Cl···O contacts at positive φ show a clear preference for a Cl interaction with just one of the nitro O atoms, as in the asymmetric motif (II).

The overall φ distribution changes significantly for Br (Fig. 4*b*), where the motif (III) peak is clearly less dominant than for Cl, and the approach of Br between the two nitro O atoms is now preferred. However, in this latter mode there is again a clear preference for asymmetric motif (II) interactions: only three of the approaching Br atoms have $D2 \leq 3.57$ Å (the limit in search *A*1 of Table 1) and have a maximum φ of 104°. It is only for iodine (Fig. 4*c*) that we observe a clear preference for the symmetrical motif (I): all five of the I···O interactions having φ in the range 90–116° also have $D2 \leq 3.70$ Å (the search limit) and three of these have the secondary contact, *D*2, less than the sum of van der Waals radii (3.50 Å). It seems that the increased size of I enables this halogen to make attractive contacts to both nitro O atoms, whereas for the smaller Cl and Br atoms a strong interaction with one nitro-O is energetically preferable to a pair of longer and weaker *X*···O interactions.

The results of Fig. 4 show that motifs (I) and (II) are arbitrary subdivisions of a broad continuum of positive φ values in the range 90–150°. Hence, it is appropriate to describe this continuum as arising from 'bifurcated' interactions, whereas those interactions in the φ range from −175 to −120° are clear examples of the mono-coordinate motif (III). Within these broad definitions, the frequency of mono-coordinate interactions falls in the order Cl (46%) > Br (35%) > I (25%), while the frequency of bifurcated interactions rises in the order Cl (35%) < Br (44%) < I (63%).

The plots of Fig. 4 also indicate that, whilst a significant number of Cl atoms approach the C—NO_2 plane at angles $|\chi| > 30°$ [$(1 - \sin|\chi|)^{1/2} > 0.71$], there is an increasing preference for the in-plane approach of Br and, especially, I. This preference can be quantified by the decreasing mean values of $|\chi|$ and by the increasing percentage of contacts from search series *B*1 (Table 1) that have $|\chi| < 30°$; these data pairs for Cl, Br and I are: 27 (2)°, 58%; 23 (3)°, 71%; 14 (7)°, 88%.

The data of Fig. 4 might be taken to imply some degree of lone-pair directionality in the *X*···O(nitro) interactions. However, the simple calculations above show that N—O···*X* angles are sterically constrained to be close to the expected positions of conventional O lone pairs in the broad range of bifurcated interactions. Further, the φ peak for the mono-coordinate motif (III) appears consistently at *ca* −150 to −160°, far from the conventional O lone-pair position. Steric interactions between the incoming halogen and other C substituents provide a simple explanation for this shift towards a more linear N—O···*X* arrangement in motif (III). Since these steric effects will increase with increasing size of the halogen, they may be partly responsible for the increasing preference of Br and I to approach the nitro O atoms in the bifurcated mode. In their study of C—*X*···O=C interactions LSTA96 found only very limited crystallographic and computational evidence for lone-pair directionality of *X* approach to O, despite the steric accessibility of both lone pairs in their case.

4. Molecular orbital calculations

4.1. *Choice of model compounds*

The simplest model system is the nitromethane–chloromethane dimer. However, it was found (LSTA96) that chloromethane was not a good model, since the computed Cl···O=C interaction was repulsive. Their conclusion stressed that the Cl atom should be in an electron-withdrawing molecular environment to promote a favourable attractive interaction. Examination of the 96 C—Cl···O(nitro) fragments from search *B*1 (Table 1) showed that 48% had C = C_{ar}, 25% had C = $C_{ethylenic}$, while only 27% arose from C = Csp^3 substituents. This observation reinforces the importance of an electron-withdrawing environment for Cl. The obvious choice of model compound, chlorobenzene, unfortunately imposes significant computational overheads in the IMPT procedure, hence we adopted the compromise employed by LSTA96 and selected chloroacetylene as the model. However, even this choice was immediately modified since the highly acidic Csp^1—H formed the most favourable dimers with O(nitro) within the IMPT procedure. This proton was replaced by a methyl group in all subsequent calculations and the final model system: nitromethane–1-chloro-2-methylacetylene is depicted in (VII). The use of nitromethane rather than, say, nitrobenzene, has already been justified above. The parameters used to describe the dimer geometry are those of Fig. 1.

