A correlation between the proton stretching vibration red shift and the hydrogen bond length in polycrystalline amino acids and peptides

Mark Rozenberg,[*a] **Gil Shoham,**[a] **Igor Reva**[b] **and Rui Fausto**[b]

[a] *Department of Inorganic and Analytical Chemistry, The Hebrew University of Jerusalem, 91904 Jerusalem, Givat Ram, Israel. E-mail: markroz@chem.ch.huji.ac.il*
[b] *Department of Chemistry, University of Coimbra, 3004-535 Coimbra, Portugal*

The FTIR spectra of pure and isotopically diluted (H/D and D/H) polycrystalline L-glutamine, L-hystidine, L-tyrosine, DL-serine, L-threonine, di-, tri-glycine and di-glycine · HCl · H$_2$O salt were measured in the range 4000–2000 cm^{-1} at temperatures from 300 to 10 K. The frequencies of decoupled proton stretching mode bands ν_1, which can be observed only at low temperature, were used for estimation of the of ν_1-bands red shift, which occurs upon formation of H-bonds involving ionized NH$_3^+$ and/or peptide HN–CO groups. The empirical correlation between the red shift and H-bond length, which was found previously for binary gas phase H-bonded complexes, carbohydrates and nucleosides [M. Rozenberg, A. Loewenschuss and Y. Marcus, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2699–2702; M. Rozenberg, C. Jung and G. Shoham, *Phys. Chem. Chem. Phys.*, 2003, **5**, 1533–1535], was now extended to H-bonded networks in polycrystalline amino acids and peptides. The energies of the different H-bonds present in the crystalline structures could also be successfully estimated from the well-established empirical correlation [A. V. Iogansen, *Spectrochim. Acta*, 1999, **A55**, 1585–1612] between this property and the red shifts of the corresponding ν_1 mode bands.

Introduction

Infrared spectroscopy represents one of main instruments for hydrogen bonding study and has been applied repeatedly for examination of amino acids and peptides structure.[1,2] To this end, the *amide I* band, which can be described essentially as the C=O stretching mode of the peptide linkage, has been used as the main source of information. On the other hand, the proton stretching (ν_1) and out of plane bending (ν_4) modes have had a very limited use to investigate the H-bonding in amino acids and peptides,[2,3] despite these being the most sensitive vibrations to the geometry and energy of the H-bonds.[4,5] In particular, the spectroscopic parameters associated with both ν_1 and ν_4 modes were never correlated with structural data, such as the H-bond length (H···B).

In this work, bands due to the ν_1 stretching modes of the NH$_3^+$ and peptide (NH–CO) moieties (and other proton containing groups present in the studied molecules: OH, NH, NH$_2$) were identified in the low temperature spectra of polycrystalline L-glutamine, L-hystidine, L-tyrosine, DL-serine, L-threonine, di-, tri-glycine and di-glycine · HCl · H$_2$O salt and assigned to the different H-bonds previously found to be present in these crystals by neutron scattering studies. This was made possible due to implementation of the novel experimental approach we have been developing for systematic studies of other biologically relevant H-bonded systems in the solid state,[6–14] which is based on the study of isotopically doped crystals at low temperature. In this technique, partially deuterated substituted molecules are diluted in crystalline samples of the compound at low 2–10% deuterium content (or unsubstituted molecules are embedded in crystals of the deuterated compound at the same doping concentration), and advantage is taken of working at a temperature of a few kelvin. Under these conditions, the vibrations of the isolated doping molecules are not affected by symmetry-related interactions and local anisotropy resulting from thermal excitation, which usually lead to a strong broadening of the vibrational bands originated in the H-bonded groups and conceal their characteristic temperature dependence.[15] At low temperature (<30 K), the bandwidth associated with these decoupled vibrations changes typically from hundreds to tens of wavenumbers and, as a result, the resolution of the IR spectra significantly improves, allowing the observation of proton-related bands ascribable to specific H-bonds in the crystal. This approach was previously applied successfully to study the H-bonding structure in crystalline carbohydrates,[6] nucleobases (cytosine,[7] uracil and thymine[8] and cocrystallized adenine···uracil pair[9]), nucleosides (cytidine, adenosine and uridine),[10–12] and simple amino acids (alanine and glycine).[13] These studies contributed to confirmation of a very general correlation between the red shift of the ν_1 proton stretching vibrations upon hydrogen bond formation and the H-bond length.[6,10,14] The results now presented confirm that this correlation between the spectral (the red shift) and structural (H-bond distance) parameters does also hold in the case of more complex amino acids and small peptides, thus appearing as a powerful experimental tool to access details of H-bonding schemes in crystals of amino acids and peptides. Furthermore, in this study the experimentally observed red shifts could also be used to estimate H-bond energies of individual H-bonds present in the crystals, taking advantage of the well-established Iogansen's empirical correlation between these spectral and thermodynamic properties.[16] Application of the methodology here used to the study of specific H-bonding interactions in selectively deuterated doped crystals of complex peptides can be foreseen.

Experimental

The IR spectra of polycrystalline pure NH- (Sigma), ≈5% ND- and ≈95% ND-amino acids or peptides in a KBr

This journal is © The Owner Societies 2005

(1 : 200) disc, attached to the cold finger of an APD Cryogenics closed-cycle helium refrigeration system with a DE-202A expander, were recorded with a Mattson Infinity 60AR series FTIR spectrometer, with spectral resolution 1 cm^{-1}. The temperature (10–300 K) was measured directly at the sample holder by a silicon diode temperature sensor connected to a Scientific Instruments (model 9650) temperature controller. The sample temperature during registration of spectra was kept stable within ±0.2 K. The temperature-induced spectral changes observed for all substances were found to be reversible and highly reproducible. Deuterated samples were obtained from the commercial substances by repetitive recrystallization from D_2O (Aldrich). The conditions of substance preparation were similar to those used in corresponding structural works.

Results and discussion

The experimental procedure described in the previous sections was applied to all substances studied and the results are summarized in Table 1, where the peak positions of decoupled proton/deuteron bands are given. The red shifts $\Delta\nu = \nu_1^H - \nu_1^0$, where the superscripts "H" and "0" refer to H-bonded and free molecules, respectively, were calculated relatively to the ν_1^0 peak positions of the NH (or OH) bands observed for the compounds either in diluted CCl_4 solutions or in low temperature inert matrices. In order to compare data obtained at room and low (10 K) temperature, the peak positions of the ν_1 bands in the low temperature spectra were corrected by assuming a linear dependence of the peak wavenumber with temperature (see ref. 14, and references therein mentioned). This correction does not exceed 10% of the observed red shift values.

In Table 1, the red shifts are given together with the corresponding experimentally measured H-bond distances as well as with the H-bond distances and energies estimated using the appropriate empirical correlations (see below). Some relevant information and references are given in the note of Table 1. In the next sections of this paper, the detailed discussion of the spectra and proposed band assignments are given for two amino acids (L-glutamine and L-histidine) and two peptides (di- and tri-glycine), as illustrative examples (see Fig. 1 for schematic graphical representations of the studied molecules).

