G. H. *Ibid.* **1969**, *50*, 719–721.

(10) For a review of relaxation theory, see: Noack, F. In "NMR: Basic Principles and Progress", Diehl, P., Ed.; Springer-Verlag: West Berlin, 1971; Vol. 3, pp 83–144.

(11) (a) King, R.; Jardetzky, O. *Chem. Phys. Lett.* **1978**, *55*, 15–18. (b) King, R.; Maas, R.; Gassner, M.; Nanda, R. K.; Conover, W. W.; Jardetzky, O. *Biophys. J.* **1978**, *6*, 103–117. (c) Jardetzky, O. In "NMR and Biochemistry: A Symposium in Honor of Mildred Cohn", Opella, S., Lu, P., Eds.; Marcel Dekker: New York, 1979; pp 141–167. (d) Jardetzky, O.; Ribeiro, A.; King, R. *Biochem. Biophys. Res. Commun.*, **1980**, *92*, 883–888. (e) Ribeiro, A.; Wade-Jardetzky, N. G.; King, R.; Jardetzky, O. *Appl. Spectrosc.*, in press.

(12) (a) Levine, Y. K.; Partington, P.; Roberts, G. C. K. *Mol. Phys.* **1973**, *25*, 497–514. (b) Levine, Y. K.; Birdsall, N. J. M.; Lee, A. G.; Metcalfe, J. C.; Partington, P.; Roberts, G. C. K. *J. Chem. Phys.* **1974**, *60*, 2890–2899.

(13) (a) Cutnell, J. D.; Bleich, H. E.; Glasel, J. A. *J. Magn. Reson.* **1976**, *21*, 43–46. (b) Vold, R. L.; Waugh, J. S.; Klein, M. P.; Phelps, D. E. *J. Chem. Phys.* **1968**, *48*, 3831–3832.

(14) Nelson, D. J.; Opella, S. J.; Jardetzky, O. *Biochemistry* **1976**, *15*, 5552–5560.

(15) Kowaleski, J.; Levy, G. C.; Johnson, L. F.; Palmer, L. *J. Magn. Reson.* **1977**, *26*, 533–536.

(16) Led, J. J.; Petersen, S. B. *J. Magn. Reson.* **1978**, *32*, 1–17.

(17) Levy, G. C.; Peat, I. R.; Rosanke, R.; Parks, S. *J. Magn. Reson.* **1975**, *18*, 205–208.

(18) R = 1.09 Å is commonly used as the internuclear distance between a carbon and its directly bonded proton. E.g., see: Allerhand, A.; Doddrell, D.; Glushko, V.; Cochran, D. W.; Werhert, E.; Lawson, P. J.; Gurd, F. R. N. *J. Am. Chem. Soc.* **1971**, *93*, 544–546. A recent note by K. Dill and A. Allerhand (*J. Am. Chem. Soc.* **1979**, *101*, 4376) points out that a 2–3% error in this value can cause an error of as much as a factor of 2 in the value of the correlation time calculated from a rigid rotor model. It should be noted that the *experimental* errors in determining T_1, T_2, and NOE on existing spectrometers are sufficiently large to cause uncertainties in the calculated correlation times of a similar order of magnitude. It is therefore doubtful that more precise values of τ_c or R_{CH} can at present be derived from relaxation measurements, as proposed by Dill and Allerhand.

(19) Berne, B.; Pecora, R. In "Dynamic Light Scattering"; Wiley: New York, 1976.

(20) (a) Tschesche, H. *Angew. Chem., Int. Ed. Engl.* **1974**, *13*, 10–28. (b) Deisenhofer, J.; Steigmann, W. *Acta Crystallogr., Sect. B* **1975**, *31*, 238–250. (c) Huber, R.; Kukla, D.; Ruhlmann, A.; Steigemann, W. *Cold Spring Harbor Symp. Quant. Biol.* **1971**, *36*, 141–150. (d) Wagner, G.; De Marco, A.; Wüthrich, K. *Biophys. Struct. Mech.* **1976**, *2*, 139–158.

(21) (a) Richarz, R.; Wüthrich, K. *Biochemistry* **1978**, *17*, 2263–2269. (b) Brown, L. R.; De Marco, A.; Richarz, R.; Wagner, G.; Wüthrich, K. *Eur. J. Biochem.* **1978**, *88*, 87–95.

(22) Kuhlmann, K. F.; Grant, D. M.; Harris, D. M. *J. Chem. Phys.* **1970**, *52*, 3429–3448. At higher magnetic field strengths, chemical-shift anisotropy (CSA) becomes the dominant relaxation mechanism for *nonprotonated unsaturated* carbons. For the hydrogen-bearing unsaturated carbons of aromatic side chains, the CSA mechanism could contribute slightly (possibly as high as 10%) to the relaxation. The measured relaxation parameters have uncertainties at least of this magnitude. The contribution of the CSA mechanism to the relaxation of protonated, saturated carbons (such as α carbons and CH_2 and CH_3 groups of a protein) is totally negligible for all magnetic field strengths available today for high-resolution NMR.

(23) (a) Snyder, G. H.; Rowen III, R.; Karplus, S.; Sykes, B. D. *Biochemistry* **1975**, *14*, 3765–3777. (b) Wagner, G.; Wüthrich, K. *J. Magn. Reson.* **1975**, *30*, 435–445.