4.2. *IMPT calculations*

The two individual molecules of (VII) were geometry optimized at the 6-31G* basis set level and interaction energies (Table 3, Fig. 5) were computed for a variety of fixed geometries for the dimer. The C—Cl···O angle (α) was kept fixed at 180° in all calculations. Each fixed-geometry IMPT calculation within *CADPAC*6.0 (Amos, 1996) took approximately 3 h c.p.u. time on a Silicon Graphics Indigo-2 workstation. Four sets of calculations were carried out.

4.2.1. *IMPT1*. Interaction energies (Fig. 5*a*) were calculated at 15 or 20° clockwise increments in φ from +100 to −100°, with *D*1 fixed at 3.27 Å (sum of van der

Table 3. *Interaction energies from IMPT calculations*

All energies are in kJ mol^{-1}, angles are in ° and distances are in Å.

Variable	E_{es}	E_{cr}	E_{pol}	E_{ct}	E_{disp}	E_t
(*a*) Cl···O: Variable φ, 6-31G* basis set						
100	−2.17	5.01	−0.80	−0.55	−3.65	−2.15
120	−2.44	2.33	−0.63	−0.33	−2.73	−3.80
135	−2.53	2.04	−0.58	−0.27	−2.60	−3.94
150	−2.39	1.83	−0.52	−0.23	−2.59	−3.90
165	−2.25	1.66	−0.46	−0.19	−2.63	−3.86
180	−2.29	1.60	−0.41	−0.17	−2.67	−3.94
−165	−2.60	1.66	−0.38	−0.18	−2.72	−4.21
−150	−3.16	1.85	−0.35	−0.20	−2.82	−4.68
−135	−3.86	2.22	−0.33	−0.25	−3.08	−5.30
−120	−4.80	3.80	−0.33	−0.35	−3.83	−5.50
−100	−13.67	33.27	−0.60	−1.82	−7.97	9.12
(*b*) Cl···O: Variable *D*1, 6-31G* basis set						
3.4	−3.10	1.27	−0.25	−0.16	−2.54	−4.78
3.3	−3.66	1.96	−0.31	−0.22	−2.95	−5.18
3.2	−4.39	2.99	−0.38	−0.32	−3.44	−5.53
3.1	−5.34	4.56	−0.48	−0.45	−4.03	−5.74
3.0	−6.62	6.92	−0.60	−0.66	−4.75	−5.70
2.9	−8.37	10.48	−0.76	−0.95	−5.65	−5.25
2.8	−10.81	15.80	−0.98	−1.38	−6.78	−4.15
(*c*) Cl···O: Variable χ, 6-31G* basis set						
0	−3.86	2.22	−0.33	−0.25	−3.08	−5.30
15	−3.83	2.25	−0.32	−0.25	−3.11	−5.26
30	−3.74	2.36	−0.31	−0.27	−3.18	−5.14
45	−3.55	2.53	−0.29	−0.30	−3.30	−4.91
60	−3.23	2.77	−0.28	−0.32	−3.45	−4.51
75	−2.69	3.06	−0.26	−0.34	−3.65	−3.89
90	−1.87	3.60	−0.26	−0.34	−4.04	−2.92
(*d*) Br···O: Variable φ, 6-31G basis set						
100	−3.25	5.05	−0.85	−0.73	−1.89	−1.67
120	−4.07	2.07	−0.72	−0.46	−1.63	−4.81
135	−4.44	1.75	−0.70	−0.39	−1.74	−5.51
150	−4.44	1.54	−0.66	−0.32	−1.88	−5.76
165	−4.33	1.38	−0.61	−0.28	−1.99	−5.83
180	−4.36	1.33	−0.56	−0.27	−2.06	−5.91
−165	−4.61	1.39	−0.51	−0.28	−2.08	−6.09
−150	−5.05	1.57	−0.45	−0.31	−2.11	−6.35
−135	−5.55	1.942	−0.38	−0.35	−2.21	−6.54
−120	−6.13	3.69	−0.32	−0.46	−2.71	−5.93
−105	−9.95	18.66	−0.44	−1.20	−4.68	2.40
(*e*) Cl···O: Variable φ, 6-31G basis set						
100	0.08	4.84	−0.78	−0.45	−1.83	1.86
120	−0.73	2.09	−0.64	−0.26	−1.60	−1.15
135	−1.04	1.78	−0.61	−0.21	−1.70	−1.79
150	−1.06	1.56	−0.56	−0.18	−1.84	−2.09
165	−1.07	1.39	−0.51	−0.16	−1.96	−2.29
180	−1.22	1.33	−0.45	−0.15	−2.02	−2.51
−165	−1.63	1.40	−0.41	−0.16	−2.05	−2.85
−150	−2.30	1.59	−0.36	−0.19	−2.07	−3.32
−135	−3.19	2.03	−0.31	−0.25	−2.17	−3.90
−120	−4.61	4.18	−0.28	−0.43	−2.64	−3.78
−105	−10.58	22.80	−0.40	−1.67	−4.54	5.60

Waals radii for Cl and O) and $|\chi|$ fixed at 0° (Cl approach in the C—NO_2 plane).

