FARADAY

Ab initio molecular orbital calculations of solvent clusters of *trans*-*N*-methylacetamide: Structure, ring cluster formation and out-of-plane deformation

Ioannis N. Demetropoulos,* Ioannis P. Gerothanassis,* Constantina Vakka and Constantinos Kakavas

Department of Chemistry, University of Ioannina, Ioannina GR-451 10, Greece

The solvation of *trans* amides has been investigated by the use of full gradient optimization *ab initio* quantum mechanical calculation techniques. The complexes have been determined at the Hartree–Fock (HF) level with a 4-31G*/4-31G** basis set and at the second-order Møller–Plesset perturbation (MP2) level. Three NMA–water clusters were investigated: *trans*-NMA with two molecules of water forming a ring cluster at the amide oxygen; *trans*-NMA with two molecules of water at the amide oxygen forming hydrogen bonds along the direction of the lone-pair electrons; *trans*-NMA with one molecule of water at the CO group and one at the NH group. In addition, 4-31G* basis set calculations for *trans*-NMA with two molecules of acetonitrile were performed. The C=O···H(W) hydrogen bond lengths, electron-density population analysis and molecular-orbital analysis of *trans*-NMA with two molecules of water at the amide oxygen demonstrate the importance of concurrent water–water and water–(carbonyl) oxygen hydrogen-bond interactions. The complex of *trans*-NMA with two molecules of water forming a ring cluster at the amide oxygen indicates the formation of a non-planar amide bond and the generation of a chiral centre at the amide nitrogen; this structure has a 5% Boltzmann distribution at room temperature at the MP2 level. Vibrational-frequency analysis shows that its hydrogen-bonded water molecules are vibrationally coupled. Orbital analysis suggests that there is a considerable solute-occupied space reorganization caused by the rearrangement of the water solvent molecules. Comparisons are made with previous theoretical studies of amide–water interactions and experimental spectroscopic, X-ray and neutron-diffraction data on the hydration of amides, peptides and proteins.

The microscopic description of solute–solvent interactions in aqueous solutions of non-ionic, but polar, species such as amides, is of particular interest since amides represent a suitable model for studying both hydrophilic and hydrophobic interactions in peptides and proteins.[1–7] The peptide bond prefers to adopt two planar configurations, *cis* and *trans*.[8,9] Both isomers have been observed in peptides and proteins but the *trans* configuration appears to be much more common. The *trans*-peptide bond ($\omega = 180°$) is favoured 10^3-fold over the *cis* form ($\omega = 0°$), since in the latter the C_i^α atom of amino acid residue (i), any side chains and residue ($i + 2$) are in too close proximity (an exception is the case of a cyclic side chain and/or substitution of the peptide bond, as found in proline and sarcosine, which display an appreciable increase in the proportion present as the *cis* isomer[10–12]).

Solvation might be one of several factors that contribute to the stability of spatial molecular structures. Therefore, solute–solvent interactions have been extensively investigated by the use of theoretical calculations. One of the earliest *ab initio* calculations was concerned with formamide (FMA)[13] and the method has been reviewed by Pullman and Pullman.[14] In this case, six minima on the potential surface were located and these correspond with those suggested by chemical intuition,[15] *i.e.* water molecules act as proton donors to the carbonyl oxygen and as proton acceptors to the N—H bond, with less strongly bound water molecules accepting a proton from C—H and donating another proton to the amine nitrogen. Similar investigations have been made on *N*-methylacetamide (NMA),[16] formamide (FMA) and NMA,[17] and FMA.[18] Owing to the size of the systems, these early calculations were performed with limited geometry optimization and small basis sets.[13,16–18] Studies in which the geometries were fully optimized used the 3-21G basis set, with single-point calculations at larger basis sets.[19,20a] Jasien and Stevens[21] performed full geometry optimizations of four FMA·H_2O complexes using double ζ and double ζ plus polarization quality basis sets. More recently, particular attention has been given to *ab initio* calculations with large basis set of amides:[22a] (i) with one water proton donor at the amide oxygen; (ii) with two waters, one acting as proton donor at the CO group and the other as a hydrogen-bond acceptor at the NH group; (iii) three water molecules, one at the amide nitrogen and two at the amide oxygen. However, *a priori*, a planar peptide bond and coplanar positions of the oxygen atoms of the bound water molecules were assumed.[22] Furthermore, density functional theory was applied on four, fully geometry optimized, FMA–H_2O complexes, one of which has water acting as a proton donor *cis* to the formyl hydrogen.[23]

An analysis of X-ray crystal structure data of a number of peptides indicates that deviations of the atoms from the root-mean-square plane drawn through C_i^α, C′, O, N and C_{i+1}^α are quite large.[24] Furthermore, comprehensive analysis of the peptide-backbone hydration of peptide crystals[2] suggested that the distribution of dihedral angles between the O(W)···OC and O—C—C^α planes indicates a preference for the water molecule bound to the carbonyl group to lie near the plane of the peptide unit but not to be restricted to it. It is therefore of interest to investigate the effect of bound molecules of water in the formation of a non-planar amide bond by the use of full gradient optimization *ab initio* quantum mechanical calculation techniques.

In this paper, the solvation of *trans* amides was investigated by the use of full gradient optimization *ab initio* quantum mechanical calculations with 4-31G* and 4-31G** basis sets on three NMA–water clusters (Fig. 1): *trans*-NMA with one molecule of water at the CO group and one at the NH group (**A**); *trans*-NMA with two molecules of water forming a ring cluster at the amide oxygen (**B**); *trans*-NMA with two molecules of water at the amide oxygen forming hydrogen bonds along the direction of the lone-pair electrons (**C**). A vibrational frequency analysis was performed for **B** using HF/4-31G*. Further calculations were performed at the second-

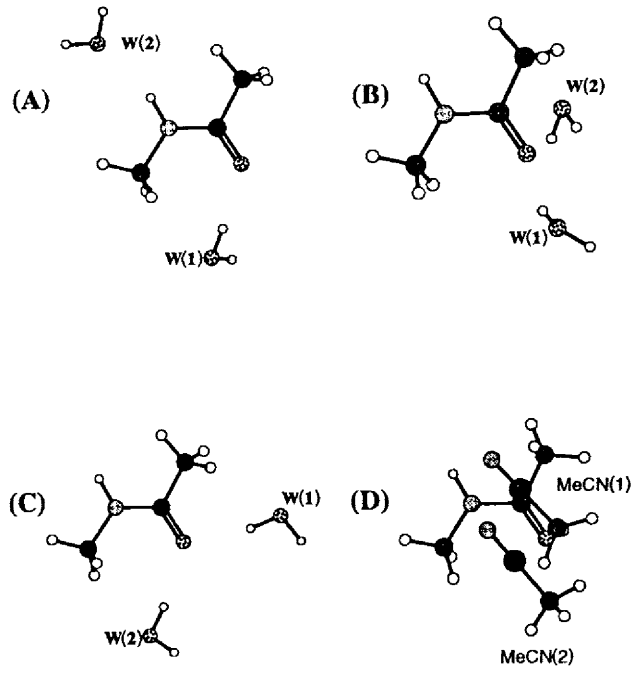

Fig. 1 Energy optimized structures, with 4-31G* basis set *ab initio* calcuations, of **A**, **B**, **C** and **D**

order Møller–Plesset perturbation MP2/4-31G*. In addition, 4-31G* basis set calculations for *trans*-NMA with two molecules of acetonitrile (**D**) were performed. The C=O···H(W) hydrogen-bond lengths, electron-density population analysis and molecular-orbital analysis of *trans*-NMA with two molecules of water at the amide oxygen atom demonstrate the importance of concurrent water–water and water–(carbonyl) oxygen hydrogen-bond interactions of the bound molecules of H_2O. Calculations, at the Hartree–Fock level, on **B** indicate the formation of a non-planar amide bond and the generation of a chiral centre at the amide nitrogen; calculations at the MP2 level indicate that there is a reversal of the energy levels. In the *trans*-NMA–CH_3CN (1 : 2) complex, the two molecules of acetonitrile[25] were found to be oriented nearly parallel (above and below) to the amide plane. Orbital analysis suggests that there is considerable reorganization caused by the rearrangement of the water solvent molecules. Comparisons are made with previous theoretical studies of amide–water interactions and experimental spectroscopic, X-ray and neutron-diffraction data on the hydration of amides, peptides and proteins.

