of the haloforms (Kawaguchi, Takashina, Tanaka & Watanabé, 1972). However, the matter cannot be disputed in the present study where no separate estimation is made of the contribution of the exchange energy term of the repulsive potentials. Despite that, a satisfactory picture of the packing forces emerges from the present mechanical models and no additional assumptions seem required to explain the striking features of the crystal packing. Moreover, as is stressed above, the respective magnitudes of the internal molecular parameters lie within the expected ranges.

References

ALLEMAND, J. & GERDIL, R. (1975). *Acta Cryst.* A**31**, S130.

ALTONA, C., KNOBLER, C. & ROMERS, C. (1963). *Acta Cryst.* **16**, 1217–1225.

BATES, J. B. & BUSING, W. R. (1974). *J. Chem. Phys.* **60**, 2414–2419.

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

COLLIN, R. L. (1956). *Acta Cryst.* **9**, 537.

DOMENICANO, A., VACIAGO, A. & COULSON, C. A. (1975). *Acta Cyst.* B**31**, 221–234.

DUNAND, A. (1977). Thesis n° 1778, Univ. de Genève.

DUNAND, A. & GERDIL, R. (1976*a*). *Acta Cryst.* B**32**, 1591–1593.

DUNAND, A. & GERDIL, R. (1976*b*). *Z. Kristallogr.* **144**, 419.

DUNAND, A. & GERDIL, R. (1981). In preparation.

GERDIL, R. & DUNAND, A. (1975). *Acta Cryst.* B**31**, 936–937.

GOMBERG, M. (1900). *Chem. Ber.* **33**, 3150–3163.

HEMILIAN, W. (1874). *Chem. Ber.* **7**, 1203–1210.

HILLIER, I. H & RICE, S. A. (1967). *J. Chem. Phys.* **46**, 3881–3889.

HOFFMANN, R. (1963). *J. Chem. Phys.* **39**, 1397–1412.

JORDAN, F. (1968). *Theoret. Chim. Acta*, **11**, 390–400.

KAWAGUCHI, T., TAKASHINA, K., TANAKA, T. & WATANABÉ, T. (1972). *Acta Cryst.* B**28**, 967–972.

LANDAIS, J. (1953). *Bull. Soc. Chim. Fr.* pp. 648–655.

MCLELLAN, A. L. (1963). *Tables of Experimental Dipole Moments*. San Francisco: Freeman.

MIRSKY, K. & COHEN, M. D. (1978). *Chem. Phys.* **28**, 193–204.

NYBURG, S. C. (1979). *Acta Cryst.* A**35**, 641–645.

RICHE, C. & PASCARD-BILLY, C. (1974). *Acta Cryst.* B**30**, 1874–1876.

SCHLODDER, R. & IBERS, J. A. (1975). *Acta Cryst.* B**31**, 708–712.

SCHOMAKER, V. & TRUEBLOOD, K. N. (1968). *Acta Cryst.* B**24**, 63–76.

STORA, C. (1953). *Bull. Soc. Chim. Fr.* pp. 1059–1063.

STORA, C. & POYER, N. (1966). *Bull. Soc. Chim. Fr.* pp. 841–849.

TRUEBLOOD, K. N. (1978). *Acta Cryst.* A**34**, 950–954.

WANG, S.-N. & LU, C.-S., (1944). *J. Am. Chem. Soc.* **66**, 1113–1114.

WILLIAMS, D. E. (1971). *Acta Cryst.* A**27**, 452–455.

WILLIAMS, D. E. (1972). *Acta Cryst.* A**28**, 629–635.

WILLIAMS, D. E. (1974). *Acta Cryst.* A**30**, 71–77.

XRAY system (1972). Version of June 1972. Tech. Rep. TR-192, Computer Science Center, Univ. of Maryland, College Park, Maryland.

YOUNG, J. (1931). BS thesis, Massachusetts Institute of Technology, unpublished. Cited in A. N. WINCHELL (1954), *The Optical Properties of Organic Compounds*, New York: Academic Press.

