Acta Cryst. (1993). B**49**, 348–356

X-ray Studies on Crystalline Complexes Involving Amino Acids and Peptides. XXV. Structures of DL-Proline Hemisuccinic Acid and Glycyl-L-Histidinium Semisuccinate Monohydrate and a Comparative Study of Amino-Acid and Peptide Complexes of Succinic Acid

BY G. SRIDHAR PRASAD AND M. VIJAYAN

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, *India*

(*Received* 28 *April* 1992; *accepted* 6 *August* 1992)

Abstract

DL-Proline hemisuccinic acid, $C_5H_9NO_2.\frac{1}{2}C_4H_6O_4$, $M_r = 174.2$, $P2_1/c$, $a = 5.254$ (1), $b = 17.480$ (1), $c = 10.230$ (1) Å, $\beta = 119.60$ (6)°, $Z = 4$, $D_m = 1.41$ (4), $D_x = 1.42$ g cm^{-3}, $R = 0.045$ for 973 observed reflections. Glycyl-L-histidinium semisuccinate monohydrate, $C_8H_{13}N_4O_3^+.C_4H_5O_4^-.H_2O$, $M_r = 348.4$, $P2_1$, $a = 4.864$ (1), $b = 17.071$ (2), $c = 9.397$ (1) Å, $\beta = 90.58°$, $Z = 2$, $D_m = 1.45$ (1), $D_x = 1.48$ g cm^{-3}, $R = 0.027$ for 1610 observed reflections. Normal amino-acid and dipeptide aggregation patterns are preserved in the structures in spite of the presence of succinic acid/semisuccinate ions. In both the structures, the amino-acid/dipeptide layers stack in such a way that the succinic acid molecules/semisuccinate ions are enclosed in voids created during stacking. Substantial variability in the ionization state and the stoichiometry is observed in amino-acid and peptide complexes of succinic acid. Succinic acid molecules and succinate ions appear to prefer a planar centrosymmetric conformation with the two carboxyl (carboxylate) groups *trans* with respect to the central C—C bond. Considerable variation is seen in the departure from and modification of normal amino-acid aggregation patterns produced by the presence of succinic acid. Some of the complexes can be described as inclusion compounds with the amino acid/dipeptide as the 'host' and succinic acid/semisuccinate/succinate as the 'guest'. The effects of change in chirality, though very substantial, are not the same in different pairs of complexes involving DL and L isomers of the same amino acid.

Introduction

The X-ray studies on crystalline complexes involving amino acids and peptides being pursued in this laboratory have led to results which are of considerable importance in relation to prebiotic polymerization and chiral discrimination during chemical evolution (Vijayan, 1980, 1988; Suresh & Vijayan, 1983*a*, 1985*a*; Suresh, Ramaswamy & Vijayan, 1986; Soman, Vijayan, Ramakrishnan & Guru Row, 1990). These studies have also resulted in the identification of specific interactions and characteristic interaction patterns which may have been important in early self-assembly processes (Salunke & Vijayan, 1981; Soman, Suresh & Vijayan, 1988). In this context, it is worthwhile to prepare and analyse complexes of amino acids and peptides with simple organic compounds which are believed to have existed on primitive Earth. Succinic acid is one such important compound (Miller & Orgel, 1974; Kvenvolden, Lawless & Ponnamperuma, 1971) and we have already analysed its complexes with DL and L forms of arginine (Prasad & Vijayan, 1990), lysine (Prasad & Vijayan, 1991) and histidine (Prasad & Vijayan, 1992). Here we report the crystal structures of succinic acid complexes with DL-proline and glycyl-L-histidine. Also presented is a comparative study of the structural features observed in the amino-acid and peptide complexes of succinic acid. In addition to their possible relevance to chemical evolution, these complexes contain a wealth of information which is of intrinsic interest in relation to molecular interactions and aggregation, and variability in ionization state and stoichiometry.

Experimental

The complexes of succinic acid with DL-proline and glycyl-L-histidine were prepared by diffusion of ethanol and propanol, respectively, into aqueous solutions of the components in molar proportions. The space group and the unit-cell dimensions were determined from X-ray diffraction photographs. The cell parameters were subsequently refined on a CAD-4 diffractometer which was also used to collect intensity data using nickel-filtered copper radiation. The densities were measured by flotation in a mixture of benzene and carbon tetrachloride. The structures were solved using the direct-methods programs *MULTAN* (Main *et al.*, 1984) and *SHELX*86 (Sheldrick, 1986) and refined using *SHELX*76

0108-7681/93/020348-09$06.00

Table 1. *Experimental details and refinement parameters*

	DL-Proline hemisuccinic acid	Glycyl-L-histidinium semisuccinate monohydrate
μ(Cu $K\alpha$) (mm^{-1})	0.948	1.034
F(000)	372	368
Crystal dimensions (mm)	0.20 × 0.22 × 0.27	0.10 × 0.10 × 0.55
Radiation used	Cu $K\alpha$	Cu $K\alpha$
θ range of 25 standard reflections used for refining lattice parameters (°)	10–45	13–46
Maximum Bragg angle (°)	60	75
Ranges of h	0 to 5	0 to 6
k	0 to 19	0 to 21
l	−11 to 11	−11 to 11
ΔF^2 for standard reflections	0.026	0.022
Number of reflections measured	1471	2091
Number of unique reflections with $I > 2\sigma(I)$	973	1610
R_{int}	0.018	0.012
R	0.045	0.027
wR	0.064	0.033
S	1.419	0.910
Weighting function	$1.0/[\sigma^2(F_o) + 0.003178(F_o)^2]$	$1.0/[\sigma^2(F_o) + 0.001627(F_o)^2]$
$(\Delta/\sigma)_{max}$	0.32	0.15
$\Delta\rho_{max}$ (e Å^{-3})	0.25	0.12
$\Delta\rho_{min}$ (e Å^{-3})	−0.22	−0.22

(Sheldrick, 1976). The non-hydrogen atoms were refined anisotropically and the hydrogen atoms, located in difference Fourier maps, isotropically using reflections with $I > 2\sigma(I)$. Details of data collection and refinement parameters are presented in Table 1. The coordinates and the equivalent isotropic thermal parameters (Hamilton, 1959) of the non-hydrogen atoms in the two complexes are given in Tables 2 and 3.*

Discussion

The bond lengths and angles in the structure, except the C(3)—C(4) length in the pyrrolidine ring in the proline complex, are normal (Allen, Kennard, Watson, Brammer, Orpen & Taylor, 1989; Benedetti, 1977; De Tar & Luthra, 1977; Vijayan, 1976). The C(3)—C(4) length is rather short which could be due to the high 'temperature' factor associated with C(4), a phenomenon observed in several other crystal structures containing proline (Ashida & Kakudo, 1974; Karle, 1974; Kartha, Ashida & Kakudo, 1974; Nair, Ramaprasad, Nagaraj, Balaram & Vijayan, 1981). The high temperature factor could have resulted from the presence of a mixture of slightly different conformations of the pyrrolidine ring.

