Theoret. chim. Acta (Berl.) 16, 309—315 (1970)

PPP Calculations of the Absorption Spectra of Purines and Pyrimidines

MARGARET L. BAILEY*
Mathematical Institute, University of Oxford

Received September 22, 1969

A variant of the PPP MO method, in which the two-centre core integrals are calculated from the gradient of the overlap integrals and the parameters are obtained by a non-arbitary method, has been used to calculate the absorption spectra of some purines and pyrimidines of biological interest. The transition energies and oscillator strengths obtained account satisfactorily for both the trends and the absolute values of the experimental spectra.

Eine Variante der PPP-MO-Methode, in der die Zweizentren-Rumpf-Integrale aus dem Gradienten der Überlappungsintegrale berechnet werden und die Parameter nach einer speziellen Methode erhalten werden, wurde benutzt, um die Absorptionsspektren einiger Purine und Pyrimidine, die von biologischem Interesse sind, zu berechnen. Die Übergangsenergien und die Oszillatorstärken geben den Gang und die absoluten Werte der experimentellen Spektren zufriedenstellend wieder.

Les spectres d'absorption de quelques purines et pyrimidines d'intérêt biologique ont été calculés par une variante de la méthode PPP MO, dans laquelle les intégrales de coeur bicentriques sont calculées à partir du gradient des intégrales de recouvrement et les paramètres sont obtenus par une méthode non arbitraire. Les énergies de transition et les forces d'oscillateur obtenues rendent compte d'une manière satisfaisante des faits expérimentaux.

Introduction

The nucleic acids, which are of prime importance in studies of biological systems, are too large for theoretical calculations with existing methods. To gain some understanding of these systems, however, studies can be made of the subunits of the nucleic acids, such as the dinucleotides, these being the smallest subunits consisting of more than one base. Before discussing the electronic properties of the dinucleotides, it is necessary to seek a good description of the individual bases. This paper therefore presents calculations of the absorption spectra of the bases; a further paper will be concerned with the calculation of the spectra of the dinucleotides. The four bases of DNA (deoxyribonucleic acid) are adenine (A), guanine (G), cytosine (C) and thymine (T), although it is usual in π-electron studies to replace thymine by uracil (U) in order to avoid the problem of including a methyl group in the calculations. Purine (P), hypoxanthine (HX) and 2,6-diaminopurine (DAP) have also been included in this study, because of their similarity to the two purine bases. Diagrams of all these molecules are given in Fig. 1.

Theoretical studies of these molecules have been performed previously by several different π-electron molecular orbital methods. In particular, a recent study has been performed by Pullman and co-workers [1, 2, 3], in which much

* Present address: Department of Chemistry, University of California, Berkeley, California 94720, USA (to March 31, 1970).

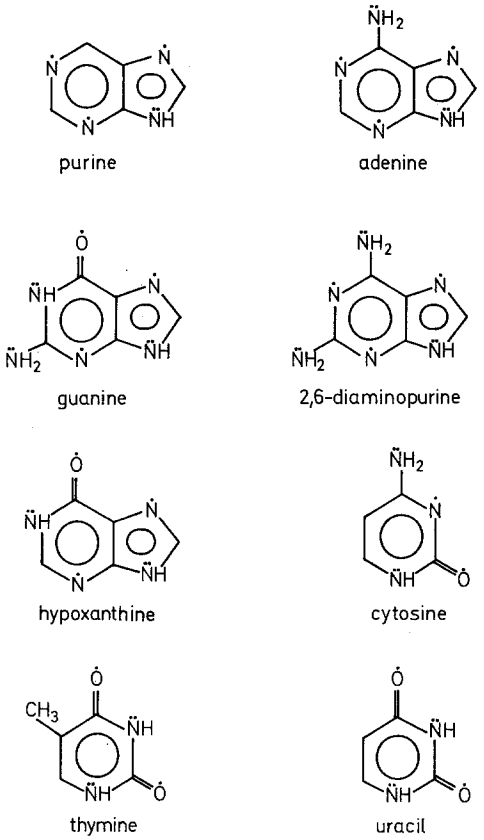

Fig. 1. Diagrams of the purines and pyrimidines

effort was spent in obtaining "a set of parameters which would reproduce as many observables as possible for a number of reference compounds" [1]. The observables used were ionization potentials and dipole moments, as well as absorption spectra. It would seem preferable, however, to calculate the absorption spectra of these molecules using a parameterisation which does not depend on recourse to experimental spectra. Such a parameterisation has been used successfully to calculate the absorption spectra of a wide range of benzene derivatives [4]; this method has therefore been applied to the calculation of the absorption spectra of these purines and pyrimidines.