L-Glutamine

The crystal and hydrogen bonding structure of L-glutamine were previously investigated by neutron diffraction.[17] In the crystal, the glutamine molecules exist in the zwitterionic keto form. The H-bonding network contains five different, rather non-linear, intermolecular H-bonds, all of the N–H$\cdots$O type, where the H-bond donor can be either the NH_2 or NH_3^+ groups. In four of these H-bonds, the $\angle$ N–H$\cdots$O angle is ca. 163–167°, whereas in the remaining one, which was found to be the longest bond, the $\angle$ N–H$\cdots$O angle is only 141°. The IR and Raman spectra of crystalline L-glutamine were first studied by Dhamelincourt and Ramirez,[18] and tentative general assignments were proposed.

The IR spectra of crystalline L-glutamine, in the 3600–1800 cm^{-1} spectral range (where N–H/D stretching ν_1 mode bands are observed), obtained in the present study at 300, 200, 100 and 10 K (traces a, b, c and d, respectively) are shown in Fig. 2, for three samples with different isotopic composition: natural isotopic composition (commercial—C-sample; frame 1) and doped with a small quantity of deuterium (ca. 5% deuterium content—D-doped sample; frame 2) or protium (ca. 95% deuterium content—H-doped sample; frame 3). The sharp bands observed in the 3000–2800 cm^{-1} spectral range undoubtedly relate to CH_2 stretching modes. The bands at 3363 and 3302 cm^{-1} are present only in the spectra of the D- and H-doped samples (frames 2 and 3), being absent in the spectrum of the commercial compound (frame 1). They can then be unequivocally assigned to the N–H stretching modes of the two non-equivalent protons of the half-substituted NHD group. The corresponding N–D stretching modes were easily assigned to the narrow bands at 2484 and 2452 cm^{-1}, which can be clearly observed in the spectrum of the D-doped sample (see frame 2) and are still discernible as weak features in the spectrum of the H-doped sample (in the latter spectrum, the two bands are partially masked by the intense group of bands observed in the 2350–2500 cm^{-1} region, which are due to ND_2 and ND_3^+ groups; see frame 3).

In the spectrum of the C-sample, the band at 3398 cm^{-1} and the doublet at 3218/3166 cm^{-1} (centered at 3192 cm^{-1}) are ascribed to the antisymmetric and symmetric stretching vibrations of the NH_2 group, respectively, and are clearly observable in all spectra. The splitting between these two NH_2 modes is ca. 205 cm^{-1}, being identical to that found previously for the same vibrations in cytosine (ca. 202 cm^{-1}).[7] On the other hand, the observed splitting of the band due to the symmetric mode results, with all probability, from a Fermi resonance interaction with the first overtone of the δNH_2 in-plane bending mode (amide II band[18]), whose fundamental band is observed at 1590 cm^{-1}. The corresponding antisymmetric and symmetric stretching vibrations of the ND_2 group give rise to the bands at 2553 and ca. 2380 observed in the spectrum of the H-doped crystal of deuterated glutamine (2551 and 2389 cm^{-1}, in the D-doped sample).

The assignment of the N–H stretching bands due the NH_3^+ group is strongly facilitated by the fact that, in the crystalline state, the three protons were found to be involved into considerably different H-bond interactions, as it is clearly reflected by their substantially different H-bond distances determined by neutron diffraction[17] (see Table 1). In this case, the three N–H oscillators are essentially uncoupled and the spectra of both the D- and C-samples (where the N–H bands are considerably more intense than in the spectrum of the H-doped sample) can also be used to identify the relevant bands associated with the three different protons. The validity of this approach was previously confirmed, for example, for alanine, where the positions of the NH_3^+ stretching bands in both the spectra of the crystal with natural isotopic content and in the H-doped deuterated alanine crystal were found to be identical.[13] For glutamine, the broad and weak bands at ca. 3112 cm^{-1} (partly overlapped with the lowest frequency component band of the Fermi doublet assigned to the symmetric stretching mode of the NH_2 group), ca. 2930 cm^{-1} and 2676 cm^{-1} (average frequency for the doublet of bands observed in the 2690–2660 cm^{-1} spectral range), which become clearly seen only at low temperature (see frames 1 and 2), are ascribable to the three independent proton stretching modes of the NH_3^+ group. The doublet structure of the lowest frequency band can be explained as being due to a Fermi resonance interaction with the first overtone of the band appearing at 1337 cm^{-1}, previously assigned to a δCH bending mode.[18]

Contrarily to the bands due to the NH_3^+ group, that are difficult to observe in the spectra of the H-doped sample, those due to the ND_3^+ group are clearly observed at 2356, 2200 and 2109 cm^{-1} in the spectrum of D-doped sample. In consonance with the hypothesis formulated above, in the spectra of the H-doped deuterated crystal these bands are observed at similar frequencies: 2369 cm^{-1} (overlapped with the strong ND_2 symmetric stretching mode), ca. 2216 cm^{-1} and ca. 2080 cm^{-1} (see frame 3). The two last frequencies correspond to the average frequencies of the two doublets ascribable to the lowest frequency ND_3^+ stretching vibrations, which result from resonance interactions with the overtones of the modes whose fundamentals are observed at 1108 and 1042 cm^{-1} (these bands are absent in the spectra of the C-sample and can then be assigned to deuteron deformational modes).

The isotopic ratios for the coupled symmetric and antisymmetric modes of the NH_2 group are 1.332 and 1.336, differing

Table 1 Peak positions of decoupled stretching ν_1/ν'_1 bands at 10 K, the red shift $\Delta\nu_1$ (cm^{-1}), corresponding $r_{AH\cdots B}$ (nm) distance (A = N, O) and H-bond energies $-\Delta H$ (kJ mol^{-1}) in amino acid and peptide crystals[a]