Ab Initio, Quantum Chemical Analysis of Noncovalent Interactions between Peptides as Modeled by Dimers and a Trimer of Formamide

Ernest L. Mehler

Contribution from the Department of Structural Biology, Biocenter, The University of Basle, CH-4056 Basle, Switzerland. *Received July 23, 1979*

Abstract: The self-consistent, nonorthogonal group function approximation has been applied to dimers and a trimer of formamide in various geometries constructed to simulate intrachain, noncovalent interactions between peptides. The interactions simulated are hydrogen-bonded and nonbonded pairs in the α helix and the 3_{10} helix and the doubly hydrogen bonded trimer in the α helix. The interactions are decomposed into Coulomb-exchange, polarization, and charge-transfer contributions, a detailed analysis of the dimers and trimer is given, and it is shown that the main source of deviations from pair additivity is the polarization effect. On the basis of this analysis estimates of the interaction energy and dipole moment are obtained for hydrogen-bonded complexes of any length. Finally, it is estimated that the positive cooperativity effect of multiple hydrogen bonds in the infinite chain increases from 10% in the purely pair-additive interaction to 23% when the deviations from pair additivity are included.

I. Introduction

The fundamental role of noncovalent interactions in stabilizing polypeptide structure has long been recognized. Nevertheless, there have been relatively few investigations, beyond peptide pairs at the ab initio quantum chemical level, aimed at analyzing the nature of these interactions and assessing the importance of their contributions to the interaction energy. Such studies are of considerable importance since they can form the basis for formulating new approaches applicable to larger polypeptide chains, or indicate where the limited success[1] of current empirical approaches can be improved.

Ab initio or near-ab initio studies of peptide interactions in single strands have been reported by Shipman and Christoffersen[2] using the fragment molecular orbital method,[3] and Kleier and Lipscomb[4] using the partial retention of diatomic differential overlap (PRDDO) approximation.[5] The results are conflicting: Shipman and Christoffersen find the α-helical structure less stable than the fully extended (FE) structure, which is also found by Kleier and Lipscomb. The latter also find that the α helix is less stable than the 3_{10} helix, although it is by far the most commonly observed helical structure in globular proteins. Finally, Scheiner and Kern[6] have used the PRDDO method to compute peptide pair interactions and have subsequently calibrated an empirical potential function based on these computations. They find the α helix to be slightly more stable than the FE structure and the 3_{10} helix.

These results indicate that a more fundamental analysis is required to clarify the various interactions. The hydrogen-bonded (H-bonded) formamide pair interaction has been extensively studied. Most of these studies have been done at arbitrary or optimal geometries,[7] but a few papers have restricted the dimeric geometry to simulate noncovalent interactions in various types of protein secondary structure.[8] Beyond the dimer very few ab initio studies have been reported, although these are useful for studying cooperative effects and deviations from pair additivity in multiply H-bonded chains. Cooperativity

0002-7863/80/1502-4051$01.00/0

effects were found for formamide chains by Hinton and Harpool[9] and Sheridan et al.[10] (SLPA) in their calculations on (formamide)$_n$, $n = 1, 5$. Neither of these papers explicitly considered the nonadditivity problem.

In this paper the recently developed nonorthogonal group function (NOGF) approximation[11] is used to study noncovalent interactions in dimers and a trimer of formamide. The geometries of the dimers and trimer have been chosen to simulate interactions in the α helix and the 3_{10} helix. A detailed analysis of the dimeric and trimeric interactions is given, and both cooperativity and deviations from pair additivity are considered. The consequences of the latter on the reliability of empirical potential functions are discussed. Finally, a method is given for estimating the interaction energy and dipole moment, and their decomposition, for an H-bonded complex of any length. For the tetramer and pentamer it is shown that our estimates are in close agreement with the ab initio results of SLPA.

II. Methods

In the NOGF approach the wave function is formulated as an antisymmetrized product of group functions[12]

$$\Psi = \mathcal{A}'[\Phi_1\Phi_2\ldots\Phi_T]$$

where each group is an antisymmetrized product of doubly occupied orbitals:

$$\Phi_R = \mathcal{A}[u_{R1}(1)\bar{u}_{R1}(2)\ldots u_{RN_R}(2N_R-1)\bar{u}_{RN_R}(2N_R)]$$

Intragroup orthogonality is assumed, but no orthogonality constraints are imposed on orbitals between groups. We thus have

$$\langle u_{Rr}|u_{Rr'}\rangle = \delta_{rr'}$$

$$\langle u_{Rr}|u_{Ss}\rangle = S_{Rr,Ss}$$

The orbital equations are obtained by defining "group energy functionals" and applying the variational method to these quantities. The derivation and form of the equations, as well as the application of the LCAO method,[13] have been reported previously.[11]

The relaxation of the orthogonality constraints permits the restriction of the LCAO expansions of the orbitals in each group to a part of the basis space used for the whole system. The groups can therefore be constructed to span only a part of the entire molecular space and their response to perturbations from other groups will then be restricted to these subspaces. This additional degree of freedom in constructing the orbitals allows a partitioning of the interactions between the subsystems which may be used to decompose the interaction energy and other interaction properties.