Computational methods

The HONDO-7[26] suite of programs was used to perform the *ab initio* quantum mechanical calculations at the Hartree–Fock level and at the second-order Møller–Plesset perturbation (MP2) level on a multiprocessor MIPS computer at the Computer Center of the University of Ioannina after being modified for a Unix environment. Geometry optimizations were carried out for *trans*-NMA in C_1 symmetry with the split valence basis set 4-31G plus polarization functions. This basis set is known to give satisfactory results for hydrogen-bond geometries[27] with small basis set superposition error.[26–28] A further exploration of the basis set size effect on the relative stability of the NMA–water complexes was to add p orbitals at the hydrogen atoms (4-31G**), as suggested by one of the referees. The initial configuration of the complex with one water molecule at the amide oxygen and one water molecule at the NH group was that reported by Mirkin and Krimm,[27,29] subsequently optimized with the total relaxation of all the parameters. The starting geometry for the two water molecules at the amide oxygen of the NMA was obtained by modifying either the previously obtained optimal structure by placing the NH-bound water at the CO site, or by using a structure similar to that reported by Guo and Karplus,[22a] subsequently optimized with total relaxation of all the parameters. An *ab initio* minimum energy search was performed and after 23 gradient evaluations, a satisfactory local minimum was reached (maximum gradient, 1×10^{-3}). The HF 4-31G* local minimum was the geometry input for the HF 4-31G** geometry optimizations. At the MP2 level of theory, the lowest-energy occupied (C, O, N, 1s-like) molecular orbitals were held doubly occupied (7 frozen cores) and the 40 lowest-lying virtual orbitals were included (HF 4-31G* gradient optimized geometry). Furthermore, calculations on *trans*-NMA in the presence of two acetonitrile molecules were performed. The initial geometries were obtained by performing a small-scale molecular dynamic calculation (MD) of an NMA molecule bathed in eight acetonitriles at the low kinetic-energy regime, after an equilibration period of 10 ps. Subsequently, a set of selected low-energy events (clusters) was subjected to energy minimization. The resulting lowest-energy complex with the two closest neighbouring acetonitrile molecules was chosen as the starting geometry for the 4-31G* *ab initio* geometry optimization.

Results and Discussion

Ab initio calculations of amide–water complexes

The optimized intra- and inter-molecular geometric parameters of our calculations for **A**, **B** and **C** are presented in Tables 1 and 2 and the energy-optimized structures are illustrated in Fig. 1. Hydration increases the CO distance in all models, but mostly in **C**. The N—C bond is shortened mostly in **A** and **C**. The N—C bond remains unchanged in **A**, it is lengthened in **B** and **C** and it is shortened in **D**. The CC distance is not affected in **A** but it is shortened in **B** and **C**. At the HF level with the 4-31G* basis set, **B**, with two molecules of water forming a loose ring cluster at the CO site, is energetically favoured by 1.43 and 0.36 kcal mol^{-1} compared with **A** and **C**, respectively (Table 3), while p orbitals on all hydrogen atoms (4-31G** basis set, 160 basis-set size, all geometrical parameters optimized) enhanced the preference of **B**. At the MP2 perturbation level, **C** is energetically favoured by *ca.* 2.45 and 0.25 kcal mol^{-1} relative to **B** and **A**, respectively. Interestingly, Evanseck and Karplus[22] noted that for the $NMA \cdot 3H_2O$ cluster (with two molecules of water at the amide oxygen and one at the NH group) the global energy minimum may correspond to a conformation in which the two water molecules interacting with the CO group are also hydrogen bonded to each other to form a non-planar, three-membered ring structure with relatively weak CO···H hydrogen bonds, as suggested by AM1 calculations.[22b] Taking into account that the equivalent out-of-plane ring structure, **B**, (Fig. 1) could be at least four-membered, by simple reasoning in terms of symmetry (*e.g.* with water molecules arranged above and below the peptide plane, and to the left and right of the plane along the CO axis normal to the peptide plane), it is expected that the Boltzmann distribution of this type of clustering, at room temperature, is *ca.* 5%. In **A**, **B** and **C**, the positions of the water molecules at the carbonyl oxygen are such that the hydrogen not involved in the hydrogen bond points away from the C-methyl group. In **A** the HOH bisector of the water molecule hydrogen bonded to the NH group makes an angle of 178.8° with the H···O direction. NMA shows a slightly non-linear hydrogen bond at the NH group, with ar. angle of 165.4° and an H···O hydrogen-bond length of 2.1 Å. At the oxygen atom, the hydrogen bond is slightly non-linear. The O···HO angle is 169.4° and the O···H hydrogen-bond length is 1.987 Å. In **B**, the primary water

Table 1 Energy optimized HF/4-31G*, HF/4-31G** intramolecular geometric parameters and out-of-plane geometric parameters of isolated *t*-NMA, **A**, **B**, **C** and **D**

		A			B		C		D
parameter[a]	isolated *t*-NMA[d]	4-31G*	4-31G**	ref. 27	4-31G*	4-31G**	4-31G*	4-31G**	4-31G*
CC	1.514 (1.513)	1.512	1.511	(1.511)	1.509	1.506	1.506	1.508	1.511
CO	1.198 (1.199)	1.209	1.208	(1.209)	1.210	1.209	1.214	1.213	1.210
CN	1.348 (1.348)	1.333	1.335	(1.336)	1.338	1.338	1.333	1.334	1.338
NH	0.991 (0.992)	0.995	0.995	(0.996)	0.994	0.991	0.992	0.991	0.991
NC	1.444 (1.444)	1.445	1.446	(1.444)	1.448	1.446	1.450	1.450	1.439
(N)CH[b]	1.080 (1.081)	1.081	1.082	(1.081)	1.081	1.082	1.079	1.081	1.082
(N)CH(1)[c]	1.082 (1.082)	1.079	1.084	(1.082)	1.080	1.084	1.081	1.082	1.081
(N)CH(2)[c]	1.082	1.082	1.081		1.081	1.082	1.078	1.080	1.081
(C)CH[b]	1.084 (1.083)	1.083	1.083	(1.081)	1.083	1.083	1.083	1.083	1.078
(C)CH(1)[c]	1.083 (1.083)	1.083	1.082	(1.083)	1.084	1.083	1.083	1.082	1.085
(C)CH(2)[c]	1.082	1.083	1.084		1.082	1.082	1.080	1.083	1.085
CCN	115.9 (116.2)	116.3	116.1	(116.5)	117.0	116.7	116.7	116.9	115.2
OCN	122.4 (122.3)	123.2	123.4	(122.2)	122.2	122.2	122.3	122.2	122.4
CNH	119.4 (119.7)	120.2	120.6	(119.7)	118.9	118.9	118.8	118.9	119.3
CNC	121.6 (121.3)	121.8	121.9	(121.3)	121.8	121.8	122.4	122.3	121.7
NCH[b]	108.5 (108.7)	108.1	108.3	(108.5)	108.3	108.6	108.3	108.4	108.5
NCH(1)[c]	110.9 (110.9)	111.5	111.5	(111.1)	111.2	111.5	111.0	111.3	111.2
NCH(2)[c]	111.1	110.8	111.2		110.6	110.5	110.9	110.9	111.2
CCH[b]	113.6 (113.7)	113.0	113.4	(113.0)	113.6	113.2	113.1	113.1	109.1
CCH(1)[c]	108.4 (108.5)	108.6	108.5	(108.6)	108.4	107.4	108.8	108.6	110.5
CCH(2)[c]	108.7	108.5	108.3		108.1	108.7	108.8	108.7	110.5
CNCH[b]	179.4 (180.0)	179.2	−172.8	(180.0)	−179.2	176.6	175.9	175.8	177.3
CNCH(1)[c]	59.6 (60.2)	59.9	67.9	(60.2)	−59.3	−63.2	−64.0	−64.4	−63.2
CNCH(2)[c]	−60.9 (−60.2)	−60.5	−52.9	(−60.2)	61.6	57.3	56.1	55.7	57.8
NCCH[b]	−1.0 (0.0)	−1.2	−7.1	(0.0)	−0.6	11.9	−5.9	−5.1	−176.8
NCCH(1)[c]	−122.8 (−121.7)	−122.9	−128.9	(−122.0)	−121.8	−108.8	116.0	−127.2	−56.5
NCCH(2)[c]	120.3 (121.7)	120.3	114.2	(122.0)	120.8	133.9	−128.0	111.7	63.1
OCNH(ω_2)	−179.1 (180.0)	−178.5	−179.8	(180.0)	−174.7	−175.9	−179.1	−179.2	179.0
OCNC(ω_3)	1.0 (0.0)	0.9	0.0	(0.0)	−2.7	−3.1	−0.3	−0.3	0.7
CCNH(ω_4)	1.3 (0.0)	1.8	1.2	(0.0)	6.7	3.9	1.3	1.2	−1.4
CCNC(ω_1)	−178.6	−178.7	−179.8		178.7	176.6	−180.0	−180.0	−179.6
θ_C		0.4	0.2		1.4	−0.3	0.3	0.3	−0.3
θ_N		0.5	1.0		8	7.3	1.4	1.2	−0.4

[a] Bond lengths in Å; bond angles in degrees. [b] In-plane H atoms. [c] Out-of-plane H atoms. [d] Present work. The values in parentheses were obtained from ref. 27.

molecule W1 (Fig. 1) exhibits a hydrogen bond with a larger deviation from linearity compared with that in **A**, with the (C)O···HO angle being 158.7° and the hydrogen-bond length being 1.963 Å. This agrees with studies which found that when water acts as the proton donor, the hydrogen bonds tend to deviate from linearity. The CO···H(W) angle is 133° and 130.9° in **B** and **A**, respectively. Since Mirkin and Krimm[27] found an angle of 115.5° for **A**, it appears that the potential energy does not change for a wide range of CO···H(W) angles. Karplus and co-workers, in a recent comprehensive analysis of the angular potential-energy curves of the monohydrated NMF, reported a broad energy minimum around 130° for this angle. They also reported that the (C)O···HO(W), NH···O(W) and (N)H···OH angles indicate a considerable degree of rotational freedom (±30°). It appears that the flattening of the potential-energy curve with respect to the CO···H(W) angle is related to the additional (bound) water molecule.

