conformations persist in solution. Diastereomer (II) has no hydrogen bonds and packs with no intermolecular distances less than van der Waals distances (Fig. 4*a*).

Unlike (II), the acid form of (I) has an acid H atom available for hydrogen bonding. This H atom makes a hydrogen bond to the ether O(11) atom, as illustrated in Fig. 4(*b*). The O(31)···O(11) distance is 2.776(9), O(31)—H(31) = 1.24(6), H(31)···O(11) = 1.59(6) Å, and the O(31)—H(31)···O(11) angle is 156.7(3.5)° (symmetry code: $\frac{3}{2}-x, \frac{1}{2}+y, 1-z$). Although an ether O atom is a less commonly observed hydrogen-bond acceptor than a hydroxyl or ketone group, the O(31)···O(11) distance and O(31)—H(31)···O(11) angle is equal to the mean $O_d\cdots O_a$ distance of 2.7667 ± 0.0700 Å and close to the mean $O_d—H\cdots O_a$ angle of 167.11 ± 6.53° observed from neutron diffraction data for 74 linear hydrogen bonds, nine involving a $C—O_a—C$ acceptor (Ceccarelli, Jeffrey & Taylor, 1981).

References

CECCARELLI, C., JEFFREY, G. A. & TAYLOR, R. (1981). *J. Mol. Struct.* **70**, 255–271.

KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.

KLAYMAN, D. L. (1985). *Science*, **228**, 1049–1055.

LIN, A. J., LI, L., KLAYMAN, D. L., GEORGE, C. F. & FLIPPEN-ANDERSON, J. L. (1990). *J. Med. Chem.* **33**, 2610–2614.

LUO, X., YEH, H. J. C., BROSSI, A., FLIPPEN-ANDERSON, J. L. & GILARDI, R. (1984). *Helv. Chim. Acta*, **67**, 1515–1522.

Qinghaosu Research Group (1980). *Sci. Sin. (Engl. Ed.)* **23**, 380–396.

SHELDRICK, G. M. (1985). *SHELXS86. Crystallographic Computing* 3, edited by G. M. SHELDRICK, C. KRÜGER & R. GODDARD, pp. 175–189. Oxford Univ. Press.

SHELDRICK, G. M. (1990). *SHELXTL-Plus88 Structure Determination Software Programs*. Nicolet Instrument Corporation, Madison, Wisconsin, USA.

Siemens (1989*a*). *P3/PC Diffractometer Program*. Version 3.13. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.

Siemens (1989*b*). *XDISK. Data Reduction Program*. Version 3.11. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.

Acta Cryst. (1995). B**51**, 1068–1081

Correlation of the Hydrogen-Bond Acceptor Properties of Nitrogen with the Geometry of the Nsp^2→Nsp^3 Transition in $R_1(X{=})C—NR_2R_3$ Substructures: Reaction Pathway for the Protonation of Nitrogen

BY FRANK H. ALLEN,* CRAIG M. BIRD AND R. SCOTT ROWLAND

Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England

AND STEPHANIE E. HARRIS AND CARL H. SCHWALBE

Pharmaceutical Sciences Institute, Aston University, Aston Triangle, Birmingham B4 7ET, England

(*Received* 24 *February* 1995; *accepted* 4 *April* 1995)

Abstract

The non-bonded N···H interactions between N atoms of $R_1(X{=})C—N(R_2R_3)$ (R_2,R_3 = Csp^3 or H) substructures and the H atoms of N—H and O—H donors have been analysed using crystallographic data and *ab initio* molecular orbital calculations. A total of 946 contacts having N···H (d_{NH}) ≤ 2.75 Å were retrieved from the Cambridge Structural Database. For the preferred all-planar (Nsp^2) conformation, d_{NH} ranges upwards from *ca* 2.5 Å and H approaches N approximately perpendicular to the plane. However, it is shown that close intramolecular steric interactions lead to major geometrical distortions [pyramidalization at N and rotation about the C—N bond: Ferretti, Bertolasi, Gilli & Gilli (1993). *J. Phys. Chem.* **49**, 13568–13574]. The N atom undergoes a transition from sp^2 to sp^3 with gradual lone-pair formation on N. If N-pyramidality is measured by χ_N (the angle between the C—N vector and the NR_2R_3 plane), then as χ_N increases beyond *ca* 35° towards the sp^3 value of *ca* 60° the N···H contacts tend to become significantly shorter (stronger), the N···H donor angle approaches linearity and H approaches N within a 20° cone that has the assumed N lone-pair vector as an axis. A plot of χ_N *versus* d_{NH} is interpreted as the reaction pathway for protonation of N and data points from $R_1(X{=})C—N^+(R_2R_3)H$ systems (the reaction product) occur at the end of this pathway. Crystallographic evidence shows that all 153 contacts ≤ 2.75 Å that have $\chi_N \geq 35°$ and a N···H donor angle above 130° are true hydrogen bonds. The evidence also suggests that the incoming H atoms track the developing N lone-pair density as χ_N increases from *ca* 35 to 60°. *Ab initio* molecular orbital calculations for aniline (6–31 G* basis

* Author to whom correspondence should be addressed.

Table 1. *Summary of CSD substructure search and non-bonded search procedures*

(i) Occurrences of substructures (I)–(V) for the R_2, R_3 substituent variants. Observed and calculated energy barriers to *cis–trans* isomerization (E_i, kJ mol^{-1}) are from Ferretti *et al.* (1993).

	R_2	R_3	(I)	(II)	(III)	(IV)	(V)
(*a*)	Csp^3	Csp^3	684	318	344	1956	297
(*b*)	Csp^3	H	254	133	188	2316	197
(*c*)	H	H	360	154	744	636	254
E_i (observed)			21.4–39.5	24.8–37.8	33.6–47.9	62.6–87.8	77.3–103.3
E_i (calculated)			24.8	34.4	55.0	82.7	111.7

(ii) (*a*) Occurrences of non-bonded search fragment (XI) in terms of contributors from parent substructures (I*a*, *b*, *c*)–(V*a*, *b*, *c*) and (*b*) instances in which the geometry at N deviates significantly from planarity.

		(I)	(II)	(III)	(IV)	(V)	Σ
(*a*)	$d_{NH} \le 2.75$ Å (N_a)	224	85	88	544	5	946
(*b*)	$\chi_N \ge 35°$, $\rho_H \ge 130°$ (N_b)	110	26	10	7	0	153
	N_b/Σ_b (%)	71.8	17.0	6.5	4.6	0	—
	N_b/N_a (%)	49.1	30.6	11.4	1.2	0	—

set) give electrostatic potential maps at incremental values of χ_N (in the range 0–48.5°) that clearly show the developing N lone-pair density.

Introduction

The generic substructure $R_1(X{=})C{-}NR_2R_3$ (R_1 = any substituent; $R_2, R_3 = Csp^3$ or H) characterizes anilines (I, $X = C_{phenyl}$), enamines (II, $X = C_{ethylenic}$), amidines (III, $X = N$), amides (IV, $X = O$) and thioamides (V, $X = S$). The stereochemical features of (I)–(V) have been the subject of considerable spectroscopic, theoretical and crystallographic interest (see *e.g.* Siddall, Stewart & Knight, 1970; Dunitz & Winkler, 1975; Dunitz, 1979; Barbieri, Benassi, Grandi, Pagnoni & Taddei, 1979; Gilli, Bertolasi, Bellucci & Ferretti, 1986; Ferretti, Bertolasi, Gilli & Gilli, 1993). Normally these substructures adopt a planar conformation due to significant contributions from the polar form (VI*a*): there is conjugation between C=X and the N lone pair and N adopts an sp^2 configuration with the C—N bond having partial double-bond character in the resonance hybrid. However, a number of crystallographic studies (see, *inter alia*, Brown, Damm, Dunitz, Eschenmoser, Hobi & Kratky, 1978; Gilli & Bertolasi, 1979; Bertolasi, Bellucci, Ferretti & Gilli, 1984; Bellucci, Bertolasi, Ferretti & Gilli, 1985) revealed significant rotations about C—N and suggested that the degree of rotation might be correlated with the changing electronegativity of X in (I)–(V). These observations are in agreement with dynamic NMR measurements of the energy barriers (E_i) to the *cis–trans* isomerization (VII)–(IX) (see *e.g.* MacKenzie & MacNicol, 1970; Siddall, Stewart & Knight, 1970); Martin, Gouesnard, Dorie, Rabiller & Martin, 1977; Barbieri *et al.*, 1979). Typical experimental E_i values range from 21.4 up to 103.3 kJ mol^{-1}, as indicated in Table 1.

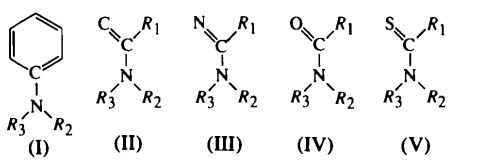
Early systematic studies of medium-ring lactams (Dunitz & Winkler, 1975; Dunitz, 1979) suggested that the out-of-plane deformations at N and the degree of rotation about the C—N bond were strongly coupled along a minimum-energy pathway that is followed during the *cis–trans* isomerization. Gilli *et al.* (1986) later presented an important survey of geometrical variations in 90 instances of (I)–(V) as observed in 68 crystal structures. They described the stereochemistry of (I)–(V) in terms of the internal coordinates τ and χ_N that are linear combinations of the four torsion angles, φ_1–φ_4, about the C—N bond (Winkler & Dunitz, 1971; Dunitz & Winkler, 1975; Bürgi & Shefter, 1975), *viz*

$$\varphi_1 = X{-}C{-}N{-}R_3$$
$$\varphi_2 = R_1{-}C{-}N{-}R_2$$
$$\varphi_3 = R_1{-}C{-}N{-}R_3$$
$$\varphi_4 = X{-}C{-}N{-}R_2$$
$$\tau = (\varphi_1 + \varphi_2)/2$$
$$\chi_N = \varphi_2 - \varphi_3 + \pi \,(\mathrm{mod}\, 2\pi)$$
$$= -\varphi_1 + \varphi_4 + \pi \,(\mathrm{mod}\, 2\pi).$$

Thus (see VIII), τ measures the degree of rotation of $-NR_2R_3$ with respect to $R_1(X{=})C{-}$, while χ_N measures the out-of-plane deformation of R_2,R_3 as N changes from a planar sp^2 configuration to a pyramidal sp^3 configuration during the isomerization (VII)→(IX). Values of χ_C, the out-of-plane deformation of R_1,X, were also calculated [as $\chi_C = \varphi_1 - \varphi_3 + \pi \,(\mathrm{mod}\, 2\pi) = -\varphi_2 + \varphi_4 + \pi \,(\mathrm{mod}\, 2\pi)$], but $|\chi_C|$ was always less than 6.0° with a mean value of 1.7° over the 90 instances.