Method

The absorption spectra were calculated using a variant of the PPP MO method described previously [4]. Its important features are the calculation of the two-centre core integrals from the gradient of the overlap integrals, as suggested by Linderberg [5], and the use of valence-state data, and in some cases experimental molecular ionization potentials, to obtain the one-centre core and repulsion integrals (the core and γ parameters). Previous calculated absorption spectra were reported as the mean of results calculated using six combinations of various

Table 1. *Parameters used in the calculations*

	$\dot{C}$	$\dot{N}$	$\ddot{N}_{am}$	$\ddot{N}_{py}$	$\ddot{N}_{la}$	$\dot{O}$
core (eV)	−11.16	−14.12	−25.73	−25.00	−26.77	−17.70
γ (eV)	11.13	12.34	16.76	16.76	16.76	15.23
S. e. c.[a]	3.25	3.90	3.90	3.90	3.90	4.55

[a] Slater effective charge.

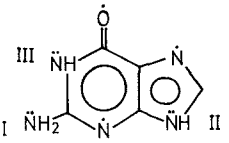

Fig. 2. The three different types of nitrogen – the amino-nitrogen (I), the pyrrolic-nitrogen (II) and the lactamic-nitrogen (III)

parameters and orbitals [4]. All the absorption spectra reported in this paper have been calculated with the combination of Slater orbitals together with the parameters given in Table 1 (set *B* and set 1 of Table 1 in [4]), this being the combination which yielded absorption spectra in closest agreement with the previous mean values.

In the bases, nitrogen atoms contributing two electrons to the π-system are found in three different environments – the amino group, the pyrrole ring and the lactam group – as illustrated in Fig. 2. The method used in the previous work [4] to calculate the core parameters of such atoms, based on an approximation discussed by Kwiatkowski [6], makes allowance for the environment of the atom through the use of the ionization potential of a molecule containing that atom in a similar environment. The experimental ionization potentials used for these core parameters were those of methylamine for the amino-nitrogen, dimethylamine for the pyrrolic-nitrogen and *N*-methylacetamide for the lactamic-nitrogen, the values being taken from Turner [7]. These core parameters are included in Table 1.

The geometries of the molecules were obtained from recent *X*-ray crystallographic studies of the appropriate crystals, instead of using the geometries proposed by Spencer [8], which are often used for theoretical calculations. In some cases two or more different crystal structures are available; for example, the crystal structure of either cytosine monohydrate [9] or anhydrous cytosine [10] is available, the latter being chosen. For uracil, calculations were performed using both the data for uracil itself [11] and that for thymine monohydrate [12]. As the results were remarkably similar, only the results using the data for uracil are reported. The geometries of purine, adenine (6-aminopurine) and 2,6-diaminopurine were taken from the data for 9-methyladenine [13]. By using the same geometry for all three molecules it was hoped that the effect of the addition of the amino groups upon the purine system could be detected. Similarly, the geometries of guanine and hypoxanthine were taken from the data for guanine hydrochloride dihydrate [14]. In the crystal all these molecules exist in the tautomeric form assumed for the calculations.

All possible configurations constructed by promoting a single electron from an occupied to a virtual level were included in the configuration interaction

calculations. The oscillator strengths were calculated by the geometric-mean method [15]. The program to perform these calculations was written in ALGOL and run on the University of Oxford English Electric KDF9 computer.