The interaction decomposition scheme obtained with the NOGF approximation is similar to that developed by Morokuma[14] and others.[8,15] Details of the method, application to a number of small dimeric systems, and comparison with other methods will be reported elsewhere.[16] Here we give a summary. Assume that monomer wave functions $\Phi_A{}^0$, $\Phi_B{}^0$, ..., with associated bases χ_A, χ_B, ..., have been determined. Define three wave functions of the form

$$\Psi_1 = \mathcal{A}'[\Phi_A{}^0\Phi_B{}^0\ldots] \quad (1a)$$

$$\Psi_2 = \mathcal{A}'[\Phi_A\Phi_B\ldots] \quad (1b)$$

$$\Psi_3 = \Phi_{AB}\ldots \quad (1c)$$

Ψ_1 contains only Coulomb and exchange (C + E) effects. In addition, Ψ_2 contains polarization (Pol) effects, while Ψ_3, the usual supermolecule SCF-MO wave function, also describes charge-transfer (CT) effects. The interaction energy is now partitioned in the following way:[14]

$$\begin{aligned} E_{\mathrm{C+E}} &= E_1 - E_0 \\ E_{\mathrm{Pol}} &= E_2 - E_1 \\ E_{\mathrm{CT}} &= E_3 - E_2 \end{aligned} \quad (2)$$

where the total interaction energy is given by

$$\Delta E = E_{\mathrm{C+E}} + E_{\mathrm{Pol}} + E_{\mathrm{CT}} \quad (3)$$

and E_0 is the sum of the monomer energies. Other properties can be similarly decomposed.

An additional type of partitioning can be made with the NOGF approximation which allows a wave function in the form of eq 1b to describe charge-transfer effects as well as the polarization, exchange, and electrostatic effects. This is accomplished by adding to the monomer's expansion set a number of basis functions from one or more of the other monomers yielding a wave function of the form

$$\Psi_4 = \mathcal{A}'[\Phi_A'\Phi_B'\ldots] \quad (4)$$

The charge-transfer contribution is now obtained from the difference $E_4 - E_2$. The importance of a given region of the supermolecule for charge transfer can be tested by including its basis functions with other groups and comparing the resulting charge-transfer energy with the total obtained from eq 2. If these two quantities are about the same, most of the important sites have been included. In addition to providing possibilities for further analysis, wave functions of the form of eq 4 can be used to calculate interaction energies with less computing than required with wave functions of the form of eq 1c. This point will be discussed further below.

The differences between the present analysis and Morokuma's[14,17] are as follows:

(i) The entire analysis is made at the wave function level. This means that all the effects, or partial effects, are defined in terms of appropriately constructed wave functions and not by manipulating elements of the Fock matrix.

(ii) The polarization term now includes exchange terms in addition to the classical polarization deformations. These are essentially Morokuma's "exchange polarization" effects.[17]

III. Details of the Calculations

1. Basis Sets. The calculations reported in this paper were made with minimum basis sets (MB) constructed from (5s,2p) Gaussian lobe functions for the heavy atoms and Huzinaga's (3s) basis for hydrogen.[18] These atomic bases were parametrized to mimic the valence shell part of Huzinaga's (7s,3p) bases,[19] and the parameters have been reported elsewhere.[20] In interpreting the results of energy decomposition with minimum basis sets, it is noted that trends tend to be relatively insensitive to basis-set quality,[21] but the values of the individual components can be quite basis set dependent.[22] In particular, the electrostatic contribution tends to be underestimated. Moreover, since polarization is characterized by a mixing of the monomer's virtual with occupied orbitals,[17] its reliability depends in part on the ratio of the number of basis to occupied orbitals. With double ζ bases this ratio is at least 2, whereas with MBs it is between 1 and 2 depending on the system and its symmetry restrictions. In $(H_2O)_2$, for example, the ratios are 1.25 for the a′ orbitals and 2.0 and 1.0 for the two a″ orbitals in Ψ_2. This lack of flexibility is reflected in the small value of E_{Pol} obtained for the water dimer with the STO-3G basis.[21] In the present calculations the ratio is 1.5 for the dimers and trimer, which suggests that E_{Pol} will be more completely accounted for in the present case than in $(H_2O)_2$.

A further test of the quality of the basis set can be obtained from the energy lowering resulting from the increase in basis-set size when the dimer is constructed. This effect appears as part of the charge-transfer contribution, and its maximum magnitude can be estimated from a counterpoise calculation.[23]

Table I. Coordinates of Formamide in the *i*th Monomer Position[a]

atom	*X*	*Y*	*Z*
C	1.6640	0.	0.
N	1.0690	1.1421	−0.4844
O	1.8992	−0.2034	1.2004
HC	1.9321	−0.7510	−0.7428
HN1	0.8860	1.2911	−1.4564
HN2	0.7923	1.8801	0.1315

	Helical Parameters	
	α helix	3_{10} helix[b]
rise/unit	1.4950	1.9284
rot/unit (rad)	1.7378	2.1216

[a] Ångstroms. [b] The radius of the 3_{10} helix is identical with that used for the α helix.

For the dimer yielding maximum H-bond stabilization the basis set effect is estimated to be 0.8 kcal/mol. This is quite small for MB calculations[22] and is probably due to the fact that the nuclei of the proton donor not directly involved in H bonding are fairly distant from the proton acceptor.

2. Coordinates, Models, and Wave Functions. The formamide monomers are assumed to be rigid and planar and the C–N distance is 1.376 Å. The calculated total energy is −166.7780 au and the dipole moment is 3.95 D, as compared to the observed value of 3.7 D.[24] The coordinates for the *i*th position and the helical parameters required to obtain the coordinates of any other position are given in Table I. The coordinates of the *i*th position may be converted to polar cylindrical form, the rise and rotation parameters for the unit desired are applied, and the result is converted back to the original coordinate system. The parameters for the α helix are obtained from refined α-helical coordinates based on fiber diffraction data from poly-L-alanine reported by Arnott and Dover,[25] and for the 3_{10} helix from theoretical formulas given by Miyazawa[26] using values of −46.0 and −26.0° for the ϕ and ψ angles.