The bond lengths and angles and the location of the hydrogen atoms indicate that the amino acid and the dipeptide molecules are zwitterionic. The proline molecule is neutral while the dipeptide molecule is positively charged because of protonation of the imidazole group. In the proline complex, the amino-acid molecule is located at a general position while the succinic acid molecule occupies an inversion centre. One of the carboxyl groups in the succinic acid molecule in the dipeptide complex is deprotonated; also the structure contains a water molecule. The dipeptide complex can therefore be described as glycyl-L-histidinium semisuccinate monohydrate.

Table 2. *Positional parameters* ($\times 10^4$) *and equivalent isotropic temperature factors* ($\times 10^3$) *of non-hydrogen atoms in* DL-*proline hemisuccinic acid*

Estimated standard deviations are given in parentheses.

$U_{eq} = (1/3)\sum_i\sum_j U_{ij}a_i^*a_j^*\mathbf{a}_i.\mathbf{a}_j$.

	x	y	z	U_{eq} (Å^2)
N(1)	−1986 (5)	2900 (1)	−535 (2)	30 (1)
O(1)	2930 (4)	2596 (1)	1976 (2)	50 (1)
O(2)	1386 (4)	2999 (1)	3520 (2)	45 (1)
C(1)	1032 (5)	2858 (1)	2235 (2)	33 (1)
C(2)	−1945 (5)	3061 (1)	918 (2)	30 (1)
C(3)	−2652 (7)	3912 (1)	862 (3)	50 (1)
C(4)	−3650 (12)	4162 (2)	−685 (5)	87 (3)
C(5)	−1879 (9)	3666 (2)	−1178 (4)	50 (2)
O(11)	5201 (4)	5409 (1)	6667 b(2)	52 (1)
O(12)	3601 (4)	6296 (1)	4872 (2)	43 (1)
C(13)	3255 (5)	5675 (1)	5511 (2)	33 (1)
C(14)	240 (5)	5354 (1)	4639 (3)	41 (1)

Table 3. *Positional parameters and equivalent isotropic temperature factors* ($\times 10^4$) *of non-hydrogen atoms in glycyl-*L*-histidinium semisuccinate monohydrate*

Estimated standard deviations are given in parentheses.

$U_{eq} = (1/3)\sum_i\sum_j U_{ij}a_i^*a_j^*\mathbf{a}_i.\mathbf{a}_j$.

	x	y	z	U_{eq} (Å^2)
N(1)	6703 (4)	5482	3121 (2)	307 (5)
C(2)	4673 (4)	6128 (2)	3157 (2)	304 (6)
C(3)	6131 (4)	6916 (1)	3257 (2)	237 (5)
O(4)	8585 (3)	6983 (1)	3017 (2)	362 (4)
N(5)	4453 (3)	7498 (1)	3633 (2)	251 (4)
C(6)	5390 (4)	8291 (1)	3907 (2)	231 (5)
C(7)	4913 (4)	8546 (2)	5457 (2)	261 (5)
O(8)	3276 (4)	8167 (1)	6196 (2)	412 (5)
O(9)	6176 (4)	9144 (1)	5853 (2)	355 (5)
C(10)	3904 (4)	8880 (1)	2932 (2)	281 (6)
C(11)	4373 (4)	8755 (1)	1380 (2)	276 (5)
N(12)	6404 (4)	9137 (2)	644 (2)	307 (5)
C(13)	6240 (5)	8938 (2)	−716 (2)	355 (6)
N(14)	4208 (4)	8435 (2)	−881 (2)	367 (5)
C(15)	2993 (5)	8306 (2)	418 (2)	354 (6)
O(21)	10155 (3)	5235 (1)	−1637 (2)	360 (4)
C(22)	9467 (4)	5280 (2)	679 (2)	452 (5)
C(23)	8888 (4)	5490 (2)	−567 (2)	301 (5)
C(24)	6542 (4)	6068 (2)	−781 (3)	359 (6)
C(25)	6400 (5)	6428 (2)	−2256 (3)	364 (4)
C(26)	8736 (4)	6985 (2)	−2489 (2)	306 (5)
O(27)	10070 (4)	7270 (2)	−1545 (2)	563 (7)
O(28)	9236 (5)	7133 (2)	−3835 (2)	550 (6)
W(1)	464 (4)	5235 (2)	5378 (2)	474 (6)

Molecular conformation

The torsion angles that describe the conformation of the molecules are given in Table 4. The pyrrolidine ring conformation in proline is intermediate between 'envelope' and 'half-chair' (Ashida &

* Lists of structure factors, anisotropic thermal parameters, bond lengths and angles, and H-atom parameters have been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 55661 (23 pp.). Copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England. [CIF reference: AL0543]

Table 4. *Conformational angles in the amino-acid and peptide molecules* (°)

Estimated standard deviations are given in parentheses.

Proline	
θ [C(5)—N(1)—C(2)—C(3)]	15.3 (3)
χ^1 [N(1)—C(2)—C(3)—C(4)]	11.4 (4)
χ^2 [C(2)—C(3)—C(4)—C(5)]	−33.2 (4)
χ^3 [C(3)—C(4)—C(5)—N(1)]	42.0 (4)
χ^4 [C(4)—C(5)—N(1)—C(2)]	−34.9 (3)
Glycyl-L-histidinium	
ψ_1 [N(1)—C(2)—C(3)—N(5)]	−164.9 (2)
ω [C(2)—C(3)—N(5)—C(6)]	175.0 (2)
φ_2 [C(3)—N(5)—C(6)—C(7)]	−116.9 (2)
ψ_2^1 [N(5)—C(6)—C(7)—O(8)]	−17.1 (3)
χ_2^1 [N(5)—C(6)—C(10)—C(11)]	−61.6 (6)
χ_2^{21} [C(6)—C(10)—C(11)—N(12)]	−93.9 (2)

Kakudo, 1974; Nair & Vijayan, 1981). As has been shown earlier (Suresh & Vijayan, 1985*a*; Vijayan, 1988), the main-chain conformation of a dipeptide is determined essentially by φ_2; the molecule is described as extended when the magnitude of φ_2 is between 120 and 180°, and folded when it is between 60 and 120°. The observed value of $\varphi_2 = -116.9\ (2)°$ in the present structure indicates that the molecule has an intermediate conformation. The side-chain imidazole group is *trans* to the terminal carboxylate group ($\chi_2^1 \simeq -60°$) and thus has an open conformation (Bhat & Vijayan, 1978; Prasad & Vijayan, 1992). χ_2^{21} (−94°) is close to one of the two expected values (±90°).