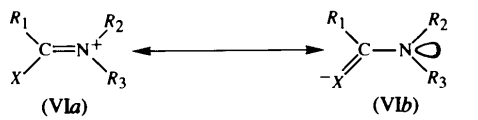
Gilli *et al.* (1986) present a χ_N–τ plot that shows a smooth curve, exemplified later (Fig. 2) for our present datasets, which can be interpreted as mapping the asymmetric portion of the isomerization (VII)→(IX). There is initial movement along the χ_N axis, *i.e.* an initial pyramidalization of N, followed by a rotation about C—N as χ_N increases further. The transition state (VIII) is reached at $\chi_N \approx 50°$, $\tau = 90°$. Further, these authors computed the potential energy surface in terms of χ_N,τ using force-field methods and show that the individual

crystallographic observations fall in a low-energy valley in that surface. More recently these workers (Ferretti *et al.*, 1993) have analysed an extended dataset of more than 300 crystallographic observations of (I)–(V), in conjunction with *ab initio* calculations of the potential energy surfaces for each separate substructure. The simplest model compounds [$R_1 = R_2 = R_3 = H$], *i.e.* aniline (I), vinylamine (II), formamidine (III), formamide (IV) and thioformamide (V), were used in the *ab initio* work and the computed values of E_i, the barrier to *cis–trans* isomerization, are included in Table 1. Ferretti *et al.* (1993) show that as E_i increases from (I) to (V), then a decreasing percentage of crystallographic observations have $\chi_N > 25°$, $\tau > 15°$, values that approximately encompass a 4.2 kJ mol^{-1} energy difference from the minimum at $\chi_N, \tau = 0°$. They also show that elongation of the C—N bond length d_{CN} is strongly correlated to increases in τ (correlation coefficient $r = 0.69$) and especially to χ_N ($r = 0.83$).

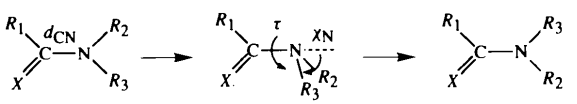


(I) (II) (III) (IV) (V)

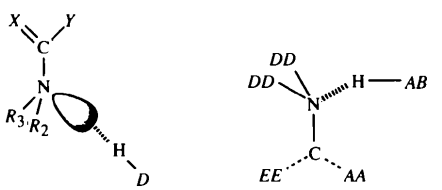
(*a*) $R_2 = R_3 = Csp^3$, (*b*) $R_2 = Csp^3$, $R_3 = H$, (*c*) $R_2 = R_3 = H$



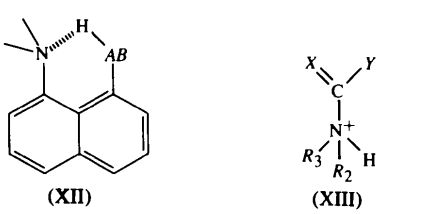
(VI*a*) (VI*b*)



(VII) (VIII) (IX)



(X) (XI)



(XII) (XIII)

The work of Gilli *et al.* (1986) and Ferretti *et al.* (1993) shows that the N$sp^2 \rightarrow$Nsp^3 transition that accompanies the decoupling of the N(lone-pair)—C=X conjugation is primarily mapped by the χ_N coordinate. Further, the change in N hybridization from sp^2 ($\chi_N = 0°$) to sp^3 ($\chi_N \sim 50$–$60°$) obviously leads to the formation of an sp^3 lone-pair orbital on N, as shown in (X). This means that N can now act as a hydrogen-bond acceptor in all R_2, R_3-substituent variants (I*a*,*b*,*c*)–(V*a*, *b*, *c*), in which the geometry is heavily distorted away from the preferred planar stereochemistry. Thus, we would expect that the hydrogen-bond acceptor ability of N would increase with increasing distortion of the geometry along the isomerization pathway towards the transition state at $\chi_N \sim 50$–$60°$, $\tau = 90°$.

It is the hydrogen-bond acceptor characteristics of N in stereochemically distorted $R_1(X=)C—NR_2R_3$ substructures that form the major subject matter of this paper. We use crystallographic information retrieved from the Cambridge Structural Database (CSD: Allen, Davies, Galloy, Johnson, Kennard, Macrae, Mitchell, Mitchell, Smith & Watson, 1991) to investigate and characterize the shorter non-bonded contacts to N in (I)–(V) and we compare these results with data from appropriate MO calculations. We begin the paper, however, with a brief overview of the available data, including an analysis of the steric origins of the geometrical distortions observed in (I)–(V).

Database methodology

The April 1993 release of Version 5 of the CSD system was used in this work. Substructure searches and data analyses were performed using *QUEST3D* and *VISTA* (CSD User Manuals, 1992, 1994) and the statistical display package *XLISP-STAT* (Tierney, 1992).

Substructure searches

Substructure searches for (I)–(V) were carried out (Table 1) with R_1 = any substituent atom, but with (*a*) $R_2 = R_3 = Csp^3$, (*b*) $R_2 = Csp^3$, $R_3 = H$ and (*c*) $R_2 = R_3 = H$ to generate 15 subsets of CSD entries. These searches were carried out using the following secondary chemical and structural criteria: (i) the N atom had a formal zero charge, (ii) connections between R groups were not permitted to avoid three-membered ring compounds, (iii) entries selected were (*a*) error-free in CSD evaluation procedures, (*b*) did not exhibit structural disorder, (*c*) were classified as organic according to CSD criteria, (*d*) had a crystallographic $R \leq 0.10$ and (*e*) had atomic coordinate data available in the CSD. A total of 6187 CSD entries yielded 8835 independent substructures distributed across the 15 subsets as shown in Table 1. The number of fragments retrieved here considerably exceeds the number used by Ferretti *et al.* (1993), who did not consider substructures having cyclic C—N bonds or permit any of the substituents R_1, R_2 or R_3 to be H atoms. For each independent substructure we calculated the C—N and C=X bond lengths (d_{CN}, d_{CX}), together

with the conformational descriptors χ_N and τ defined above. We note that both χ_N and τ can assume positive and negative vaiues [see Gilli *et al.* (1986), Ferretti *et al.* (1993)], *i.e.* they map the four possible quadrants of the conformational space. We adopt the earlier convention of placing all observations in the (++)-quadrant: χ_N, τ refer to positive values throughout this paper.

Non-bonded N···H *contact searches*

The non-bonded search module of *QUEST3D* was applied to a virtual database comprising the 6187 CSD entries that survived the substructure search procedures. *QUEST3D* was used to locate all examples of the generalized non-bonded fragment (XI) where: (*a*) the element group *DD* was restricted to Csp^3 or H, (*b*) the element group *EE* was restricted to C, N, O or S and the C—*EE* bond was double or aromatic only, and (*c*) *AA* indicates any atom, including H, and the C—*AA* bond was single or aromatic only. Definitions (*a*)–(*c*) merely generalize the five separate substructure searches carried out earlier. For the non-bonded search, the donor-H element *AB* was restricted to N or O and a combined search was carried out to locate (i) intermolecular N···H—*AB* contacts, where *AB*—H and N are in separate bonded units, and (ii) intramolecular N···H—*AB* contacts, in which *AB*—H and N are in the same bonded unit but with the acceptor N atom separated from the donor atom *AB* by at least four contiguous covalent bonds within that unit, as in structure (XII) (four-bond separation). All *AB*—H donor groups were geometrically normalized prior to the search: the H atom was repositioned on the *AB*—H vector at a distance from *AB* corresponding to the appropriate mean O—H,N—H bond length (0.983, 1.009 Å, respectively) determined from neutron diffraction studies (Allen, Kennard, Watson, Brammer, Orpen & Taylor, 1987). A non-bonded contact criterion of N···H ≤ 2.75 Å, the sum of van der Waals radii (N = 1.55, H = 1.20 Å: Bondi, 1964), was then used to establish search hits. A total of 946 contacts were located in 633 CSD entries and a breakdown of those hits for substructures (I)–(V) is included in Table 1.

For each instance of (XI), we calculated d_{CN}, χ_N and τ as before, together with additional parameters (see Fig. 1): the N···H contact distance (d_{NH}), the Csp^2—N···H angle (ρ_N) and the N···H—*AB* angle (ρ_H). To study the directionality of approach of H to the N atom, we calculate the spherical polar coordinates θ, φ (Fig. 1) by first constructing the plane that is perpendicular to the N,R_2,R_3 plane. For perfect T_d symmetry about N, this plane will then contain Csp^2,N and the assumed sp^3 lone pair, and forms the equatorial plane of a sphere with N at its centre. The angle θ is then the angle of elevation of N···H from the equatorial 'lone-pair plane' and φ is the angle of rotation of N···H projected onto this plane. Angle θ can be positive or negative and we use the absolute value, $|\theta|$, since differences between R_2,R_3 are not considered. Values of $\theta \sim 0$ and $\varphi \sim 120°$ represent close alignment of the N···H and N lone-pair vectors.

Preliminary data analysis for substructures (I)–(V)

In order to visualize the very large datasets summarized in Table 1(*a*), we plotted χ_N–τ maps (Gilli *et al.*, 1986; Ferretti *et al.*, 1993) for each of the 15 individual substructures. These maps are presented as a matrix of scattergrams in Fig. 2, in which the rows correspond to the generalized substructures (I)–(V), and the columns correspond to the R_2, R_3 substituent variants (*a*, *b*, *c*). Fig. 2 (I*a*) shows the classic *cis–trans* isomerization pathway described by Gilli *et al.* (1986): an initial increase in χ_N from 0° to *ca* 30–35° (pyramidalization of N), followed by a smooth rotation about C—N taking τ towards 90° with a further increase in χ_N towards a transition state value of *ca* 50–55°. It is visually obvious from column (*a*) of Fig. 2 that the higher χ_N, τ points on the pathway are decreasingly represented as we go down the column from (I) to (V), *i.e.* the crystallographic observations reflect the increasing height of the *cis–trans* isomerization barrier E_i and there is a gradual shrinkage of the extent of the distortions towards the planar form at $\chi_N < 20°$, $\tau < 20°$. This effect can also be seen in columns (*b*) and (*c*), despite a scattering of points due, presumably, to aberrant X-ray H-atom positions in some of these substructures. The decreasing population of higher-order points in the pathway is also observed in moving from column (*a*) to columns (*b*) and (*c*). This is due to the significantly reduced steric bulk of terminal H atoms by comparison to that of Csp^3 atoms and their associated substituents, *i.e.* the presumed driving force for distortion away from the preferred planar form – steric overcrowding – is dramatically reduced. We return to this topic in the next section.