The monomers are positioned so that the *z* axis is coincident with the helical axis and three of the pairs are shown in Figure 1. In addition to these, the α-helical *i*, *i* + 6 pair and the *i*, *i* + 3, *i* + 6 trimer are reported. In the trimer the *i* + 3rd monomer is H bonded to both of the other monomers. The α-helical and 3_{10}-helical H bonds deviate from linearity by 5.9 and 13.4°, respectively, and their O. . . N separations are 2.83 and 2.86 Å. The α-helical *i* + 2 monomer is across from and approximately parallel to the *i*th monomer.[27] The O. . . N distance is 3.43 Å and no possibility for H bonding exists. Finally, the O. . . N distance between the *i*th and *i* + 6th monomers is 7.56 Å.

The wave functions used for the calculations are constructed in accordance with the definitions given in eq 1 and 4. Thus Ψ_1 and Ψ_2 are two or three group antisymmetrized products for the dimers and trimer, respectively. The orbitals in each group are expanded in the monomer's basis, which results in their having a dimension of 18. Ψ_3 is the usual SCF-MO wave function. Finally, Ψ_4 is constructed so that each proton-accepting group also includes the basis functions of the donated proton. For the minimum basis set case each proton-accepting group is of dimension 19, whereas the group which only acts as a proton donor remains of dimension 18. These types of wave functions will be referred to as "limited charge transfer" (LCT) wave functions.

IV. Results and Discussion

1. Dimers. From Table II it is seen that the total stabilization energy of the α-helical H-bonded interaction is nearly twice as great as the 3_{10}-helical interaction. This extra stabilization is a result of the geometrical differences between the two pairs and helps to confirm the more favored conformation of the α helix. The nonbonded *i*, *i* + 2 pair interaction is slightly repulsive in the present model, but it is of interest to note that the *i*, *i* + 6 pair interaction still retains a residual stabilization even though the monomers are quite far apart.

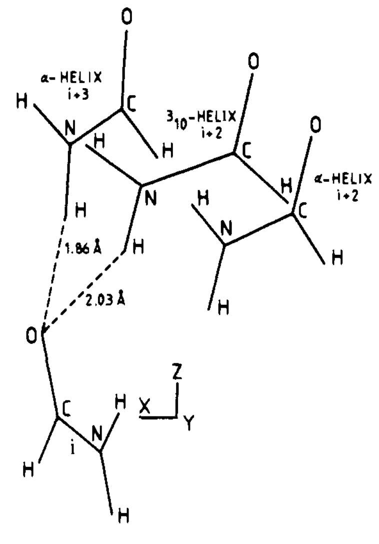

Figure 1. Diformamide configurations simulating noncovalent peptide pair interactions in the α helix and in the 3_{10} helix.

Table II. Energy Decomposition of Formamide Dimers Simulating Intrachain Peptide Pair Interactions[a]

type of contribution	α helix *i*, *i* + 3	α helix *i*, *i* + 6	α helix *i*, *i* + 2	3_{10} helix *i*, *i* + 2
C + E	−2.20	−0.53	0.71	−0.91
Pol	−0.89	−0.00	−0.08	−0.45
CT	−2.01	0.00	−0.21	−1.50
LCT	−1.72			−1.17
ΔE	−5.11	−0.54	0.42	−2.86
$\Delta E'$[b]	−4.82			−2.53

[a] Energies in kcal/mol. [b] Interaction energies from limited charge transfer.

The large drop in stability of the 3_{10}-helical H-bonded structure does not seem entirely rationalized in terms of the nonlinearity of the H bond since it is only 8° greater than in the α-helical case. On the other hand, the shifts in geometry make increased repulsion between other parts of the monomers more likely. Thus the H-bonding proton in the 3_{10}-helical model moves closer to the proton acceptor's carbon atom and, as Figure 1 illustrates, is less shielded from it by the oxygen than in the α-helical model. These geometric considerations are also valid for the α-helical *i*, *i* + 2 dimer, where an even greater interaction between the proton and carbon is possible. In both cases the proton–carbon separation is slightly less than 3 Å.

The energy decomposition of the dimeric interactions is tabulated in Table II. The electrostatic and exchange effects have not been further separated, but, since they are defined in the same way as in Morukama's analysis,[14] the results in the excellent papers of Umeyama and Morukuma[21] (UM) and Kollman[28] (K) can be used to evaluate our calculations. It should be noted that these investigators were interested in the optimal geometries of the dimeric interactions, whereas the geometries chosen here are determined by the structural parameters of the α helix and the 3_{10} helix.