Acta Cryst. (1982). B**38**, 575–580

The Structure of 9-[3-(3-Indolyl)propyl]adenine. A Model for Protein/Nucleic Acid Interactions*

BY GERARD BUNICK† AND DONALD VOET

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, *USA*

(*Received* 17 *March* 1981; *accepted* 6 *August* 1981)

Abstract

The X-ray crystal structure of 9-[3-(3-indolyl)propyl]-adenine [$C_{16}H_{16}N_6$, space group $P2_1/n$, $Z = 4$, $a = 14{\cdot}751\ (4)$, $b = 8{\cdot}239\ (2)$, $c = 12{\cdot}160\ (2)$ Å, $\beta = 97{\cdot}94\ (3)°$, $V = 1463{\cdot}7$ Å^3, $D_c = 1{\cdot}339$ Mg m^{-3}] has been solved and refined to a final R factor of 0·073 based on 1469 observed reflections. This molecule, which serves as a model for protein/nucleic acid associations, crystallizes in an extended conformation. The adenine residues in the structure form endless hydrogen-bonded chains. However, the indole and adenine groups are out of contact with one another. A similar lack of indole/adenine associations is observed in the two other known crystal structures containing

* Research supported by National Institutes of Health Grant GM23966.

† Present address: University of Tennessee, Chemistry Division, PO Box X, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, USA.

0567-7408/82/020575-06$01.00

indole and adenine moieties. These structural data suggest that an indole group cannot compete effectively with thymine for the hydrogen-bonding sites of adenine.

Introduction

Protein/nucleic acid associations provide many of the specificities required in the molecular processes concerned with the expression and transmission of genetic information. Yet there is presently very little direct structural data available concerning the nature of these interactions due to the difficulties that have been encountered in growing crystals of protein/nucleic acid complexes that are suitable for X-ray crystallographic analysis. These difficulties, which may stem from common structural properties of proteins that bind nucleic acids, have all but eliminated the possibility that the high-resolution structure of a protein/nucleic acid complex will be visualized for at least the next several years. Therefore, in order to acquire structural information on how proteins and nucleic acids associate, it is necessary to determine the structures of model compounds designed to exhibit these specific interactions.

Experimental

9-[3-(3-Indolyl)propyl]adenine (Ind^3-C_3-Ade^9), which was provided by Professor Nelson Leonard, had been prepared according to the method of Mutai, Gruber & Leonard (1975). Crystals of Ind^3-C_3-Ade^9 were grown with difficulty by the slow evaporation of a solution of Ind^3-C_3-Ade^9 in 1% aqueous ethanol. Crystals grew as bundles of small plates. The crystal that was used for the X-ray analysis, which was cut out of such a bundle, was 0·15 × 0·23 × 0·08 mm. The symmetry of preliminary precession photographs and the systematic absences of their recorded reflections indicated the space group of the crystal to be $P2_1/n$. These photographs also revealed that the crystal had several smaller crystals attached to it. However, the difficulty in growing even this relatively flawed crystal dictated that the crystal be used for diffraction data collection.

All subsequent X-ray measurements were performed using a Picker FACS-I diffractometer equipped with a graphite monochromator and employing Cu $K\alpha$ radiation (λ = 1·5418 Å). The unit-cell parameters, as determined by the least-squares analysis of 12 independent reflections, are given in the *Abstract*. X-ray reflection intensities were measured using the ω-scan mode in order to minimize the inclusion of satellite reflections because it was found that these reflections were more likely to interfere with an intensity measurement when the more conventional θ–2θ scan mode was used. The scan rate was 0·25° min^{-1} due to the small size of the crystal. A total of 1620 unique reflections were measured to the limit $2\theta = 125°$.

The three standard reflections that were monitored after every fiftieth reflection exhibited significant decay over the course of the data collection such that the last set of standards measured averaged 87% of their initially measured intensities. The reflection intensities were corrected for this decay by a least-squares linear equation relating the sequence in which the reflections were measured to the average remaining fraction of the initial intensities of the three standard reflections.

Structure determination and refinement

The intensities, I, were corrected for Lorentz–polarization effects. Their standard deviations, $\sigma(I)$, were calculated according to counter statistics using an instrumental instability factor of 0·03 (Stout & Jensen, 1968). A total of 151 reflections had $I < 2{\cdot}33\sigma(I)$ and were therefore considered to be unobserved.