The centrosymmetric succinic acid molecule in the proline complex is planar with a *trans* conformation about the central C—C bond. In the dipeptide complex, however, the torsion angle about this bond is −70.3 (3)° and the semisuccinate ion has a folded conformation. Also the C—COO^- and C—COOH groups each deviate substantially from planarity. In both the complexes, the carboxyl hydrogen atom in the succinic acid molecule (semisuccinate ion) is *cis* with respect to the double-bonded oxygen atom. A more detailed discussion of succinic acid conformation is given in the next section.

Molecular aggregation and hydrogen bonding

The crystal structure of the proline complex is illustrated in Fig. 1 and the parameters of the hydrogen bonds that stabilize it are listed in Table 5. The proline molecules aggregate into hydrogen-bonded layers parallel to the *ac* plane at $y = \frac{1}{4}$ and $\frac{3}{4}$. Such a layer as seen along the *b* axis is shown in Fig. 2. The molecules within the layer are related to one another by **c** glide and cell translations. The hydrogen bonds that interconnect the molecules in the layer give rise to two DL-type head-to-tail sequences (Suresh & Vijayan, 1983*a*), a DL2 sequence parallel to **c** and a DL1 sequence along the [201] direction. The arrangement is remarkably similar to the aggregation patterns stabilized by *S*1 and *S*2 sequences found in the crystals of most hydrophobic L-amino acids (Suresh & Vijayan, 1983*a*).

The proline layers are stacked along the *b* axis (Fig. 1) in such a way that hollow cylinders parallel to **a** centred around $y = z = 0$ and $y = z = \frac{1}{2}$ are created. The succinic acid molecules are situated in these cylinders. The cross section of the cylinder around $y = z = \frac{1}{2}$ is clearly discernible in Fig. 1. Each cylinder is made up of two halves. The proline molecules in each half are hydrogen bonded among themselves. The two halves are held together by van der Waals interactions involving C(3) and C(4) and their symmetry equivalents. Each succinic acid molecule forms a hydrogen bond with O(2) in one half of the hollow cylinder and a symmetry-related hydrogen bond with O(2) in the other half. Thus the complex could be considered as an inclusion compound with proline as the host and succinic acid as the guest.

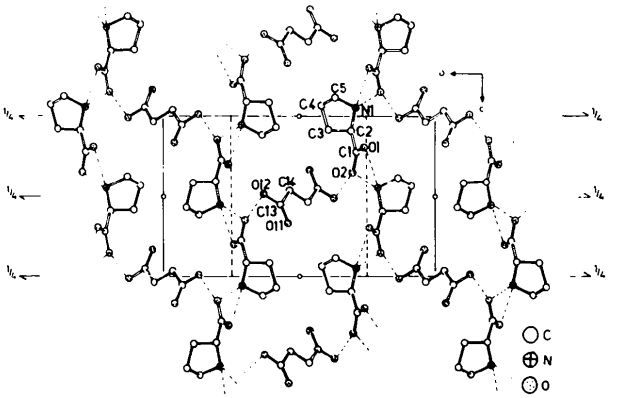


Fig. 1. Crystal structure of DL-proline hemisuccinic acid as projected on to the *bc* plane. In this and in subsequent figures hydrogen bonds are indicated by broken lines.

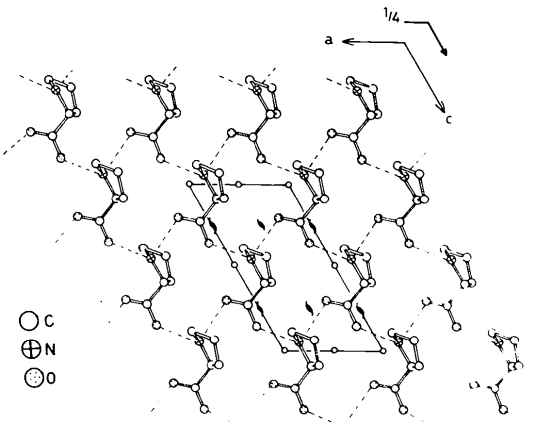


Fig. 2. The arrangement of molecules in the proline layer.

Table 5. *Hydrogen-bond parameters in* DL-*proline hemisuccinic acid*

Estimated standard deviations are given in parentheses.

A—H···*B*	*A*···*B* (Å)	H—*A*···*B* (°)	Symmetry of atom *B*
N(1)—H1N(1)···O(2)	2.870 (5)	26 (2)	$x, 0.5-y+1, 0.5+z-1$
N(1)—H2N(1)···O(1)	2.766 (2)	18 (2)	$x+1, 0.5-y+1, 0.5+z-1$
O(12)—HO(12)···O(2)	2.623 (3)	5 (3)	$-x+1, -y+1, -z+1$

Fig. 3 illustrates the crystal structure of the dipeptide complex. The parameters of the hydrogen bonds in this complex are listed in Table 6. The dipeptide molecules and the water molecules are arranged in layers parallel to the *ab* plane, with the histidyl side chains protruding alternately on either side of each layer. The arrangement of molecules in the layer is illustrated in Fig. 4.