Summary statistics for the 15 substructures presented in Table 2 highlight the energetic and steric influences on

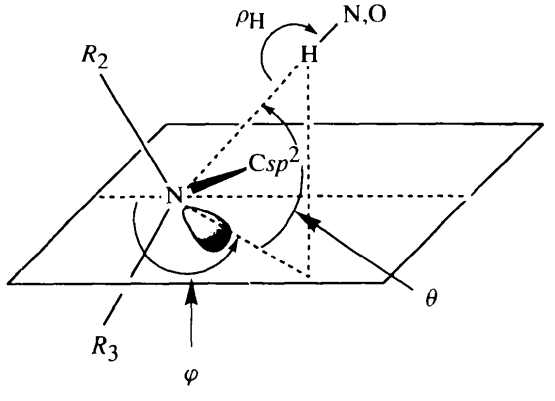


Fig. 1. Geometrical parameters used to describe the non-bonded interactions in substructure (XI).

the geometrical distortions. For the (*a*) series ($R_2 = R_3 = Csp^3$), mean values of χ_N, τ decrease as the E_i values (Table 1) increase. [Values of $\langle\tau\rangle$ for (II*a*) and (V*a*) are anomalously high due to the presence of a small number of outliers of the general formula C,S=C—(NR_2R_3)$_2$. Here, one or both of the —NR_2R_3 groups shows quite large rotations about the appropriate C—N bond without a concomitant increase in χ_N. These outliers are noted and discussed by Ferretti *et al.* (1993).] Gilli *et al.* (1986) and Ferretti *et al.* (1993) also quantify the correlation that exists between d_{CN} and χ_N, τ: the C—N bond increases in length as χ_N, τ increase, *i.e.* as the conjugative interaction between C=*X* and the N lone-pair is decoupled. Correlation coefficients for our (*a*) series datasets are comparable to those reported earlier for their composite dataset. Values of $\langle d_{CN}\rangle$ decrease as E_i increases, reflecting the increasing proportion of substructures that adopt the preferred planar conformation. Finally, in Table 2 we report *P*, the percentage of substructures for which $\chi_N \geq 30°$ and $\tau \geq 20°$. Again, the inverse correlation with E_i is obvious. Despite the problems caused by the use of low-precision H-atom coordinates, the overall effects of increasing E_i can also be seen in the mean geometrical data for both the (*b*) and (*c*) series substructures.

The significant reduction in steric overcrowding in the (*b*) and (*c*) series substructures, due to the substitution of one (*b*) or both (*c*) Csp^3 N substituents by H, is clearly revealed in the data of Table 2. The two most sensitive indicators, $\langle\tau\rangle$ and *P*, both show dramatic reduction for the (*b*) and (*c*) series. The results of Fig. 2 and Table 2 confirm that unfavourable steric interactions, particularly in the (*a*) series substructures with $R_2, R_3 = Csp^3$, are the driving force for the observed geometrical distortions. Thus, in certain molecules, the close non-bonded interactions that would have occurred in the planar form are avoided by nitrogen pyramidalization and/or rotation about the C—N bond. An analysis of these steric factors for substructure (I*a*) is presented below.

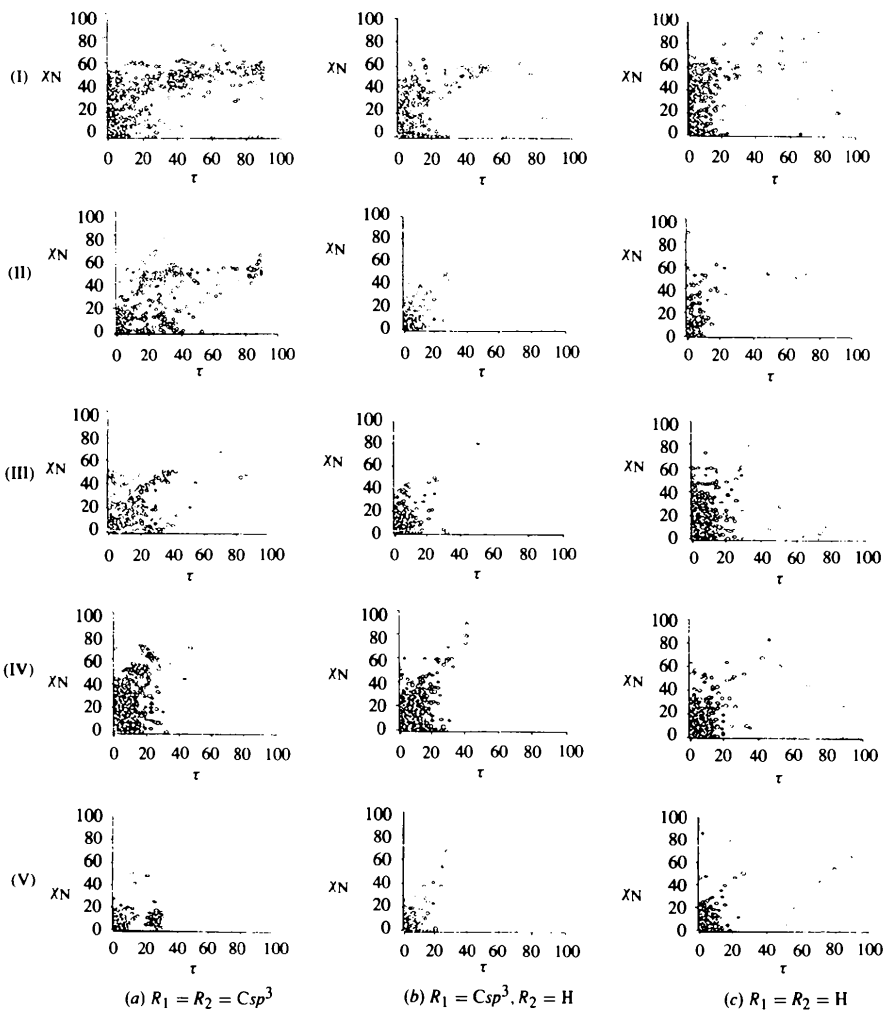


Fig. 2. χ_N–τ mappings (angles in °) of the *cis–trans* isomerization pathway. Rows (I)–(V) represent the individual substructures, columns (*a*), (*b*) and (*c*) refer to the R_2, R_3 variants of each substructure.

Table 2. *Mean values of geometrical descriptors for substructures* (I)–(V), *distances in* Å *and angles in* ° *with e.s.d.'s in parentheses*

P is the percentage of substructures having $\chi_N \geq 30°$, $\tau \geq 20°$. $r(\chi_N)$ and $r(\tau)$ are correlation coefficients linking d_{CN} with χ_N, τ, respectively (see text).

	$\langle d_{CN} \rangle$	$\langle \chi_N \rangle$	$\langle \tau \rangle$	*P*	$r(\chi_N)$	$r(\tau)$
(I*a*)	1.395 (1)	27.6 (8)	19.9 (9)	29.7	0.789	0.700
(II*a*)	1.365 (2)	23.9 (11)	23.6 (13)	27.7	0.776	0.686
(III*a*)	1.346 (1)	15.8 (8)	11.2 (7)	11.6	0.789	0.498
(IV*a*)	1.349 (1)	12.2 (2)	5.1 (1)	2.4	0.495	0.350
(V*a*)	1.331 (1)	6.2 (4)	7.8 (3)	0.8	0.488	0.295
(I*b*)	1.383 (2)	26.1 (12)	12.6 (9)	12.6	—	—
(II*b*)	1.336 (2)	11.1 (10)	6.0 (6)	3.0	—	—
(III*b*)	1.334 (1)	12.2 (9)	6.1 (5)	2.7	—	—
(IV*b*)	1.336 (1)	7.9 (2)	4.8 (1)	1.2	—	—
(V*b*)	1.322 (1)	5.9 (7)	4.4 (3)	1.5	—	—
(I*c*)	1.377 (1)	29.5 (11)	9.7 (8)	7.5	—	—
(II*c*)	1.340 (2)	18.7 (13)	7.1 (10)	4.5	—	—
(III*c*)	1.329 (1)	12.2 (5)	5.5 (3)	1.6	—	—
(IV*c*)	1.328 (1)	10.6 (5)	5.2 (4)	1.7	—	—
(V*c*)	1.320 (1)	9.9 (2)	5.8 (7)	1.9	—	—

Steric origins of geometrical distortions

The steric analysis for substructure (I*a*) made use of local code which performed the following operations. (i) The atomic coordinates of the $—NC_2$ group in each molecule were replaced by coordinates generated for an $—NC_2$ group that was coplanar with the phenyl ring. The N—C distance was set to 1.47 Å [the mean N—Csp^3 value in our sample for (I*a*)], but other distances and angles were retained from each parent molecule. (ii) The shortest intramolecular distance from either $—NC_2$ C atom of the dummy group to any other atom *Z* (including H atoms) was calculated as d_{CZ}. The only constraint applied in this calculation was that the C and *Z* atoms were separated by a minimum of five contiguous covalent bonds. (iii) The shortest d_{CZ} distance was normalized with respect to the sum of van der Waals radii V_C, V_Z to generate $\Delta = d_{CZ} - (V_C + V_Z)$, where V_C was taken as 1.50 Å and V_Z were taken from tables given by Bondi (1964).