4.2.2. *IMPT2.* Interaction energies (Fig. 5*b*) were calculated at 0.1 Å increments of *D*1 from 2.8 to 3.4 Å, with $|\chi|$ fixed at 0° and φ fixed at −135° and the position of the energy minimum obtained in IMPT1 above.

4.2.3. *IMPT3.* Interaction energies (Fig. 5*c*) were calculated at 15° increments in $|\chi|$, with *D*1 fixed at 3.27 Å and φ fixed at −135°.

4.2.4. *IMPT4.* Interaction energies (Fig. 5*d*) were calculated for φ increments exactly as for IMPT1, but with 1-bromo-2-methylacetylene as the model and the fixed *D*1 now at 3.37 Å. The [6*s*4*p*1*d*] basis set (Binning & Curtiss, 1990) was used for Br and 6-31G basis sets for other atoms. The IMPT1 calculations for Cl were rerun at the 6-31G basis set level (Fig. 5*e*) to provide a close comparison with the Br results.

4.3. *Energy variation with φ (IMPT1, Fig. 5a)*

There is a small decrease in the total energy E_t as φ increases from 90°. This is almost entirely due to a reduction in E_{er}, a reduction that is partially offset by increasing E_{disp}. However, as φ increases further, E_{er} also begins to rise; the incoming Cl atom begins to make unfavourable contacts with the methyl protons. Beyond $\varphi \simeq -125°$, the repulsive E_{er} becomes totally dominant, despite reductions in the attractive E_{es} and E_{disp} terms. A φ value of −135° was selected as representing the minimum interaction energy in the φ scan and was used as a fixed value in the IMPT2 and IMPT3 calculations.

The absence of a discrete minimum in E_t or E_{es} at $\varphi \simeq 120°$ shows that O lone-pair involvement is not significant. Indeed, the approximately constant low density of crystallographic observations in the clockwise φ range +100 to −160° (Fig. 4*a*) correlates well with E_t in Fig. 5(*a*). Further, the decrease in E_t at $\varphi \simeq -135°$ is well correlated with the major crystallographic peak of Fig. 4(*a*) ($\varphi \simeq -150$ to −140°). That there is a small difference in these two minima is merely an artefact of the computational model, which has only small H substituents on the C—Cl carbon atom, by comparison with the varied C substitution patterns that exist in the crystal structure data and, hence, increase the repulsive interaction as φ moves past −150°.

4.4. *Energy variation with D1 (IMPT2, Fig. 5b)*

There is a very shallow dip in E_t as Cl penetrates the O sphere. The minimum value occurs at *ca* 3.1 Å, which is 0.17 Å shorter than the sum of van der Waals radii. Although E_{es} and E_{disp} become more attractive as *D*1 decreases below 3.0 Å, this is compensated by rapidly increasing E_{er}. In the crystallographic observations of *D*1 (search *B*1, Table 1), there are only eight examples of *D*1 < 3.0 Å and 32 < 3.1 Å. The remaining 64 instances are evenly distributed up to the *D*1 search limit of 3.27 Å.

4.5. *Energy variation with χ (IMPT3, Fig. 5c)*

Fig. 4(*a*) indicates that, although lower values of $|\chi|$ are preferred in crystal structures, there is, nevertheless, quite a broad overall $|\chi|$ distribution. The IMPT results indicate

that this is to be expected, since E_t is almost constant over the $|\chi|$ range 0–45° and 74 of the 96 crystallographic substructures (77%) of search *B*1 (Table 1) fall into this broad range. For $|\chi| > 45°$, E_{es} and E_{er} begin to rise, albeit slowly, to yield a maximum E_t at $|\chi| = 90°$. Again, these results indicate an insignificant involvement of the O lone pairs in the Cl···O interaction.