The geometry optimization at 4-31G* and at 4-31G** basis set levels produced minor intramolecular differences. There is

Table 2 Intermolecular geometric parameters of isolated *t*-NMA, **A**, **B**, **C** and **D**

		A[b]		B[b]		C[c]		D
parameter[a]		4-31G*	4-31G**	4-31G*	4-31G**	4-31G*	4-31G**	4-31G*
O···H(W)	(1.957)	1.987	1.983	1.963 [W(1)]	1.960 [W(1)]	1.983 [W(1)]	2.014 [W(1)]	
				2.480 [W(2)]	2.588 [W(2)]	2.025 [W(2)]	2.026 [W(2)]	
O···O(W)	(2.902)	2.913	2.912	2.872 [W(1)]	2.871 [W(1)]	2.923 [W(1)]	2.940 [W(1)]	
				3.163 [W(2)]	3.205 [W(2)]	2.971 [W(2)]	2.967 [W(2)]	
(N)H···O(W)	(2.089)	2.102	2.123					
N···O(W)	(3.085)	3.076	3.100					
CO···H(W)	(115.5)	130.9	135.2	133.0 [W(1)]	130.0 [W(1)]	112.8 [W(1)]	113.5 [W(1)]	
				91.6 [W(2)]	94.0 [W(2)]	131.8 [W(2)]	133.6 [W(2)]	
O···HO(W)	(172.5)	169.4	165.8	158.7 [W(1)]	160.1 [W(1)]	168.0 [W(1)]	165.2 [W(1)]	
				128.9 [W(2)]	123.3 [W(2)]	172.1 [W(2)]	171.3 [W(2)]	
NH···O(W)	(179.3)	165.4	116.8					
(N)H—N[(CH$_3$CN(1)][d]								3.546
(N)H—N[CH$_3$CN(2)][d]								3.432
(C)O—C[CH$_3$CN(1)][d]								3.275
(C)O—C[CH$_3$CN(2)][d]								3.405

[a] Bond lengths in Å; bond angles in degrees. [b] Present work. The values in parentheses were obtained from ref. 27. [c] Present work. Starting geometry similar to that of ref. 22(*a*). [d] See Fig. 1. Method of geometry optimization as in Table 1.

Table 3 Calculated optimized total energies in E_h (a) and relative energies in kcal mol^{-1} (b) resulting from HF *ab initio* calculations with 4-31G* and 4-31G** (values in parentheses) basis sets of **A**, **B** and **C**

	HF		MP2	
cluster	(a)	(b)	(a)	(b)
A	−398.666 027 (−398.708 658)	1.430 (1.455)	−399.121 244	0.252
B	−398.668 308 (−398.710 977)	0.000 (0.000)	−399.117 744	2.447
C	−398.667 734 (−398.709 755)	0.360 (0.766)	−399.121 645	0.000

The MP2 energies are at HF/4-31G* energy optimized geometries.

an elongation of N(CH) by the out-of-plane H atoms of *ca.* 0.044 Å (**A** and **B**), an elongation of C(CH) by the out-of-plane H atoms of *ca.* 0.033 Å (**C**), a maximum change of ±0.6° in angles and some important differences related to the N-methyl and C-methyl hydrogen torsion angles (Table 1); these effects are due to the p orbitals at the hydrogen atoms of the 4-31G** basis set. There is an elongation of the O···H(W) distance of 0.03 Å of **C**, while for **A** and **B**, the primary hydrogen-bonding distance is unaffected. The intermolecular angles CO···H(W) and O···HO(W), as presented in Table 2, have a maximum deviation of 6°. A partial optimization of the C=O distance, keeping all other parameters frozen at the MP2/4-31G* level, yielded an elongation of 0.02 Å as a measure of the electron correlation effect on the C=O bond (**B**) in good agreement with a recent carbonyl—water hydrogen-bonding study of formaldehyde.[31]

Table 4 contains the Mulliken population results and the electronic charge redistribution within NMA, resulting from the presence of bound molecules of water. The MP2 Mulliken populations at 4-31G* HF optimized geometries are similar to 4-31G* HF population results to within 0.002 Å. The amide oxygen appears to gain a small amount of charge when functioning as a proton acceptor. The carbonyl carbon becomes more electronegative when a hydrogen bond is formed with the CO group. This is also the case for the nitrogen atom in **A**, where the NH group is involved in hydrogen bonding, but not for **B** and **C**. As expected, the overall electron transfer takes place from the proton acceptor (NMA) to the proton donor (W1). The overall electron transfer is small and appears to increase very little in the presence of the second molecule of water (W2) hydrogen bonded to W1; it is 5% and 3% for **C** and **B**, respectively. In the presence of water there is a reversal in the relative basicity between the in-plane and out-of-plane hydrogens of the $(CH_3)C$ group. The overall electron transfer of 4-31G** basis set Mulliken populations is similar to 4-31G* results, although the intramolecular distribution is markedly different.

Fig. 2 shows the gradient following path (GFP) resulting from full gradient HF/4-31G* optimization of **A**, **B** and **C** as a function of the indicated distances. The top diagram of Fig. 2 shows the monotonic energy decrease against the sequence of the conformations of **A** and **B** (**C** is omitted for clarity). The region explored is *ca.* 1 kcal above the minimum achieved for all structures, **A**, **B** and **C**. The ability of the weakly bound

Table 4 Mulliken electron populations of solvent clusters of NMA

		A			B		C		
atoms (method)	NMA (isolated) HF(1)	HF(2)	MP2	A[a] HF(1)	HF(2)	MP2	HF(2)	MP2	D HF(1)
(C)C	6.559	6.366	6.556	6.578	6.375	6.563	6.372	6.571	6.580
C	5.254	5.264	5.239	5.243	5.241	5.215	5.248	5.223	5.239
N	7.783	7.709	7.812	7.815	7.681	7.770	7.681	7.770	7.744
H	0.627	0.671	0.558	0.555	0.712	0.617	0.711	0.617	0.614
O	8.617	8.665	8.665	8.660	8.679	8.681	8.678	8.679	8.672
(N)C	6.257	6.091	6.280	6.259	6.088	6.270	6.095	6.283	6.248
H[b][(CH_3)C]	0.841	0.859	0.801	0.829	0.822	0.800	0.832	0.780	0.798
H[b][(CH_3)C]	0.841	0.859	0.803	0.829	0.856	0.774	0.837	0.774	0.802
H[c][(CH_3)C]	0.798	0.889	0.838	0.791	0.900	0.834	0.898	0.835	0.807
H[c][(CH_3)N]	0.839	0.901	0.840	0.836	0.890	0.830	0.890	0.830	0.830
H[b][(CH_3)N]	0.814	0.877	0.816	0.820	0.862	0.803	0.859	0.800	0.827
H[b][(CH_3)N]	0.814	0.852	0.793	0.820	0.867	0.811	0.847	0.786	0.829
C(CN)[CH_3CN(1)][d]									6.514
C(N)[CH_3CN(1)][d]									5.679
N[CH_3CN(1)][d]									7.487
H[c][CH_3CN(1)][d]									0.789
H[b][CH_3CN(1)][d]									0.769
H[b][CH_3CN(1)][d]									0.763
C(CN)[CH_3CN(2)][d]									6.521
C(N)[CH_3CN(2)][d]									5.677
N[CH_3CN(2)][d]									7.487
H[b][CH_3CN(2)][d]									0.781
H[b][CH_3CN(2)][d]									0.753
H[c][CH_3CN(2)][d]									0.776
O[W(2)]		8.619	8.913		8.644	8.895	8.660	8.910	
O[W(1)]		8.664	8.862		8.666	8.916	8.662	8.914	
H[W(2)]		0.678	0.521		0.685	0.563	0.660	0.527	
H[W(2)]		0.678	0.588		0.684	0.560	0.702	0.584	
H[W(1)]		0.653	0.556		0.653	0.521	0.702	0.532	
H[W(1)]		0.706	0.559		0.695	0.577	0.666	0.585	

[a] Ref. 27. [b] Out-of-plane. [c] In-plane. [d] See Fig. 1. HF(1) ⇒ HF/4-31G*. HF(2) ⇒ HF/4-31G**.