Plots of Δ *versus* χ_N and τ are shown in Figs. 3(*a*) and (*b*) to a Δ limit of +1.0 Å. Fig. 3(*a*) is characterized by a line of points close to, but less than, $\Delta = 0.0$ Å and stretching upwards from $\chi_N = 0$ to $\sim 30°$. The Δ position of this line may be indicative of a slight overestimation of V_C, V_Z values in the Bondi (1964) tabulation. Points at low χ_N and Δ approaching −1.0 Å arise from H atoms and must be ascribed to their positioning in X-ray structures and to an overlong Bondi (1964) radius of 1.20 Å for H. However, there is no doubt that highly significant negative Δ's occur beyond $\chi_N \sim 30°$ and the Δ values reach a minimum as χ_N approaches 55–60°. Fig. 3(*b*) at low τ is very similar to Fig. 3(*a*) at low χ_N, but this behaviour ceases at $\tau \sim 20°$ when there is almost linear increase in τ towards the transition state at 90° as Δ assumes more negative values.

The results of Fig. 3 indicate clearly that intramolecular non-bonded interactions are responsible for the distortions of fragment geometry along the mixed χ_N, τ coordinate that maps the *cis–trans* isomerization pathway, *i.e.* for values of $\chi_N \geq 30°$, $\tau \geq 20°$. The present rather crude calculations do not explain the initial movement along the χ_N coordinate that is evident from Fig. 2(I*a*). These initial movements correspond to a simple vibrational coordinate for which the force constants, calculated as increasing from *ca* 15 to 40 kJ mol^{-1} rad^{-2} for substructures (I)–(V) by Ferretti *et al.* (1993), are rather small by comparison to the *cis–trans* isomerization barriers (E_i, Table 1). There is no reason to believe that these vibrational movements are not also associated with close non-bonded interactions, but further and more precise modelling studies are required to quantify these effects.

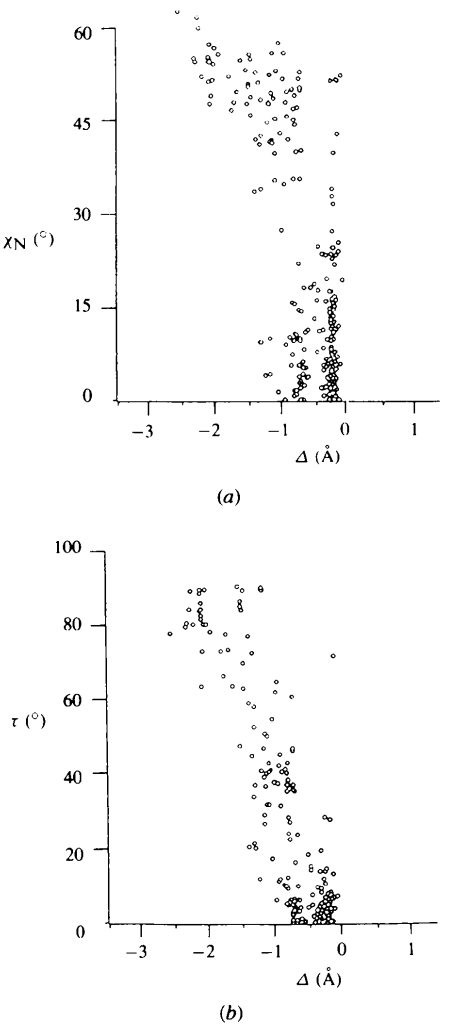


Fig. 3. Steric origins of geometric deformations: (*a*) plot of χ_N *versus* Δ and (*b*) plot of τ *versus* Δ.

Analysis of N···H non-bonded contacts

General survey: all data ($d_{NH} \leq 2.75$ Å)

The ability of N to act as a hydrogen-bond acceptor in distorted (non-planar) examples of (I)–(V) was examined by plotting χ_N, the parameter that best maps the $Nsp^2 \rightarrow Nsp^3$ transition, *versus* the N···H contact distance, d_{NH}. The χ_N–d_{NH} plot for all 926 values of $d_{NH} \leq 2.75$ Å is shown in Fig. 4(*a*). There is a very large population of fragments having $\chi_N < 20°$ (essentially Nsp^2) and having d_{NH} ranging from *ca* 2.30 Å to the limiting value. However, as the $Nsp^2 \rightarrow Nsp^3$ transition proceeds further and χ_N increases from *ca* 35°, then d_{NH} values of < 2.30 Å are consistently observed, with the shorter intermolecular d_{NH} (≤ 2.0 Å) being associated with the higher χ_N (~ 55°). The lowest d_{NH}, those in the range 1.5–1.8 Å, all arise from intramolecular interactions that occur in 1,8-disubstituted naphthylamines (XII).

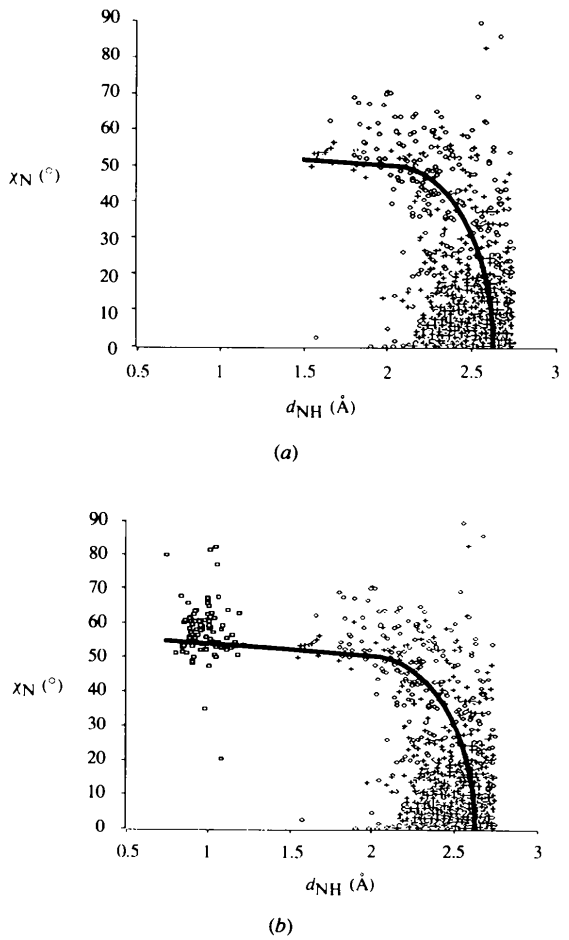


Fig. 4. Non-bonded N···H interactions (all $d_{NH} \leq 2.75$ Å). Plots of χ_N *versus* d_{NH} showing the pathway for the protonation of N: (*a*) for non-bonded N···H interactions only and (*b*) including bonded (N^+—H) interactions from substructure (XIII) (see text). Open circles represent intermolecular N···H contacts, + symbols represent intramolecular N···H contacts and open squares represent N^+—H bonded distances.

Compounds of type (XII) have been implicated in previous crystallographic studies of reaction pathways for the incipient attack of a nitrogen nucleophile on carbonyl or diazonium groups (see *e.g.* Bürgi, Dunitz & Shefter, 1973; Dunitz, 1979; Wallis, Easton & Dunitz, 1993). In our case, the shape of the χ_N–d_{NH} plot of Fig. 4(*a*) can be interpreted in terms of the reaction pathway for protonation of the acceptor N atom in $R_1(X=)C-NR_2R_3$ systems. The product of such a reaction would be the protonated species (XIII) having a N^+—H bond at tetragonal nitrogen. An additional CSD search was carried out to locate examples of (XIII) in which the substituents at C and N, and the chemical bond types, were specified as for (X) (see *Database Methodology* section). A total of 98 examples of (XIII) were located in 90 crystal structures. Their χ_N and d_{NH} (bonded) values are included in Fig. 4(*b*) to generate the full protonation pathway.

The direction of approach of the *AB*—H proton to the N atom can be mapped in a variety of ways. Fig. 5(*a*) shows a plot of the C—N···H angle (ρ_N) *versus* d_{NH} for all 926 $d_{NH} \leq 2.75$ Å, together with the 98 bonded d_{NH} values from (XIII). The large cluster of points in Fig. 5(*a*) with $d_{NH} > ca$ 2.30 Å corresponds to χ_N in the 0–20° range (see Fig. 4). Here the complete $R_1(X=)C-NR_2R_3$ system is close to planarity and, for obvious steric reasons, the most appropriate direction for any non-bonded approach to N is at right angles to this plane, *i.e.* at $\rho_N = 90°$ (XIV). As the sp^3 lone pair develops on N (XV), ρ_N moves towards the central 109.8° line on Fig. 5(*a*) and, for the reaction product (XIII), values of ρ_N (now a C—N—H valence angle) are scattered close to this line. This gradual movement of ρ_N from *ca* 90° to *ca* 110° is most clearly illustrated in Fig. 5(*b*). This plot is derived from the data of Fig. 5(*a*) by dividing the d_{NH} range into five bands of approximately equal populations and then averaging ρ_N within each band.

Figs. 6(*a*) and (*b*) shows plots of d_{NH} *versus* (*a*) the C—N bond lengths, d_{CN}, and (*b*) the rotation angle, τ. The obvious strong correlation between d_{CN} and d_{NH} (Fig. 6*a*) is a consequence of the d_{CN}–χ_N dependency (Ferretti *et al.*, 1993, see also Table 2). Fig. 6(*a*) is essentially another representation of the protonation pathway of Fig. 4(*b*): as conjugation between the N lone-pair and C=*X* is decoupled, then the C—N bond gradually lengthens as Nsp^2 changes into Nsp^3. This lengthening (Fig. 6*a*) is associated with decreasing N···H contact distances due to hydrogen-bond formation and, finally, with protonation of N in species such as (XIII) for which d_{CN} shows its longest values. The τ–d_{CN} plot (Fig. 6*b*) does show the expected high concentration of longer d_{NH} values at $\tau \sim 0°$, together with a gradual movement to shorter d_{NH} at higher τ values. However, it also shows appreciable density of shorter d_{NH} values for

$\tau \leq 30°$ indicating that, as expected, hydrogen-bond formation is dependent on χ_N alone.

Detailed survey: all substructures with $d_{NH} \leq 2.5$ Å

Further evaluation of the effective minimum χ_N for hydrogen-bond formation, together with an investigation of the directionality of the N···H non-bonded interactions, was carried out using those 526 examples of substructure (X) having $d_{NH} \leq 2.5$ Å. Figs. 4–6 show that this distance cut-off primarily eliminates fragments with low χ_N, but still provides a dataset that fully represents all geometrical variations for visualization purposes. The representative nature of the reduced dataset is confirmed by a comparison of the mean parameters given in columns *A* and *B* of Table 3.