Substance	ν_1/ν'_1(10 K)	$(\Delta\nu_1^{10K})\,\Delta\nu_1^{300K}$	$r_{AH\cdots B}^{cal}$ [b]	$r_{AH\cdots B}^{exp}$ AH	$-\Delta H$ [c]
Glycine	3154/2341	(235) 211	0.199	0.2121 NH$_3^{+(3)}$	18
	2902/2185	(481) 433	0.176	0.1832 NH$_3^{+(2)}$	27
	2600/2030	(780) 700	0.163	0.1728 NH$_3^{+(1)}$	36
L-Alanine	3076/2287	(304) 274	0.190	0.1861 NH$_3^{+(1)}$	21
	2990/2240	(390) 350	0.183	0.1828 NH$_3^{+(3)}$	24
	2677/2088	(713) 642	0.166	0.178 NH$_3^{+(2)}$	34
L-Glutamine	3363/2484	(113) 102	0.224	0.2088 NH$_2^{(22)}$	10
	3302/2452	(174) 157	0.208	0.1919 NH$_2^{(21)}$	15
	3112/2356	(268) 241	0.194	0.1941 NH$_3^{+(2)}$	19
	2930/2200	(450) 405	0.178	0.1854 NH$_3^{+(1)}$	26
	2676/2109	(704) 634	0.166	0.1752 NH$_3^{+(3)}$	34
DL-Serine	3074/2300	(306) 275	0.190	0.1844 NH$_3^{+(2)}$	21
	3005/2238	(375) 337	0.184	0.1787 NH$_3^{+(1)}$	24
	2927/2200	(453) 408	0.178	0.1814 NH$_3^{+(3)}$	27
	2780/2150	(840) 756	0.161	0.1692 OH	37
L-Tyrosine	3197/2393	(183) 165	0.207	0.212 NH$_3^{+(1)}$	15
	3095/2294	(285) 256	0.192	0.1853 NH$_3^{+(3)}$	20
	3025/2180	(355) 319	0.185	0.1789 NH$_3^{+(2)}$	23
	2700/2065	(910) 819	0.159	0.1689 OH	39
L-Histidine	3100/2301	(280) 252	0.193	0.1899 NH$_3^{+(2)}$	20
	3005/2256	(375) 337	0.184	0.1841 NH$_3^{+(1)}$	24
	2933/2205	(447) 402	0.176	0.1786 NH$_3^{+(3)}$	26
	2856/2180	(610) 550	0.170	0.172 NH	31
L-Threonine[d]	3160/2338	(260) 220	0.201	0.223 NH$_3^{+(2)}$	19
	3068/2296	(330) 300	0.188	0.203 NH$_3^{+(1)}$	22
	2927/2201	(470) 420	0.178	0.197 NH$_3^{+(3)}$	27
	2870/2148	(750) 675	0.164	0.182 OH	35
Di-glycine	3280/2443	(200) 180	0.203	0.1967 NH$\cdots$OC	16
	3066/2293	(314) 283	0.188	0.184 NH$_3^{+(1)}$	22
	2990/2224	(390) 351	0.183	0.1806 NH$_3^{+(3)}$	24
	2770/2191	(610) 549	0.169	0.1723 NH$_3^{+(2)}$	31
Tri-glycine[e]	3312/2463	(164) 148	0.211	0.212 ^{1}N$^{(2)}$H$^{(2)}\cdots^2$O$^{(4)}$C	14
	3281/2432	(195) 175	0.205	0.207 ^{1}N$^{(3)}$H$^{(3)}\cdots^1$O$^{(2)}$C	16
	3111/2300	(269) 242	0.194	0.198 ^{1}N$^{(1)}$H$_3^{+(11)}\cdots^1$O$^{(4)}$	20
	2861/2160	(519) 467	0.174	0.177 ^{2}N$^{(2)}$H$_3^{+(13)}\cdots^1$O$^{(3)}$	29
	2760/2100	(620) 558	0.169	0.171 ^{2}N$^{(1)}$H$_3^{+(11)}\cdots^1$O$^{(3)}$	33
Di-glycine$\cdot$HCl$\cdot$H$_2$O	3273/2381	(107) 96	0.226	0.2268 NH$_3^{+(2)}$Cl$^-$	10
	3336/2462	(144) 130	0.215	0.2295 N^{2}H$^6\cdots$Cl$^-$	13
	3303/2322	(357) 321	0.185	0.1764 O^{4}H$^{10}\cdots$O$^{1'''}$	23
	3217/2285	(163) 147	0.211	0.207 NH$_3^{(3)}\cdots$O^4	14
	2830/2180	(704) 634	0.166	0.1658 O^{3}H$^9\cdots$O$^{4'''}$	24

[a] $\Delta\nu_1^{300K} = \Delta\nu_1^{10K} \times 0.9$; spectra of alanine published in ref. 13; $r_{H\cdots B}^{exp}$ distances and H-numbering: in alanine—data ref. 39; in glycine—ref. 40; in L-glutamine—ref. 17; in DL-serine—ref. 41; L-tyrosine—ref. 42; L-histidine—ref. 21; L-threonine—ref. 43; di-glycine—ref. 20; tri-glycine—ref. 27; glycylglycine$\cdot$HCl$\cdot$H$_2$O—ref. 31. The red shift for NH$_3^+$ is taken relatively to 3380 cm^{-1}, the calculated frequency for free neutral alanine,[44] and observed at 3422 and 3395 for neutral valine isolated in a low temperature argon matrix;[45] for the OH group of serine: 3620 cm^{-1}, of n-butanol in CCl$_4$;[16] for the OH group of tyrosine: 3610 cm^{-1}, of phenol in CCl$_4$;[16] for the free NH$_2$ group of a primary and secondary amide (peptide) group: 3476 cm^{-1}, of acetamide in CCl$_4$,[46] which coincides with 3480 cm^{-1}, of dipeptides in a matrix;[26] for NH in the imidazole ring: 3466 cm^{-1} of unsubstituted imidazole in CHCl$_3$ solution;[16] for free HOD in monohydrate: 3660 cm^{-1}.[16] [b] $\Delta\nu = 0.011r^{-6.1}$. [c] $(\Delta H)^2 = 1.92[(\Delta\nu)-40]$. [d] L-Threonine is the single system where the experimental H$\cdots$B distances systematically exceed the calculated by a value of $ca.$ 0.02 nm.[43] For other acids the deviation of the experimental data from the curve looks like random. [e] From the structure data it follows that there are two molecules in the crystal unit cell and the two molecules form slightly different sets of H-bonds.[27] It is impossible to differentiate all these bonds and only those with the $\angle$ AH$\cdots$B angle greater than 150° were chosen for comparison.

considerably more from the harmonic value than those of the two decoupled vibrations of the NHD group, which are 1.354 and 1.347. On the other hand, the protons of NH$_3^+$ group participate in rather strong H-bonds, and the isotopic ratios associated with the NH$_3^+$ stretching vibrations follow the well-known[19] general trend of decreasing with the H-bond strength (or red shift). Thus, the two high frequency modes, which are relatively close in frequency, have isotopic ratios of 1.324 and 1.332, while the low frequency vibration has a considerably lower isotopic ratio (1.269).

In general, our results confirm the assignment of the bands due to the NH$_2$ group made in ref. 18, but the bands due to the decoupled modes of the NHD group were now observed for the first time. In addition, the present assignments for the proton stretching bands originated in the NH$_3^+$ group differ substantially from those previously proposed,[18] since, as shown in Fig. 2, the N–H stretching bands of L-glutamine appearing below 3000 cm^{-1} can only be clearly observed at the low temperature conditions now used for the first time to study this compound.