Suresh & Vijayan (1985*a*) have previously predicted the presence of seven idealized aggregation patterns for the L-alanyl-L-alanyl dipeptide, six of them planar. They have also demonstrated that most dipeptides follow these aggregation patterns in their crystal structures irrespective of the nature of the side chains and in spite of disrupting influences such

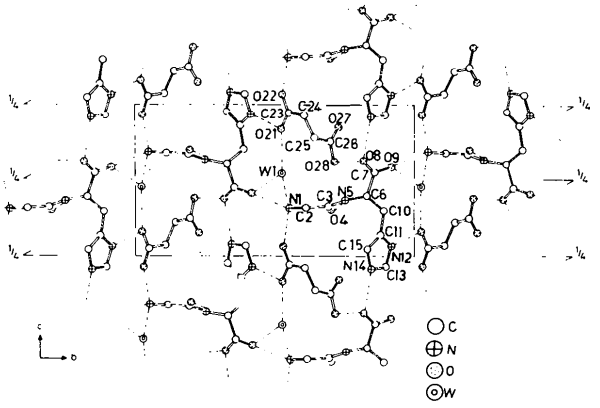


Fig. 3. The crystal structure of glycyl-L-histidinium semisuccinate monohydrate as projected on to the *bc* plane.

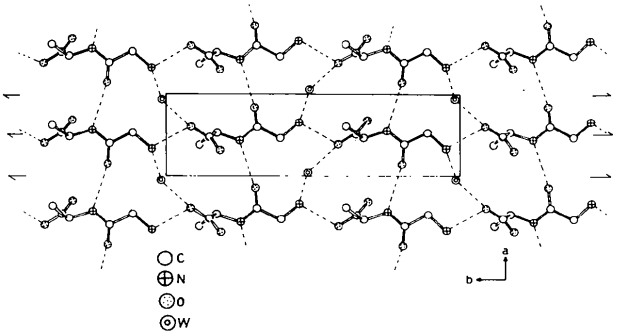


Fig. 4. The arrangement in the layer containing the peptide and water molecules.

Table 6. *Hydrogen-bond parameters in glycyl-*L-*histidinium semisuccinate monohydrate*

Estimated standard deviations are given in parentheses.

A—H···*B*	*A*···*B* (Å)	H—*A*···*B* (°)	Symmetry of atom *B*
N(1)—H1N(1)···O(9)	2.850 (2)	9 (2)	$-x+1, y-\frac{1}{2}, -z+1$
N(1)—H2N(1)···O(22)	2.694 (3)	10 (2)	x, y, z
N(1)—H3N(1)···*W*(1)	2.818 (3)	17 (3)	$x+1, y, z$
N(5)—HN(5)···O(4)	3.036 (2)	17 (2)	$x-1, y, z$
N(12)—HN(12)···O(21)	2.678 (3)	6 (3)	$-x+2, y+\frac{1}{2}, -z$
N(14)—HN(14)···O(8)	2.817 (3)	20 (2)	$x, y, z-1$
O(28)—HO(28)···O(8)	2.640 (3)	8 (3)	$x+1, y, z-1$
W(1)—H1*W*(1)···O(21)	2.810 (3)	7 (4)	$x-1, y, z+1$
W(1)—H2*W*(12)···O(9)	2.741 (3)	11 (3)	$-x+1, y-\frac{1}{2}, -z+1$

as those caused by the presence of water molecules. The aggregation of dipeptide molecules in the present structure is similar to one of the idealized patterns, termed EZA (Fig. 7 in Suresh & Vijayan, 1985*a*). In the idealized pattern, both carboxylate oxygen atoms hydrogen bond to the amino nitrogen of the neighbouring screw-related molecule. These hydrogen bonds give rise to a *Z*1*Z*2 head-to-tail sequence with a periodicity of 18 Å (Suresh & Vijayan, 1985*a*). In the present structure one of the two hydrogen bonds does not form and the head-to-tail sequence is of type *Z*2. The periodicity of the sequence ($b = 17.1$ Å) is, however, close to the value in the idealized pattern. In the idealized and the observed patterns, the peptide nitrogen is hydrogen bonded to the peptide carbonyl group of a neighbouring molecule. This hydrogen bond gives rise to a *S*5→4 sequence (Suresh & Vijayan, 1985*a*) perpendicular to the 2_1 screw axis. The periodicity of this sequence in the predicted idealized Ala-Ala pattern is 4.6 Å. The corresponding periodicity in the complex is $a = 4.9$ Å. The third hydrogen bond formed by the amino nitrogen atom in the idealized pattern gives rise to a *Z*2 sequence parallel to the *Z*1*Z*2 sequence with the same periodicity. This hydrogen bond, and hence the head-to-tail sequence, does not exist in the dipeptide layer in the complex. However, it is interesting to note that the N—H···O hydrogen bond in the idealized pattern is replaced by an N—H···O(*W*)···O water bridge in the complex, as can be seen from Fig. 4. Thus the geometrical relationship among the molecules remains nearly the same.

As indicated earlier, the idealized dipeptide aggregation pattern was derived from simple geometrical and hydrogen-bonding considerations for L-Ala-L-Ala. Both the residues are different in the complex. One of the residues is now histidine, the side chain of which can, unlike that of alanine, form hydrogen bonds. Furthermore, the structure contains semisuccinate ions and water molecules which might normally be expected to influence molecular aggregation strongly through ionic interactions and hydrogen bonding. Yet it is remarkable that the aggregation of the dipeptide in its layer in the complex is substantially similar to that in a pattern

predicted for L-Ala-L-Ala. This reinforces the earlier conclusion (Suresh & Vijayan, 1985*a*,*b*) that dipeptide aggregation is essentially controlled by mainchain interactions and is largely unaffected by disturbing influences.

The terminal carboxylate and the side-chain imidazole groups of each molecule point at opposite directions from the dipeptide–water layer. The layers are stacked along the **c** direction in the crystal in such a way as to bring each carboxylate group in one layer at hydrogen-bonding proximity to an imidazole group in an adjacent layer, as can be seen in Fig. 3. The aggregation is such that it creates two voids per unit cell. The semisuccinate ions are located in these voids. Thus, the dipeptide molecules may be considered as host molecules and the semisuccinate ions as the guest molecules in the structure. Each ion is surrounded by four dipeptide molecules. As shown in Fig. 3, the carboxylate group in the ion interacts through hydrogen bonds with the terminal amino group of one dipeptide molecule, N^{δ} in the imidazole group of another and the water molecule associated with the dipeptide layer. The carboxyl group in the ion forms an O—H···O hydrogen bond, as a donor, with the terminal carboxylate group of yet another of the four surrounding dipeptide molecules.