The directionality of the N···H interactions was investigated *via* histograms (Fig. 7) of the parameters ρ_N, θ and φ defined in Fig. 1. The ρ_N distribution (Fig. 7*a*) spans the effective range 75–130° with a major peak at 90–100° and an obvious shoulder at *ca* 105–110°. These two features represent purely non-bonded N···H interactions (XIV) and true hydrogen-bonded interactions (XV), respectively, *i.e.* Fig. 7(*a*) is a restatement of Fig. 5. More than 90% of the θ values (Fig. 7*b*) lie in the narrow range 0–25°, *i.e.* H approaches N in, or close to, the lone-pair plane of Fig. 1. This is to be expected: non-bonded interactions prefer this direction of approach for steric reasons, true hydrogen bonds are assumed to prefer it for electrostatic reasons. Given that θ is close to zero in most cases, then φ and ρ_N are closely related parameters, as shown in (XIV) and (XV). Thus, we would expect a major peak in the φ-distribution close to 90° for interactions of the type (XIV) and a smaller peak close to 120° for true hydrogen bonds aligned close to the

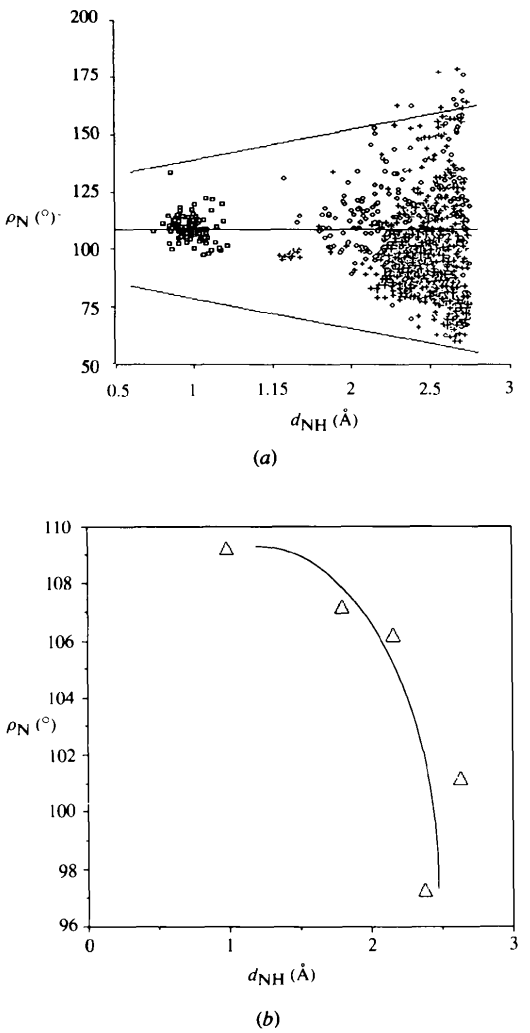


Fig. 5. Non-bonded N···H interactions (all $d_{NH} \leq 2.75$ Å). Direction of H-atom approach to N: (*a*) plot of ρ_N *versus* d_{NH} (all values: symbolic identifiers as in Fig. 4) and (*b*) plot of mean ρ_N values for ranges of d_{NH} (see text).

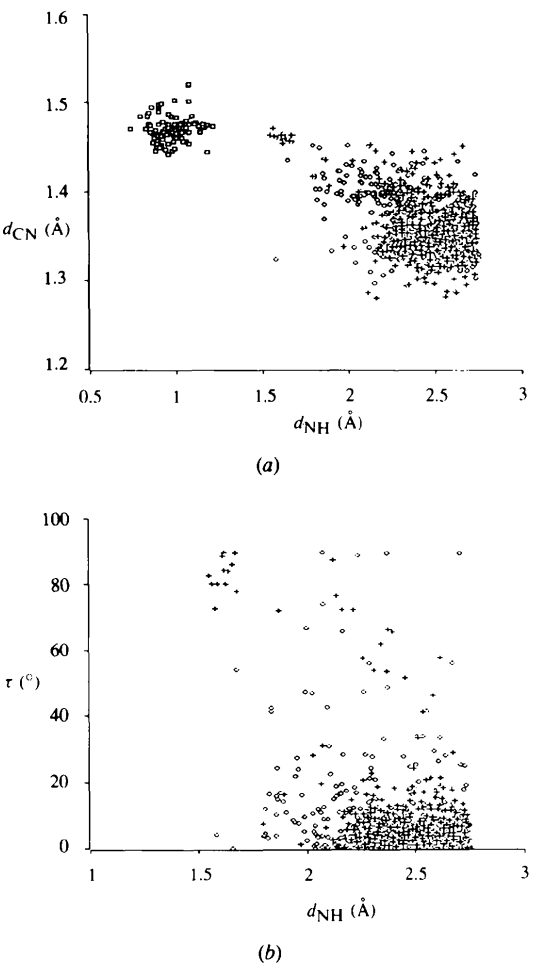
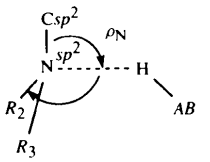
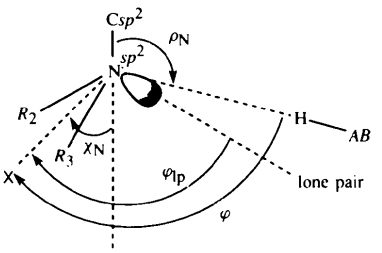


Fig. 6. Non-bonded N···H interactions (all $d_{NH} \leq 2.75$ Å). Plots of (*a*) d_{NH} *versus* d_{CN} and (*b*) d_{NH} *versus* τ. Symbols used on the plots are defined in Fig. 4.

lone-pair direction (XV). These expectations are confirmed in Fig. 7(*c*), which highlights the obvious similarities of the φ and ρ_N distributions.



(XIV)



(XV)

Linearity of the N···H—*AB* system (XV) is investigated in Fig. 8. The ρ_H histogram of Fig. 8(*a*) is uncorrected for geometrical factors (Kroon, Kanters, van Duijneveldt-van der Rijdt, van Duijneveldt & Vliegenthart, 1975) and shows a major peak at *ca* 105° and appreciable density in the 130–180° range that is typical for strong hydrogen bonding. The scatterplot of ρ_H *versus* d_{NH} (Fig. 8*b*) shows that the large peak in the ρ_H distribution arises from a coherent cluster that comprises the longer N···H distances. The distribution of N···H distances for $\rho_H > 140°$ does span the complete range out to the limit of 2.5 Å [which is still 0.25 Å less than the sum of van der Waals radii given by Bondi (1964)], but the majority of the shorter N···H values are also observed in this section of the map. A few shorter (1.8–2.2 Å) d_{NH} values do occur for $\rho_H = 120$–130°, but all these arise from intramolecular hydrogen bonds of type (XII), where linearity of N···H—*AB* is impossible to achieve.

Examination of the protonation pathway presented in Fig. 4 provides a visual estimate of 30–35° for the minimum effective χ_N for hydrogen-bond formation. This point is further investigated *via* the χ_N–φ and χ_N–ρ_H plots of Figs. 9(*a*) and (*b*), which both show two major clusters of observations that are identified on each plot. Fig. 9(*a*) shows a major cluster at low χ_N and centred around $\varphi \sim 90°$ that corresponds to non-bonded contacts exemplified by (XIV). The second cluster occurs at high χ_N with $\varphi \sim 110$–130° and corresponds to hydrogen-bonded contacts (XV). The χ_N–ρ_H plot of Fig. 9(*b*) shows a similar pattern: cluster 1 at $\chi_N < 20$ and $\rho_H = 95$–110° corresponds to (XIV), cluster 2 at $\chi_N > 40$ and $\rho_H > 140°$ corresponds to (XV).

The inter-relationships between the geometrical parameters that are visualized in Figs. 7–9 are further examined and quantified *via* the mean geometries collected in Table 3, columns *C*, *D* and *E*. There are highly significant differences between the mean values obtained for subset *C* ($d_{NH} \leq 2.5$ Å, $\chi_N \leq 30°$) and those for the higher χ_N subsets *D* ($d_{NH} \leq 2.5$ Å, $\chi_N \geq 20°$) and *E* ($d_{NH} \leq 2.5$ Å, $\chi_N \geq 30°$). Thus, in moving from subset *C* to subset *E*, we observe increases in mean values for ρ_N, ρ_H and φ that are symptomatic of the change from simple non-bonded N···H interactions (XIV, subset *C*) to true hydrogen bonds (XV, subset *E*).

Selection and characterization of the hydrogen-bonded subset

In the light of the evidence presented above, we have used the angular criteria $\chi_N \geq 35°$ $\rho_H \geq 130°$ to define a subset of true hydrogen bonds. Further, we have applied

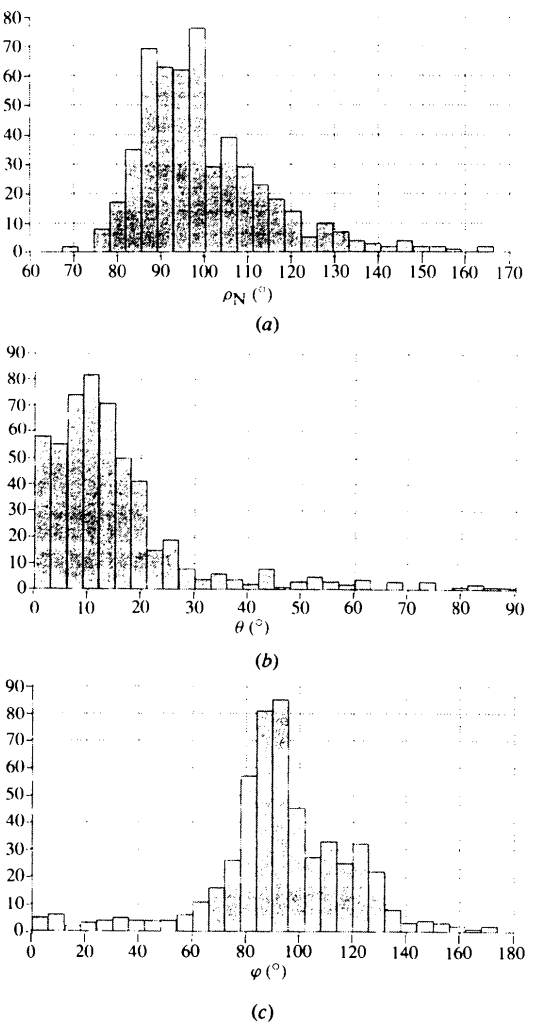


Fig. 7. Detailed study of all $d_{NH} \leq 2.5$ Å. Directionality of the N···H interaction: (*a*) the ρ_N distribution, (*b*) the θ distribution and (*c*) the φ distribution.

these criteria to the full set of N···H contacts, *i.e.* out to the original distance limit of 2.75 Å. A total of 153 contacts (16% of the total) satisfied these criteria and is denoted as dataset *F* in Table 3. This dataset differs from dataset *E* through (*a*) the elimination of 17 observations having $\rho_H < 130°$ in *E*, and (*b*) the addition of 21 observations with $d_{NH} > 2.5$ Å to form dataset *F*. The geometrical characteristics of the hydrogen-bonded contacts of dataset *F* are displayed in Fig. 10.