L-Histidine

The crystal structure of orthorhombic crystalline L-histidine was solved by neutron diffraction.[20] In the crystal, the histidine molecule exists in the zwitterionic form, with four labile hydrogen atoms involved in hydrogen bonds. One of the

2378 *Phys. Chem. Chem. Phys.*, 2005, **7**, 2376–2383

This journal is © *The Owner Societies 2005*
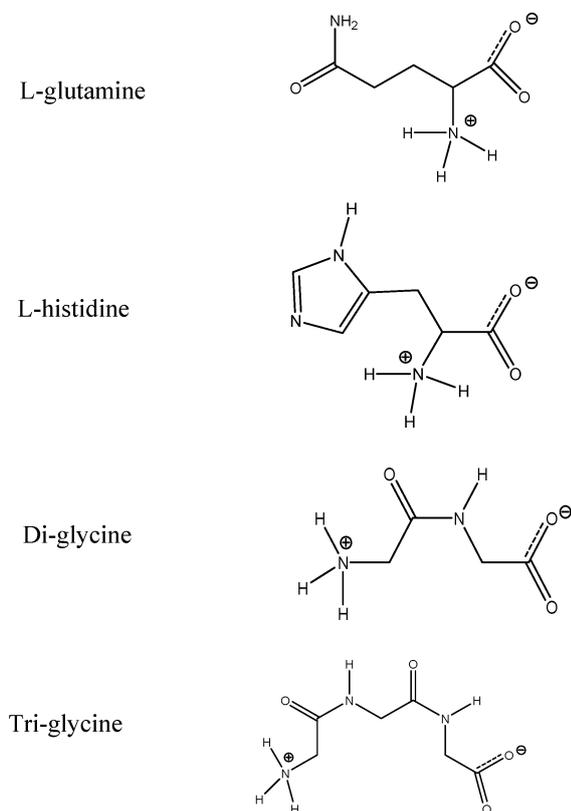
Fig. 1 Schematic structures of L-glutamine, L-histidine, di-glycine and tri-glycine.

NH_3^+ hydrogens makes part of a long, rather non-linear intramolecular N–H···N bond, with the $\angle$ N–H···N angle of *ca.* 143°; the other two hydrogen atoms of the NH_3^+ group and the imidazole hydrogen are involved in N–H···O intermolecular hydrogen bonds with $\angle$ N–H···O angles of 158 and 172°. To the best of our knowledge, no studies on the vibrational spectra of crystalline L-histidine in the ν_1 range were undertaken hitherto.

The IR spectra of L-histidine, in the N–H and N–D stretching ranges, are shown in Fig. 3, for commercial (C, natural
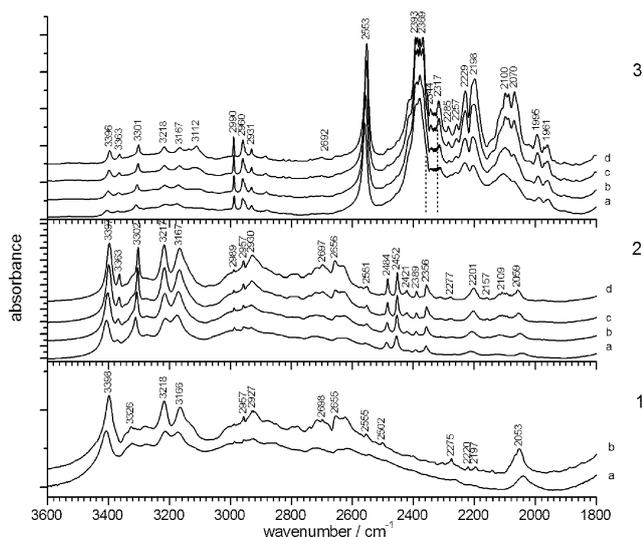
Fig. 2 FTIR spectra in the range of the NH stretching mode of L-glutamine: 1—naturally occurring isotopic content (nearly 100% NH) at the temperatures 300 K (a) and 10 K (b); 2—D-doped sample (*ca.* 95% NH) and 3—H-doped sample (*ca.* 95% ND) at the temperatures 300 K (a), 200 K (b), 100 K (c) and 10 K (d). In frame 3 the range distorted by atmospheric CO_2 absorption is marked by dotted vertical lines.
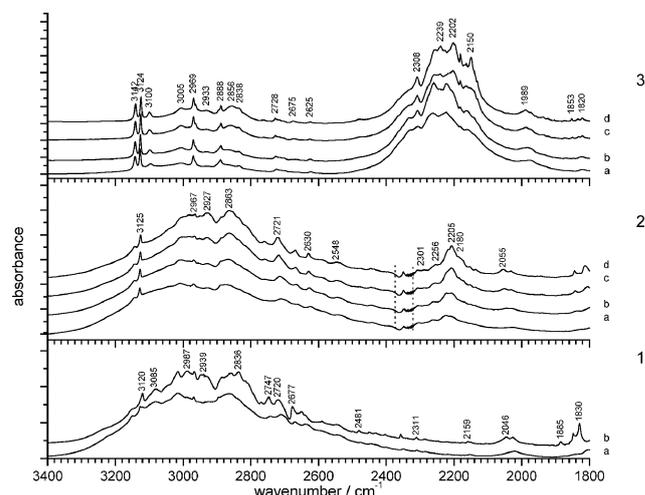
Fig. 3 FTIR spectra in the range of the NH stretching mode of L-histidine: 1—naturally occurring isotopic content (nearly 100% NH) at the temperatures 300 K (a) and 10 K (b); 2—D-doped sample (*ca.* 95% NH) and 3—H-doped sample (*ca.* 95% ND) at the temperatures 300 K (a), 200 K (b), 100 K (c) and 10 K (d). In frame 2 the range distorted by atmospheric CO_2 absorption is marked by dotted vertical lines.

isotopic content, frame 1), D-doped (frame 2) and H-doped (frame 3) samples, at different temperatures 300, 200, 100 and 10 K (traces a, b, c and d, respectively).

The C–H stretching modes, which are also expected to absorb in the discussed spectral region, could be easily ascribed, since they shall give rise to narrow bands weakly sensitive to temperature changes. The corresponding bands are better seen in the spectrum of the H-doped deuterated L-histidine crystal and are observed at 3142 and 3124 cm^{-1} (assigned to the C–H stretching modes of the two imidazole hydrogen atoms) and around 2969 cm^{-1} (doublet of bands ascribed to the two stretching modes of the CH_2 group). The weak, relatively narrow band at 2888 cm^{-1} is probably due to the first overtone of the vibration giving rise to the strong band at 1450 cm^{-1} that is present in the spectra of both C- and D-doped L-histidine samples.