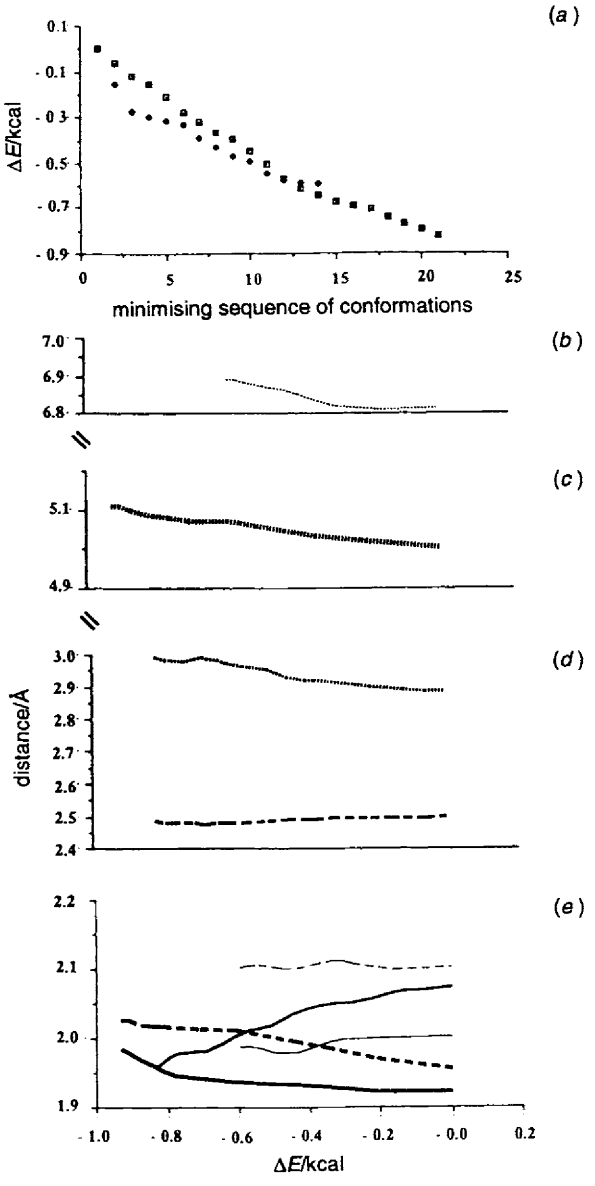

Fig. 2 Gradient following path (GFP) resulting from full gradient optimization of **A**, **B** and **C** as a function of (*a*) the minimizing sequence of conformations and (*b*)–(*e*) the indicated distances. (*a*) **A** (◆), **B** (⊡). (*b*) **A** O(W1)···O(W2). (*c*) **C** O(W1)···O(W2). (*d*) **B** O(W1)···O(W2) (dotted line), O(C)···H(W2) (thick dashed line). (*e*) **A** H(N)···O(W2) (thin dashed line), O(C)···H(W1) (thin continuous line); **B** O(C)···H(W1) (continuous line); **C** O(C)···H(W2) (thick dashed line), O(C)···H(W1) (thick continuous line).

molecule of water W2 to enhance the hydrogen bond of the primary molecule of water W1 is shown in Fig. 2 where the E[CO···H(W1)] and E[CO···H(W2)] curves of **B** are depicted. While E[CO···H(W1)] is strongly attractive, E[CO···H(W2)] is nearly flat at a distance of *ca.* 2.5 Å. There is a compensating effect as the W1–W2 interaction is repulsive in this region around the stationary point of the potential-energy hypersurface. **C** exhibits a close to minimum repulsive region for the hydrogen bonding of the two water molecules towards the carbonyl oxygen. Similarly, the E(water–water) curve favours a maximization of the intermolecular distance. The potential curves of Fig. 2 for **A** do not show a cooperative or anti-cooperative relationship between the W1 and W2 as they approach the (N)H and (C)O sites of NMA, respectively. All NMA–water clusters studied exhibit overall repulsive water–water interactions (the curves for **A** and **C** being near monotonically repulsive).

The effects of the bound molecules of water in the optimized out-of-plane geometric parameters of the amide bonds are clearly demonstrated in Table 1. The special arrangement of a non-planar amide group can quantitatively be described by the following four torsion angles defined by Ramachandran and co-workers:[24,32]

$$\omega(C_i^\alpha C'N_{i+1}C_{i+1}^\alpha) = \omega_1$$
$$\omega(OC'N_{i+1}H) = \omega_2$$
$$\omega(OC'N_{i+1}C_{i+1}^\alpha) = \omega_3$$
$$\omega(C_i^\alpha C'N_{i+1}H) = \omega_4 \quad (1)$$

The angle ω_1 is the conventional ω rotation, which has the value 180° for a planar *trans*-peptide unit. Furthermore, following Winkler and Dunitz,[33] we may denote the non-planarity of the three bonds at the C′ and N atoms by the descriptive notations:

$$\theta_{C'} = \theta(C_i^\alpha, O; C'N) = \omega_1 - \omega_3 + \pi \quad (2)$$

$$\theta_N = \theta(C_2^\alpha, H; NC') = -\omega_1 + \omega_4 + \pi \quad (3)$$

$\theta_{C'}$ is the angle of rotation (positive in the clockwise sense when looking from N towards C′) from the NC′C_1^α plane to the NC′O plane, the angle being less than 180°. Likewise, θ_N is the angle of rotation (positive when clockwise, looking from C′ towards N) from the C′NC_2^α plane to C′NH.

A shows very small pyrimidicity parameters $\theta_{C'}$ and θ_N and dihedral angle ω_2. **B**, with two molecules of water forming a loose ring cluster at the amide oxygen, exhibits a very small $\theta_{C'}$ value, while θ_N and ω_2 (which are of the opposite sign) have moderate but significant values (Table 1, Fig. 3). The basis set size does not produce significant differences for θ_N and θ_C values. An analysis of crystal structure data of a limited number of peptides[24] indicates that the pyrimidicity parameter θ_N and the dihedral angle ω_1 have significant values and are correlated (θ_N is most likely to have a value equal to $-1.5--2.0\omega_1$), while $\theta_{C'}$ is small and is uncorrelated with ω_1. Furthermore, our results are consistent with the general observation of crystallographers, who find that the deviations of the atoms, from the root mean square C_i^α, C′, O, N, C_{i+1}^α plane, are large for the hydrogen atom of H(N), which is on the same side of the plane as atom C_{i+1}^α. These results are in agreement with our *ab initio* calculations for the complexes *t*-NMF–(H_2O) (1 : 3)[34] and *t*-NMF–(H_2O) (1 : 8) with initial coordinates taken from MM2(87) geometry optimization. Again, θ_N and ω_2 parameters have significant values,

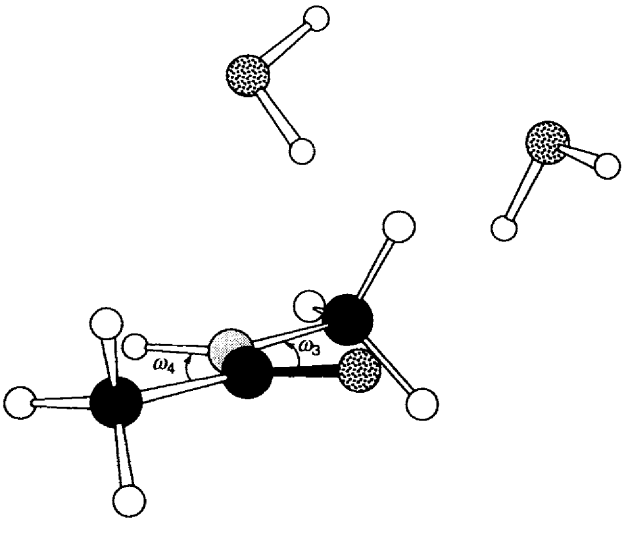

Fig. 3 Newman projection along C′—N amide bond of the *t*-NMA–(H_2O) (1 : 2) complex (**B**)

Table 5 Geometries for computed amide–water complexes and experimental main chain C=O···H_2O hydrogen bonds

complex	method	C=O···O(W) angle/degrees	O···O(W) distance/Å
1*t*-NMA + $2H_2O$ (one proton donor + one proton acceptor)[a] (**A**)	4-31G*	129	2.91
1*t*-NMA + $2H_2O$ (proton donors)[a] (**B**)	4-31G*	130 (81)	2.87 (3.16)
1*t*-NMA + $2H_2O$ (proton donors)[a] (**C**)	4-31G*	108.9	2.923
		129.3	2.971
1FMA + $1H_2O$ (proton acceptor)	MBS	92	2.81
(*cis* to the formyl hydrogen)[b]	3-21G	121	2.83
	6-31G*(d)	121	2.83
	OPLS	144	2.73
1FMA + $1H_2O$ (proton acceptor)	MBS	108	2.81
(*trans* to the formyl hydrogen)[b]	3-21G	104	2.77
	6-31G*	104	2.77
	OPLS	105	2.73
1NMA + $1H_2O$ (proton donor)	MBS	110	2.81
(*cis* to the formyl hydrogen)[b]	OPLS	142	2.72
1NMA + $1H_2O$ (proton donor)	MBS	140	2.81
(*trans* to the formyl hydrogen)[b]	OPLS	142	2.72
1*t*-NMF + $1H_2O$ (proton donor)	6-31G*	115	2.94
(*cis* to the formyl hydrogen)[c]	OPLS	147	2.73
	MOD	150	2.72
	6-31G*	134	2.94
	OPLS	142	2.73
	MOD	142	2.72
1*c*-NMF + $1H_2O$ (proton donor)	6-31G*	114	2.94
(*trans* to the formyl hydrogen)[c]	OPLS	143	2.73
	MOD	144	2.72
	6-31G*	111	2.90
	OPLS	107	2.73
	MOD	108	2.72
1*t*-NMA + $2H_2O$ [1 H_2O proton donor (CO) + 1 H_2O proton acceptor (NH)][d]	4-31G*	117	2.902
1*c*-NMA + $2H_2O$ [1 H_2O proton donor (CO) + 1 H_2O proton acceptor [NH)][d]	4-31G*		2.910
1*t*-NMA + $1H_2O$ (proton donor)[e]	6-31G*		2.77
1*c*-NMA + $1H_2O$ (proton donor)[e]	6-31G*		2.75
C=O bound to 1 H_2O [f]	X-ray	137 ± 16	2.85 ± 0.12
C=O bound to 1 H_2O [g]	X-ray and neutron diffraction	133 ± 14	2.91 ± 0.26
C=O bound to 2 H_2O [h]	X-ray and neutron diffraction	130 ± 16	2.94 ± 0.33

[a] Present work. The values in parentheses refer to the secondary molecule of water W2 (Fig. 1). [b] Ref. 20(*a*). [c] Ref. 20(*b*). [d] Ref. 27. [e] Ref. 39; 6-31G* *ab initio* electrostatic potential derived charges in molecular mechanics force fields. [f] 58 cases from 37 peptide crystal structures (ref. 2). [g] 308 cases from 15 protein crystal structures (ref. 38). [h] 258 cases from 15 protein crystal structures (ref. 38).

particularly for the *t*-NMF–(H_2O) (1 : 8) complex.† Interestingly, the vibrational analysis of Mirkin and Krimm[27] indicated an out-of-plane CO bent force constant increase of 3% in going from the isolated NMA to dihydrated NMA–H_2O (1 : 1), while the NH out-of-plane bent force constant increases by 30%. Williams[35] found a change in the out-of-plane CO bent force constant of *ca.* 2% and an increase of nearly 45% for the out-of-plane NH bent force constant between the isolated and the trihydrated NMA (4-31G scaled to 4-31G* calculation). Recently Sulzbach *et al.*[36] produced NMR and theoretical studies supporting the fact that non-planarity of the amide moiety is related to amide nitrogen pyramidalization. Their work excluded consideration of solvation effects.