The structure was solved uneventfully by application of the direct-methods program *MULTAN* (Germain, Main & Woolfson, 1971). All 22 expected non-H atoms were found in the E map based on the phase set with the highest combined figure of merit. The structure was refined by full-matrix least-squares methods in which the quantity minimized was $\sum w(|F_o| - |F_c|)^2$ and $w = 1/\sigma^2(F_o)$. The atomic scattering factors for non-H atoms were taken from Cromer & Waber (1965) whereas those for H atoms were taken from Stewart, Davidson & Simpson (1965). The refinement of the scale factor, the atomic positional parameters and the isotropic thermal parameters, with the later refinement of the anisotropic thermal parameters reduced $R = \sum ||F_o| - |F_c|| / \sum |F_o|$ from its initial value of 0·281 to 0·115. All H atoms were then located in a difference Fourier map. Further refinement was carried out in which non-H atoms were treated anisotropically and H atoms had only their positional parameters refined. The H-atom isotropic thermal parameters were fixed at the isotropic equivalent of the atom to which they were covalently bonded. The refinement converged with R = 0·073 based on the 1469 observed reflections. The highest peak in the final difference Fourier map had a peak density of 0·18 e $Å^{-3}$.

Results

The atomic coordinates in the structure of Ind^3-C_3-Ade^9 are given in Table 1.* The molecular structure of Ind^3-C_3-Ade^9, as well as the atomic numbering

* Lists of structure factors and anisotropic thermal parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 36345 (11 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

system used in this report, are illustrated in Fig. 1. It can be seen there that Ind^3-C$_3$-Ade^9 exhibits a conformation such that the adenine and indole moieties of the same molecule are out of contact.

The covalent bond distances and angles found in Ind^3-C$_3$-Ade^9 are presented in Fig. 2. The dimensions of the adenine moiety are in good agreement with the corresponding quantities found in other crystal structures containing adenine (Voet & Rich, 1970). Likewise, the bond parameters of the propyl and the indole groups are all within their expected ranges. The r.m.s. deviation of the nine atoms of the purine group from their least-squares plane is 0·012 Å. Similarly, the nine atoms of the indole moiety have an r.m.s. deviation of

Table 1. *Atomic parameters for* Ind^3-C$_3$-Ade^9

The atomic positions are given as fractions of a unit-cell edge. The quantities in parentheses are the standard deviations of their corresponding parameters as estimated from the final cycle of least-squares refinement. The prefixes *A* and *I* refer to the adenine and indole moieties, respectively. B_{eq} is calculated according to $B_{eq} = \frac{4}{3}\sum_i \sum_j \beta_{ij}\,\mathbf{a}_i.\mathbf{a}_j$.