For this compound, the straight assignment of N–H stretching bands was complicated due to the extensive overlap and complexity of the N–H stretching region in the two samples with high H-content (see frames 1 and 2 in Fig. 3) and because the spectrum of the H-doped deuterated L-histidine crystal shows in this spectral region mostly weak and broad bands, which in some cases are not easily discernible even at low temperature (see frame 3 in Fig. 3). Then, after a preliminary identification of the possible candidate bands to be assigned to the NH_3^+ and imidazole N–H stretching vibrations (the N–H bands must be relatively broad and sensitive to temperature), the empirical correlation between the H-bond distance and the N–H stretching red shift (relatively to a free NH_3^+ or N–H group)[14] was used in order to estimate the band positions from the known H-bond distances previously obtained by neutron diffraction.[20] Among the group of bands initially selected as possible candidates, four bands were found to lie in the vicinity of the predicted wavenumbers, and then assigned to the decoupled proton modes of the NH_3^+ group (at 3100, 3005 and 2933 cm^{-1}) and imidazole N–H stretching (2856 cm^{-1}) (see frame 3, in Fig. 3). Other (weak) bands, at 2728, 2675 and 2625 cm^{-1} can be assigned to overtones and related with the intense bands observed at 1419, 1344 and 1335 cm^{-1}. In the spectrum of the D-doped sample (frame 2), the ND_3^+ and deuterated imidazole N–D stretching modes are assigned to the bands at 2301, 2256, 2205 and 2180 cm^{-1}. Note that the three ND_3^+ bands are observed at nearly the same frequencies in the D-doped L-histidine crystals and H-doped deuterated L-histidine crystals, thus giving further support to the above extracted

conclusion that, in general, in the amino acid crystals the three NH_3^+ (or ND_3^+) stretching vibrations are essentially uncoupled.

Di-glycine

The crystal structure of α-glycylglycine (di-glycine) has also been previously studied by neutron diffraction.[21] In the crystal, the molecules exist in the zwitterionic form and establish four intermolecular H-bonds through the NH_3^+ (with ∠N–H···O angles of *ca.* 147–160°) and the peptide N–H groups (the longest H-bond, with an ∠N–H···O angle of *ca.* 158°).

The IR spectra now obtained for this compound, in the range of the N–H stretching modes, are shown in Fig. 4.
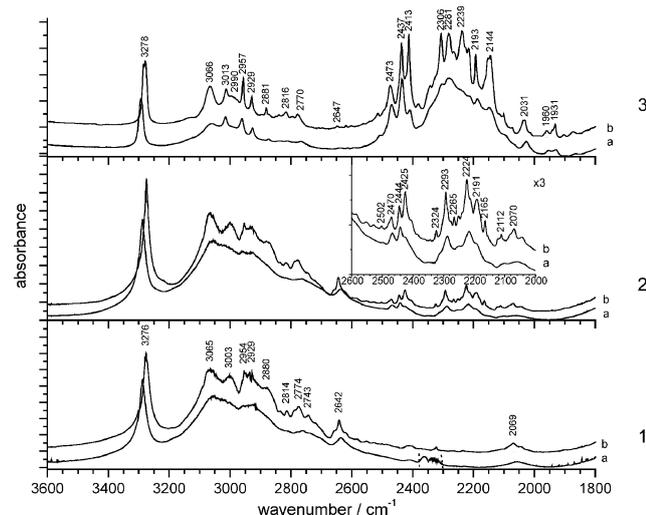
Like for all other molecules studied, spectra were obtained at different temperatures and for C- (commercial; natural isotopic contents), D-doped and H-doped samples.

The four expected C–H stretching bands are easily recognizable in the low temperature spectrum of the H-doped sample (frame 3 in Fig. 4), appearing at 3013, 2957, 2929 and 2881 cm^{-1}.

The assignment of the features near 3278 cm^{-1} to the peptide N–H stretching mode is doubtless. Comparison of the spectra of di-, tri- and tetra-glycines shows that the intensity of the bands appearing at this frequency increases with the number of peptide bonds, as expected, and close to peak position of 3288 cm^{-1} observed in ref. 22. The cooling leads to the splitting of the band in two components (at 3284 and 3278 cm^{-1}) which can be either due to Fermi resonance with the antisymmetric stretching vibration of the carboxylate group (COO$^-$), which strongly absorbs around 1650 cm^{-1}, or crystal splitting effects (due to presence of two slightly different crystal sites). The latter seems less probable, because there is no observable splitting of other narrow bands.

The deuterated peptide N–D stretching mode gives rise to the triplet observed in the 2475–2410 cm^{-1} spectral range (see frames 2 and 3 in Fig. 4) with an average frequency of 2443 cm^{-1}, also quite close to the 2434 cm^{-1} observed in ref. 22. The origin of the observed multiple splitting cannot be unequivocally interpreted, but it might be due to a simultaneous occurrence of crystal and Fermi resonance splittings.

The bands due to the stretching modes of the three H-bonded NH_3^+ hydrogen atoms can only be clearly seen at low temperature and are not discussed in ref. 22 at all. In the spectra at high temperatures, the bands due to these vibrations

extensively overlap and produce the broad, almost structureless, band absorbing in the range 3100–2600 cm^{-1}. Upon cooling, this broad band develops a clear structure that is seen better in the spectrum of the H-doped sample (frame 3 in Fig. 4), enabling the identification of the bands due to the three NH_3^+ modes at 3066, 2990 and 2770 cm^{-1}. A similar situation is observed for the vibrations due to the ND_3^+ group, which give rise to the bands at 2293, 2224 and 2191 cm^{-1} (see frame 2 in Fig. 4).

It is difficult to compare our spectra with literature data, since dipeptides have been mostly studied with the main interest of examining the possible conformations of peptide groups and usually substituted amino acid residues were considered (di-glycine is a very special unstrained dipeptide of less general interest for elucidation of the conformational preferences of the peptide group and its influence on the structure of more complex peptides and proteins).[23–26] Moreover, the bands observed in our low temperature spectra of the isotopically-diluted samples that relate to the NH_3^+ (or ND_3^+) stretching modes have never been observed previously, because, as noticed above, they are very broad and weak at room temperature. Previous calculations predicted the NH_3^+ vibrations as absorbing in the 3240–3000 cm^{-1} spectral range,[25] and only the band at *ca.* 3060 cm^{-1} (corresponding to the main maximum of the broad and unstructured band observed in this spectral region at high temperature) was previously assigned to an NH_3^+ vibration.

Tri-glycine

The crystal structure of tri-glycine obtained by neutron diffraction[27] shows two molecules per unit cell, which form two slightly different systems of hydrogen bonds. There are both inter- and intra-molecular (including bifurcated) hydrogen bonds, and several of them have very similar lengths.

The IR spectra of tri-glycine are shown in Fig. 5. For this compound, some of the bands due to the CH_2 groups are difficult to observe, due to extensive overlap with more intense bands due to the N–H stretching modes. However, the narrow bands observed at 3026, 2965 and 2946 cm^{-1} in the spectra of the three samples studied as well as the two bands observed in the low temperature spectra of the H-doped crystal around 2920 cm^{-1} can be undoubtedly ascribed to C–H stretching modes.

Fig. 4 FTIR spectra in the range of the NH stretching mode of di-glycine at the temperatures of 300 K (a) and 10 K (b). 1—naturally occurring isotopic content (nearly 100% NH); 2—D-doped sample (*ca.* 95% NH); 3—H-doped sample (*ca.* 95% ND). Inset—part of spectrum with the ordinate scale enlarged by factor 3. In frame 1—trace (a), the range distorted by atmospheric CO_2 absorption is marked by dotted vertical lines.