It is worthwhile to compare our results with those obtained from previous *ab initio* calculations and experimental data of hydrogen bonds (Table 5). Yang *et al.*,[2] in their comprehensive analysis of the geometry of peptide backbone hydration of peptide crystals, suggested that the most probable value of the C=O···O(W) angle is *ca.* 138°. In an idealized hydrogen-bond model, these angles would correspond to the angle between the oxygen–carbon bond and the axis of the sp^2 orbital containing the lone pair involved in a hydrogen bond, which is *ca.* 120°. The mean value for the 58 water–amide carbonyl bonds is significantly greater, *ca.* 138°; this value should be compared with our calculated value of *ca.* 130° for **A**, **B** and **C** (Table 5). The distribution of dihedral angles between the O(W)···OC and O—C—C^{α} planes indicates a preference for water molecules bound to the carbonyl group to lie near the plane of the peptide unit, as the idealized hydrogen-bond model predicts. Among 93 examples, 41% are within 20° of the peptide plane and only 15% are within 20° of the plane perpendicular to it. There are 77 examples of O(W)···OC hydrogen bonds in the Mitra and Ramakrishnan[37] study, and of these, 57% have the water oxygen within 20° of the plane of the carbonyl group. The water oxygen W1 of **A** is at 23.7° to the amide plane and the two water oxygens of **C** are at −5.0° and −170.1° to the plane of the amide group. The C—N···O(W) angle, which would be the same as the C—N—H angle for a linear N—H···O(W) bond, has a mean value of 128° compared with 123° for the C—N—H angle of the standard peptide bond; this is in agreement with our value of 129.6° in **A**. Furthermore, the distribution shows a minimum in the range 120–134°, a result similar to the observations of Mitra and Ramakrishnan[37] for O—H···O(W) hydrogen bonds. As expected, the water tends to be near to the amide plane. In 12 out of 20 cases, the O(W)···N—C and N—C—C^{α} planes are within 20° of each other.

Analysis of experimental hydrogen bonds between water molecules and protein main-chain functional groups[38] indicate that water molecules hydrogen bonded to the mainchain NH···O(W) angles in proteins are slightly non-linear (156 ± 15°) with 90% in the range 140–180°; the average H···O(W) and N···O(W) distances are 2.06 ± 0.2 and

† Manuscript in preparation.

2.97 ± 0.21 Å, respectively. These average experimental values are in good agreement with our theoretical data and with previous *ab initio* calculations with 6-31G*, 4-31G* and 4-31G** basis sets for amides (Table 5). Water molecules hydrogen bonded to the C=O groups associate preferentially in the direction of the oxygen lone pair orbitals but are not restricted to it. This is in agreement with our calculations. Furthermore, the experimental mean C=O···O(W) angle and O···O(W) distances (Table 5) are similar to our calculated values. In accord with the general trend observed in the hydrogen bonding from a crystal data base[38] and theoretical calculations of Guo and Karplus[22] and Williams,[35] the hydrogen-bond distance involving a single acceptor, *i.e.* the CO group that only accepts one hydrogen bond, is generally shorter than that involving multiple acceptors. Owing to the crystal-field distortion, the hydrogen-bond lengths tend to be shorter in crystals than those observed experimentally or predicted theoretically in hydrogen-bonded dimers. (The most obvious example of this difference is that between the values for the H···O distances of 1.7 to 1.8 Å, observed in ice, and 2.0 Å for the water dimer.[38]) The agreement, therefore, between our results and experimental X-ray data indicates that the deviation from linearity of the N—H···O(W) bond angles and the slight deviation from the directions of the conventionally viewed oxygen sp^2 lone pairs of the C=O···H(W) bond, might not be due exclusively to crystal-field distortion and sterical constraints imposed by the folding of the polypeptide chain. Very probably, the intrinsic hydrogen-bond properties of the bound molecules of water to the amide oxygen play an additional role.

A force-field calculation using the HF 4-31G* basis set was performed for **B**. The complex was found to yield no imaginary frequencies. Table 6 summarizes the results of the frequency calculation. The scale factors are taken from Mirkin and Krimm[27] since they used 4-31G* basis set on their **A** complex. **A** and **C** (of C_s symmetry) are reported to be at true minima.[22,27] The zero-point vibrational energy (ZPVE) of **B** is calculated to be 102.115 kcal mol^{-1}. The C=O stretch is concentrated into two modes, both at 1636 cm^{-1}, as experimentally found by Williams,[35] one mainly an H_2O bend with a C=O stretch perturbation (calculated frequency, 1634 cm^{-1}) and the second mostly a C=O stretch contaminated with an H_2O bend contribution (calculated frequency, 1633 cm^{-1}). The associated eigenvectors of the 1636 cm^{-1} modes were inputted to the Winvib[40] visualization package; Fig. 4 provides a picture of these two vibrational modes. It seems that the findings of Chen *et al.*[41] can be extended to be non-coplanar to the peptide bond water molecules, since the two out-of-plane water molecules of **B** are strongly vibrationally coupled to the C=O group of *t*-NMA.

Ab initio calculations of the amide–CH_3CN (1 : 2) complex

The energy-optimized structure of our calculation for the **D** complex is illustrated in Fig. 1 and the resulting intramolecu-

Table 6 Frequencies for *trans*-NMA in water (**B**)

frequency/cm^{-1}			
observed[a]	calculated[b]	deviation	description[c]
—	10	—	$2H_2O$, CC t
—	16	—	$2H_2O$, NC t
—	74	—	$2H_2O$, CC t
—	94	—	$2H_2O$, CC t
—	126	—	CC t
—	157	—	$2H_2O$, NC t
171	171	0	NC t, H_2O
—	177	—	NC t, CO ob, NH ob, $2H_2O$
195	204	9	CN t, CC t, $2H_2O$, NH ob
297	275	−22	$2H_2O$, CNC sd, CO sd, CO r
297	282	−17	CNC sd, CO sd, CO r, $2H_2O$
443	438 (ν_1)	−5	CO sd, CO r, CNC sd, NH ob, CN t, NC t, CCH_3 r1
443	445 (ν_2)	2	NH ob, CO sd, CO r, CNC sd, CN t, NC t, CCH_3 r1
630	625 (ν_3)	−5	CO r, CC s, NC s, CNC sd, CCH_3 r1, NCH_3 r1, H_2O
—	648 (ν_4)	—	CO ob, CCH_3 r2, NH ob, $2H_2O$
883	878 (ν_5)	−5	NCH_3 r1, CC s, CN s, NC s, CO r, CCH_3 r1, CNC sd
994	983 (ν_6)	−11	NC s, CCH_3 r1, CC s, CO s
1044	1049 (ν_7)	5	CCH_3 r2, CO ob
1095	1092 (ν_8)	−3	CCH_3 r1, NC s, NH ib, NCH_3 r1, CO r, CC s
1129	1125	−4	NCH_3 r2
1164	1164 (ν_9)	0	NCH_3 r1, CNC sd, NH ib
1312	1308 (ν_{10})	−4	NH ib, CN s, CCH_3 sd, CC s, CO r
1378	1378	0	CCH_3 sd
1416	1425 (ν_{11})	11	NCH_3 sd, CO s
1428	1435 (ν_{12})	7	CCH_3 ad1, NH ib
1440	1434	−6	CCH_3 ad2
1452	1446	−6	NCH_3 ad2
1462	1470	8	NCH_3 ad1
1577	1567 (ν_{13})	−10	CN s, NH ib, CO r, NCH_3 r1, NC s, CC s
1636	1634 (ν_{14})	−2	$2H_2O$ b (out of phase), CO s
1636	1633 (ν_{15})	−3	CO s, NH ib, CO sd, CN s, CCH_3 r1, NCH_3 sd, $2H_2O$ b
2925	2929	4	CCH_3 ss
2942	2928	−14	NCH_3 ss
2973	2988	15	NCH_3 as2
2992	2994	2	CCH_3 as
3009	3012	3	CCH_3 as
3020	3008	−12	NCH_3 as1
3120	3120 (ν_{16})	0	NH s

[a] From ref. 35. [b] HF/4-31G*. [c] t, torsion; r, in-plane bend; r1, rock (A′); r2, rock (A″); b, bend; ob, out-of-plane bend; ib, in-plane bend; s, symmetric; ss, symmetric stretch; as1, antisymmetric stretch (A′); as2, antisymmetric stretch (A″); sd, symmetric deformation (A′); ad1, antisymmetric deformation (A′); ad2, antisymmetric deformation (A″); A′, in-plane; A″, out-of-plane. For frequencies ν_1–ν_6, ν_{10}, ν_{12}, the scale factors (s_f) are averages of the contributions (as in ref. 27); the H_2O was ignored. For frequencies ν_{14}, ν_{16}, s_f are based on the analysis of ref. 35.