4.6. *Overall Cl···O(nitro) interaction energies*

Given the level of computational intensity, it is not appropriate to use IMPT calculations to search for local or global energy minima. Rather, we use the crystallographic data to indicate appropriate areas of parameter space for investigation by IMPT and refine these indications as the calculations proceed. Under these

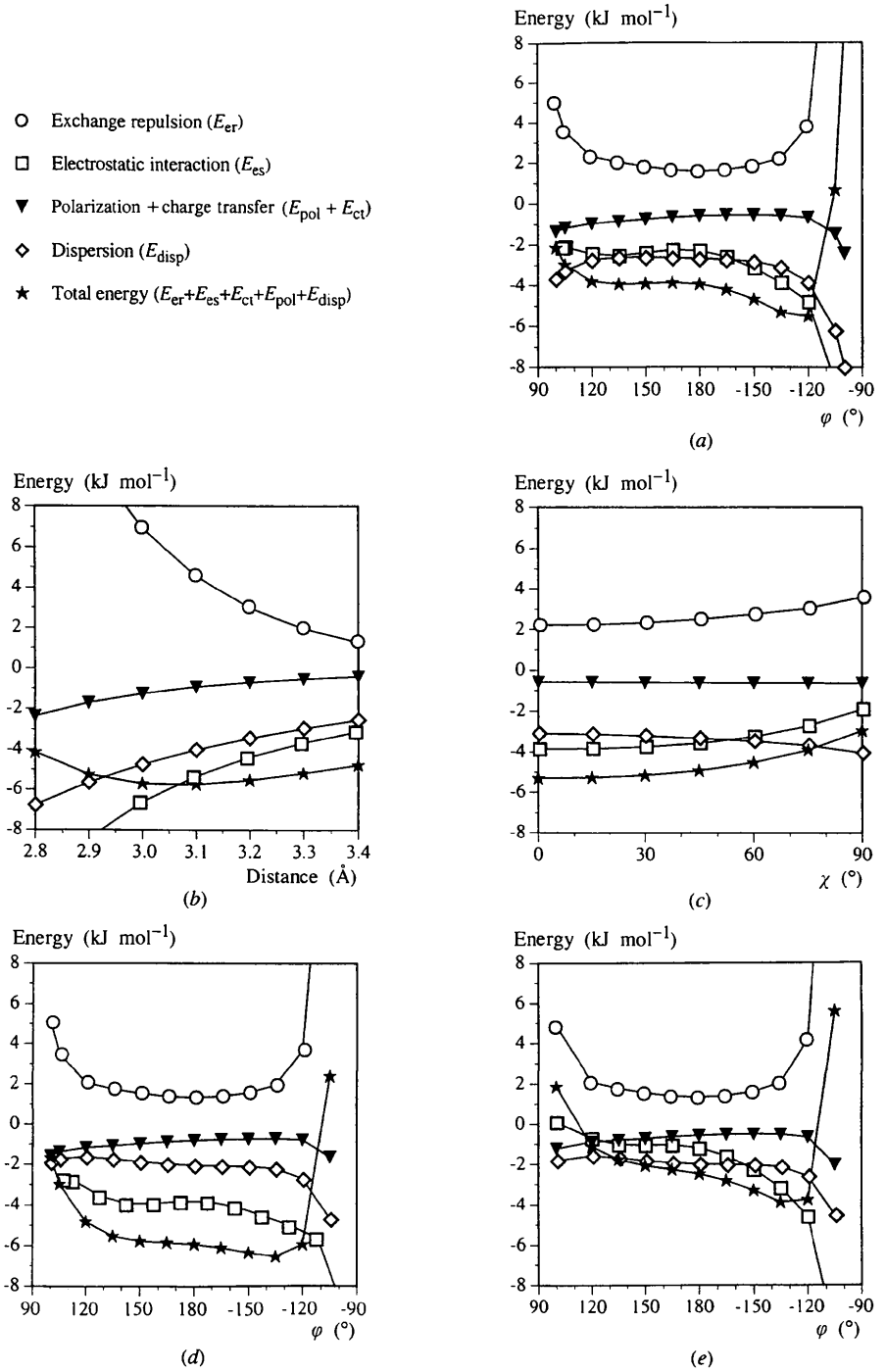


Fig. 5. Interaction energies (kJ mol^{-1}) calculated for model dimers using the IMPT method. Frames (*a*)–(*c*) are for the nitromethane···1-chloro-2-methylacetylene model using 6-31G* basis sets and show energy variations with changes in (*a*) φ, (*b*) *D*1 and (*c*) χ. Frame (*d*) shows the energy variation with φ for the nitromethane···1-bromo-2-methylacetylene dimer, using the [6*s*4*p*1*d*] basis set for Br and 6-31G basis sets for other atoms. Frame (*e*) shows comparative φ results for the 'chloro' dimer, but using 6-31G basis sets. Since the division between E_{pol} and E_{ct} is not completely basis-set independent, and their values are small, these energy contributions have been combined in the plots.

