Theoretical Study of the Electronic Spectroscopy of Peptides. III. Charge-Transfer Transitions in Polypeptides

Luis Serrano-Andrés and Markus P. Fülscher

Contribution from Departamento de Química Física, Universitat de València, Dr. Moliner 50, Burjassot, E-46100 Valencia, Spain, and Department of Theoretical Chemistry, Chemical Centre, University of Lund, P.O.B. 124, S-221 00 Lund, Sweden

Received April 6, 1998

Abstract: The valence excited singlet states of several model polypeptides have been studied using the complete active space (CAS) SCF and multiconfigurational second-order perturbation (CASPT2) methods. It has been found that polypeptides, in addition to the well-known $n \rightarrow \pi^*$ (W) and $\pi \rightarrow \pi^*$ (NV_1) intrapeptide excitations near 220 nm (5.6 eV) and 190 nm (6.5 eV), respectively, have a characteristic band at 165−170 nm (7.5−7.3 eV) corresponding to charge transfers between neighboring peptide units. The intensity of these transitions depends on the conformation of the polypeptide. The present results are consistent with experimental information.

1. Introduction

Electronic spectra of proteins can be considered as the superposition of the spectra of the individual building blocks, provided all chromophores are well-isolated and behave as independent units. Due to the tight coupling of the monomers, however, electronic spectra of enzymes expose interesting changes as compared to the monomers which are controversial and are actively studied experimentally and theoretically. In an effort to bring new information to bear on these questions, earlier we presented papers on the spectroscopic properties of the aromatic amino acids in electronically excited states. This publication is the third contribution in a row of reports on the spectroscopic properties of the protein backbone of which the first paper[1] (hereafter referred to as paper I) focused on amides. The second report[2] included a theoretical study of electronically excited states in glycine and *N*-acetylglycine (hereafter referred to as paper II). Finally, herein we present ab initio calculations on the spectroscopic properties of dipeptides and tripeptides.

Hunt and Simpson[3] reported in 1953 gas-phase ultraviolet spectra of formamide and *N*,*N*-dimethylformamide. In addition to a rich structure of Rydberg states, they identified in formamide the first $\pi \rightarrow \pi^*$ (NV_1) transition near 7.3 eV and placed the second $\pi \rightarrow \pi^*$ (NV_2) transition higher than 9.2 eV. In *N*,*N*-dimethylformamide, of which the electronic spectrum is much broader, they identified the NV_1 band around 6.3 eV and the NV_2 band near 7.7 eV, assuming that the latter shifted as much as 1.5 eV to lower energies upon alkyl substitution. Four years later, Peterson and Simpson[4] reported the electronic spectrum of myristamide, and in accord with the previous work, they considered the 6.7- and 7.7-eV bands to be due to the NV_1 and NV_2 $\pi \rightarrow \pi^*$ valence excited states, respectively. In 1968, however, Basch and coworkers[5] found that the 7.7-eV band of *N*,*N*-dimethylformamide could not be observed in condensed phase experiments and therefore reassigned it to the 3s Rydberg state. In 1975, Robin[6] discarded the assignment of the 165−170-nm (7.5−7.3-eV) band once more in favor of an $n \rightarrow \sigma^*$ transition for the following reasons. By considering exclusively the π orbitals of the amide group (π_1, π_2, and π^*), the NV_1 transition can be described as a $\pi_2 \rightarrow \pi^*$ one-electron promotion (charge transfer between the nitrogen and the carbonyl π^* orbital) and the NV_2 transition as a $\pi_1 \rightarrow \pi^*$ excitation (mainly an intracarbonyl transition). Thus, the NV_2 excitation can be compared to the $\pi \rightarrow \pi^*$ state of ketones and aldehydes, which has never been detected experimentally and is placed by theory at energies higher than 9.0 eV.[7,8] Extending such similarity considerations, one may expect that the $n_O \rightarrow \pi^*$ and $n_O \rightarrow \sigma^*$ transitions will be of lower frequency than the $\pi_2 \rightarrow \pi^*$ band. In addition, Bensing and Pysh[9] noted experimentally that the sum of the intensities of the 7.3−7.5-eV and the NV_1 bands in polypeptides is conserved. In random-coil polypeptides and simple amides, most of the intensity is carried in the NV_1 band where as in helical polypetides, the intensity of the higher band is enhanced at the expense of the NV_1 transition. This observation has been confirmed by McMillin et al.[10] who also showed that both bands have approximately similar intensities in extended β-polypeptides. As the observed polarization of the 7.7-eV band is not compatible with an out-of-plane transition and also not with an in-plane polarization perpendicular to the C═O axis and furthermore, rejecting the appearance of Rydberg states in condensed media, the suggested $n_O \rightarrow \sigma^*$ assignment does not contradict experiments and is in agreement with early all-electron calculations.[6] Recently however, Clark[11] presented crystal spectra of propanamide and *N*-acetylglycine. The observed excitation energies were 6.85 and 6.63 eV, respec-