Results and Discussion

The absorption spectra of the purines and pyrimidines have been studied experimentally many times, including a study by Clark and Tinoco [16] in which the bands were classified into three types, labelled B_{2u}, B_{1u} and E_{1u} in analogy with the three bands appearing in the benzene spectrum. As the molecules do not have D_{6h} symmetry, it is perhaps better to give them some other classification; the three types are therefore labelled as I, II and III respectively. The calculated

Table 2. *Transitions energies (eV) and oscillator strengths for the purines and pyrimidines*

a	b	expt		
Adenine		[16]*	[18]	[19]
4.69 (0.02)	4.8 (0.1)	4.76 (0.3)	4.61 (10^{-3})	4.64 ($\sim$0.0)
4.87 (0.20)	5.0 (0.1)		4.73 (0.27)	4.79 (0.30)
5.89 (0.40)	5.8 (0.2)	5.96 (0.4)	5.98 (0.40)	5.99 (0.37
6.01 (0.08)	6.3 (0.1)			
6.24 (0.10)	6.5 (0.5)	6.70	6.77 (0.26)	
6.83 (0.17)	6.9 (0.2)			
6.89 (0.07)				
7.28 (0.09)			7.56	
7.79 (0.08)				
Guanine		[16]	[16]*	[19]
4.37 (0.17)	4.3 (0.3)	4.51 (*s*)	4.51 (*s*0.1)	4.49 (0.08)
4.85 (0.20)	5.1 (0.3)	4.84	4.92 (0.3)	4.92 (0.29)
5.80 (0.02)	5.7 (0.1)			
6.14 (0.33)	5.9 (0.1)	6.11 (*s*)	6.05 (*s*)	
6.29 (0.05)	6.2 (0.3)		(1.1)	
6.49 (0.31)	6.6 (.03)	6.53	6.59	6.63 (1.13)
Uracil		[16]	[17]	
4.96 (0.14)	4.8 (0.3)	4.81 (0.2)	5.08	
5.83 (0.14)	5.4 (.06)			
5.96 (0.22)	5.8 (0.2)	6.11 (0.3)	6.05 (*s*)	
6.58 (0.34)	6.2 (0.6)	6.85	6.63	
7.56 (0.03)	7.2 (.04)			
Cytosine		[16]	[19]	
4.27 (0.05)	4.1 (0.1)	4.48 (0.2)	4.57 (0.18)	
5.28 (0.11)	5.1 (0.1)	5.23 (*s*0.2)	5.39 (0.19)	
5.93 (0.54)	5.8 (0.4)	6.08		
6.41 (0.14)	6.3 (0.3)	(0.6)	6.29 (0.60)	
7.03 (0.12)	6.9 (0.2)	6.70		
Purine		[16]	[21]	
4.72 (0.05)	4.9 (0.1)	4.68 (0.10)	4.72	
5.06 (0.05)	5.5 (0.2)	5.17 (*s*0.05)	5.12 (*s*)	
6.21 (0.50)	6.3 (0.3)	6.20		
6.61 (0.18)	6.7 (0.3)	(0.6) 6.59		
6.80 (0.11)				
6.88 (0.04)				

Table 2 (continued)

a	b	expt	
Hypoxanthine		[16]	[16]*
4.36 (0.11)	4.2 (0.2)	4.46 (0.05)	4.59 (*s*)
5.02 (0.14)	5.1 (0.2)	5.02 (0.2)	4.98
5.85 (0.08)	5.8 (0.1)		
6.05 (0.01)	5.9 (0.2)		
6.25 (0.29)	6.6 (0.2)	6.26 (*s*) (0.9)	6.20
6.46 (0.23)	6.8 (0.5)	6.70	
7.06 (0.14)			
2,6-diaminopurine		[16]*	[22]
4.35 (0.07)		4.43	4.43 – 4.44
4.89 (0.13)		5.12	5.02 – 5.04
5.63 (0.54)		5.77	
5.93 (0.08)			
5.96 (0.16)		6.14	
6.34 (0.11)			

a Present work.

b Theoretical values taken from [2].