	x	y	z	B_{eq}/B (Å^2)
*A*N(1)	0·6682 (3)	0·8698 (4)	1·1516 (3)	4·5 (1)
*A*C(2)	0·6070 (4)	0·9032 (5)	1·0622 (4)	4·9 (2)
*A*N(3)	0·5718 (3)	0·8068 (4)	0·9792 (3)	4·7 (1)
*A*C(4)	0·6110 (3)	0·6593 (5)	0·9932 (3)	3·9 (1)
*A*C(5)	0·6753 (3)	0·6062 (4)	1·0786 (3)	3·6 (1)
*A*C(6)	0·7040 (3)	0·7190 (4)	1·1631 (3)	3·9 (1)
*A*N(7)	0·6977 (3)	0·4440 (4)	1·0648 (3)	4·4 (1)
*A*C(8)	0·6455 (4)	0·4056 (5)	0·9724 (3)	4·6 (1)
*A*N(9)	0·5913 (3)	0·5263 (4)	0·9246 (3)	4·2 (1)
*A*N(6)	0·7633 (3)	0·6850 (4)	1·2526 (3)	4·9 (1)
C(1′)	0·5241 (4)	0·5213 (6)	0·8239 (4)	4·8 (2)
C(2′)	0·5566 (4)	0·6086 (6)	0·7276 (3)	4·4 (2)
C(3′)	0·4785 (4)	0·6349 (6)	0·6335 (4)	4·8 (2)
*I*N(1)	0·5839 (4)	0·8542 (6)	0·4189 (4)	5·8 (2)
*I*C(2)	0·5886 (5)	0·7739 (7)	0·5191 (4)	5·4 (2)
*I*C(3)	0·5049 (5)	0·7222 (6)	0·5357 (3)	4·4 (2)
*I*C(3α)	0·4435 (5)	0·7730 (5)	0·4414 (3)	4·2 (2)
*I*C(4)	0·3504 (5)	0·7576 (6)	0·4110 (4)	5·2 (2)
*I*C(5)	0·3106 (5)	0·8288 (8)	0·3112 (5)	6·6 (2)
*I*C(6)	0·3634 (6)	0·9148 (7)	0·2438 (5)	6·5 (2)
*I*C(7)	0·4547 (6)	0·9299 (6)	0·2712 (4)	5·9 (2)
*I*C(7α)	0·4946 (5)	0·8574 (6)	0·3702 (4)	4·7 (2)
*A*H(2)	0·582 (3)	1·016 (5)	1·058 (3)	4·9
*A*H(8)	0·645 (3)	0·308 (5)	0·936 (3)	5·3
*A*H(6*A*)	0·788 (3)	0·589 (5)	1·262 (3)	5·2
*A*H(6*B*)	0·774 (3)	0·757 (5)	1·312 (3)	5·2
H(1′*A*)	0·515 (3)	0·414 (5)	0·807 (3)	5·1
H(1′*B*)	0·469 (3)	0·588 (5)	0·837 (3)	5·1
H(2′*A*)	0·611 (3)	0·551 (5)	0·705 (3)	4·9
H(2′*B*)	0·580 (3)	0·713 (5)	0·751 (3)	4·9
H(3′*A*)	0·451 (3)	0·530 (5)	0·612 (3)	5·4
H(3′*B*)	0·425 (3)	0·692 (5)	0·661 (3)	5·4
*I*H(1)	0·622 (4)	0·895 (6)	0·394 (4)	5·5
*I*H(2)	0·642 (4)	0·764 (6)	0·563 (4)	5·6
*I*H(4)	0·310 (3)	0·690 (5)	0·458 (3)	5·6
*I*H(5)	0·252 (4)	0·821 (8)	0·297 (5)	8·2
*I*H(6)	0·330 (3)	0·966 (6)	0·171 (4)	7·6
*I*H(7)	0·496 (3)	0·989 (6)	0·230 (4)	6·8

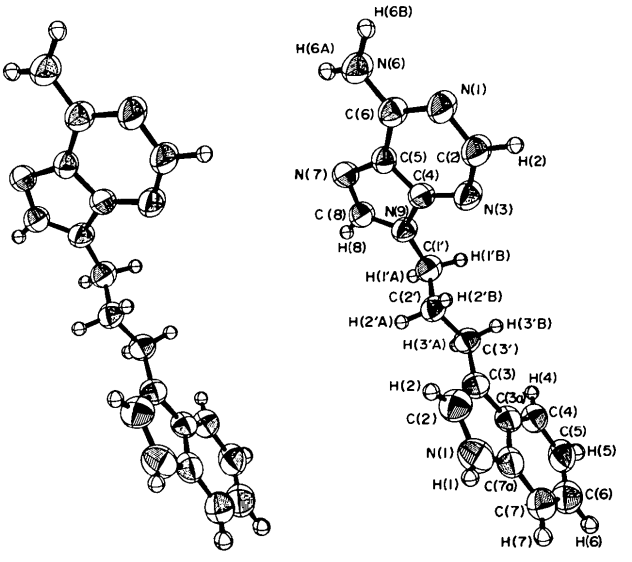


Fig. 1. A stereodrawing of a molecule of Ind^3-C$_3$-Ade^9. Non-H atoms are represented by thermal ellipsoids at the 50% probability level. H atoms are represented by spheres of arbitrary radii. The atomic numbering scheme used in this report is presented.