In the H-bonded dimers it is seen that the C + E contribution loses the most stabilization. From UM and K this loss is probably due to a strong decrease in electrostatic stabilization, which is probably a result of the proton–carbon interaction mentioned above. This is probably also true for the other two

Table III. Axial Component of the Dipole Moment of Formamide Dimers Simulating Helical Peptide Pair Interactions[a]

type of contribution	α helix			3_{10} helix
	$i, i + 3$	$i, i + 6$	$i, i + 2$	$i, i + 2$
μ_Z (monomers)	−6.871	−6.871	−6.871	−6.872
C + E	−0.029	0.000	0.003	−0.014
Pol	−0.611	−0.047	−0.041	−0.392
CT	−0.180	0.000	−0.002	−0.077
LCT	−0.306			−0.161
$\Delta\mu_Z$	−0.821	−0.047	−0.046	−0.483
$\Delta\mu'_Z$[b]	−0.946			−0.567
μ_Z (dimer)	−7.691	−6.919	−6.919	−7.355
	(−0.410)[c]	(−0.024)	(−0.022)	(−0.242)
μ'_Z (dimer)	−7.817			−7.439
	(−0.473)			(−0.284)

[a] Dipole moments in debye. [b] Primed quantities are from LCT calculation. [c] Quantities in parentheses give enhancement of dipole-moment component per monomer.

Table IV. Energy and Dipole Moment Decomposition of Trimer and Deviations from Pair Additivity

	$-E$, kcal/mol	dev[a]		$-\mu_Z$,[b] D	dev[a]
			(monomers)	10.307	0.0[d]
C + E	4.97	−0.04		0.059	−0.000
Pol	2.35	−0.57		1.306	−0.037
CT	4.21	−0.18		0.392	−0.032
LCT	3.63	−0.19		0.658	−0.046
Δ	11.52	−0.77		1.758	−0.069
Δ'[c]	10.95	−0.78		2.023	−0.084
			(trimer)	12.064	−0.585[d]
			(trimer′)	12.330	−0.674[d]

[a] Deviation from pair additivity is defined in eq 5. [b] Axial component of the dipole moment. [c] Prime refers to LCT calculation. [d] Enhancement of the axial component/monomer.

dimers. The remaining contributions in the H-bonded dimers also lose significantly in stability, the combined loss comprising about 40% of the total decrease in stabilization energy. It is also seen that the LCT wave function yields a substantial part of the charge-transfer energy. Evidence will be given elsewhere[16] that the difference between ΔE and $\Delta E'$ is almost purely basis set effects so that in a certain sense $\Delta E'$ may be more reliable than ΔE. The residual stabilization in the $i, i + 6$ dimer is almost certainly due to electrostatic attraction since the separation between the monomers is 7 Å or more. On the other hand, the $i, i + 2$ dimer is slightly repulsive since the polarization and charge-transfer effects are not large enough to overcome the large positive value of the C + E effect.

The peptide units in the α helix are positioned in such a way that their dipole moments are nearly parallel and directed along the helical axis. From the present calculations, four independent formamide monomers, positioned to simulate one turn of an α helix, have an axial dipole moment component of −13.7 D and a total dipole moment of 13.8 D. This is a contribution of about 3.4 D per monomer, in close agreement with the value assumed by Hol et al.[29] in their discussion of the α-helix dipole. In addition, it has been suggested that, owing to cooperative effects, intrachain interactions in helices can increase the monomer's contribution to the dipole moment by up to 30%.[30]

The paper of Hol et al.[29] assigns an important functional role to the α helix's large dipole moment, and in particular to the axial component. Because of this we have applied our analysis to the dimers' and trimer's dipole moments, and in Table III the axial component and its decomposition are tabulated for the dimers. We only give the axial component since it is the quantity of interest, and in all cases it comprises at least 90% of the total dipole moment.

Table III shows that the polarization enhances the dipole moment of all four dimers, but that it is nonnegligible only for the H-bonded cases. For these latter two dimers the charge transfer also makes a smaller but substantial contribution. The importance of the polarization seems to be due to the large and global changes in the charge density caused by this effect, whereas both antisymmetrization and charge transfer are much smaller and more localized. This has been shown by Dreyfus and Pullman for formamide dimers simulating peptide pair interactions,[8] as well as for other systems by other workers.[31] From Table III it is seen that the axial component of the dipole moment is enhanced about 12% in the α-helical $i, i + 3$ dimer and about 7% in the 3_{10}-helical dimer. Dreyfus and Pullman obtained an enhancement of 13% in the dipole moment, while Yamabe and Morokuma[31] find 18% for $(H_2O)_2$. The observed enhancement for the water dimer is 23%.[32]

The trends presented in Tables II and III for the dimeric interactions appear to be in good agreement with previous results.[8,21,28] The relative sensitivity of the energy components to variation in geometry seems to be typical for H-bonded dimers. Also, the interaction energy of the α-helical H-bonded dimer computed here is in excellent agreement with the value of 5.2 kcal/mol obtained by SLPA using similar model assumptions but different atomic basis sets and ab initio methods. In addition, this dimer appears to lose about 2 kcal/mol stability with respect to the optimum geometry formamide dimer,[9,10] showing that H bonding is not the only factor determining the optimum conformation of α helices in proteins, but that the other interactions compensate for the loss of H-bond stability.

2. Trimer. The results for the α-helical double H-bonded trimer are tabulated in Table IV. The stabilization energy per H bond is 5.76 kcal/mol, giving an enhancement of 13% relative to the dimer. The enhancement of the dipole moment has increased from 12% in the dimer to 17% in the trimer. The results obtained with the LCT wave function are similar, i.e., 13 and 20% for the interaction energy and dipole moment, respectively. The usefulness of the LCT calculation for analysis has already been mentioned, and the results obtained here show that the charge-transfer interaction is energetically localized to the regions around the H-bonded proton. There is, however, a more important advantage. In the dimer the full SCF wave function requires the calculation and diagonalization of a 36 × 36 Fock matrix, whereas in the LCT wave function a 19 × 19 and 18 × 18 matrix have to be processed. The time required per iteration is about the same for these two wave functions because the matrix elements of the LCT Fock matrix are more complicated.[11] In the trimer the SCF matrix is 54 × 54, whereas the LCT wave function requires two 19 × 19 matrices and one 18 × 18 matrix. The time required per iteration for the determination of the LCT wave function is about 60% as long as for the SCF wave function. As discussed previously,[11,33] this computational advantage will increase as the system gets larger and the number of groups increases.