conditions, the minimum value of E_t for Cl···O(nitro) interactions in Table 3 is the -5.74 kJ mol^{-1} that occurs at $\chi = 0°$, $\varphi = -135°$ and $D1 = 3.1$ Å. This attractive energy is about 0.20–0.25 of the strength of a strong hydrogen bond involving O or N acceptors and O—H or N—H donors, but is similar in strength to a typical C—H···O bond (see *e.g.* Jeffrey & Saenger, 1991; Desiraju, 1996; for typical energy values). The Cl···O(nitro) interaction energy is certainly less attractive than the -10 kJ mol^{-1} obtained by LSTA96 for Cl···O = C interactions, although they used an even stronger electron withdrawing environment for Cl in their calculations. Nevertheless, it would be reasonable to presume that nitro O atoms are weaker hydrogen-bond acceptors than *e.g.* keto O atoms and this point is currently under investigation (Baalham, 1996; Allen *et al.*, 1997) using the combined crystallographic and computational IMPT approach employed here.

4.7. *IMPT comparison of Br···O(nitro) and Cl···O(nitro) interactions*

Calculated energies for the φ scans of IMPT4 for the Br···O(nitro) interaction are shown in Fig. 5(*d*) and listed in Table 3(*d*). These data are generated at the [6*s*4*p*1*d*]/6-31G basis set level and while they may be compared with the 6-31G* φ scan results for Cl (Fig. 5*a*), a more direct comparison is afforded by Fig. 5(*e*) and Table 3(*e*), which contain IMPT φ scan results for the Cl···O(nitro) interaction that have been rerun at the 6-31G basis set level.

For Br, the minimum in E_t occurs at *ca* $-135°$, the same φ value as for Cl at both basis set levels. However, the actual [6*s*4*p*1*d*]/6-31G-based energy levels are more attractive for Br (minimum $E_t = -6.54$ kJ mol^{-1}) than they are for Cl even at the 6-31G* level (min. $E_t = -5.74$ kJ mol^{-1}) and significantly more attractive than for the directly comparable 6-31G-based Cl results (min $E_t = -3.90$ kJ mol^{-1}). If we consider the known underestimation of the attractive E_{disp} component at the 6-31G basis level, apparent from a comparison of Tables 3(*a*) and 3(*e*), then we can predict that the minimum E_t for the Br···O(nitro) interaction would be significantly lower if calculations at the 6-31G* level were possible with Br, perhaps as low as *ca* -8.0 kJ mol^{-1}.

Importantly, visual comparisons of Figs. 5(*d*) (Br) and 5(*e*) (Cl) indicate a greater difference between E_t ($\varphi = 120°$) and E_t ($\varphi = -135°$) for Cl than for Br. These differences can be quantified from Tables 3(*d*) and 3(*e*) as 2.8 kJ mol^{-1} for Cl and 1.7 kJ mol^{-1} for Br. Thus, one might expect Br to form more bifurcated ($\varphi \simeq 120°$) interactions than Cl, a fact that is apparent from the crystallographic results depicted in Figs. 4(*a*) and (*b*). Given these trends, we would tentatively predict that for iodine we might obtain an I···O(nitro) interaction energy of *ca* -10 kJ mol^{-1} at the 6-31G* basis set level and an energy difference between $\varphi = 120$ and $\varphi = -135°$ that is even more favourable to the formation of true bifurcated I···O interactions.

5. Conclusion

Crystallographic and computational results for C—*X*···O(nitro)—C interactions are in excellent agreement at a very detailed level. The essential conclusions of the combined study are: (*a*) the C—*X*···O angle prefers to be linear, (*b*) the angle of *X* approach to the C—NO_2 plane is preferentially less than 45°, with an increasing tendency to planarity in the order Cl < Br < I, (*c*) within the C—NO_2 plane, *X* either forms a monocoordinate interaction with one nitro O atom or approaches both nitro O atoms in a bifurcated manner; the tendency to form bifurcated motifs increases in the order Cl < Br < I, with only I consistently forming short interactions to both nitro O atoms to yield a symmetrical bifurcated motif and (*d*) the IMPT interaction energies for *X*···O are in the range -6 to -10 kJ mol^{-1} and almost certainly in the order Cl < Br < I, thus representing interactions that are comparable in strength to C—H···O hydrogen bonds.