Protonated positively charged amino groups and deprotonated negatively charged carboxyl groups have been shown to give rise to characteristic interaction patterns (Vijayan, 1988). One such pattern involves an infinite array of alternating amino and carboxylate groups (Salunke & Vijayan, 1982, 1983, 1984; Suresh & Vijayan, 1983*b*; Soman, Suresh & Vijayan, 1988; Soman & Vijayan, 1988; Vijayan, 1988). The present structure contains a variant of such a pattern. This variant has water molecules, in addition to amino and carboxylate groups, in the array and runs parallel to **c** at $x \simeq 1.0$ and $y \simeq 0.5$ (Fig. 3).

Comparative study of amino-acid and peptide complexes of succinic acid

Ionization state and stoichiometry

Succinic acid is a dicarboxylic acid and it can, in principle, exist in the three ionization states [(I), (II) and (III)] illustrated in Fig. 5. In addition to these expected species, a complex dimeric species with a double negative charge (IV), illustrated in Fig. 6, was also observed in one of the complexes (Prasad & Vijayan, 1991). The two molecules in the dimer are connected to each other by a symmetric O···H···O hydrogen bond.

The composition of the nine complexes is given in Table 7. Disregarding differences in ionization states, the stoichiometry of succinic acid in the complexes with DL-arginine, DL-histidine and DL-proline is 1:2 whereas in the remaining six complexes it is 1:1. No reason for this difference is immediately obvious. It may, however, be noted that all three complexes with a 1:2 stoichiometry involve DL-amino acids although the stoichiometry in the DL-lysine complex is 1:1. Only one ionic species exists in each of the 1:2 complexes, the succinate ion in the DL-arginine and the DL-histidine complexes, and neutral succinic acid in the DL-proline complex. Three of the 1:1 complexes, all involving L-amino acids, contain only semisuccinate ions while each of the remaining three has equal proportions of succinate ions and succinic acid molecules. It is interesting to note that the semisuccinate ion does not occur in any complex involving DL-amino acids. In such complexes it is found that succinate ions and/or succinic acid molecules are located at inversion centres. The semisuccinate ion, however, cannot have an inversion centre. It could well be that a centrosymmetric arrangement is favoured in molecular packing wherever possible and this could be one factor responsible for the absence of semisuccinate ions in the complexes involving DL-amino acids.

The neutral succinic acid molecules, the singly charged semisuccinate ions and the doubly charged

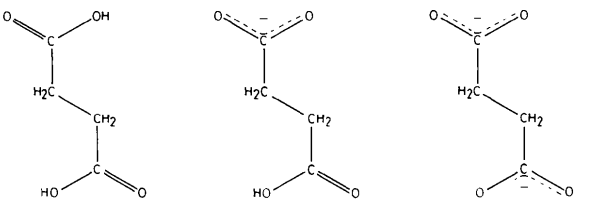


Fig. 5. Possible ionization states of succinic acid.

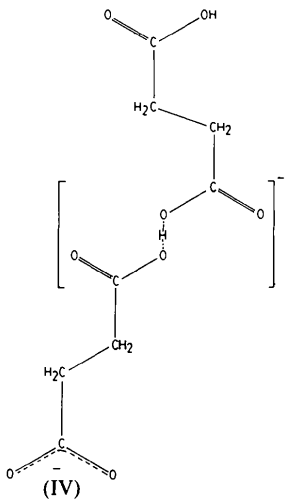


Fig. 6. Complex dimeric species of succinic acid.

Table 7. *Composition of complexes involving succinic acid*

Complex with	(I)*	(II)*	(III)*	(IV)*	H_2O	Ref.
DL-Arginine			$\frac{1}{2}$		2	(*a*)
L-Arginine	$\frac{1}{2}$		$\frac{1}{2}$		1	(*a*)
DL-Lysine	$\frac{1}{2}$		$\frac{1}{2}$			(*b*)
L-Lysine [form (I)]		1				(*b*)
L-Lysine [form (II)]	$\frac{1}{4}$		$\frac{1}{4}$	$\frac{1}{4}$		(*b*)
DL-Histidine			$\frac{1}{2}$		2	(*c*)
L-Histidine		1			3	(*c*)
DL-Proline	$\frac{1}{2}$					(*d*)
Glycyl-L-histidine		1			1	(*d*)

References: (*a*) Prasad & Vijayan (1990); (*b*) Prasad & Vijayan (1991); (*c*) Prasad & Vijayan (1992); (*d*) present study.

* Please see Figs. 5 and 6.

succinate ions occur with nearly equal frequency in the complexes. Two or more ionization states coexist in three crystals. Thus small environmental changes are sufficient to produce different ionization states in succinic acid. This is a situation similar to that found in the imidazole group in histidine which very often occurs in the catalytic sites of enzymes. Thus, the possible catalytic role of succinic acid in prebiotic processes may deserve further examination.

Succinic acid conformation

The torsion angles that define the molecular conformation of neutral or ionized succinic acid are shown in Fig. 7. O(2) and O(2′) are protonated in the neutral molecule. O(2) or O(2′) is deprotonated in the semisuccinate ion while both are deprotonated in the succinate ion. θ_1 and θ_1' essentially describe individually the conformations of the two carboxyl (carboxylate) groups with respect to the rest of the molecule. θ_2 defines the mutual orientation of the carboxyl (carboxylate) groups. The succinic acid molecule (ion) is not asymmetric and hence there is no intrinsic difference between the two carboxyl (carboxylate) groups contained therein. Therefore, at least to a first approximation, θ_1 and θ_1' can be considered simultaneously. The distribution of θ_1 and θ_1' in the complexes is given in Fig. 8. In a majority of cases, the magnitude of the torsion angle is less than 10°. There are only two instances where it is greater than 20°.

The correlation between θ_1 and θ_1' in the different structures is illustrated in Fig. 9. When the molecule/ion is centrosymmetric $\theta_1' = -\theta_1$ and the points in the plot illustrated in the figure would fall on a straight line with slope -1. Among the 14 molecules/ions in the complexes only five are crystallographically centrosymmetric. However, as can be seen from the figure, the joint distribution of θ_1 and θ_1' indicates the tendency of other molecules/ions also to be centrosymmetric.

The distribution of θ_2 in the complexes is shown in Fig. 10. As expected, an overwhelming majority of the values are in the close neighbourhood of 180°. At

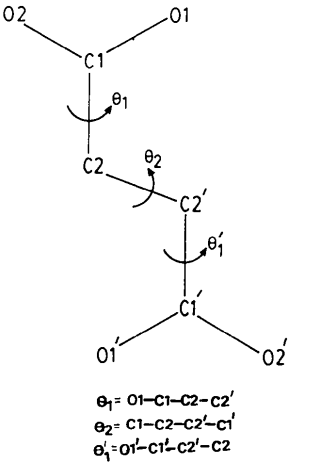


Fig. 7. Torsion angles that define the conformation of the succinic acid molecule/semisuccinate ion/succinate ion.