(1) Serrano-Andrés, L.; Fülscher, M. P. *J. Am. Chem. Soc.* **1996**, *118*, 12190.
(2) Serrano-Andrés, L.; Fülscher, M. P. *J. Am. Chem. Soc.* **1996**, *118*, 12200.
(3) Hunt, H. D.; Simpson, W. T. *J. Am. Chem. Soc.* **1953**, *75*, 4540.
(4) Peterson, D. L.; Simpson, W. T. *J. Am. Chem. Soc.* **1957**, *79*, 2375.
(5) Basch, H.; Robin, M. B.; Kuebler, N. A. *J. Chem. Phys.* **1968**, *49*, 5007.
(6) Robin, M. B. *Higher Excited States of Polyatomic Molecules*; Academic Press: New York, 1975; Vol. II.
(7) Merchán, M.; Roos, B. O. *Theor. Chim. Acta* **1995**, *92*, 227.
(8) Merchán, M.; Roos, B. O.; McDiarmid, R.; Xing, X. *J. Chem. Phys.* **1996**, *104*, 1791.
(9) Bensing, J. L.; Pysh, E. S. *Macromolecules* **1971**, *4*, 659.
(10) McMillin, C. R.; Rippon, W. B.; Walton, A. G. *Biopolymers* **1973**, *12*, 589.
(11) Clark, L. B. *J. Am. Chem. Soc.* **1995**, *117*, 7974.

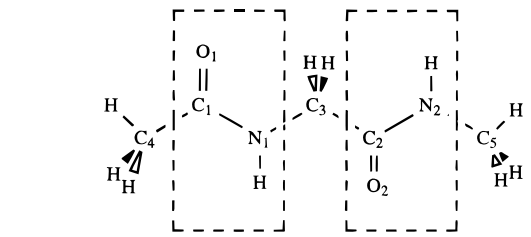
10.1021/ja981148+ CCC: $15.00
Published on Web 10/08/1998

tively, for the NV_1 band and 9.76 and 8.92 eV, respectively, for the NV_2 band. At energies close to 7.7 eV, the signal of the spectrum of propanamide was barely detectable by Clark, whereas the signal of the spectrum of *N*-acetylglycine in this region is dominated by the NV'_1 band localized at the carboxyl group. Thus, Clark could not reach any conclusion regarding the presence of a $n_O \rightarrow \sigma^*$ band.