The decomposition of the interaction energy and dipole moment, given in Table IV, shows that the importance of the contributions in the trimer is qualitatively the same as in the dimer. The C + E and charge transfer energy contributions are largest with the polarization being smaller but still substantial. The enhancement of the axial component of the dipole moment is primarily due to the polarization effects as in the dimer, while the charge transfer makes a small but nonnegligible contribution. The axial component comprises 94% of the total dipole moment.

The cooperativity effect obtained here for the interaction energy is in good agreement with the result already obtained by SLPA. The increased enhancement of the trimer's dipole moment shows that for this property cooperative effects make an important contribution. Of equal interest is the extent of

Table V. Further Decomposition of the Polarization Contribution of the Trimer and *i*, *i* + 3 Dimer

group polarized	$-E$, kcal/mol	dev[a]	μ_Z, D	dev[a]
		Trimer		
Pol (*i*)	0.34	−0.05	−0.304	0.000
Pol (*i* + 3)	1.16	−0.35	−0.550	0.001
Pol (*i* + 6)	0.58	−0.07	−0.293	0.000
ΣPol (*j*)	2.07	−0.47	−1.147	0.002
Pol	2.35	−0.57	−1.306	−0.037
		i, *i* + 3 Dimer		
Pol (*i*)	0.29		−0.281	
Pol (*i* + 3)	0.51		−0.270	
ΣPol (*j*)	0.81		−0.551	
Pol	0.89		−0.611	

[a] Deviation is calculated as the difference between the trimeric contribution and the appropriate dimeric contributions.

nonadditivity in the trimer interaction. From the practical point of view, it is perhaps the more important consideration since most empirical potential functions assume pair additivity.[34] We have seen from Tables II and III that all three dimers which can be constructed from the trimer exhibit interactions, although for the *i*, *i* + 6 pair they are small. The "true" three-body part of the interaction energy or deviation from pair additivity is therefore defined as

$$\Delta E(i,i+3,i+6) = \Delta E_{i,i+3,i+6} - 2\Delta E_{i,i+3} - \Delta E_{i,i+6} \qquad (5)$$

where the quantities on the right-hand side of eq 5 are the interaction energies defined in eq 3. A similar definition applies to the components of the dipole moment.

The deviations from pair additivity of the interaction energy and the axial component of the dipole moment, reported in Table IV, show that the former is lowered an additional 0.8 kcal/mol, which is 7% of the sum of the pair energies. The dipole moment also exhibits nonadditivity effects, but to a lesser extent, comprising 4% of the sum of the pair contributions. The deviations from pair additivity computed from the LCT wave function are not very different. The decomposition of the nonadditivity effects is also tabulated in Table IV. For the geometrical arrangement of the trimer studied in this paper the main source of nonadditivity is the polarization effect for the interaction energy and the dipole moment. Charge transfer deviates to a lesser degree while the C + E deviation is very small. Kollman has analyzed the nonadditivity effects in trimers of water[28] and it is of interest to compare these with the present results. Compared to $(H_2O)_3$ there is an important structural difference in the formamide trimer in that a carbon atom separates the H bonds of the multiply bonded monomer. This feature should attenuate the movement of density due to double H-bond formation, as confirmed by the total deviation from pair additivity being half as large as in the water trimer.[28] A second difference is that the non-H-bonded pair is more separated here, so that its exchange interaction has essentially vanished. Thus it seems reasonable to assume that the small deviation of the C + E term is not due to cancellation of two large numbers with opposite sign, but shows that this term is nearly pairwise additive.

It is of interest, therefore, to make a more detailed decomposition of the polarization effect in order to gain additional insight into the response of the monomers' charge distribution when its orbitals are allowed to relax in the field of the environment. This type of decomposition is easily made with the NOGF method since a wave function of the form Ψ_2, eq 1b, can be determined with all the groups except one frozen. Thus the polarization of the *I*th group in the field of the remaining electrons is defined as

$$\Psi_2(I) = \mathcal{A}'[\Phi_A{}^0 \ldots \Phi_I \ldots] \qquad (6)$$

and the associated polarization energy is given by

$$E_{\mathrm{Pol}}(I) = E_2(I) - E_1 \qquad (7)$$

and similarly for other properties.

The results of applying this analysis to the trimer and *i*, *i* + 3 dimer are presented in Table V. The contributions to the interaction energy and the dipole moment obtained from each monomer and the deviation from pair additivity are tabulated. The deviations are obtained as the difference between the contribution in the trimer and the analogous contributions in the dimer. The *i*, *i* + 6 pair has not been included since its polarization energy is negligible, and half of its small contribution to the dipole moment is assigned to each monomer.