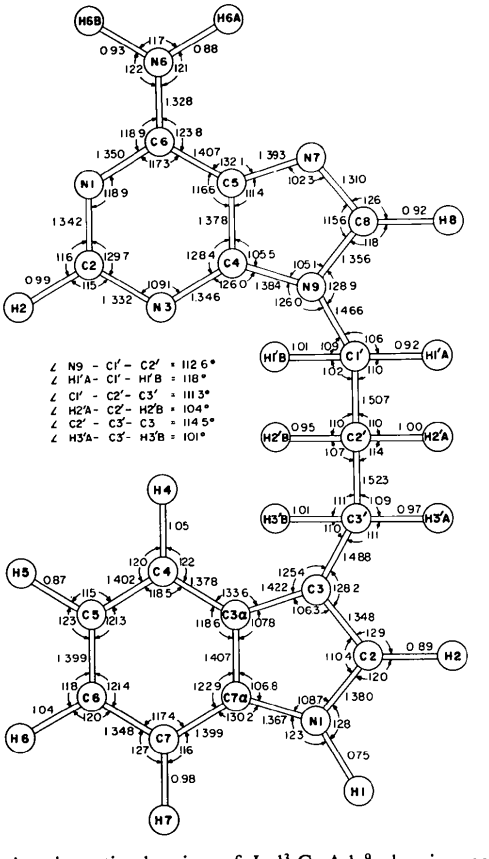


Fig. 2. A schematic drawing of Ind^3-C$_3$-Ade^9 showing covalent distances (Å) and angles (°). The average estimated standard deviations of these quantities are 0·006 Å and 0·4°, respectively, for distances and angles involving non-H atoms. The corresponding quantities involving H atoms are 0·04 Å and 3°, respectively.

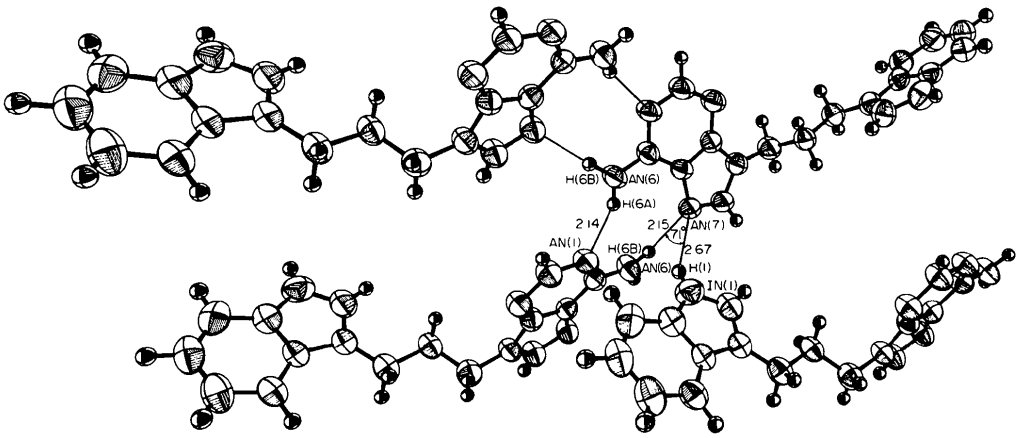


Fig. 3. A perspective drawing illustrating the intermolecular associations in the crystal structure of Ind^3-C_3-Ade^9.

Table 2. *Hydrogen-bond parameters*

D—H···*A*	*D*···*A* (Å)	H···*A* (Å)	D—H···*A* (°)
*A*N(6)—*A*H(6*A*)···*A*N(1^i)	2·966 (5)	2·14 (4)	156 (4)
*A*N(6)—*A*H(6*B*)···*A*N(7^i)	3·076 (6)	2·15 (4)	173 (4)
*I*N(1)—*I*H(1)···*A*N(7^i)	3·285 (7)	2·67 (6)	141 (5)

Symmetry code: none x,y,z; (i) $\frac{1}{2}-x, -\frac{1}{2}+y, \frac{1}{2}-z$.

0·016 Å from their least-squares plane. Hence, each of these groups is essentially planar. The dihedral angle between the least-squares planes of the adenine and indole moieties is 83·5°.

The hydrogen-bonding interactions in the structure of Ind^3-C_3-Ade^9 are illustrated in Fig. 3. The adenine groups associate in an endless chain of doubly hydrogen-bonded molecules associating through *A*N(6)—*A*H(6*A*)···*A*N(1) and *A*N(6)—*A*H(6*B*)···*A*N(7) hydrogen bonds. The hydrogen-bond distances and angles in the structure are tabulated in Table 2. It can be seen there that these hydrogen-bonding associations have normal geometries (Voet & Rich, 1970).