* The experimental maxima taken from [16] are measured in TMP (trimethylphosphate) except for the starred values which are measured in water at pH 6 to pH 7; the oscillator strengths are those reported for these experimental spectra in [2].

s Shoulder

and experimental absorption spectra, together with the calculated spectra from [2], are reported in Table 2.

All seven molecules show two strong bands in their experimental absorption spectra belonging to type III, the first appearing at about 5.9 to 6.3 eV, the second at about 6.5 to 6.9 eV for all molecules except DAP for which both bands are somewhat lower. The first band (type I) occurs in the range 4.4 to 4.8 eV with P, A and U at the upper end of the range, and HX, DAP, G and C at the lower end. This separation of the molecules into two groups becomes more pronounced for spectra obtained in the vapour phase [17], the first bands of A and U occurring at 4.92 eV and 5.08 eV, but those of G and C occurring at 4.23 eV and 4.28 eV. The second band (type II) is not readily seen in the spectra of A and U, occurs as a shoulder at 5.2 eV for C and P, and is visible in the spectra of HX, G and DAP at 5 eV. More detailed measurements of the spectrum of adenine [18, 19] have shown that the first band is in fact a composite band consisting of two bands. The circular dichroism (CD) spectrum of uridine (the nucleoside derived from uracil) shows three bands [20], at 4.63, 5.30, and 5.71 eV. By comparing the CD spectrum and the absorption spectrum of uridine (4.75 and 6.05 eV), which is very similar to the absorption spectrum of uracil (4.81 and 6.11 eV) [20], the presence of a hidden transition in the region 5.4 to 5.7 eV is deduced. This band may be hidden by the onset of the strong band belonging to type III. Beyond these strong bands, spectra have been measured using evaporated films on quartz plates [18] for adenine and thymine, the bands being at 7.56 and 7.6 eV respectively.

The general trends in the absorption spectra are correctly calculated by the present work. The first transition occurs in the range 4.3 to 4.9 eV; DAP, HX, C and G occur at less than 4.4 eV, and A, U and P at greater than 4.6 eV. The oscillator

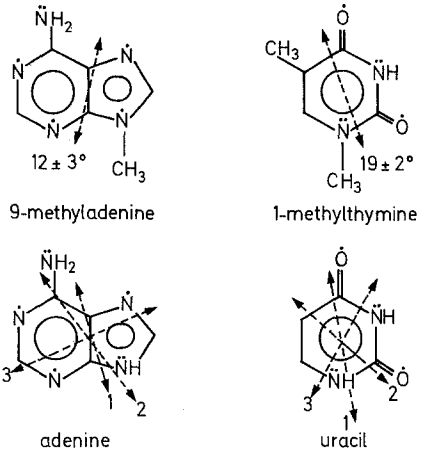

Fig. 3. Experimental and calculated transition moment directions-1, 2, 3 refer to the first, second and third transition, as given in Table 2

strengths are approximately 0.1, signifying a weak transition. A very low oscillator strength is calculated for the first transition of adenine; the first component of the first band in adenine has been shown experimentally to be very weak. The second transition in adenine is calculated to be close to the first transition and of higher oscillator strength, thus predicting a composite band. In uracil, the second transition is predicted to be very close to the third transition and could therefore be masked. For the other molecules, the second transition is correctly calculated to correspond to the second band. Several transitions are calculated within the range of the absorption bands of type III; some of these transitions are of low oscillator strength and would be masked by the strong transitions with oscillator strengths in the range 0.3 to 0.5.

The results from the present work are in significantly better agreement with experiment than those obtained in an earlier study by Pullman *et al.* [2]; however, the trends calculated for this series of molecules are similar. An important feature of the present method is that experimental absorption spectra are not required to parameterise the calculations, as was necessary in [2]. The present work leads to very good agreement of calculated and experimental spectra for purine and adenine, and good agreement for the other molecules.