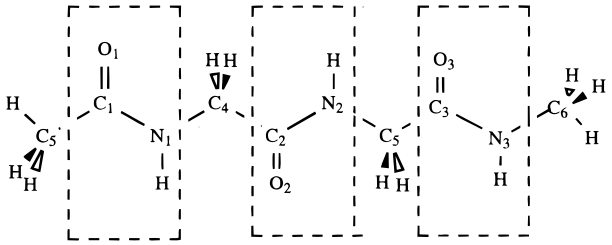
The CASSCF/CASPT2 calculations on a series of primary, secondary, and tertiary amides[1] presented in paper I are in agreement with Clark's crystal spectra and gas-phase spectra.[3,12] The calculations support the Rydberg character of the intermediate bands. In *N*,*N*-dimethylformamide, for instance, the NV_1 and NV_2 transitions are computed at 6.50 and 9.73 eV, respectively, as the most intense bands of the spectrum, while the group of sharp bands near 7.7 eV are computed to be 3p, 3d Rydberg states originated from both the π_2 and the oxygen lone-pair orbitals of the peptide group.[1] We found large effects on the excitation energies upon alkyl substitution at the nitrogen. In contrast, substitutions at the carbon atom of the carbonyl group were computed to shift the excitation energies hardly at all. This is consistent with the character of the NV_1 transition mentioned above. In contrast, substitutions on the nitrogen cannot be expected to affect largely the energy of the NV_2 transition which can be characterized as an intracarbonyl excitation. The NV_2 transition of the simple amides from formamide to *N*,*N*-dimethylformamide was always computed vertically above 9.6 eV. Our calculations, however, questioned the proposed $n \rightarrow \sigma^*$ character of the 7.5−7.3-eV transition in polypeptides.

In a recent theoretical study of glycine (cf. paper II), we identified the most intense features of the vacuum spectra of amino acids[10] to correspond to the following states: At energies ranging from 6.9 to 7.0 eV and close to 9.0 eV, we identified the R'_1 and R'_2 Rydberg transitions, respectively. The NV'_1 and NV'_2 states of the carboxylic group are placed at 8.10 and 10.21 eV, respectively, and the $n_N \rightarrow \sigma^*$ state from the lone-pair of the amide is found close to 8.5 eV. None of these transitions can be expected to be observed in long polypeptides because they are related to the terminal groups. In *N*-acetylglycine,[2] the $n_O \rightarrow \sigma^*$ transition from the peptidic oxygen was found at 9.14 eV with low intensity. It is, however, unlikely that such a transition correspond to the 7.3−7.5-eV band in polypeptides. In addition, hydrogen bonding to proximal proton donors and other solvation effects will possibly lead to a blue shift and weakening of the intensity of the $n_O \rightarrow \sigma^*$ transition.

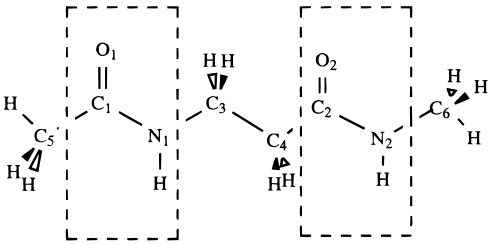
Lacking any reasonable explanation of the 7.7-eV band, we included in paper II an exploratory study of a simple dipeptide, the planar *N*-methyl-acetamide dimer. Within this model, the CASSCF/CASPT2 calculations predicted two charge-transfer (CT) transitions approximately 1 eV higher in energy than the two NV_1 transitions. The CT states included a charge shift between neighboring peptide units. In agreement with experimental values, the charge-transfer bands were found to be long-axis polarized, and the computed oscillator strengths were approximately $^1/_3$ of the NV_1 type transitions. Actually, charge-transfer transitions from the terminal groups to the peptide backbone groups have been reported earlier,[13,14] but to the authors knowledge, they have never been considered as an explanation of the structure of electronic spectra of enzymes. Another experimental observation explained by the present suggestion is that for nylons or, generally speaking, polypeptides separated by long carbonyl chains, the 7.5-eV band is not observed.[6] In short peptides such as diglycine, the band is not clearly observed possibly because it is still too weak. In proteins, however, the number of peptide units in a conformation favorable to charge transfer may increase rapidly.