The indole group is hydrogen-bonded to the adenine of a neighboring molecule through a distorted *I*N(1)—*I*H(1)···*A*N(7) hydrogen bond. This interaction is unusual in that the *A*N(7) atom serves as the acceptor group for two hydrogen bonds. However, the hydrogen bond involving the indole ring lies considerably out of the plane of the adenine ring, with an *A*N(6)···*A*N(7)···*I*N(1) angle of 67·7°. Furthermore, the *I*H(1)···*A*N(7) distance of 2·67 Å is quite long for a hydrogen bond. This is reduced to the still relatively long value of 2·51 Å if it is assumed that the *I*N(1)—*I*H(1) covalent distance is lengthened from its refined value of 0·75 Å to the more reasonable value of 0·95 Å. Nevertheless, it is apparent that the *I*N(1)—*I*H(1)···*A*N(7) hydrogen bond is relatively weak.

The structure of Ind^3-C_3-Ade^9 is unusual in that there are no stacking associations between any of its component rings. The structure exhibits no intermolecular contacts closer than van der Waals interactions.

Discussion

Small-molecule structural studies that serve as models for intermolecular associations in other environments are predicated on the supposition that these associations are not significantly influenced by crystal-packing forces. However, crystal-packing forces probably have energies similar to those of the interactions under study. Hence, if the intermolecular interactions observed in a crystal structure are to be taken as meaningful it must be assumed that crystal-packing forces are isotropic, that is, the same in all directions so as to have no net influence on intermolecular associations.

Present-day theoretical descriptions of crystal-packing forces are not of sufficient accuracy to demonstrate unambiguously whether or not these forces are isotropic. However, the observation that related molecules often exhibit similar interactions in different crystal structures is strongly suggestive that crystal-packing forces are, in fact, isotropic. For example, crystal structures of nucleic acid bases exhibit recurrent patterns of hydrogen bonding (Voet & Rich, 1970) and base stacking (Bugg, Thomas, Sundaralingam & Rao, 1971). Likewise, globular-protein molecules with similar amino acid sequences assume nearly identical molecular structures. Nevertheless to establish convincingly that a particular intermolecular association mode is the normal tendency of the groups involved requires that this interaction be observed in several different crystal structures.

In Ind^3-C_3-Ade^9, the trimethylene group serves as an inert and flexible tether between the indole and adenine residues so as to force these groups to crystallize

together and yet allow them to interact with minimal interference from other influences. The observation that compounds of the sort Ind^3-C$_3$-X, where X = Ade^9, Gua^9, Cyt^1 or Thy^1, form stacked intramolecular complexes in aqueous solutions (Mutai, Gruber & Leonard, 1975) establishes the flexibility of the trimethylene bridge. Nevertheless, in the solid state the existence of any preferred conformations of the trimethylene group which might destabilize intramolecular interactions is of little consequence in the exposition of specific association affinities. This is because, assuming crystal-packing forces to be isotropic, the association of groups on different molecules should be equivalent to the association of these groups on the same molecule.

The supposition that the intermolecular associations, or lack of them, observed in the crystal structure of Ind^3-C$_3$-Ade^9 are characteristic of adenine/indole associations is corroborated by the crystal structure of the complex 9-ethyladenine–indole (Kaneda & Tanaka, 1976) and that of 3-(9-adenyl)propionyltryptamine monohydrate (Ohki, Takenaka, Shimanouchi & Sasada, 1977). In both of these structures, as in that of Ind^3-C$_3$-Ade^9, there are no indole/adenine-ring stacking associations although such interactions have often been postulated to be of biological importance in the associations of proteins and nucleic acids. Likewise, the hydrogen-bonding patterns in all three of these structures are quite similar. In 9-ethyladenine–indole the adenine groups form endless hydrogen-bonded chains identical to that observed in the present study. In 3-(9-adenyl)propionyltryptamine the adenines also form endless hydrogen-bonded chains but such that neighboring adenine residues alternately associate about centers of symmetry through pairs of N(6)–H···N(1) and N(6)–H···N(7) hydrogen bonds. In both of these latter structures, the adenine and indole groups associate through the formation of IN(3)–H···AN(3) hydrogen bonds whose respective lengths of 3·03 and 3·06 Å are rather long for N–H···N associations.