As expected, the doubly H-bonded monomer makes the largest contribution to the polarization energy and dipole moment. The deviation from pair additivity of this monomer is computed as the difference between $E_{\mathrm{Pol}}(i+3)$ of the trimer and $E_{\mathrm{Pol}}(i) + E_{\mathrm{Pol}}(i+3)$ of the dimer. Thus it is not surprising that this term yields the largest deviation since no doubly H-bonded monomer appears in the dimer. It is also seen that the nonadditivity is due to electron-electron interactions only, which is why the individual terms $\mu_Z(I)$ show no deviation since they are a one-electron property. The sum of the individual polarization terms is also given in Table V, and it is seen that higher order corrections, i.e., simultaneous relaxation of all groups, yield an additional 14% to E_{Pol} and $\mu_Z(\mathrm{Pol})$. The nonadditivity is also increased.

3. Higher Order Complexes. In the previous sections we have given a detailed analysis of the noncovalent interactions in dimeric and trimeric complexes of formamide. This analysis has shown that the trimer interaction contains important nonadditive contributions and that the main source of these nonadditivities is the polarization effect. We now show how the analysis can be used to obtain an estimated interaction decomposition for the higher complexes of formamide simulating H bonding in the α helix. The present estimate naturally assumes that there are no nonadditivities beyond the trimer level.

Table VI tabulates the contribution per H bond of each type of interaction to the energy and dipole moment for a formamide chain of *n* monomers. The polarization energy is obtained by assuming that each doubly H-bonded monomer makes a contribution $E_{\mathrm{Pol}}(i+3)$ scaled by the factor 2.35/2.07 to allow for the relaxation of the other monomers. There are

Table VI. Estimated Energy and Dipole Moment Decomposition for *n* Formamide Monomers Simulating H Bonding in an α Helix

property	C + E	Pol	CT	total
$-E^a$	$\frac{2.78(n-2)+2.20}{n-1}$	$\frac{1.32(n-2)+1.04}{n-1}$	$\frac{2.20(n-2)+2.01}{n-1}$	$\frac{6.30(n-2)+5.25}{n-1}$
$\mu_Z{}^b$	−0.059	$-\frac{0.695(n-2)+0.611}{n-1}$	$-\frac{0.212(n-2)+0.180}{n-1}$	$-\frac{0.966(n-2)+0.850}{n-1}$

[a] Energies are kcal/mol per H bond. [b] Dipole moments are debye/H bond.

Table VII. Estimated Interaction Analysis of High-Order Complexes

no. monomers	C + E	Pol	CT	Δ
		Energy[a] ($-E$)		
4	2.59	1.23	2.14	5.96 (6.00)[c]
5	2.63	1.25	2.15	6.03 (6.13)
10	2.72	1.29	2.18	6.18
∞	2.78	1.32	2.20	6.30 (6.57)
		Dipole Moment[b]		
4	−0.059	−0.667	−0.201	−0.927
5	−0.059	−0.674	−0.204	−0.937
10	−0.059	−0.685	−0.208	−0.952
∞	−0.059	−0.695	−0.212	−0.966

[a] Energies are kcal/mol per H bond. [b] Dipole moments are axial components in debye/H bond. [c] Results in parentheses are from ref 10.

$n - 2$ such monomers for a chain of length n. Similarly, $E_{\mathrm{Pol}}(i) + E_{\mathrm{Pol}}(i + 6)$ scaled by the same factor is the contribution from the two end monomers. Since the nonadditivities of the remaining terms are much smaller a cruder method is used for their estimation. For the C + E term one sums the $n - 1$ nearest neighbors and $n - 2$ next-to-nearest neighbor contributions given in Table II and takes a nonadditivity contribution of 0.04 kcal/mol for the $n - 2$ nearest neighbor triplets. Finally, for the charge-transfer term there are $n - 1$ nearest-neighbor contributions and $n - 2$ nonadditivity contributions of 0.18 kcal/mol. Table II shows, however, that the next-to-nearest neighbor contributions are negligible. The dipole moment contributions are estimated in exactly the same way from the results given in Table III for the α-helical, i, $i + 3$ and i, $i + 6$ pairs, and from the nonadditivities given in Table IV.

The result of extrapolating to an infinite chain can be read off directly from the entries in Table VI. This case and a few intermediate finite chain lengths are tabulated in Table VII. The results of the ab initio calculations by SLPA for the interaction energies of the tetramer and pentamer, and their extrapolation to infinite chain length, are also given. The present estimate for the tetramer is in very close agreement with SLPA's result, whereas for the pentamer the discrepancy is somewhat greater. This is probably due to neglecting new terms, especially residual electrostatic stabilization between distant pairs. For the infinite chain this discrepancy has increased somewhat, but considering the simplicity of our approach the overall agreement is very satisfactory. The extrapolation to infinite chain length of the dipole moment yields an enhancement of 28% of the monomer's axial component. This result is in excellent agreement with Wada's[30] estimate of 30% obtained from different considerations.

This paper may be concluded by noting that the above extrapolation can be carried out neglecting the nonadditivity effects. The infinite chain then yields 5.64 kcal/mol per H bond, which gives an enhancement of about 10%. SLPA also carried out calculations using an electrostatic model which is pair additive by definition. Their extrapolation to an infinite chain gave an enhancement of about 10%. Since both the ab initio calculations and the extrapolation made here are rather different from SLPA's approach, the close agreement of these results and those obtained with inclusion of the nonadditivity effects suggests that potentials which assume pair additivity will only yield about 10% enhancement per H bond for an infinite chain. With inclusion of the nonadditivities the enhancement will be about 25%. From the present calculations this amounts to about 0.6 kcal/mol per H bond, which can become a substantial energy for a long polymer. At the same time, SLPA have also given evidence that the energy surface obtained from an electrostatic potential may be incorrect. These points together suggest that empirical potential functions which are essentially pair additive have a serious defect which adversely affects their reliability. This may contribute to the limited success in predicting protein structure discussed by Hagler and Honig.[1] It certainly indicates that it would be advantageous to include a term for the trimeric nonadditivity effects in empirical potential functions.