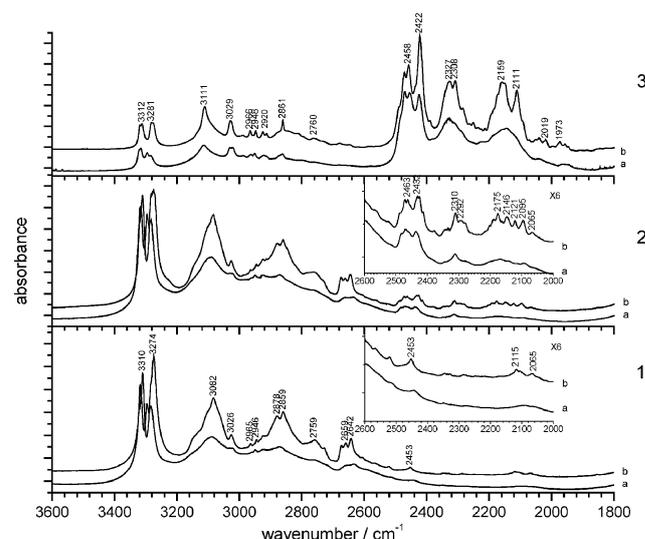**Fig. 5** FTIR spectra in the range of the NH stretching mode of tri-glycine at the temperatures of 300 K (a) and 10 K (b). 1—naturally occurring isotopic content (nearly 100% NH); 2—D-doped sample (*ca.* 95% NH); 3—H-doped sample (*ca.* 95% ND). Insets—parts of spectra with ordinate scale enlarged by a factor of 6.

From the comparison of the spectra of all the three samples studied—natural isotopic contents (frame 1), D-doped (frame 2) and H-doped (frame 3)—and also those of di-glycine, previously discussed, the assignment of the bands at 3312 and 3281 cm^{-1} to the N–H stretching vibrations of the peptide groups seems also doubtless. On the other hand, as it was already mentioned above, a comparison of the spectra of di- tri- and tetra-glycine reveals that the intensity of the bands appearing at this frequency increases with the number of peptide bonds. In the single previously reported study of the IR spectra of crystalline di- and tri-peptides suitable for comparison with ours,[28] two bands appearing at *ca.* 3320 and 3280 cm^{-1} were observed for the Ala–Ser–Leu tri-peptide (but without exhibiting any fine structure), close to those observed in our study for tri-glycine. The corresponding N–D stretching bands in the deuterated molecules can also be ascribed straightforwardly, by analysis of the spectra of the D-doped sample, to the features observed at 2463 and 2432 cm^{-1}.

In the case of the stretching vibrations originated in the NH$_3^+$ group of crystalline tri-glycine, the assignments were more difficult to undertake. The two structured bands that become considerably more intense upon cooling (observed in the spectrum of the H-doped sample with absolute maxima at 3111 and 2861 cm^{-1}) are ascribable to these vibrations without any doubt. The assignment of the third mode is less secure, but it can be suggested to correspond to the band observed around 2760 cm^{-1} in the low temperature spectra of both C- and D-doped samples, which, in the spectra of the H-doped sample, appears as a very weak band. The large relative intensity and structure shown by the two first bands (at *ca.* 3111 and 2861 cm^{-1}) can be explained by the contribution to these features of several overlapping bands from NH$_3^+$ groups occupying different positions in the crystal unit cell but establishing energetically equivalent H-bonds (at least in the case of the lowest frequency band, bands due to C–H modes might also contribute to the observed band profile). With all probability, the fine structure that is seen in the ND$_3^+$ spectral region in the spectrum of the D-doped sample (frame 2 in Fig. 5; inset), can also be attributed to the same causes (in Table 1, only the average frequency values of the groups of bands assigned to the ND$_3^+$ stretching modes are presented).

Correlation between the red shifts and hydrogen bond lengths

All experimentally determined frequency red shifts ($\Delta\nu$/cm^{-1}) of N–H hydrogen-bonded groups in the various crystals here studied relatively to the corresponding frequency of the "free" species were plotted as a function of the hydrogen bond length ($r_{\mathrm{H\cdots B}}$/nm) in Fig. 6.

The data obtained for the whole investigated set of systems obey fairly well to the equation

$$\Delta\nu/\mathrm{cm}^{-1} = 0.011 \, [r_{\mathrm{H\cdots B}}/\mathrm{nm}]^{-6.1} \qquad (1)$$

which was found previously for other H-bonded systems.[10,14] The points plotted in Fig. 6 refer not only to the H-bonds made by the NH$_3^+$ groups, but include also data relative to the H-bonds established by the peptide groups for all systems studied, the amine group of L-glutamine, the hydroxyl groups of DL-serine, L-tyrosine and water in the di-glycine·HCl·H$_2$O salt, comprising a range of H-bond distances of *ca.* 0.07 nm.

For generality, Fig. 6 also includes recent data relative to ≡CH···N and S–H···phenyl hydrogen bonds measured in crystal state.[29,30] These weak and lengthy H-bonds do also reasonably fit the same correlation, extending the distances limit evaluated up to *ca.* 0.1 nm.

It is worth to note that all fluctuations found in the data plotted in Fig. 6 can be considered as random for all substances but one: L-threonine. The data relative to the four hydrogen bonds existent in the crystal of L-threonine show a large systematic deviation. Because the red shifts are with all
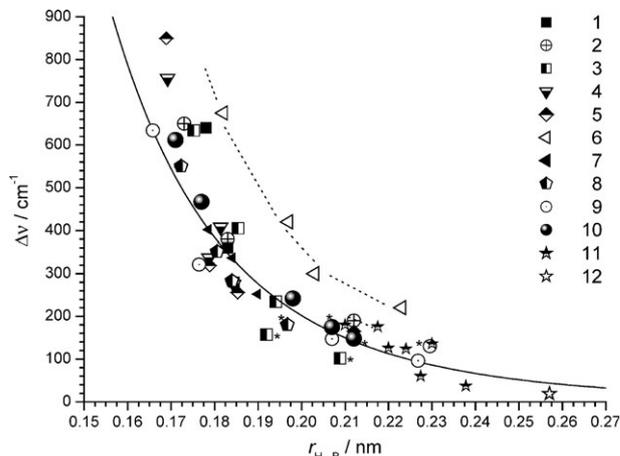
Fig. 6 Correlation of red shift, $\Delta\nu$/cm^{-1}, with the H···B distance $r_{\mathrm{H\cdots B}}$/nm. L-Alanine (1), glycine (2), L-glutamine (3), DL-serine (4), L-tyrosine (5), L-threonine (6), L-histidine (7), di-glycine (8), di-glycine·HCl·H$_2$O (9), tri-glycine (10), triple CH···N (11),[29] S–H···phenyl (12).[30] Relate to HN–CO peptide bond (*); the line is the function of the correlation (1)—see text.

certainty not affected by errors as large as 300–100 cm^{-1}, the reason for the observed deviation from the general correlation might indicate an inadequate crystal structure H-bonding data determination.