The results obtained here provide a clear rationale for the importance of the I···O(nitro) interaction observed in our recent crystal engineering experiments (Allen *et al.*, 1994; Thalladi *et al.*, 1996) and for the classification of this interaction pattern as a robust and important supramolecular synthon (Desiraju, 1995). Finally, this work provides a further illustration of the value of combining crystallographic database analyses with *ab-initio*-based molecular orbital calculations in the detailed study of molecular recognition processes.

We thank Drs Anthony Stone and Roger Amos (University of Cambridge) for access to CADPAC6.0 and for assistance and advice on the IMPT method. We also thank the European Communities for financial support to JPML under the Training and Mobility of Researchers Programme, contract number ERB-CHR-XT-94-0469 (Molecular Recognition Network).

References

Allen, F. H., Baalham, C. A., Lommerse, J. P. M., Raithby, P. R. & Sparr, E. (1997). *Acta Cryst.* **B53**, 1017–1024.

Allen, F. H., Davies, J. E., Galloy, J. J., Johnson, O., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M. & Watson, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187–204.

Allen, F. H., Goud, B. S., Hoy, V. J., Howard, J. A. K. & Desiraju, G. R. (1994). *Chem. Commun.* pp. 2729–2730.

Amos, R. D. (1996). *CADPAC6.0. The Cambridge Analytical Derivatives Package. Issue 6.0. A Suite of Quantum Chemistry Programs.* University of Cambridge, England.

Baalham, C. A. (1996). M.Phil. Thesis. University of Cambridge, England.

Binning, R. C. & Curtiss, L. A. (1990). *J. Comput. Chem.* **11**, 1206–1216.

Bondi, A. (1964). *J. Phys. Chem.* **68**, 441–451.

Cambridge Structural Database (1994). *User's Manual. Getting Started with the CSD System.* Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.

Cambridge Structural Database (1995). *VISTA*2.0 *User's Manual.* Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.

Desiraju, G. R. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 2311–2327.

Desiraju, G. R. (1996). *Acc. Chem. Res.* **29**, 441–449.

Desiraju, G. R. & Pedireddi, V. R. (1989). *Chem. Commun.* pp. 1112–1113.

Desiraju, G. R., Pedireddi, V. R., Sarma, J. A. R. P. & Zacharias, D. E. (1993). *Acta Chim. Hung.* **130**, 451–465.

Fisher, N. I., Lewis, T. & Embleton, B. J. J. (1987). *Statistical Analysis of Spherical Data.* Cambridge University Press.

Hayes, I. C. & Stone, A. J. (1984). *J. Mol. Phys.* **53**, 83–105.

Jeffrey, G. A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Systems.* Berlin: Springer-Verlag.

Lommerse, J. P. M., Stone, A. J., Taylor, R. & Allen, F. H. (1996). *J. Am. Chem. Soc.* **118**, 3108–3116.

Nyburg, S. C. & Faerman, C. H. (1985). *Acta Cryst.* B**41**, 274–282.

Pedireddi, V. R., Reddy, D. S., Goud, B. S., Craig, D. C., Rae, A. D. & Desiraju, G. R. (1994). *J. Chem. Soc. Perkin Trans. 2*, pp. 2353–2360.

Pedireddi, V. R., Sarma, J. A. R. P. & Desiraju, G. R. (1992). *J. Chem. Soc. Perkin Trans. 2*, pp. 311–315.

Price, S. L., Stone, A. J., Lucas, J., Rowland, R. S. & Thornley, A. E. (1994). *J. Am. Chem. Soc.* **116**, 4910–4918.

Stone, A. J. (1993). *Chem. Phys. Lett.* **211**, 401–409.

Stone, A. J. & Amos, R. D. (1996). Personal Communication.

Thalladi, V. R., Goud, B. S., Hoy, V. J., Allen, F. H., Howard, J. A. K. & Desiraju, G. R. (1996). *Chem. Commun.* pp. 401–402.

Tsirelson, V. G., Zou, P. F., Tang, T. H. & Bader, R. F. W. (1995). *Acta Cryst.* A**51**, 143–153.