Acknowledgment. I would like to thank the University of Basle Computing Center for generous grants of computing time to carry out this research. I am also grateful to the referees for their critical analysis of the first version of the manuscript. The partial support of the Swiss National Science Foundation (Grant 3.991-0.78) is gratefully acknowledged.

References and Notes

(1) A. T. Hagler and B. Honig, *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 554 (1978).
(2) L. L. Shipman and R. E. Christoffersen, *J. Am. Chem. Soc.*, **95**, 1408, 4733 (1973).
(3) R. E. Christoffersen, D. W. Genson, and G. M. Maggoria, *J. Chem. Phys.*, **54**, 239 (1971).
(4) D. A. Kleier and W. N. Lipscomb, *Int. J. Quantum Chem., Quantum Biol. Symp.*, **4**, 73 (1977).
(5) T. A. Halgren and W. N. Lipscomb, *J. Chem. Phys.*, **58**, 1569 (1973); T. A. Halgren, D. A. Kleier, J. H. Hall, Jr., L. D. Brown, and W. N. Lipscomb, *J. Am. Chem. Soc.*, **100**, 6595 (1978).
(6) S. Scheiner and C. W. Kern, *J. Am. Chem. Soc.*, **100**, 7539 (1978).
(7) A. Johansson, P. Kollman, S. Rothenberg, and J. McKelvey, *J. Am. Chem. Soc.*, **96**, 3794 (1974); T. Ottersen and H. H. Jensen, *J. Mol. Struct.*, **26**, 355 (1975); P. Otto, *Chem. Phys.*, **33**, 407 (1978).
(8) M. Dreyfus and A. Pullman, *Theor. Chim. Acta*, **19**, 20 (1970); H. Berthod and A. Pullman, *Chem. Phys. Lett.*, **14**, 217 (1972).
(9) J. F. Hinton and R. D. Harpool, *J. Am. Chem. Soc.*, **99**, 349 (1977).
(10) R. P. Sheridan, R. H. Lee, N. Peters, and L. C. Allen, *Biopolymers*, **18**, 2451 (1979). This paper is referred to as SLPA below.
(11) E. L. Mehler, *J. Chem. Phys.*, **67**, 2728 (1977).
(12) R. McWeeny, *Proc. R. Soc. London, Ser. A*, **253**, 242 (1959); *Rev. Mod. Phys.*, **32**, 335 (1960); M. Klessinger and R. McWeeny, *J. Chem. Phys.*, **42**, 3343 (1965).
(13) C. C. J. Roothaan, *Rev. Mod. Phys.*, **23**, 69 (1951).
(14) K. Morokuma, *J. Chem. Phys.*, **55**, 1236 (1971).
(15) P. A. Kollman and L. C. Allen, *J. Chem. Phys.*, **52**, 5085 (1970).
(16) E. L. Mehler, *J. Chem. Phys.*, to be submitted.
(17) K. Kitaura and K. Morokuma, *Int. J. Quantum Chem.*, **10**, 325 (1976).
(18) S. Huzinaga, *J. Chem. Phys.*, **42**, 1293 (1965).
(19) S. Huzinaga, D. McWilliams, and B. Domsky, *J. Chem. Phys.*, **54**, 2283 (1971).
(20) E. L. Mehler and C. Paul, *Chem. Phys. Lett.*, **63**, 145 (1979).
(21) H. Umeyama and K. Morokuma, *J. Am. Chem. Soc.*, **99**, 1316 (1977).
(22) P. Schuster, "Intermolecular Interactions: From Diatomics to Biopolymers", B. Pullman, Ed., Wiley, New York, 1978, pp 363–432.
(23) S. F. Boys and F. Bernardi, *Mol. Phys.*, **19**, 558 (1970).
(24) R. J. Kurland and E. B. Wilson, *J. Chem. Phys.*, **27**, 585 (1957).
(25) S. Arnott and S. D. Dover, *J. Mol. Biol.*, **30**, 209 (1967).
(26) T. Miyazawa, *J. Polym. Sci.*, **55**, 215 (1961).
(27) R. E. Dickerson and I. Geis, "The Structure and Action of Proteins", Harper and Row, New York, 1969, p 29.
(28) P. Kollman, *J. Am. Chem. Soc.*, **99**, 4875 (1977).
(29) W. G. J. Hol, P. T. van Duijnen, and H. J. C. Berendson, *Nature (London)*, **273**, 443 (1978).
(30) A. Wada, *Adv. Biophys.*, **9**, 1 (1976).
(31) S. Yamabe and K. Morokuma, *J. Am. Chem. Soc.*, **97**, 4458 (1975).
(32) T. R. Dybe and J. S. Muenter, *J. Chem. Phys.*, **60**, 2929 (1974).
(33) E. L. Mehler, *Int. J. Quantum Chem., Suppl.*, **12**, 407 (1978).
(34) G. Némethy and H. A. Scheraga, *Q. Rev. Biophys.*, **10**, 239 (1977).