Cl$^-$ anion as proton acceptor

From the present study, it could also be concluded that the correlation expressed by eqn. (1) does not hold when Cl$^-$ acts as proton acceptor in crystalline amino acid·HCl salts. Since the correlation can be applied successfully when proton acceptors are oxygen, nitrogen or a π-electron cloud, even when the H-bonds are considerably weak and long (ref. 14, and new data on Fig. 6), the reason for its failure when Cl$^-$ is the acceptor group cannot be the weakness and length of the observed Cl$^-$···H hydrogen-bonds (*i.e.* small shifts and long distances).

As an example, the data now obtained for di-glycine·HCl·H$_2$O are shown in the Table 1. Among the seven hydrogen bonds with angles not smaller than 147° found in the crystal,[31] all three H-bonds where oxygen acts as the proton acceptor satisfy the correlation, whereas only one from the four H-bonds where Cl$^-$ is the acceptor group reasonably fits the correlation (and, indeed, shall be treated as an accidental coincidence). All the remaining Cl$^-$···H bonds, including the shortest one, do not fit relationship (1) at all. This deviation from such a rather broad empirical relationship may underline the distinction of the more ionic and less directional interaction involving the Cl$^-$ species, where the interaction with proton stretching mode is minimal, from those with a greater covalent and directional character established either with nitrogen, oxygen or a π-electron cloud, where the electron sharing within the A–H···B moiety is more effective and leads to a more important perturbation of the proton stretching vibration. Based on this result correlation (1) will not be suitable for systems with proton acceptors as F$^-$, Cl$^-$, Br$^-$, I$^-$ and other complex inorganic ions in hydrate crystals, which were discussed in refs. 32–34.

Estimation of the H-bond energies. Comparison with computational results

The last column of Table 1 shows the H-bond energies estimated using the well-established[16] empirical correlation between the red shifts ($\Delta\nu$/cm^{-1}) and the H-bond energies (ΔH/kJ mol^{-1}):

$$(\Delta H)^2 = 1.92[(\Delta\nu) - 40]. \qquad (2)$$

It is seen that in crystalline amino acids and peptides the H-bond energies ($-\Delta H$) vary within the limits 10–40 kJ mol^{-1}. The consistency of the H-bond energies obtained using correlation (2) with experimental data obtained independently was demonstrated previously for alanine.[13]

There is a lack of previous systematic studies on H-bond energies in amino acid crystals obtained using other techniques to compare with our results. However, a general evaluation of the relative capability of the methodology here used in predicting H-bond energies in this type of systems can be done by comparing data obtained with empirical relationships from spectra of similar compounds in diluted solutions with recent computational results and those resulting from direct experimental measurements. For example, the H-bond energy of formation of the *cis*-dimer of the cyclic amide ε-caprolactam in CCl$_4$ solution was estimated as 17.9 kJ mol^{-1} with Iogansen's empirical "rule of intensity",[16] in excellent agreement with the experimental value of 17.1 kJ mol^{-1}.[35] For the H-complex of ε-caprolactam with *N*-methyl pyrrolidone, which can be considered as a model of the *trans*-dimer, the H-bond energy was estimated by the empirical relationship as 14.3 kJ mol^{-1}.[35] On the other hand, *ab initio* molecular orbital calculations predict H-bond energies of 24.5 and 26 kJ mol^{-1} for *cis*- and *trans*-dimers of the similar compound, *N*-methylacetamide.[36] Hence, the computational results exceed the experimental values for both types of dimers by approximately 10 kJ mol^{-1}, whereas the empirically derived correlation yields estimations in close agreement with experiment. The same tendency was also observed recently, when an estimation of the contribution of the C–H···O hydrogen bonds to the stabilization energy in a transmembrane peptide was made.[37] Arbely and Arkin[37] measured the peak positions of the C–D bands ascribable to the groups directed to and out of the closest C=O group, and using the empirical correlation (2), between the red shift (recalculated for C–H group)[16] and the H-bond energy, estimated the H-bond energy as being *ca.* 3.7 kJ mol^{-1}, which is 2–4 times lower than the *ab initio* calculated value for this type of interaction.[37] The different values obtained certainly result partially from the circumstance that a comparison is being made between data obtained from experimental measurements in condensed phases and theoretical results derived for free molecules (dimers) in the gas phase. However, it can always be sustained that the refinements that are necessary to introduce to the theoretical calculations in order to make them competitive with the empirical relationships here used, even not taking into consideration computational time, are still not foreseen for the near future. This fact attributes great significance to the methodology here followed to the study of complex systems in condensed phases and, in particular, to biologically relevant species establishing H-bond networks, like amino acids and peptides.

Conclusion

An empirical correlation, $\Delta\nu \propto r^{-6}$, between spectral and structural parameters, which had been found to be applicable to H-bonds in gas phase diatomic hydrogen bonded complexes[14] as well as to more than a hundred H-bonds with $\angle$ A–H···B $\geq 140°$ in solid state carbohydrates, nucleobases and nucleosides,[6–14] is now extended to amino acids and small peptides. Considering its demonstrated wide applicability, it can then be foreseen that this correlation might also be used successfully to investigate H-bonding in other, more complex molecular systems of biological interest, including derivatives of nucleic acids, peptides and proteins.

Together with other well-established empirical correlations, such as those relating the hydrogen bond enthalpy with the intensity enhancement of the stretching vibration upon hydrogen bond formation (Iogansen's "rule of intensity"),[16] the red shift of the stretching mode of the H-bonded group [relationship expressed by eqn. (2)] or the blue shift of the out of plane deformation vibration of the H-bonded group,[38] the spectra–structure correlation (1) forms a powerful "toolbox" for the investigation of complex H-bonded systems, otherwise inaccessible to detailed investigation. It shall be noted that none of the mentioned empirical correlations was predicted or explained by theoretical methods. All these correlations thus provide important information for refinement of theoretical models, as it was recognized, for example, in the above mentioned spectroscopic studies on membrane proteins.[37] One not less important result follows from combining the mentioned empirical correlations—that the energy of an AH···B bond is a function of $r_{\mathrm{H}\cdots\mathrm{B}}^{-3}$.[14]

On the other hand, when the spectra of a complex network of H-bonds in the solid state cannot be easily interpreted, correlation (1) can be used in a reversed way for prediction of the positions of the bands associated with the fundamental vibrations originated in the H-bonded groups taking as input data the available experimental H-bond distances. The expected accuracy in the predictions made this way, resulting from analysis of its previous systematic application to different systems, is better than 50 cm^{-1} (corresponding to an average relative precision of *ca.* 2%). This can be considered more than enough to positively identify the bands due to the fundamental stretching and out of plane bending vibrations of groups involved into H-bonds in most of the relevant chemical systems, particularly those with biological relevance, from which amino acids and peptides are certainly paradigmatic examples.