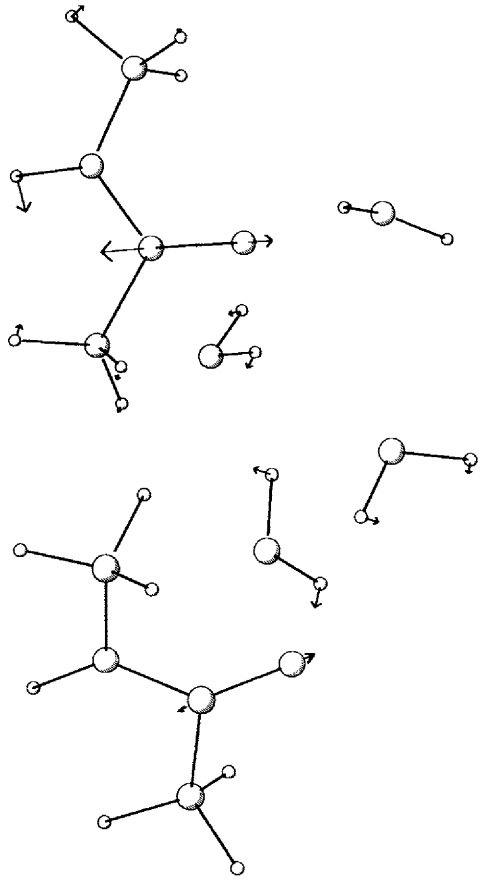

Fig. 4 A view of the C=O stretching modes of the *t*-NMA–H_2O (1 : 2) complex (**B**)

lar and intermolecular geometric parameters in Tables 1 and 2. The two molecules of acetonitrile are oriented in a nearly parallel way above and below the amide plane. However, the two sites are not stereochemically equivalent. Note that this structure corresponds to the most favourable orientation of the overall dipoles. In contrast, for NMA–H_2O complexes, the directional influence of the amide-oxygen lone pairs is evident and the orientation of the molecules of bound water does not correspond to the most favourable orientation of the overall dipoles.

There are two thermodynamically stable phases for solid acetonitrile. In the high-temperature phase, the nearest neighbours are antiparallel to each other[42] whereas in the low-temperature phase the neighbouring molecules have a parallel orientation.[43] Furthermore, in the course of gas-phase deposition of CH_3CN onto a cold target (13 K), a solid is formed which consists exclusively of randomly oriented antiparallel dimers.[44] Experiments in liquid acetonitrile are less conclusive; however, a study of intermolecular associations from *ab initio* calculations of IR frequencies and intensities indicated the presence of cyclic and antiparallel clusters.[45] The intermolecular distances between the sp hybridized carbons of CH_3CN vary little in the linear (L) and antiparallel (A) clusters: from 3.29 (L_4) to 3.33 Å (L_2) and from 3.55 (A_2) to 3.62 Å (A_4), respectively. The corresponding distances for the cyclic (C) configuration are more sensitive to the size of the cluster: 5.80 Å for C_3 and 6.02 Å for C_4. Our intermolecular distances between the sp hybridized carbon of acetonitrile and the carbonyl carbon of NMA (3.35 and 3.32 Å) are in reasonable agreement with the intermolecular distances between the sp carbons of CH_3CN in antiparallel clusters.

Table 4 indicates that the overall electron transfer between NMA and CH_3CN is only 0.5% compared with 5 and 3% for **C** and **B**, respectively. Furthermore, in **D** the charges of the $(CH_3)_3C$ hydrogens become very similar, contrary to the case of **A**, **B** and **C**; the same applies for the $N(CH_3)$ hydrogens.

Molecular orbital analysis

The first canonical SCF HOMO orbital is mainly a two-centre (perpendicular to the peptide plane) π-antibonding orbital interaction between O and N of the NMA with some orbital contribution from the methyl group of the nitrogen. The second water molecule in the vicinity of the carbonyl oxygen of **B** induces, in addition to this [N, O, nitrogen's methyl] hybrid molecular orbital, the appearance of a π-antibonding orbital component coplanar to the peptide bond (Fig. 5). This effect is coupled to a moderate but significant departure from planarity of the HNCO angle ($\omega = 174.7°$, Table 1). The orbital energy of the first HOMO orbital for **D** and **A** is invariant to within 10^{-2} kcal which, presumably, can be attributed to the negligible interference of the solvent provided that the relative position of the solvent molecules towards the NMA molecule is similar. The proximity of two water molecules to the amide oxygen decreases the HOMO orbital energy. Spiro *et al.*,[46] using extensive ultraviolet resonance Raman (UVRR) spectroscopic studies of NMA, showed that the amide bond is strongly affected by the electron-acceptor (or hydrogen-bond donor) properties of its environment in both the ground and first $\pi \rightarrow \pi^*$ excited states. Furthermore, an extended Hückel theory model suggested that the HOMO, a three-centre N—C—O π-antibonding orbital, somehow alters its composition by diminishing the p-orbital carbonyl carbon contribution and by a subsequent lowering of the HOMO energy upon hydrogen-bond donor interactions. The present orbital analysis confirms the lowering of the HOMO energy when the water molecules approach the amide oxygen. In addition, the LUMO–HOMO energy gap for **B** and **C** is widened in comparison with **A**. The first HF HOMO of **C** has an orbital energy almost the same as the second HOMO of the isolated NMA molecule (Fig. 6). Its composition is similar to **A** except that it is in-plane with the amide bond.

The second canonical SCF HOMO orbital, Fig. 5, is again a solute hybrid orbital of p character mostly between the nuclei of the plane of the amide bond. There is a small orbital energy differentiation between the **D** and **A** complexes. The second HOMO of **B** has a significant σ-peptide bond orbital component plus a three-centre N—C—O π bond perpendicular to the peptide bond orbital. For **C** this is a hybrid four-centre orbital on the N—(CO)C fragment perpendicular to the peptide bond.

The third canonical SCF HOMO orbital is the first direct evidence of the solvent effect for **A**, **B**, **C** and **D**. This orbital is a combination of atomic orbitals centred on the solvent molecules. The fourth canonical SCF HOMO orbital is a combination of solvent atomic orbitals for **B**, **C** and **D**; for **A**, this is a solute π-bonding orbital. The fifth canonical SCF HOMO orbital shows solvent–solute interactions for **A** and **B**. The complex of NMA with CH_3CN is a solvent orbital, while **B** is a solute orbital. The sixth canonical SCF HOMO orbital is for all cases a solvent orbital. The seventh canonical SCF HOMO is a solute π-bonding orbital for the **D** complex. For **A** and **B** this is a solvent–solute orbital interaction, and it is a solvent orbital for **C**.

The solute-occupied π-bonding molecular orbital of the carbonyl fragment perpendicular to the peptide plane turns out to be a five-centre hybrid moiety, bonded between the N—C=O segment, plus a π-antibonding component on the two methyl carbons. It is located as the fourth HOMO for **A**, the seventh HOMO for **D** and fifth HOMO for **B** and **C**. The vertical orbital energy difference between the occupied π-