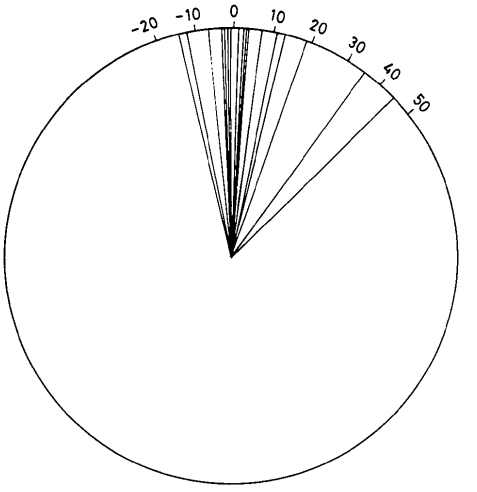


Fig. 8. Distribution of θ_1 and θ_1' in the complexes.

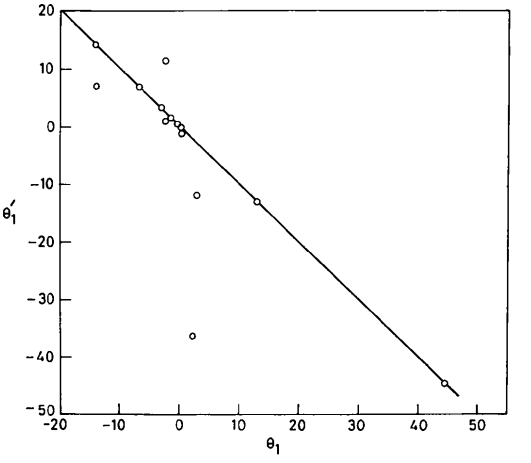
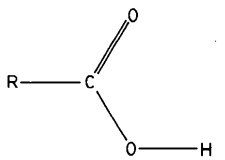


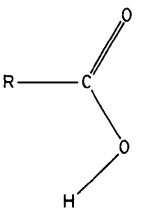
Fig. 9. Joint distribution of θ_1 and θ_1' in the complexes.

$\theta_2 = 180°$, the two carboxyl (carboxylate) groups in the molecule (ion) are kept as far apart as possible. Obviously, steric and ionic considerations favour such a conformation. The other possibilities cannot be ruled out however. There is one instance, namely in the dipeptide complex, where a semisuccinate ion assumes a folded conformation with $\theta_2 = -70°$.

All the succinic acid molecules and the succinate ions in the complexes are exactly or very nearly centrosymmetric, with the two carboxyl (carboxylate) groups *trans* with respect to the central C—C bond. Semisuccinate ions cannot be exactly centrosymmetric and indeed departures from inversion symmetry are most substantial in these ions. The folded ion in the dipeptide complex does not have even an approximate inversion centre. Substantial departures from centrosymmetric geometry are exhibited by the semisuccinate ion in form (I) of the L-lysine complex (Prasad & Vijayan, 1991) and, to a lesser extent, in the L-histidine complex (Prasad & Vijayan, 1992).



(I)



(II)

The remaining conformational aspect pertaining to succinic acid to be considered is the location of the hydrogen atom in the carboxyl group. The O—H bond may be either synplanar (*cis*) to the C=O bond or antiplanar (*trans*). From infrared studies of gaseous monomeric formic acid (Miyazawa & Pitzer, 1959), it has been shown earlier that the *cis* conformation is energetically more favourable than the *trans* conformation to the extent of 8.4 kJ mol^{-1}. A subsequent study indicated the difference to be as much as 16.8 kJ mol^{-1} (Lide, 1964). The results of the present investigations agree with the finding that the *cis* conformation is more favourable than the *trans* conformation, but not overwhelmingly so. In the eight crystallographically independent carboxyl groups in the complexes in which the hydrogen atom has been located, five have the *cis* conformation while the remaining three have the *trans* conformation. The hydrogen atom is also located in the *cis* position in the two crystal structures containing succinic acid reported earlier (Huang, Leiserowitz & Schmitt, 1973; Leviel, Auvert & Savariault, 1981).

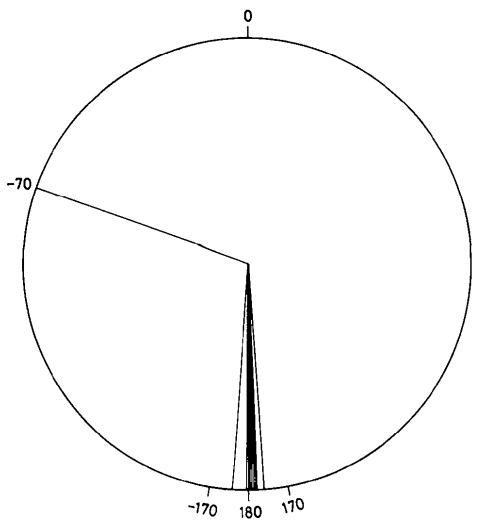


Fig. 10. Distribution of θ_2 in the complexes.

Molecular aggregation

As shown earlier, DL-proline hemisuccinic acid and glycyl-L-histidinium semisuccinate monohydrate represent situations in which the normal amino-acid aggregation patterns are preserved in spite of the presence of succinic acid molecules/semisuccinate ions. This is also partially true for the histidine complexes. The aggregation pattern in L-histidine (Madden, McGandy, Seeman, Harding & Hoy, 1972; Madden, McGandy & Seeman, 1972), as in the crystals of most hydrophilic L-amino acids, is stabilized by an *S*2 head-to-tail sequence as well as a *Z*2 sequence (Suresh & Vijayan, 1983*a*). The *S*2 sequence is disrupted in L-histidine semisuccinate trihydrate; however, the *Z*2 sequence is retained with nearly the same periodicity and similar structural characteristics (Prasad & Vijayan, 1992). In DL-histidine (Edington & Harding, 1974) as well as in DL-histidine hemisuccinate dihydrate (Prasad & Vijayan, 1992) the L- and D-amino-acid molecules dimerize across inversion centres through hydrogen bonding and they form double ribbons through intermolecular interactions between imidazole and carboxylate groups. However, hydrogen bonds involving succinate ions and water molecules interconnect the double ribbons in the complex instead of the DL2-type head-to-tail hydrogen bonds in DL-histidine. The resemblance of the amino-acid aggregation patterns observed in the succinic acid complexes of arginine and lysine (Prasad & Vijayan, 1990, 1991) to those observed earlier (Vijayan, 1988) is not particularly striking although they contain common elements such as head-to-tail sequences. Thus the departures from and modifications of normal amino-acid aggregation patterns produced by the presence of succinic acid show considerable variation.