(1)



(2)



(3)

Figure 1. Model dipeptide (**1**), tripeptide (**2**), and β-dipeptide (**3**) in their all-trans conformation and atom labeling. The peptide groups are signaled by the dashed squares.

In the present paper, we explore in more detail the presence of CT states as prominent features in electronic spectra of proteins. Evidently, CT excitations are strongly dependent on the conformation of the peptide. For these reasons we computed the spectra of the *N*-methyl-acetamide dimer in 15 different conformations. The present study also includes the *N*-methyl-acetamide trimer (tripeptide) and a simple model of a β dipeptide. Figure 1 displays the structures of the studied molecules in the planar, all-trans conformation with respect to the peptide groups and the atomic labeling.

In addition to the expected $n \rightarrow \pi^*$ (W) and $\pi \rightarrow \pi^*$ (NV_1) transitions localized at each one of the peptide units, CT transitions between neighboring peptide units have been found in all of the cases at energies ranging from 7.3 to 7.5 eV. In a planar conformation, the first CT transition is much more intense than the second and has $\pi \rightarrow \pi^*$ character, involving a charge shift from the nitrogen of one peptide group to the closest carbonyl group of the neighboring peptide. The second CT transition involves excitation from the oxygen lone-pair orbital to the π^* orbital of the neighboring peptide.

(12) Kaya, K.; Nakagura, S. *Theor. Chim. Acta* **1967**, *7*, 117.
(13) Chen, X. G.; Li, P.; Holtz, J. S.; Chi, Z.; Pajcini, V.; Asher, S. A.; Kelly, L. A. *J. Am.Chem. Soc.* **1996**, *118*, 9705.
(14) Pajcini, V.; Chen, X. G.; Bormett, R. W.; Geib, S. J.; Li, P.; Asher, S. A.; Lidiak, E. G. *J. Am. Chem. Soc.* **1996**, *118*, 9716.

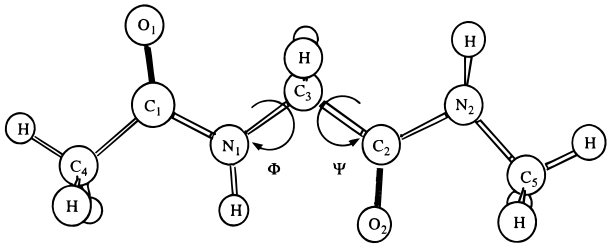


Figure 2. Model dipeptide atom labeling and definition of the rotation angles Φ and Ψ. The all-trans conformation means $\Phi = -180°$ and $\Psi = -180°$.

2. Methods and Computational Details

The present study includes calculations on electronically excited states of the following systems (see Figure 1): the dipeptide **1** model in 15 conformations, the tripeptide **2** model, and the β-dipeptide **3** model. The ground state geometries in a planar, all-trans configuration were optimized at the MP2 level using the 6-31G* basis set. The remaining conformations of the dipeptide model were then generated by rotation about the angles Φ and Ψ (Ramachandran angles), as specified in Figure 2, in steps of 60°, keeping the geometry of the monomers fixed. The nomenclature follows that of ref 15. The coordinates of the optimized and rotated geometries of the three model compounds can be obtained by contacting the authors.

To compute excited-state properties, we used the CASSCF/CASPT2 method and ANO-type basis sets[16] contracted to a split-valence plus polarization quality for C, N, and O [3s2p1d] and [2s] for hydrogen. No Rydberg functions have been added to the basis sets. Their presence is not required as evident by comparison of the present results with data given in paper I. Moreover, we are only interested in the valence states and the influence of the conformation on the computed excited-state properties. The carbon, nitrogen, and oxygen 1s electrons were kept frozen in the form determined by the ground state HF-SCF wave function and were not included in the calculation of the correlation energy.