The inspection of numerous adenine-containing crystal structures reveals that only its N(1) and N(7) positions are commonly utilized as hydrogen-bond acceptor groups. For example, the modes of adenine/adenine hydrogen bonding described above are the only three types that have been observed between adenine residues in the solid state (Voet & Rich, 1970). Adenine position N(3) apparently participates in hydrogen bonding only when all other hydrogen bonding acceptor positions on adenine are occupied. In fact, as the present structure demonstrates, it is occasionally observed that a hydrogen-bond donor is left unsatisfied even though adenine N(3) is available. Hence it has been concluded that adenine N(3) is a weak hydrogen-bond acceptor (Shieh & Voet, 1976). Therefore, the observation of IN(3)–H···AN(3) hydrogen bonding in the structures of 9-ethyladenine–indole and 3-(9-adenyl)propionyltryptamine monohydrate further indicates that the IN(3)–H group can only weakly hydrogen-bond to adenine.

In the crystal structure of Ind^3-C$_3$-Thy^1, neighboring indole and thymine residues associate by both stacking interactions as well as an IN(3)–H···TO(2) hydrogen bond of normal geometry (Voet, 1980). In addition neighboring thymine residues associate through TN(3)–H···TO(4) hydrogen bonds about a twofold axis to form an endless hydrogen-bonded helix. It is these latter atoms that participate in Watson–Crick A·T base pairing. It has been noted that an indole group could hydrogen-bond to a thymine or uracil residue that is simultaneously involved in Watson–Crick base pairing (Voet, 1980).

In chloroform solution at 298 K nucleic acid bases form hydrogen-bonded base pairs. Under these conditions 9-ethyladenine has a self-association constant of 2·0 M^{-1} and that of 1-cyclohexyluracil is 10·0 M^{-1} whereas the association constant of the complex between these two bases is 110 M^{-1} (Kyogoku, Lord & Rich, 1967; Nagel & Hanlon, 1972). Therefore, the A·T base-pairing association is much stronger than those of A·A or T·T. The analyses of structures containing both indole and adenine residues indicate that indole/adenine hydrogen bonding is much weaker than adenine/adenine hydrogen bonding. Similarly the structure of Ind^3-C$_3$-Thy^1 suggests that indole/thymine hydrogen-bonding affinity is approximately equal in strength to that of thymine/thymine hydrogen bonding (Voet, 1980). Hence it seems unlikely that the indole residue of tryptophan will be effective in competing for the Watson–Crick hydrogen-bonding sites of an A·T base pair.

The authors wish to thank Professor Nelson Leonard for his generous gift of Ind^3-C$_3$-Ade^9.

References

BUGG, C. E., THOMAS, J. M., SUNDARALINGAM, M. & RAO, S. T. (1971). *Biopolymers*, **10**, 175–219.

CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104–109.

GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A**27**, 368–376.

KANEDA, T. & TANAKA, J. (1976). *Bull. Chem. Soc. Jpn*, **49**, 1799–1804.

KYOGOKU, Y., LORD, R. C. & RICH, A. (1967). *J. Am. Chem. Soc.* **89**, 496–504.

MUTAI, K., GRUBER, B. A. & LEONARD, N. J. (1975). *J. Am. Chem. Soc.* **97**, 4095–4104.

NAGEL, G. M. & HANLON, S. (1972). *Biochemistry*, **11**, 816–823, 823–830.

OHKI, M., TAKENAKA, A., SHIMANOUCHI, H. & SASADA, Y. (1977). *Acta Cryst.* B**33**, 2956–2958.

SHIEH, H.-S. & VOET, D. (1976). *Acta Cryst.* B**32**, 2361–2367.

STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.

STOUT, G. H. & JENSEN, L. H. (1968). *X-ray Structure Determination*, pp. 456–458. New York: Macmillan.

VOET, D. (1980). *J. Am. Chem. Soc.* **102**, 2071–2074.

VOET, D. & RICH, A. (1970). *Prog. Nucleic Acid Res. Mol. Biol.* **10**, 183–265.