Some of the complexes may be described as inclusion compounds. In those involving DL-arginine and DL-histidine, the amino-acid molecules aggregate into columns or ribbons stabilized by main-chain–main-chain as well as main-chain–side-chain interactions. In the DL-arginine complex, the amino-acid columns and water molecules pack in such a way that the succinate ions are trapped in the voids created in packing (Fig. 1 in Prasad & Vijayan, 1990). In the DL-histidine complex, however, the succinate ions are essentially trapped between two adjacent amino-acid columns or ribbons (Fig. 5 in Prasad & Vijayan, 1992). In DL-proline hemisuccinic acid and glycyl-L-histidinium semisuccinate monohydrate, the amino-acid/dipeptide layers stack along the largest crystallographic axis in such a way as to enclose succinic acid molecules/semisuccinate ions among them (Figs. 1 and 3). Most of the remaining complexes follow the more familiar pattern in binary complexes involving amino acids (Vijayan, 1988) in which dissimilar molecules aggregate into separate alternating layers.

Complexes involving DL- and L-amino acids have been prepared and analysed in the case of arginine, lysine and histidine. Initially it appeared that these complexes could possibly provide some broad generalizations regarding the effect of chirality on molecular aggregation and other structural features. This expectation was not entirely realized as indeed was the case with the LL and LD complexes among amino acids (Vijayan, 1988). The effect of change in chirality of half the amino-acid molecules in the structure (as happens when going from a complex involving a DL-amino acid to that involving only the L isomer of the same amino acid) is indeed substantial, but it differs from amino acid to amino acid.

As has been shown earlier (Salunke & Vijayan, 1981; Soman, Suresh & Vijayan, 1988; Vijayan, 1988) different functional groups found in proteins have the propensity to take part in specific interactions or characteristic interaction patterns. Earlier work in this laboratory (Vijayan, 1988; Soman, Rao, Radhakrishnan & Vijayan, 1989; Soman & Vijayan, 1989) and the present work also demonstrate that amino acids, peptides, and other relevant molecules in different combinations give rise to different, often predictable, aggregation patterns with characteristic features. The elucidation of the geometrical features of several possible specific interactions, characteristic interaction patterns and aggregation patterns is a precondition for understanding the self-assembly processes that may have led to the first multimolecular systems. This would also be highly relevant to present-day biological organization.

The authors thank the Department of Science and Technology, India, for financial support.

References

ALLEN, F. H., KENNARD, O., WATSON, D. G., BRAMMER, L., ORPEN, A. G. & TAYLOR, R. (1989). *J. Chem. Soc. Perkin Trans.* 2, pp. S1–S19.

ASHIDA, T. & KAKUDO, M. (1974). *Bull. Chem. Soc. Jpn*, **47**, 1129–1133.

BENEDETTI, E. (1977). *Peptides.* In *Proceedings of the Fifth American Peptide Symposium*, edited by M. GOODMAN & J. MEIENHOFER, pp. 257–273. New York: John Wiley.

BHAT, T. N. & VIJAYAN, M. (1978). *Acta Cryst.* B**34**, 2556–2565.

DE TAR, D. F. & LUTHRA, N. P. (1977). *J. Am. Chem. Soc.* **99**, 1232–1244.

EDINGTON, P. & HARDING, M. M. (1974). *Acta Cryst.* B**30**, 204–206.

HAMILTON, W. C. (1959). *Acta Cryst.* **12**, 609–610.

HUANG, C. M., LEISEROWITZ, L. & SCHMITT, G. M. J. (1973). *J. Chem. Soc. Perkin Trans.* 2, pp. 503–508.

KARLE, I. L. (1974). *Biochemistry*, **13**, 2155–2162.

KARTHA, G., ASHIDA, T. & KAKUDO, M. (1974). *Acta Cryst.* B**30**, 1861–1866.

KVENVOLDEN, K. A., LAWLESS, J. G. & PONNAMPERUMA, C. (1971). *Proc. Natl Acad. Sci. USA*, **68**, 486–490.

LEVIEL, J.-L., AUVERT, G. & SAVARIAULT, J.-M. (1981). *Acta Cryst.* B**37**, 2185–2189.

LIDE, D. R. JR (1964). *Ann. Rev. Phys. Chem.* **15**, 225–250.

MADDEN, J. J., MCGANDY, E. L. & SEEMAN, N. C. (1972). *Acta Cryst.* B**28**, 2377–2382.

MADDEN, J. J., MCGANDY, E. L., SEEMAN, N. C., HARDING, M. M. & HOY, A. (1972). *Acta Cryst.* B**28**, 2382–2389.

MAIN, P., FISKE, S. J., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J.-P. & WOOLFSON, M. M. (1984). *MULTAN*84. *A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.* Univs. of York, England, and Louvain, Belgium.

MILLER, S. L. & ORGEL, L. E. (1974). *The Origins of Life on Earth.* New Jersey: Prentice-Hall.

MIYAZAWA, T. & PITZER, K. S. (1959). *J. Chem. Phys.* **30**, 1076–1078.

NAIR, C. M. K., RAMAPRASAD, S., NAGARAJ, R., BALARAM, P. & VIJAYAN, M. (1981). *Acta Cryst.* B**37**, 597–601.

NAIR, C. M. K. & VIJAYAN, M. (1981). *J. Indian Inst. Sci.* **63**, 81–103.

PRASAD, G. S. & VIJAYAN, M. (1990). *Int. J. Pept. Protein Res.* **35**, 357–364.

PRASAD, G. S. & VIJAYAN, M. (1991). *Acta Cryst.* B**47**, 927–935.

PRASAD, G. S. & VIJAYAN, M. (1992). *Biopolymers.* In the press.

SALUNKE, D. M. & VIJAYAN, M. (1981). *Int. J. Pept. Protein Res.* **18**, 348–351.