The CASSCF/CASPT2 method[17−19] is a two-step procedure which calculates the first-order wave function and the second-order energy with a multiconfigurational wave function constituting the reference function. The latter are determined at the CASSCF level of approximation[20] and are used to calculate molecular properties. The CAS state interaction method (CASSI) is then used to compute transition properties[21] and transition moments. Finally, the dynamic correlation energy contributions to the state energies are obtained by the second-order perturbation treatment. Recently, a level-shift technique has been introduced,[22,23] the so called LS-CASPT2 approach, which allows the avoidance of the effect of intruder states common in many calculations on excited states. Here, a value of 0.35 au has been used for the level shift in all of the computed states, after testing the stability of the computed excitation energies within a range of level-shift values.

The selection of the proper active space is the crucial step in the CASSCF/CASPT2 approach. In general, the active space should include all of the orbitals with occupation numbers appreciably different from two or zero in any of the excited states under consideration. This means that all near-degeneracy effects are included in the CASSCF reference function, and consequently, there will be no large terms in the perturbation expansion. To compute the electronic spectra of the present molecules, the active space should initially include for each peptide unit the three π, π^* orbitals and the σ lone-pair orbital on the oxygen together with the six corresponding electrons. This is the minimal space for calculations on the valence excited states. For the dipeptide **1** and the β-dipeptide **3** models, an active space of 8 orbitals and 12 electrons has been employed. For the tripeptide **2** model, the active space consists of 12 orbitals and 18 electrons.

The calculations have been performed with the MOLCAS-4[24] program package on IBM RS/6000 workstations, except for the MP2 optimizations which used the MULLIKEN[25] program.

3. Results and Discussion

This section is structured as follows: In the first subsection, which is further divided into two parts, we shall discuss the dipeptide model. The calculations on the tripeptide and the β-dipeptide models are presented in the second subsection.

3.1. Electronic Spectra of the Dipeptide Model. 3.1.1. Electronic Spectrum of the Planar Species. Considering the π and the oxygen lone-pair orbitals localized at a peptide group, we can expect the following types of transitions at low energies: (1) The W band due to the excitation from oxygen σ lone-pair to the π^* orbital of the amide, (2) the NV_1 band, a $\pi_2 \rightarrow \pi^*$ excitation in origin, and (3) the NV_2 band due to a $\pi_1 \rightarrow \pi^*$ electron promotion. It is interesting to note that the π_1 and π^* orbitals are mainly localized at the carbonyl group, whereas the π_2 orbital is strongly localized at the nitrogen atom. In addition to these valence excitations, one may also expect excitations corresponding to the R_1 and R_2 Rydberg bands found in the isolated molecules. However, in the following we ignored these states since they will be shifted to very high energies in polymers and are therefore of limited interest with respect to the present discussion. Preliminary calculations also show that, after polymerization, the π orbitals of the peptide groups remain localized to a large extent. The NV_2 transitions found at energies above 9 eV are not shifted to low energies, and therefore, we have not included them in the present study.

As shown in paper II, upon polymerization one may expect new features, i.e., charge-transfer bands due to the displacement of electronic charge between neighboring peptide units. The CT_1 charge-transfer excitation can be characterized as a one-electron promotion from the π_2 orbital of group 1 to the π^* orbital of group 2. In contrast, the CT_2 state involves charge transfer from the lone-pair orbital of the oxygen of the peptide group 1 to the π^* orbital of the peptide group 2 or, finally, the opposite charge-transfer excitations from group 2 towards group 1. The ordering of the CT states can be understood qualitatively by geometric arguments and by considering the localization of the orbitals, i.e., the distance between the N_2 atom and the C_1O_1 group is larger than the distance between N_1 and C_2O_2 (cf. Figure 1). To this end, we focus on the W(1) and W(2) and

(15) Adamo, C.; Dillet, V.; Barone, V. *Chem. Phys. Lett.* **1996**, *263*, 113.

(16) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291.