Acta Cryst. (1982). B**38**, 580–583

Structure of (13αH,14βH)-Jeunicin

BY A. J. WEINHEIMER,* J. A. MATSON, M. POLING AND DICK VAN DER HELM

Department of Chemistry, University of Oklahoma, Norman, Oklahoma 73019, *USA*

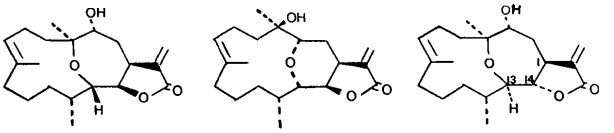
(*Received* 23 *March* 1981; *accepted* 6 *August* 1981)

Abstract

$C_{20}H_{30}O_4$, M_r = 334·5, is orthorhombic, $P2_12_12_1$, with a = 8·921 (2), b = 18·076 (4), c = 22·647 (5) Å, V = 3652·0 Å^3 [at 113 (2) K]; a = 8·962 (4), b = 18·30 (1), c = 22·96 (1) Å, V = 3765·5 Å^3, Z = 8, D_m = 1·184, D_x = 1·180 Mg m^{-3} (at 295 K). The compound is a toxic cembranolide isolated from the gorgonian *Eunicea mammosa*. The molecular structure and absolute configuration have been determined from 4191 X-ray intensity data. The final R factor is 0·063. The compound is an epimer of jeunicin.

Introduction

The diterpenes jeunicin (I) (van der Helm, Enwall, Weinheimer, Karns & Ciereszko, 1976) and eunicin (II) (Hossain, Nicholas & van der Helm, 1968) have been isolated as the major cembranolides in the gorgonian *Eunicea mammosa* from Jamaica and Bimini, respectively. Both are found to be cytotoxic against the National Cancer Institute's KB cell line. Because the crude extract of the Bimini gorgonian showed confirmed antineoplastic activity in NCI's *in vivo* bioassay against P-388 lymphocytic leukemia, we have subjected it to systematic fractionation with guidance at each stage by bioassays in NCI's *in vitro* KB and LE cell lines. This approach has led to the isolation of, in addition to eunicin, a second cytotoxic cembranolide in minor amounts. It is shown to have the structure of (13αH,14βH)-jeunicin (III).



(I) (II) (III)

The bioactive material is contained in the chloroform phase. This fraction is further refined by three successive partitions using hexane, carbon tetrachloride and chloroform *vs* 20, 25 and 35% water in methanol respectively.

Chromatography of the residue from this phase on Florisil in benzene/ethyl acetate, and re-chromatography of the active fraction on silica gel using acetone/hexane affords (13αH,14βH)-jeunicin, followed closely by eunicin.

(13αH,14βH)-Jeunicin (III) melts at 420–420·5 K; $[\alpha]_D^{20°C} = -31·1°$ (c = 3·0 g dm^{-3}, $CHCl_3$); $C_{20}H_{30}O_4$ (calc.: C 71·82, H 9·04%; found: C 71·74, H 8·93%), m/e 334 (M^+), IR (KBr): 3560, 3520 (–OH), 1760 (C=O), and 1665 cm^{-1} (C=C). Its ^{1}H NMR spectrum ($CDCl_3$) shows methyl signals at δ 1·01, d, J = 7 Hz (Me–C–H); 1·25, s (Me–C–O) and 1·57, bs (Me–C=C). One proton absorption appears at δ 3·22, m (H-1); 3·27, t, J = 8 Hz (H-13, coupled to H-14 by NMDR); 3·69, dd, J = 8, 10 Hz (H-14, coupled to H-1 by NMDR); 3·90, dd, J = 3, 5 Hz (H-3); 5·26, bd, J = 10 Hz (H-7, coupled to 8-Me by NMDR); 5·41, d, J = 3 Hz and 6·13, d, J = 3 Hz (H_2C=C, both coupled to H-1 by NMDR).

The spectral data demonstrate the presence of an α-methylenic lactone, a methyl-substituted olefin, a cyclic ether and an alcohol function. A 14-membered cembrane ring is suggested by the nature of the six C

* Present address: Department of Medicinal Chemistry and Pharmacognosy, University of Houston, Houston, Texas 77004, USA.

0567-7408/82/020580-04$01.00