SALUNKE, D. M. & VIJAYAN, M. (1982). *Acta Cryst.* B**38**, 1328–1330.

SALUNKE, D. M. & VIJAYAN, M. (1983). *Int. J. Pept. Protein Res.* **22**, 154–160.

SALUNKE, D. M. & VIJAYAN, M. (1984). *Biochim. Biophys. Acta*, **798**, 175–179.

SHELDRICK, G. M. (1976). *SHELX*76. Program for crystal structure determination. Univ. of Cambridge, England.

SHELDRICK, G. M. (1986). *SHELX*86. Program for crystal structure determination. Univ. of Göttingen, Germany.

SOMAN, J., RAO, T., RADHAKRISHNAN, R. & VIJAYAN, M. (1989). *J. Biomol. Struct. Dyn.* **7**, 269–277.

SOMAN, J., SURESH, C. G. & VIJAYAN, M. (1988). *Int. J. Pept. Protein Res.* **32**, 352–360.

SOMAN, J. & VIJAYAN, M. (1988). *Acta Cryst.* C**44**, 1794–1797.

SOMAN, J. & VIJAYAN, M. (1989). *J. Biosci.* **14**, 111–125.

SOMAN, J., VIJAYAN, M., RAMAKRISHNAN, B. & GURU ROW, T. N. (1990). *Biopolymers*, **29**, 533–542.

SURESH, C. G., RAMASWAMY, J. & VIJAYAN, M. (1986). *Acta Cryst.* B**42**, 473–478.
SURESH, C. G. & VIJAYAN, M. (1983*a*). *Int. J. Pept. Protein Res.* **22**, 129–143.
SURESH, C. G. & VIJAYAN, M. (1983*b*). *Int. J. Pept. Protein Res.* **22**, 617–621.
SURESH, C. G. & VIJAYAN, M. (1985*a*). *Int. J. Pept. Protein Res.* **26**, 311–328.
SURESH, C. G. & VIJAYAN, M. (1985*b*). *Int. J. Pept. Protein Res.* **26**, 329–336.
VIJAYAN, M. (1976). In *Handbook of Biochemistry and Molecular Biology*, Vol. 2, *Proteins*, 3rd ed., edited by G. D. FASMAN. Cleveland: CRC Press.
VIJAYAN, M. (1980). *FEBS Lett.* **112**, 135–137.
VIJAYAN, M. (1988). *Prog. Biophys. Mol. Biol.* **52**, 71–99.

Acta Cryst. (1993). B**49**, 356–366

Structure and Conformation of the Muscarinic Agonists 3-(3-Amino-1,2,4-oxadiazol-5-yl)-1-azabicyclo[2.2.2]octane and 1,2,5,6-Tetrahydro-1-methyl-3-pyridinecarboxaldehyde Oxime and Related Tertiary Amine, Quaternary Ammonium and Sulfonium Analogues

BY H. KOOIJMAN, J. A. KANTERS AND J. KROON

Vakgroep Kristal- en Structuurchemie, Rijksuniversiteit Utrecht, Padualaan 8, 3584 *CH Utrecht, The Netherlands*

AND J. KELDER

Scientific Development Group, Organon International BV, PO Box 20, 5340 *BH Oss, The Netherlands*

(*Received* 21 *May* 1992; *accepted* 3 *August* 1992)

Abstract

The crystal structures of two muscarinic agonists, 3-(3-amino-1,2,4-oxadiazol-5-yl)-1-azabicyclo[2.2.2]-octane [L-660863, $C_9H_{14}N_4O$, $M_r = 194.24$, triclinic, $P\bar{1}$, $a = 6.607$ (1), $b = 8.2157$ (8), $c = 9.287$ (1) Å, $\alpha = 105.52$ (1), $\beta = 93.88$ (1), $\gamma = 91.79$ (1)°, $V = 484.0$ (1) Å^3, $Z = 2$, $D_x = 1.33$ g cm^{-3}, λ(Mo $K\alpha$) = 0.71073 Å, $\mu = 0.9$ cm^{-1}, $F(000) = 208$, $R = 0.041$ for 1935 reflections with $I > 2.5\sigma(I)$] and 1,2,-5,6-tetrahydro-1-methyl-3-pyridinecarboxaldehyde oxime monohydrochloride [Org 31956, $C_7H_{13}N_2O^+.Cl^-$, $M_r = 176.64$, triclinic, $P\bar{1}$, $a = 6.843$ (3), $b = 6.997$ (3), $c = 9.837$ (4) Å, $\alpha = 89.72$ (3), $\beta = 87.78$ (4), $\gamma = 75.69$ (4)°, $V = 456.0$ (3) Å^3, $Z = 2$, $D_x = 1.286$ g cm^{-3}, λ(Mo $K\alpha$) = 0.71073 Å, $\mu = 3.7$ cm^{-1}, $F(000) = 188$, $R = 0.046$ for 2195 reflections with $I > 2.5\sigma(I)$] have been determined. A model, based on these crystal structure determinations, Cambridge Structural Database statistics and molecular mechanics calculations, is presented in which the muscarinic agonists L-660863, Org 31956, 1-azabicyclo[2.2.2]octane-3-carboxylic acid methyl ester, arecoline, sulfonium-arecoline, sulfonium-isoarecoline and *N*-methylisoarecoline are matched. A common interaction mode for these reverse ester bioisosteres of acetylcholine is found, provided that the different interaction geometries of quaternary ammonium, protonated tertiary amino and sulfonium groups with a negatively charged receptor site are taken into account. The low muscarinic activity of *N*-methylarecoline and isoarecoline can also be explained using this model. Acetylcholine cannot be fitted into this model, which suggests that the discussed compounds bind to the muscarinic receptor site in a mode that is different from that of acetylcholine.

Introduction

The muscarinic cholinergic receptors mediate in a large number of physiological functions, such as smooth muscle contraction and relaxation, glandular secretion and a number of cardiac functions (Burgen, 1990). The main interest is nowadays directed towards the role of muscarinic receptors in the central nervous system (CNS). Over the past decades it has become clear that a number of muscarinic receptor subtypes exist. In the CNS the M_1 receptor subtype appears to be primarily involved with higher brain functions like learning and memory. The M_2 subtype is mainly involved with sensory and motor functions and vegetative processes (Cortés, Probst, Tobler & Palacios, 1986). Two widely occurring forms of senile cognitive decline, Alzheimer's disease

0108-7681/93/020356-11$06.00