The d_{NH} distribution (Fig. 10*a*) ranges up to 2.72 Å. If we ignore the shortest N···H contacts, *i.e.* those arising from intramolecular hydrogen bonds of type (XII), then the distribution is nearly normal with a median $d_{NH} \sim 2.35$ Å. The ρ_H distribution is shown in its observed and geometrically corrected (Kroon *et al.*, 1975) forms in Figs. 10(*b*) and (*c*). The raw data have a mean of 160(1)° with 78% of the observations in the narrow range 150–174°. Angles in the 130–140° range arise from constrained intramolecular interactions (XII). The geometrically corrected distribution (Fig. 10*c*) shows the expected tendency of intermolecular hydrogen bonds towards linear arrangements.

The remaining histograms of Fig. 10 examine the directionality of the N···H interaction with respect to the N lone-pair. Fig. 10(*d*) shows a circular histogram of θ, the angle of elevation of the N···H vector from the 'lone-pair plane' defined in Fig. 1. The θ values in Fig. 10(*d*) have been symmetized about $\theta = 0°$, *i.e.* enantiomorphic arrangements are included in the displayed distribution. The vast majority of donor H atoms approach N in, or close to, the lone-pair plane: 146 (86%) of the $|\theta|$ values lie in the narrow range 0–20° and the overall mean $|\theta|$ is 12(1)°.

The distribution of φ, the angle of approach of H to N in the lone-pair plane [Fig. 1 (XV)], is shown in Fig. 10(*e*). The effective range of φ, from *ca* 85 to 135°, is

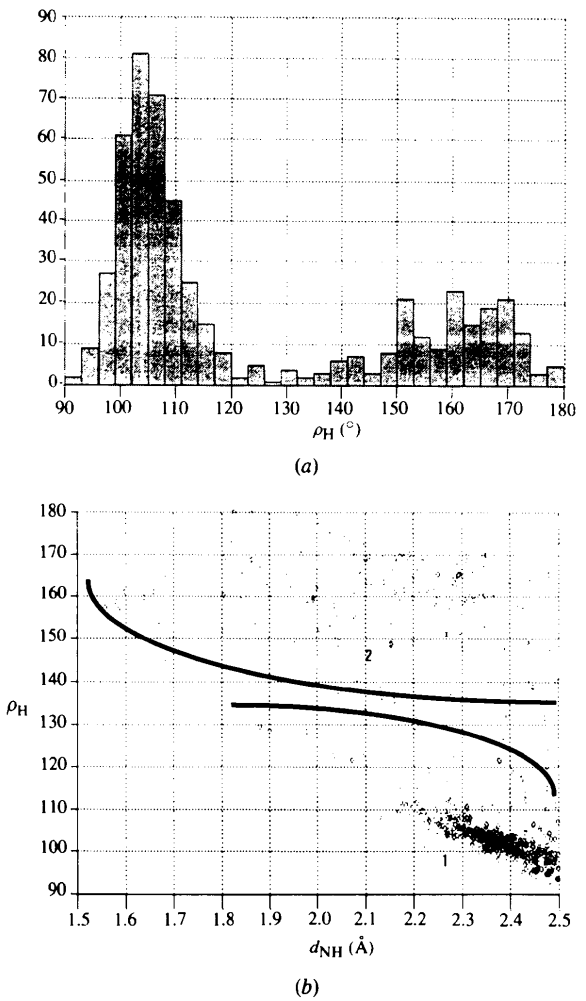


Fig. 8. Detailed study of all $d_{NH} \leq 2.5$ Å. Linearity of the N···H—*AB* system: (*a*) the uncorrected ρ_H distribution and (*b*) plot of ρ_H *versus* d_{NH}.

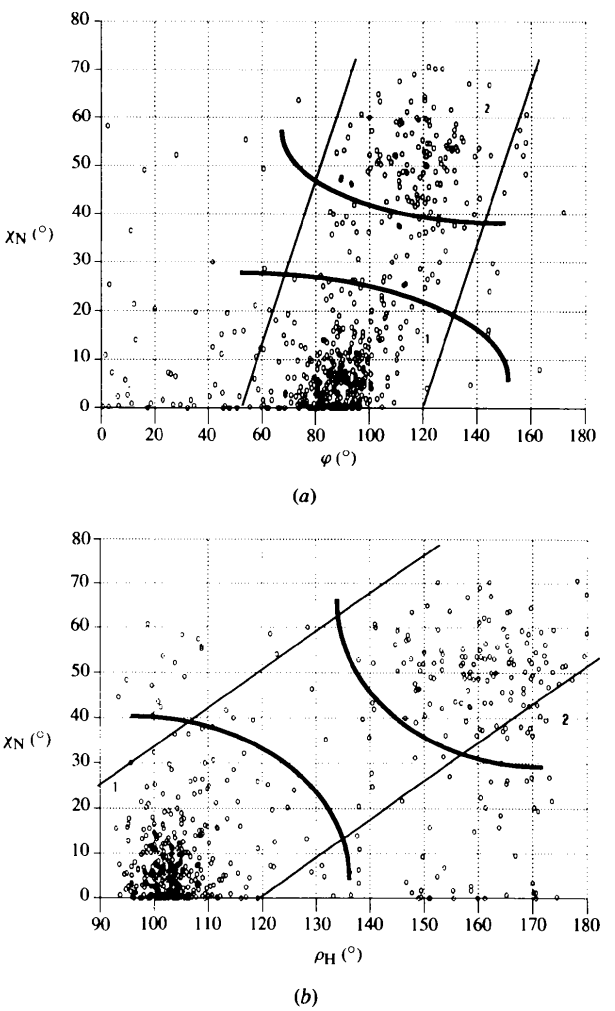


Fig. 9. Detailed study of all $d_{NH} \leq 2.5$ Å. The effective minimum χ_N for hydrogen-bond formation: (*a*) plot of χ_N *versus* φ and (*b*) plot of χ_N *versus* ρ_H.

quite broad. There is a peak at ~120°, the position that might be expected for an interaction between H and an sp^3 lone pair, but the distribution is skewed towards lower values, with an overall mean of 114 (2)°. While we can ascribe some of the scatter in φ to inaccuracies in H-atom positions, it is possible to explain both the spread of φ values and the skewness of the distribution in terms of the changing position of the N lone-pair vector as χ_N increases from 35° to the tetrahedral value of 60°.

Reference to (XV) shows that both χ_N and φ are measured in the lone-pair plane and from a common vector: the bisector of the R_2—N—R_3 angle [N—X in (XV)]. If we assume that the lone pair forms along the bisector of the obtuse angle C—N—X (XV), then we may derive a value for the X—N—lone-pair angle [φ_{lp} in (XV)] that defines the assumed lone-pair position with respect to the vector N—X. By inspection of (XV), we have $\varphi_{lp} = (180 + \chi_N)/2$, *i.e.* φ_{lp} increases from 107.5° at $\chi_N = 35°$ to 120° at $\chi_N = 60°$. Fig. 9(*a*) shows that χ_N and φ are positively correlated in both the non-bonded and hydrogen-bonded regions and thus suggests that the donor H atom tracks the developing lone pair. If this is true, *i.e.* $\varphi \sim \varphi_{lp}$, then we would expect to see a φ distribution for true hydrogen bonds that exhibits a broad spread (from below 107.5° to above 120°) and which is skewed towards the lower values of this range. The φ