Acknowledgements

References

1 S. Krimm and J. Bandekar, *Adv. Prot. Chem.*, 1986, **38**, 181–354.
2 J. Bandekar, *Biochim. Biophys. Acta*, 1992, **1120**, 123–143.
3 Y. Masuda, K. Fukushima, T. Fujii and T. Miyazawa, *Biopolymers*, 1969, **8**, 91–99.
4 *The Hydrogen Bond*, ed. P. Schuster, G. Zundel and C. Sandorfy, North-Holland, New York, 1976.
5 *Vodorodnaya Svyaz (The Hydrogen Bond)*, ed. N. D. Sokolov, Izd. Nauka, Moscow, 1981.
6 M. Rozenberg, A. Loewenschuss and Y. Marcus, *Carbohydr. Res.*, 2000, **328**, 307–319.
7 M. Rozenberg, G. Shoham, I. Reva and R. Fausto, *Spectrochim. Acta A*, 2004, **60A**, 463–470.
8 M. Rozenberg, G. Shoham, I. Reva and R. Fausto, *Spectrochim. Acta A*, 2004, **60A**, 2323–2336.
9 M. Rozenberg, G. Shoham, I. Reva and R. Fausto, *Spectrochim. Acta A*, 2005, DOI: 10.1016/j.saa.2004.12.037.
10 M. Rozenberg, C. Jung and G. Shoham, *Phys. Chem. Chem. Phys.*, 2003, **5**, 1533–1535.
11 M. Rozenberg, C. Jung and G. Shoham, *Spectrochim. Acta A*, 2004, **60A**, 2369–2375.
12 M. Rozenberg, C. Jung and G. Shoham, *Spectrochim. Acta A*, 2005, **61A**, 733–741.
13 M. Rozenberg, G. Shoham, I. Reva and R. Fausto, *Spectrochim. Acta A*, 2003, **59A**, 3253–3266.
14 M. Rozenberg, A. Loewenschuss and Y. Marcus, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2699–2702.
15 A. V. Iogansen, *Izv. Akad. Nauk, SSSR, Ser. Fiz.*, 1989, **53**(9), 1741–1745 (Russ.), Engl. Transl. by Allerton Press Inc., 1989.
16 A. V. Iogansen, *Spectrochim. Acta A*, 1999, **55**, 1585–1612.
17 T. F. Koetzle, M. N. Frey, M. S. Lehmann and W. C. Hamilton, *Acta Crystallogr. Sect B: Struct. Sci.*, 1973, **B29**, 2571–2575.
18 P. Dhamelincourt and F. J. Ramirez, *Appl. Spectrosc.*, 1993, **47**(4), 446–451.
19 A. Novak, *Struct. Bond.*, 1974, **18**, 177–216.
20 S. Lehmann, T. F. Koetzle and W. C. Hamilton, *Int. J. Pept. Protein. Res.*, 1972, **4**, 229–239.

21 A. Kvick, A. R. Al-Karaghouli and T. F. Koetzle, *Acta Crystallogr. Sect B: Struct. Sci.*, 1977, **B33**, 3796–3801.

22 C. Destrade, E. Dupart, M. Joussot-Dubien and C. Carrigou-Lagrange, *Can. J. Chem.*, 1974, **52**, 2590.

23 M. Miyazawa, Y. Kyogoku and H. Sugeta, *Spectrochim. Acta A*, 1994, **50A**, 1505–1511.

24 T. C. Cheam and S. Krimm, *J. Mol. Struct.*, 1989, **193**, 1–34.

25 P. Lagant, G. Vergoten, M. H. Loucheux-Lefebvre and G. Fleury, *Biopolymers*, 1983, **22**, 1267–1283.

26 Y. Grenie, M. Avignon and C. Garrigou-Lagrange, *J. Mol. Struct.*, 1975, **24**, 293–307.

27 T. Spikrishnan, N. Winiewicz and R. Parthasarathy, *Int. J. Pept. Protein Res.*, 1982, **19**, 103–113.

28 M. Narita, M. Doi and H. Takegahara, *Bull. Chem. Soc. Jpn.*, 1987, **60**, 2445–2451.

29 S. Boryczka, M. S. Rozenberg, A. M. M. Schreurs, J. Kroon, E. B. Starikov and T. Steiner, *New J. Chem.*, 2001, **25**, 1111–1113.

30 M. S. Rozenberg, T. Nishio and T. Steiner, *New J. Chem.*, 1999, **23**, 585–586.

31 T. F. Koetzle, W. C. Hamilton and R. Parthasarathy, *Acta Crystallogr. Sect B: Struct. Sci.*, 1972, **B28**, 2083–2090.

32 W. Mikenda and S. Steinböck, *J. Mol. Struct.*, 1996, **384**, 159–163.

33 W. Mikenda and S. Steinböck, *J. Mol. Struct.*, 1994, **326**, 123–130.

34 W. Mikenda, *J. Mol. Struct.*, 1986, **147**, 1–15.

35 L. A. Dementjeva, A. V. Iogansen and G. A. Kurchi, *Zh. Prikl. Spektrosk., SSSR*, 1969, **10**, 625–629 (Russ.), Engl. Transl.

36 H. Torii, T. Tatsumi, T. Kanazavwa and M. Tasumi, *J. Phys. Chem. B*, 1998, **102**, 309–314.

37 E. Arbely and I. T. Arkin, *J. Am. Chem. Soc.*, 2004, **126**, 5362–5363.

38 M. Sh. Rozenberg, A. V. Iogansen, A. A. Mashkovsky and S. E. Odinokov, *Spectrosc. Lett.*, 1972, **5**, 75–80.

39 M. S. Lehmann, T. F. Koetzle and W. C. Hamilton, *J. Am. Chem. Soc.*, 1972, **94**, 2657–2660.

40 P.-G. Jonsson and A. Kvick, *Acta Crystallogr. Sect B: Struct. Sci.*, 1972, **B28**, 1827–1833.

41 M. N. Frey, M. S. Lehmann, T. F. Koetzle and W. C. Hamilton, *Acta Crystallogr. Sect B: Struct. Sci.*, 1973, **B29**, 876–884.

42 M. N. Frey, T. F. Koetzle, M. S. Lehmann and W. C. Hamilton, *J. Chem. Phys.*, 1973, **58**, 2547–2556.

43 J. Janczak, D. Zobel and P. Luger, *Acta Crystallogr. Sect. C: Cryst. Struct. Commun.*, 1997, **C53**, 1901–1904.

44 S. G. Stepanian, I. D. Reva, E. D. Radchenko and L. Adamowicz, *J. Phys. Chem. A*, 1998, **102**, 4623–4629.

45 S. G. Stepanian, I. D. Reva, E. D. Radchenko and L. Adamowicz, *J. Phys. Chem. A*, 1999, **103**, 4404–4412.

46 A. V. Iogansen, G. A. Kurkchi and L. A. Dementjeva, *J. Mol. Struct.*, 1976, **35**, 101–114.