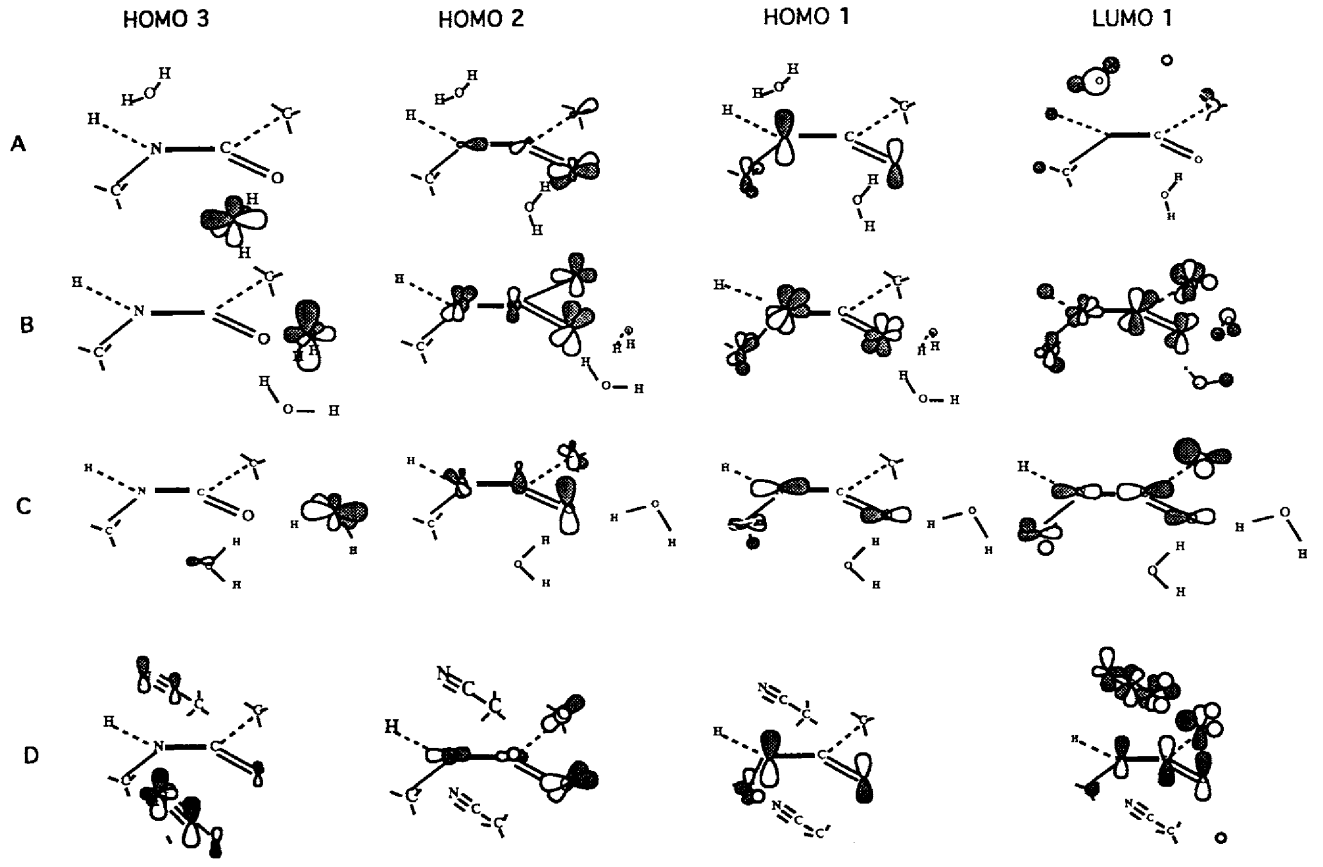

Fig. 5 Orbital diagram of three HOMOs (numbered 1 to 3) and of one LUMO of NMA illustrating the altered composition of the complexes discussed in the text

bonding {O, C and N} hybrid three-centre molecular π_1 orbital and the π-antibonding {O, C and N} π_2 orbital of *trans*-NMA seems to be independent of the nature of the solvent since, for both **A** and **D** complexes, it is 75 kcal. However, the equivalent bonding–antibonding π MO energy difference, $\pi_1 - \pi_2$, between **B** and **C** is narrowed to 71 and 70.3 kcal, respectively, mainly owing to interference of in-plane orbital interactions. The isolated NMA $\pi_1 - \pi_2$ MO energy separation is 72.7 kcal while for formamide it is reported to be 98.23 kcal.[45] The energy level of this π-bonding orbital is approximately equal for the **A** and **D** moieties but it is lowered for both *t*-NMA–(H_2O) (1 : 2) complexes. Hence, we can argue that this π-bonding orbital is affected by the geometry of the solvent both in composition and in energetic terms, while Spiro *et al.*[46] suggested that this remains unchanged by solvent hydrogen-bond interactions.

The lower virtual orbital space of **D** (except for the first LUMO) is dominated by the solvent orbitals. The lowest LUMO of **D** is the π^*-antibonding molecular orbital on the NCO segment and its equivalent is the second LUMO for **A** and is of more or less the same MO energy. This π^*-antibonding MO is located at the first LUMO for **B** and **D**, which energetically is lower than that of its counterparts for **A** and **D** (Fig. 6), and as the second LUMO of **A**. In addition, the second LUMO of isolated NMA is not present within the five lowest LUMOs of **A** and **D**, while it is located as the second LUMO for **B** and **C**. The π^* state, as depicted by the LUMO (Fig. 4), is more delocalized than previously suspected in accordance with the findings of Chen *et al.*[41] Pre-resonance Raman enhancement of the in-plane amide II and III modes and *cis*/*trans* isomerization led to the proposal that the amide π–π^* excited state resembles the π–π^* excited state of ethene, which is known to have a single minimum twisted 90° relative to the planar ground state.[47] Li *et al.*[48] have found, from *ab initio* 6-31G* calculations, that the excited states of NMA have pyramidal carbonyl and amine groups as well as multiple energy minima. They analysed the molecular orbitals of formamide which resemble the π orbitals of allyl systems; their π_2 is concentrated more heavily on N than on O, while π^* is a slightly perturbed carbonyl π^* orbital.

It is of interest to compare the composition of the π^* LUMO for all clusters studied. The relative ratios of the components on methyl(C), C, O, N and methyl(N) are: 2 : 2 : 1.5 : 1 : 0.5, for **A**; 6 : 5 : 4 : 2.3 : 2 for isolated NMA; 3.5 : 3 : 1.5 : 1 : 2 for **B**; and 3.5 : 3 : 2 : 1 : 1 for **D**.

The emerging conclusions are: (i) the π^* excited MO (LUMO) is more delocalized in agreement with the hypothesis of Chen *et al.*[41] (ii) The relative positions of the solvent molecules enhance the role of C or N methyl substitution. (iii) The directionality of the p orbitals of this π^* LUMO is dependent on the topology of the solvent molecules. Geometrical relaxation results in pyramidization of both the carbonyl and the amine groups for the excited states of the NMA;[48] numerous excited conformers have been investigated which have the common feature of a considerable C=O lengthening, although one of them simultaneously exhibited a C—N shortening. The pyramidal carbonyl group of NMA, studied *in vacuo*,[48] destabilizes the interaction between the carbonyl and amine groups present in the ground state, which becomes repulsive and, as a consequence, the amine group pyramidalizes. The ring structure of water molecules around the carbonyl group in **B**, induces a 6° pyramidalization on the amine group (Table 1), which might act as a precursor catalyst for the *trans*/*cis* isomerization.

From the above orbital analysis, it is evident that, first, the NMA solute-occupied orbital space is essentially unaffected both in terms of orbital compositions and orbital energies by the nature of the solvent in the case of the complexes **A** and **D**. Secondly, there is considerable solute-occupied orbital space reorganization caused by the rearrangement of the

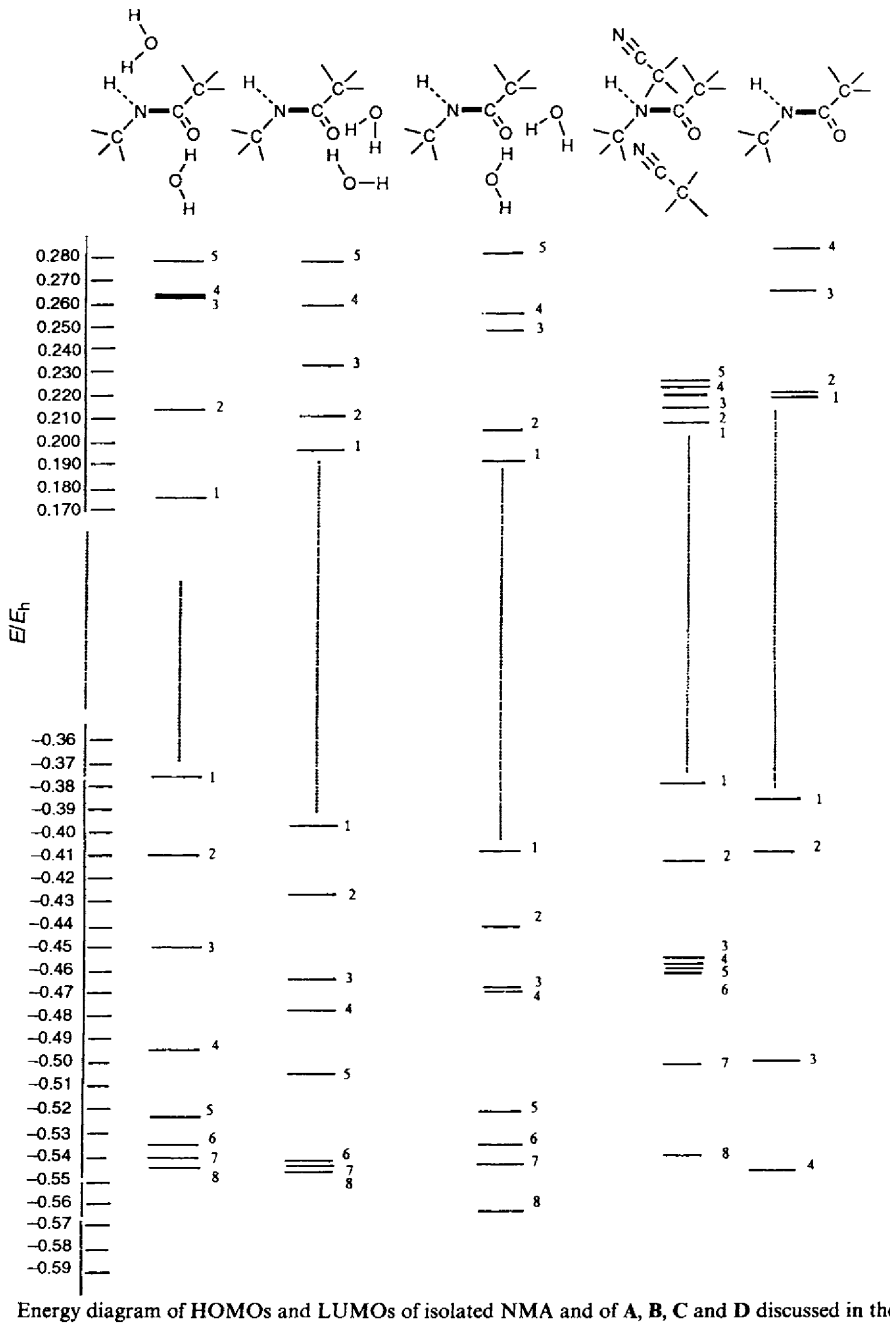

Fig. 6 Energy diagram of HOMOs and LUMOs of isolated NMA and of **A**, **B**, **C** and **D** discussed in the text

water solvent molecules. Thirdly, the solvent orbital space is well separated from the solute space in the case of acetonitrile, while the water orbitals spread evenly and mix well with solute orbitals. Finally, the unoccupied orbital space is subject to severe reordering that affects the first LUMO composition.