In the series P, A and DAP, the addition of the amino group has considerable effect on the properties of the molecule. Indeed, the addition of the second amino group in the 2-position causes the spectra of DAP to resemble that of G, which is quite different from A. This resemblance between DAP and G has been used in experiments on the binding of actinomycin to DNA [23], which have shown that the binding is specific to guanine because of the amino group in the 2-position. dAT does not bind actinomycin, but a dAT copolymer synthesized containing DAP bases in place of some of its A bases bound actinomycin. In this case the steric relationship is probably the most important factor, although the similarity of the electronic properties of G and DAP may have some influence.

The polarised absorption spectra of 1-methylthymine and 9-methyladenine have been measured [24], the transition moment directions being obtained for

the first band. The direction of the transition moment for the second band is given as approximately perpendicular to the first band. In Fig. 3 are shown the calculated transition moment directions for the first three transitions of adenine and uracil, together with the experimental values for the first band of the methylated derivatives. (For both these molecules the third transition corresponds to the second band; for adenine the first band consists of two transitions.) Allowing for the perturbations to the directions caused by the methyl groups, it is seen that the calculated directions are in reasonable agreement with the experimental values.

Conclusion

The method based on calculating the two-centre core integrals from the gradient of the overlap integrals has led to the satisfactory interpretation of the experimental absorption spectra of a number of purines and pyrimidines. The present work has shown that the use of a non-arbitrary parameterisation can lead to successful results for molecules of biological interest, such as the purines and pyrimidines.

Acknowledgements. This work is taken from a D. Phil. thesis submitted to the University of Oxford. The author wishes to thank Professor C. A. Coulson, F. R. S., for his comments on this work, and acknowledges financial assistance from a Shirtcliffe Fellowship.

References

1. Berthod, H., Giessner-Prettre, C., Pullman, A.: Theoret. chim. Acta (Berl.) **5**, 53 (1966).
2. – – – Int. J. Quantum Chem. **1**, 123 (1967).
3. Pullman, A., Pullman, B.: Adv. Quantum Chem. **4**, 267 (1968).
4. Bailey, M. L.: Theoret. chim. Acta (Berl.) **13**, 56 (1969).
5. Linderberg, J.: Chem. Physics Letters **1**, 39 (1967).
6. Kwiatkowski, S.: Acta physica polon. **29**, 477 (1966).
7. Turner, D. W.: Advan. physic. org. Chem. **4**, 31 (1966).
8. Spencer, M.: Acta crystallogr. **12**, 59 (1959).
9. Jeffrey, G. A., Kinoshita, Y.: Acta crystallogr. **16**, 20 (1963).
10. Barker, D. L., Marsh, R. E.: Acta crystallogr. **17**, 1581 (1964).
11. Stewart, R. F., Jensen, L. H.: Acta crystallogr. **23**, 1102 (1967).
12. Gerdil, R.: Acta crystallogr. **14**, 333 (1961).
13. Stewart, R. F., Jensen, L. H.: J. chem. Physics **40**, 2071 (1964).
14. Iball, J., Wilson, H. R.: Proc. Roy. Soc. (London) A **288**, 418 (1965).
15. Hansen, A. E.: Molecular Physics **13**, 425 (1967).
16. Clark, L. B., Tinoco, I., Jr.: J. Amer. chem. Soc. **87**, 11 (1965).
17. –, Peschel, G. G., Tinoco, I., Jr.: J. physic. Chem. **69**, 3615 (1965).
18. Tanaka, M., Nagakura, S.: Theoret. chim. Acta (Berl.) **6**, 320 (1966).
19. Devoe, H., Tinoco, I., Jr.: J. molecular Biol. **4**, 518 (1962).
20. Voelter, W., Records, R., Bunnenberg, E., Djerassi, C.: J. Amer. chem. Soc. **90**, 6163 (1968).
21. Drobnik, J., Augenstein, L.: Photochem. Photobiol. **5**, 13 (1966).
22. Mason, S. F.: J. chem. Soc. (London) **1954**, 2071.
23. Waring, M. J.: Nature **219**, 1320 (1968).
24. Stewart, R. F., Davidson, N.: J. chem. Physics **39**, 255 (1963).

Dr. Margaret L. Bailey
Chemistry Department,
Victoria University of Wellington,
P. O. Box 196,
Wellington, New Zealand