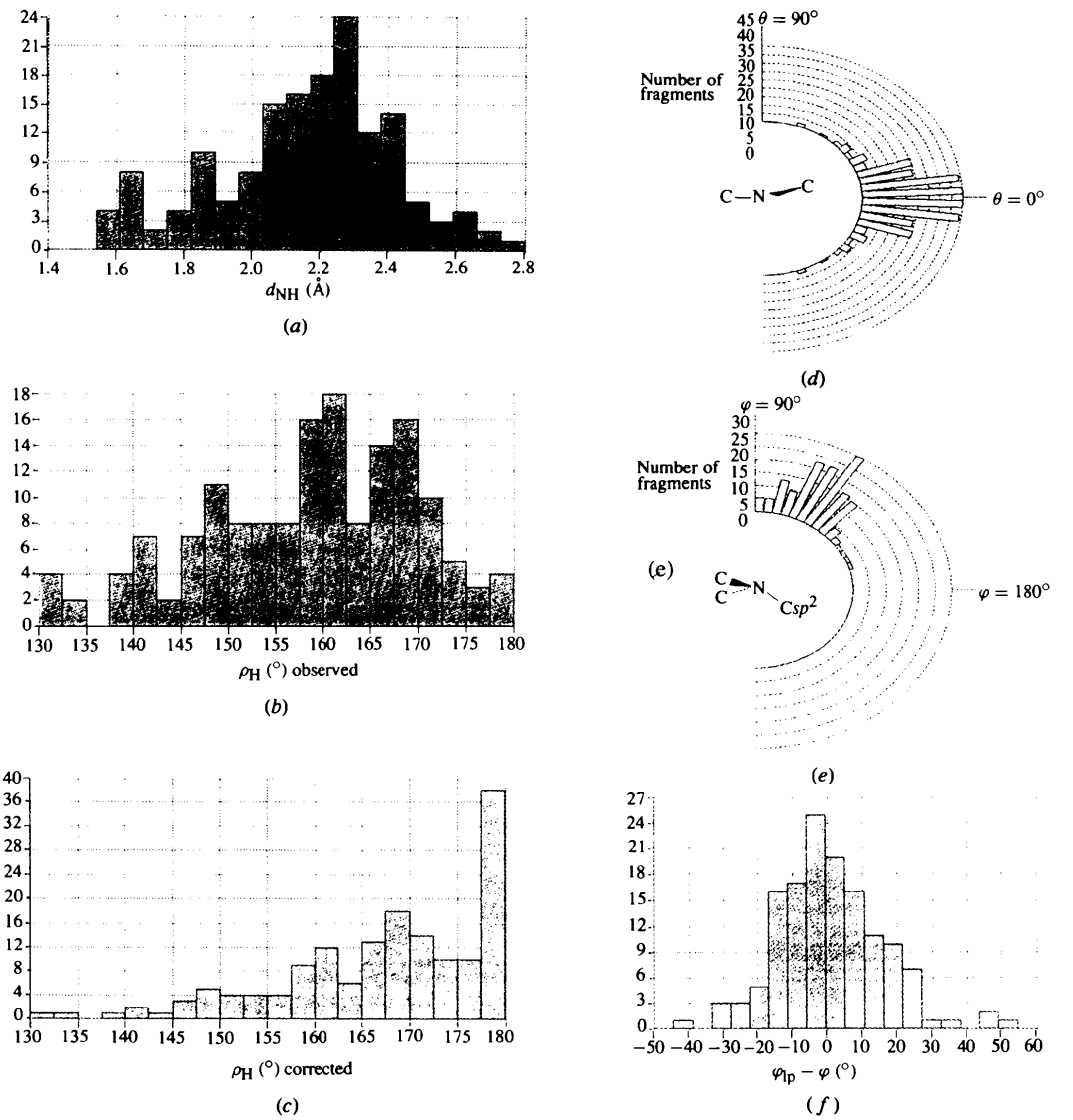


Fig. 10. Hydrogen-bonded N···H interactions ($d_{NH} \leq 2.75$ Å, $\chi_N \geq 35°$, $\rho_H \geq 130°$). Geometrical characteristics: (*a*) the d_{NH} distribution, (*b*) the uncorrected ρ_H distribution, (*c*) the corrected ρ_H distribution (see text), (*d*) the θ distribution, (*e*) the φ distribution and (*f*) the distribution of $\varphi_{lp}-\varphi$ for intermolecular hydrogen bonds (see text).

Table 3. *Mean values for geometrical descriptors (see Methodology) for various data subsets for the non-bonded substructure* (XI)

Distances in Å and angles in ° with e.s.d.'s in parentheses. N_{obs} is the number of fragments in each subset.

Subset	*A*	*B*	*C*	*D*	*E*	*F*
d_{NH} limit	≤ 2.75	≤ 2.5	≤ 2.5	≤ 2.5	≤ 2.5	≤ 2.75
χ_N limit			< 30	≥ 20	≥ 30	≥ 35
ρ_H limit						≥ 130
d_{CN}	1.361 (1)	1.365 (2)	1.349 (1)	1.397 (2)	1.405 (2)	1.402 (2)
d_{NH}	2.448 (7)	2.309 (8)	2.366 (5)	2.200 (19)	2.165 (22)	2.206 (25)
ρ_N	100.7 (7)	100.0 (7)	97.1 (7)	105.7 (12)	107.2 (13)	108.0 (11)
χ_N	18.1 (6)	19.6 (9)	7.5 (4)	45.5 (10)	50.4 (7)	51.2 (8)
τ	8.7 (5)	10.1 (8)	5.6 (5)	20.0 (18)	21.6 (20)	20.4 (20)
θ	20.4 (7)	15.4 (6)	15.5 (7)	16.0 (12)	15.4 (12)	12.0 (10)
φ	91.8 (11)	92.6 (11)	83.5 (12)	111.6 (20)	115.4 (18)	114.0 (17)
ρ_H	118.2 (8)	123.2 (11)	111.2 (9)	148.2 (18)	153.2 (17)	159.7 (8)
N_{obs}	946	526	377	184	149	153

histogram of Fig. 10(*e*) shows both of these properties. Additional evidence for the close association of the direction of H approach with the N lone-pair vector is provided by the histogram (Fig. 10*f*) of $(\varphi_{lp}-\varphi)$ for the sterically unconstrained intermolecular hydrogen bonds. This distribution is near normal, with lower and upper decile values of −17 and +20°, respectively, and mean [1 (1)°] and median (−0.9°) values that are essentially equal to each other and to zero. The shape of the $(\varphi_{lp}-\varphi)$ distribution lends credence to our rather crude estimate of the N lone-pair vector direction. The decile ranges, taken together with the effective θ limit (Fig. 10*d*) of 20°, indicate that some 80% of the donor H atoms in our hydrogen-bonded subset approach N within a 20° cone that has the N lone-pair vector as an axis.

Finally, we consider the composition of dataset *F* (Table 3) in terms of contributions from individual substructures (I)–(V). Earlier (Fig. 1; Table 2) we showed that the degree of deformation of (I)–(V) from the preferred planar geometry ($\chi_N \sim 0°$, $\tau \sim 0°$) is correlated with E_i, the energy barrier to the *cis–trans* isomerization (VII)–(IX). Since we have now shown that true hydrogen bonds to N acceptors in (I)–(V) occur only for $\chi_N > 35°$, it is no surprise to find (Table 1) that dataset *F* is dominated by contributions from the 'aniline' substructure (I) that has the lowest E_i. Further (Table 1), we note that the percentage of non-bonded contacts located from (I)–(V) that actually satisfy the hydrogen-bonding criteria of dataset *F* decreases almost linearly as E_i increases.

Molecular orbital calculations

Ab initio molecular orbital calculations were carried out for the aniline molecule in order to compute geometries, energies and electrostatic potential maps for various values of χ_N. Calculations were performed with the *GAMESS-UK* package (Guest, van Lenthe, Kendrick, Schoeffel, Sherwood & Harrison, 1993) using the HF/SCF method and the 6–31 G* basis set. Calculations were carried out for planar (C_{2v}, $\chi_N = 0°$) aniline and for a further five C_s symmetric conformations, having N—C in a mirror plane that is perpendicular to the phenyl ring, *i.e.* τ is fixed at 0°. Each C_s conformation was initially characterized by an increasing fixed pyramidality at N (χ_N, Table 4) and full geometry optimization of the unconstrained parameters was carried out at the 6–31 G* level. In a final run, the torsional and angular constraints were removed so as to establish χ_N for the minimum-energy C_s symmetric conformation.

Table 4. *Results for ab initio molecular orbital calculations for aniline*

Final energies for each value of χ_N are given as E_f in atomic units. Relative energies with respect to the minimum energy conformation at $\chi_N = 43.8°$ (marked with an asterisk) are given as E_r in kJ mol^{-1}. The minimum value of the electrostatic potential is given as E_p in kJ mol^{-1}. Calculated and experimental d_{CN} values are in Å. The crystallographic means have [σ(mean), number of observations] in parentheses.

χ_N	E_f	E_r	E_p	d_{CN}(calc)	d_{CN}(expt)
0	−285.728227	6.82	−453	1.374	1.363 [4,58]
8.75	−285.728417	6.32	−524	1.375	1.365 [2,47]
18.75	−285.729048	4.66	−603	1.378	1.371 [2,47]
32.90	−285.730295	1.38	−710	1.387	1.379 [3,32]
40.94	−285.730776	0.12	−772	1.394	1.389 [3,41]
43.78*	−285.730822	0.00	−794	1.397	—
48.52	−285.730673	0.39	−836	1.403	1.397 [3,42]

The *ab initio* results for energies and geometries, summarized in Table 4, are closely comparable to earlier calculations (Bock, George & Trachtman, 1986) that used *GAUSSIAN*80 and both 6–31 G and 6–31 G* basis sets. These authors find $\chi_N = 42.5°$ for the minimum-energy C_s symmetric form, in good agreement with the value of 37.5° obtained by microwave spectroscopy (Lister, Tyler, Hoeg & Larsen, 1974). Our present calculations yield $\chi_N = 43.8°$ for the C_s symmetric energy minimum and Table 4 lists energy differences from this minimum for the planar conformation (6.82 kJ mol^{-1}) and for each of the five conformations having constrained non-zero values of χ_N. The calculations reproduce the systematic increases in the C—N bond length (d_{CN}) with increasing χ_N and the *ab initio* d_{CN} values are compared, in Table 4, with mean d_{CN}

values obtained from the crystallographic results retrieved from the CSD for substructure (Ic). For each of the six χ_N values reported in Table 4, we averaged d_{CN} for instances of (Ic) having $\chi_N \pm 4.0°$. The incremental upward trend in the crystallographic mean values closely mirrors the trend in *ab initio* d_{CN} values, but the general level of experimental bond lengths is ~0.01 Å lower than the *ab initio* values.

We then calculated electrostatic potential maps for planar aniline and for each of the five χ_N constrained conformers. These maps were computed using the *MEPMAP* program of the *GAMESS-UK* package and are shown in Fig. 11. Each map shows the electrostatic potential in the mirror plane that bisects the phenyl ring and contains the C—N bonded vector. Positive potential is represented by solid contours, negative potential by broken contours, with the first broken contour representing the zero level. Contours are drawn at intervals of 42 kJ mol^{-1}.

The maps of Fig. 11 clearly show the development of the sp^3 lone pair on N as the H atoms gradually bend away with increasing χ_N. Thus, in the C_{2v} symmetric form (Fig. 11*a*) the largest negative potential at N is −453 kJ mol^{-1}, but this value becomes more negative by *ca* −8 kJ mol^{-1} for every 1° increment in χ_N. The negative potential at our final χ_N of 48.5° is −836 kJ mol^{-1} (Fig. 11*f*). The *ab initio* results are entirely consistent with the crystallographic observations and particularly with Fig. 4(*a*) where the N···H interactions are seen to become shorter and stronger with increasing χ_N.