Conclusions

Ab initio molecular-orbital calculations of *trans*-NMA with two molecules of water at the amide oxygen atom demonstrate the importance of concurrent water–water and water–(carbonyl) oxygen hydrogen-bond interactions. The directional influence of the oxygen lone pairs in the geometry of C=O···H hydrogen-bond interactions is evident for water complexes, while for complexes with acetonitrile the geometry of the cluster is determined by the overall dipole–dipole interactions. Orbital analysis suggests that there is substantial solute-occupied space organization caused by the rearrangement of the water solvent molecules. Therefore, it may be advantageous to take HOMO–LUMO interactions into consideration, particularly when discussing hydration phenomena. Calculations at the HF level with 4-31G* and 4-31G** basis sets of *trans*-NMA with two molecules of water forming a ring cluster at the amide oxygen atom indicate the formation of a non-planar amide bond and the generation of a chiral centre at the amide nitrogen atom which is more stable than the formation of a nearly planar amide bond with two molecules of water forming hydrogen bonds along the direction of the oxygen lone-pair electrons; the stability of the molecular clusters is reversed at the MP2 level, which implies a Boltzmann distribution of the ring water cluster, at room temperature, of *ca.* 5%. The frequency vibrational analysis verified that this is a true minimum, since no imaginary frequencies have been found. The water molecules hydrogen-bonded to NMA are vibrationally coupled. The question of the generality of ring water clusters around the amide oxygen atom and the ready occurrence of a non-planar amide bond is, at present, difficult to answer. Nevertheless, such a phenomenon might be of considerable interest since it could influence the formation of regular protein structures in which a greater number of moderate non-planar amide groups mutually interact and the overall effect could be considerably enlarged.[49,50]

Furthermore, since changes in ω and θ_N values are usually not included in the empirical force field, this raises the question of whether the parameterization should include changes in ω and θ_N values or, alternatively, a force field that changes the internal parameters as a function of external perturbations.

Financial support from the Greek General Secretary of Research and Technology is gratefully acknowledged. Thanks are also due to the referees that made important suggestions incorporated in the revised manuscript.

References

1 I. D. Kuntz Jr and W. Kauzmann, *Adv. Protein Chem.*, 1974, **28**, 239.
2 C-H. Yang, J. N. Brown and K. D. Kopple, *Int. J. Pept. Protein Res.*, 1979, **14**, 12.
3 E. N. Baker and R. E. Hubbard, *Prog. Biophys. Mol. Biol.*, 1984, **44**, 97.
4 A. A. Kossiakoff, *Ann. Rev. Biochem.*, 1985, **54**, 1195.
5 L. Packer, *Methods in Enzymology*, 1986, **127**, 1.
6 W. Saenger, *Annu. Rev. Biophys. Biophys. Chem.*, 1987, **16**, 93.
7 I. P. Gerothanassis, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1994, **26**, 171.
8 L. A. La Planche and M. T. Rogers, *J. Am. Chem. Soc.*, 1964, **86**, 337.
9 W. E. Stewart and T. H. Sidall, *Chem. Rev.*, 1970, **70**, 517.
10 F. A. Bovey and F. P. Hood, *J. Am. Chem. Soc.*, 1966, **88**, 2326.
11 G. N. Ramachandran and C. M. Venkatachalam, *Biopolymers*, 1968, **6**, 1255.
12 R. Deslaurier and I. C. P. Smith, *Biol. Magn. Reson.*, 1980, **2**, 243.
13 G. Alagona, A. Pullman, E. Scrocco and J. Tomasi, *Int. J. Pept. Protein Res.*, 1973, **5**, 251.
14 A. Pullman and B. Pullman, *Quart. Rev. Biophys.*, 1975, **7**, 505.
15 W. C. Richards, in *Water: A Comprehensive Treatise*, ed. F. Franks, Plenum Press, New York, 1979, vol. 6, ch. 3, p. 123.
16 A. Pullman, G. Alagona and J. Tomasi, *Theor. Chim. Acta*, 1974, **33**, 87.
17 A. Johansson, P. Kollman, S. Rothenberg and J. McKelvey, *J. Am. Chem. Soc.*, 1974, **96**, 3794.
18 J. F. Hinton and R. D. Harpool, *J. Am. Chem. Soc.*, 1977, **99**, 349.
19 T. J. Zienlinski, R. A. Poirier, M. R. Peterson and I. G. Csizmadia, *J. Comput. Chem.*, 1983, **4**, 419.
20 (*a*) W. L. Jorgensen and C. J. Swenson, *J. Am. Chem. Soc.*, 1985, **107**, 1489; (*b*) W. L. Jorgensen and J. Gao, *J. Am. Chem. Soc.*, 1988, **110**, 4212.
21 P. G. Jasien and W. I. Stevens, *J. Chem. Phys.*, 1986, **84**, 3271.
22 (*a*) H. Guo and M. Karplus, *J. Phys. Chem.*, 1992, **96**, 7273; (*b*) J. Evanseck and M. Karplus, unpublished results.
23 F. Sim, A. St-Amant, I. Papai and D. R. Salahub, *J. Am. Chem. Soc.*, 1992, **114**, 4391.
24 G. N. Ramachandran and A. S. Kolaskar, *Biochim. Biophys. Acta*, 1973, **303**, 385.
25 G. LaManna, *Chem. Phys. Lett.*, 1983, **103**, 55.
26 HONDO 7 Dupuis M. IBM corporation; M. Dupuis, J. D. Watts, H. D. Villar and G. J. B. Hurst, *Comput. Phys. Commun.*, 1989, **52**, 415.
27 N. G. Mirkin and S. Krimm, *J. Am. Chem. Soc.*, 1991, **113**, 9742.
28 S. F. Boys and F. Bernandi, *Mol. Phys.*, 1970, **19**, 553.
29 N. G. Mirkin and S. Krimm, *J. Mol. Struct.*, 1991, **242**, 143.
30 S. Nakagawa, H-A. Yu, M. Karplus and H. Umeyama, *Proteins*, 1993, **16**, 172.
31 T. H. Ramelot, C. H. Hu, J. E. Fowler, B. J. DeLeeuw and H. F. Schaefer III, *J. Chem. Phys.*, 1994, **100**, 4347.
32 G. N. Ramachandran, A. Y. Lakshminarayanan and A. S. Kolaskar, *Biochim. Biophys. Acta*, 1973, **303**, 8.
33 F. K. Winkler and J. D. Dunitz, *J. Mol. Biol.*, 1971, **59**, 169.
34 I. P. Gerothanassis, I. N. Demetropoulos and C. Vakka, *Biopolymers*, 1995, **36**, 415.
35 R. W. Williams, *Biopolymers*, 1992, **32**, 829.
36 H. M. Sulzbach, P. v. R. Schleyer and H. F. Schaefer III, *J. Am. Chem. Soc.*, 1995, **117**, 2632.
37 J. Mitra and C. Ramakrishnan, *Int. J. Pept. Protein Res.*, 1977, **9**, 27.
38 G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures*, Springer–Verlag, Berlin, 1991.
39 P. Cielpak and P. Kollman, *J. Comput. Chem.*, 1991, **123**, 1232.
40 Winvib, S. Lane, Vibrations for windows visualisation package, 1995.
41 X. G. Chen, R. Schweitzer-Stenner, S. Krimm, N. G. Mirkin and S. A. Asher, *J. Am. Chem. Soc.*, 1994, **116**, 11141.
42 M. Barrow, *Acta Crystallogr., Sect. B*, 1981, **37**, 2239.
43 O. K. Antson, K. J. Tilli and N. H. Anderson, *Acta Crystallogr., Sect. B*, 1987, **43**, 296.
44 W. Langel, H. Kollhoff and E. Knözinger, *Ber. Bunsen-Ges. Phys. Chem.*, 1985, **89**, 927.
45 M. Defranceschi, D. Peeters, D. Mathieu, J. Delhalle and G. Lecayon, *J. Mol. Struct. (Theochem)*, 1993, **287**, 153.
46 Y. Wang, R. Purello, S. Georgiou and T. G. Spiro, *J. Am. Chem. Soc.*, 1991, **113**, 6368.
47 S. Song, S. A. Asher, S. Krimm and K. D. Shaw, *J. Am. Chem. Soc.*, 1991, **113**, 1155.
48 Y. Li, R. L. Garrell and K. N. Houk, *J. Am. Chem. Soc.*, 1991, **113**, 5895.
49 D. W. Weatherford and F. R. Salemme, *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 19.
50 T. Blundell, D. Barlow, N. Borkakoti and J. Thornton, *Nature (London)*, 1983, **306**, 281.

Paper 5/01761K; *Received* 20*th March*, 1995