(17) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. In *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics*; Prigogine, I., Rice, S. A., Eds., John Wiley & Sons: New York, 1996; Vol. XCIII, p 219.

(18) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej; A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.

(19) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.

(20) Roos, B. O. In *Advances in Chemical Physics; Ab Initio Methods in Quantum Chemistry II*; Lawley, K. P., Ed.; John Wiley & Sons Ltd.: Chichester, England, 1987; Chapter 69, p 399.

(21) Malmqvist, P. Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.

(22) Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215.

(23) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M.; Molina, V. *J. Mol. Struct. (Theochem)* **1996**, *388*, 257.

(24) Andersson, K.; Blomberg, M. R. A.; Fülscher, M. P.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Neogrády, P.; Olsen, J.; Roos, B. O.; Sadlej, A. J.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Widmark, P.-O. *MOLCAS Version 4.0.* Department of Theoretical Chemistry, Chemical Center, University of Lund, P.O.B. 124, S-221 00 Lund, Sweden, 1997.

(25) Rice, J. E.; Horn, H.; Lengsfield, B. H.; McLean, A. D.; Carter, J. T.; Replogle, E. S.; Barnes, L. A.; Maluendes, S. A.; Lie, G. C.; Gutowski, M.; Rudge, W. E.; Sauer, S. P. A.; Lindh, R.; Andersson, K.; Chevalier, T. S.; Widmark, P.-O.; Bouzida, D.; Pacanski, G.; Singh, K.; Gillan, C. J.; Carnevali, P.; Swope, W. C.; Liu, B. *MULLIKEN Version 1.1.0.* Almaden Research Center, IBM Research Division, 6500 Harry Road, San Jose, CA, 95120-6099, 1994.

Table 1. Excitations Detected in the Absorption Spectra of Different Polypeptides in Various Conformations. Excitation Energies (ΔE, eV) and Intensities (f)[a]

β-extended			α-helix			random		
	ΔE	f		ΔE	f		ΔE	f
polyglycine	5.8	.08	polyleucine	5.7	.02	poly(alanyl)-lysine	6.4	.04
	6.5	.13		6.5	.06		7.3	.26
	7.3	.18		7.2	.21			
polyalanine	5.8	.02	poly-γ-benzylglutamate	5.7	.03			
	6.4	.11		6.7	.19			
	7.4	.17		7.6	.08			
polyvaline	5.7	.06						
	6.3	.14						
	7.2	.21						
polyserine	6.4	.06						
	7.6	.16						
poly(glycyl)-(alanyl)-glutamate	6.4	.08						
	7.3	.18						

[a] References 6, 10. Intensities normalized to one.

the $NV_1(1)$ and $NV_2(2)$ bands localized at the individual peptide groups as well as on the CT_1 and CT_2 excitations.

Table 1 compiles the excitations observed in the absorption spectra of different polypeptides and for various conformations.[6,10] Although the data are not homogeneous for the different compounds, it is clear that three bands can be identified, a weak band at 5.7−5.8 eV, a medium intensity band at 6.3−6.5 eV, and finally a band at 7.2−7.6 eV with an intensity which strongly depends on the conformation of the polypeptide. It is noticeable from the data in Table 1[10] that for polypeptides in β-extended conformations the two highest bands have similar intensities.

Table 2 contains the excitation energies and oscillator strengths computed at the CASSCF/CASPT2 level of approximation for the model dipeptide at the different conformations resulting from twisting the molecule according to the angles defined in Figure 2. However, before entering a general discussion on the effect of conformational changes on the excitation energies and oscillator strengths, we shall focus on the planar structure characterized by the angles $\Phi = -180°$ and $\Psi = 0°$ (structure **1d**). Figure 3 displays charge density difference plots computed at the CASSCF level between the ground and each of the excited states.