Concluding remarks

The primary focus of this work has been to investigate and characterize the non-bonded N···H contacts made by N in crystal structures containing the $R_1(X=)C-N(R_2R_3)$ substructure. Although the all-planar (conjugated) form is preferred, a significant number of crystallographic examples show geometrical distortions that can be described by the N-pyramidalization parameter (χ_N) and by rotation about the C—N bond (τ). In this paper, we have used simple model calculations on dialkylanilines to show that these distortions are steric in origin. More importantly, and uniquely, we have been able to study changes in the geometrical characteristics of the non-bonded N···H interactions as N gradually changes from the planar sp^2 configuration ($\chi_N = 0$) to the pyramidal sp^3 configuration ($\chi_N \sim 50$–$60°$), *i.e.* as the sp^3 lone-pair gradually develops on N.

The work has shown that the longer N···H contacts consistently occur at χ_N values that are below 30–35° and that here the H atom prefers to approach N from the least sterically hindered direction, perpendicular to the all-planar conformation. For higher values of χ_N, the N···H contacts become shorter (stronger), the N···H—donor angles (ρ_H) tend to become linear and H approaches N within a 20° cone that has the N lone-pair vector as an axis. In particular, it is shown that the χ_N–d_{NH} plot maps a reaction pathway for protonation of N, a pathway that is extended by the inclusion of bonded N—H distances in $Csp^2-N^+(R_2R_3)H$ systems.

Ab initio molecular orbital calculations for aniline at the 6–31G* level show an increasingly negative

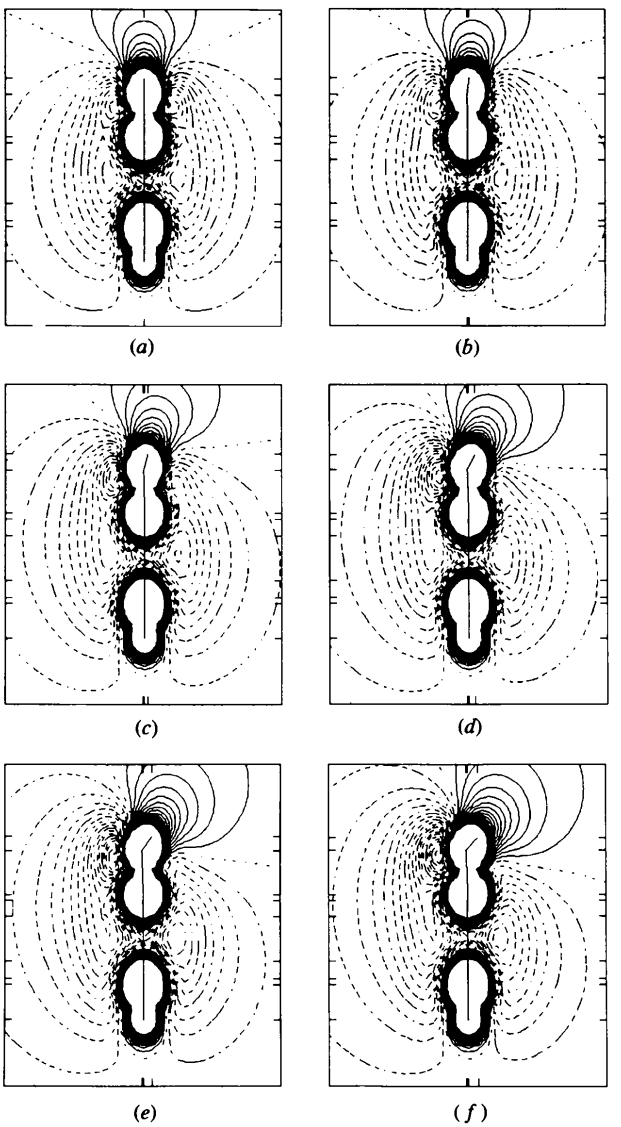


Fig. 11. Electrostatic potential maps for aniline computed by *ab initio* methods (*GAMESS-UK*: Guest *et al.*, 1993). The maps are edge-on views of the molecule in the plane containing the phenyl ring and the C—N bond. Frames (*a*)–(*f*) were computed with fixed χ_N values of 0, 8.75, 18.75, 32.90, 40.94 and 48.52°, respectively: the increased out-of-plane bending of the amino H atoms can be seen (top right) in frames (*b*)–(*f*). The maps are contoured at 42 kJ mol^{-1} intervals; full contours represent positive potentials, broken contours represent negative potentials; the first broken contour is the zero level.

electrostatic potential in the expected N lone-pair position as χ_N is increased incrementally from 0 to 48.5°. These calculations are consistent with the crystallographic evidence that all N···H contacts within the sum of van der Waals radii (2.75 Å) that have $\chi_N \geq 35°$ and $\rho_H \geq 130°$ may be classified as hydrogen bonds. These N···H contacts are closely aligned with the assumed N lone-pair vector and such evidence is normally taken to infer an interaction between H(δ+) and the lone-pair density. However, by analogy with the perpendicular approach of truly 'non-bonded' H to planar Nsp^2 systems (XIV), the lone-pair direction is also the approach direction of minimal steric hindrance to sp^3 hybridized N atoms (XV). Despite this analogy, we feel that the crystallographic results presented here provide tentative evidence that the incoming donor H atom tracks the changing direction of the N lone-pair vector as χ_N increases from 35° to ~60° during the latter phases of the N$sp^2 \rightarrow$ Nsp^3 transition.

We are indebted to Professor G. Gilli and his colleagues at the Universita di Ferrara, Italy, for preprints of their 1993 paper (Ferretti *et al.*, 1993) and for helpful discussions. We also thank referees for constructive suggestions on initial versions of this paper, Dr Neil Stewart (CCDC) for assistance with the illustrations and Dr Olga Kennard OBE FRS for her encouragement of this work.

References

ALLEN, F. H., DAVIES, J. E., GALLOY, J. J., JOHNSON, O., KENNARD, O., MACRAE, C. F., MITCHELL, E., MITCHELL, G. F., SMITH, J. M. & WATSON, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187.

ALLEN, F. H., KENNARD, O., WATSON, D. G., BRAMMER, L., ORPEN, A. G. & TAYLOR, R. (1987). *J. Chem. Soc. Perkin Trans.* 2, pp. S1–S19.

BARBIERI, G., BENASSI, R., GRANDI, R., PAGNONI, U. M. & TADDEI, F. (1979). *J. Chem. Soc. Perkin Trans.* 2, pp. 330–336.

BELLUCCI, F., BERTOLASI, V., FERRETTI, V. & GILLI, G. (1985). *Acta Cryst.* C**41**, 544–546.

BERTOLASI, V., BELLUCCI, F., FERRETTI, V. & GILLI, G. (1984). *Acta Cryst.* A**40**, C107.

BOCK, C. W., GEORGE, P. & TRACHTMAN, M. (1986). *Theor. Chim. Acta*, **69**, 235–245.

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

BROWN, K. L., DAMM, L., DUNITZ, J. D., ESCHENMOSER, A., HOBI, R. & KRATKY, C. (1978). *Helv. Chim. Acta*, **61**, 3108–3135.

BÜRGI, H.-B. & SHEFTER, E. (1975). *Tetrahedron*, **31**, 2976–2981.

BÜRGI, H.-B., DUNITZ, J. D. & SHEFTER, E. (1973). *J. Am. Chem. Soc.* **95**, 5065.

Cambridge Structural Database User's Manual (1992). Version 5.1. Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.

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

DUNITZ, J. D. (1979). *X-ray Analysis and the Structure of Organic Molecules*, p. 333. Ithaca, NY: Cornell Univ. Press.

DUNITZ, J. D. & WINKLER, F. K. (1975). *Acta Cryst.* B**31**, 251–263.

FERRETTI, V., BERTOLASI, V., GILLI, P. & GILLI, G. (1993). *J. Phys. Chem.* **49**, 13568–13574.

GILLI, G. & BERTOLASI, V. (1979). *J. Am. Chem. Soc.* **101**, 7704–7711.

GILLI, G., BERTOLASI, V., BELLUCCI, F. & FERRETTI, V. (1986). *J. Am. Chem. Soc.* **108**, 2420–2424.

GUEST, M. F., VAN LENTHE, J. H., KENDRICK, J., SCHOEFFEL, K., SHERWOOD, P. & HARRISON, R. J. (1993). *GAMESS-UK User's Guide and Reference Manual*. Computing for Science Ltd., SERC Daresbury Laboratory, Daresbury, Warrington, England.

KROON, J., KANTERS, J. A., VAN DUIJNEVELDT-VAN DE RIJDT, J. G. C. H., VAN DUIJNEVELDT, F. B. & VLIEGENTHART, J. A. (1975). *J. Mol. Struct.* **24**, 109–129.

LISTER, D. G., TYLER, J. K., HOEG, J. H. & LARSEN, W. (1974). *J. Mol. Struct.* **23**, 253–264.

MACKENZIE, R. & MACNICOL, D. D. (1970). *Chem. Commun.* pp. 1299–1300.

MARTIN, G. L., GOUESNARD, J. P., DORIE, J., RABILLER, C. & MARTIN, M. L. (1977). *J. Am. Chem. Soc.* **99**, 1381–1384.

SIDDALL, T. H., STEWART, W. E. & KNIGHT, F. D. (1970). *J. Phys. Chem.* **74**, 3580–3583.

TIERNEY, L. (1992). *The XLISP-STAT Package*. School of Statistics, Univ. of Minnesota, Minneapolis, USA.

WALLIS, J. D., EASTON, R. J. C. & DUNITZ, J. D. (1993). *Helv. Chim. Acta*, **76**, 1411–1424.

WINKLER, F. K. & DUNITZ, J. D. (1971). *J. Mol. Biol.* **59**, 169–182.

Acta Cryst. (1995). B**51**, 1081–1097

Electrostatic Properties of β-Cytidine and Cytosine Monohydrate from Bragg Diffraction

BY LIRONG CHEN AND B. M. CRAVEN*

Department of Crystallography, University of Pittsburgh, Pittsburgh, PA 15260, *USA*

(*Received* 4 *January* 1995; *accepted* 27 *April* 1995)

Abstract

The charge-density distribution in the crystal structure of the nucleoside β-cytidine at 123 K has been determined from X-ray diffraction data (Ag $K\alpha$, $\lambda = 0.5608$ Å) using all 7233 reflections with $\sin\theta/\lambda \leq 1.14$ Å^{-1}. Maps of electrostatic potential for the cytosine base in cytidine are similar to those derived from previous charge-density studies of cytosine monohydrate and 1-β-D-arabinosylcytosine, after taking chemical differences into account.

* Author for correspondence.