The W(2) and W(1) states are almost degenerate and computed at 5.54 and 5.50 eV, respectively, with a small oscillator strength of 0.001 for both of them. The small intensity is typical for an $n \rightarrow \pi^*$ transition in aromatic systems where the overlap between the lone-pair and the π orbitals is almost negligible. Figure 3 shows that charge is transferred from the σ lone-pair orbital of the oxygen (brighter zones) to the peptide π system. The transitions to the 2^1A and 3^1A states can undoubtably be assigned to the weak 5.7−5.8-eV band described in the solution spectra of polypeptides[6,10] which are shifted somewhat to higher energies due to solvation effects.

The valence excited $NV_1(1)$ and $NV_1(2)$ states are predicted to follow the W bands. These states are localized within the peptide groups (cf. Figure 3). The computed vertical excitation energies are 6.54 and 6.58 eV, respectively. The computed oscillator strength is 0.31 for both states. The charge density difference plots also show that charge is displaced in both cases from the nitrogen to the carbonyl group of the same peptide unit. The assignment of these transitions to the prominent band observed in the electronic spectra of polypeptides at 6.4−6.7 eV with high intensity is straightforward. This is the band typically observed for all of the peptides ranging from small amides[1,11] to large polymers such as nylons.[6] The band follows the additivity rule[26] and was shown to be insensitive to conformational changes.[10]

Table 2. Computed CASPT2 Excitation Energies (ΔE, eV) and Oscillator Strengths (f) for the Model Dipeptide at Different Conformations[a]

	Φ(−180)		Φ(−120)		Φ(−60)	
	ΔE	f	ΔE	f	ΔE	f
Ψ(180)	**1a**		**1f**		**1k**	
W(2)	5.79	.001	5.65	.001	5.52	.001
W(1)	5.62	.001	5.69	.001	5.59	.001
$NV_1(1)$	6.39	.321	6.16	.313	6.42	.324
$NV_1(2)$	6.49	.303	6.19	.339	6.44	.278
CT_1	7.18	.134	7.23	.050	7.25	.057
CT_2	8.07	.000	7.54	.166	7.47	.120
Ψ(120)	**1b**		**1g**		**1l**	
W(2)	5.62	.001	5.66	.001	5.65	.001
W(1)	5.43	.001	5.56	.002	5.53	.002
$NV_1(1)$	6.32	.304	6.16	.568	6.42	.157
$NV_1(2)$	6.34	.356	6.13	.176	6.35	.310
CT_1	7.12	.040	7.10	.084	7.19	.089
CT_2	7.45	.184	8.71	.012	7.25	.200
Ψ(60)	**1c**		**1h**		**1m**	
W(2)	5.48	.006	5.64	.001	5.66	.001
W(1)	5.26	.001	5.35	.001	5.50	.001
$NV_1(1)$	6.38	.356	6.08	.457	6.24	.165
$NV_1(2)$	6.40	.315	6.07	.153	6.19	.439
CT_1	7.36	.031	7.59	.032	7.56	.019
CT_2	8.23	.007	7.64	.108	8.76	.115
Ψ(0)	**1d**		**1i**		**1n**	
W(2)	5.54	.001	5.57	.001	5.37	.001
W(1)	5.50	.001	5.49	.001	5.44	.001
$NV_1(1)$	6.54	.314	6.31	.361	6.16	.285
$NV_1(2)$	6.58	.313	6.60	.377	6.68	.213
CT_1	7.57	.139	7.86	.018	7.73	.005
CT_2	8.32	.000	8.47	.057	7.88	.054
Ψ(−60)	**1e**		**1j**		**1o**	
W(2)	5.55	.001	5.59	.001	5.47	.001
W(1)	5.32	.001	5.26	.002	5.22	.002
$NV_1(1)$	6.28	.278	6.25	.219	6.19	.110
$NV_1(2)$	6.38	.289	6.38	.478	6.28	.323
CT_1	7.15	.070	7.25	.109	7.02	.211
CT_2	7.61	.144	8.43	.008	7.56	.039

[a] See Figure 2 and the text for further explanation. Structures **1a**−**1o** in Figure 5.

By the application of the exciton theory,[27] it was predicted that the NV_1 band of helical polypeptides should split into two

(26) Demchenko, A. P. *Ultraviolet Spectroscopy of Proteins*; Springer-Verlag: Berlin, 1986.
(27) Moffitt, W. *Proc. Natl. Acad. Sci. U.S.A.* **1956**, *42*, 736.

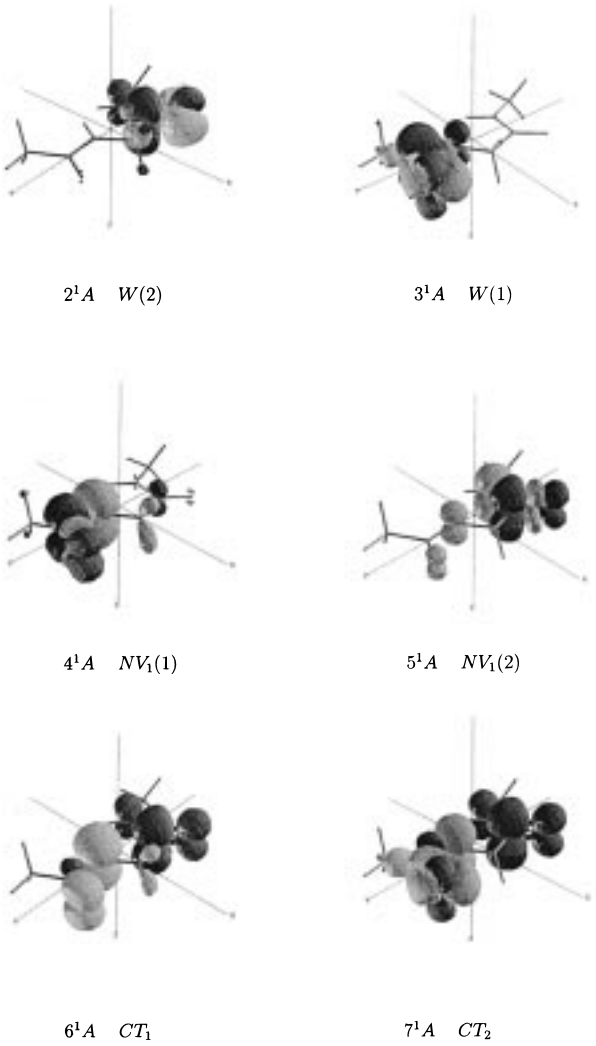


Figure 3. Electron density difference between the ground state and the $^1A'$ excited states in the dipeptide ($\Phi = -180°, \Psi = 0°$). The surface value is 0.004 au. Brighter zones have larger density in the ground state. Darker zones have larger density in the excited state. See text.

components polarized in perpendicular directions. Similarly, the W band was also predicted to be composed of two components. The splitting has been confirmed by experiment.[6] Our calculated polarization directions are in agreement with the predictions of the exciton theory, e.g., the polarization angles of the near degenerate pairs of NV_1 transitions are 64 and $-25°$ for the $NV_1(2)$ and $NV_1(1)$ bands, respectively, with respect to the inertial axes of the molecule.

The 6^1A and 7^1A excited states were computed at 7.57 and 8.32 eV, respectively. The 6^1A state is the CT_1 state described above. As shown by the density difference plots (cf. Figure 3), charge is shifted from the nitrogen and oxygen π orbitals of the peptide group 1 towards the π system of the peptide group 2. The computed oscillator strength of the CT_1 transition is 0.14. The intensity is somewhat weaker than that of the NV_1 band, and as we will show later, it depends on the conformation. The CT_2 transition, on the other hand, is computed with an oscillator strength smaller than 10^{-6} and at high energy, 8.32 eV. This type of transition ($n \rightarrow \pi^*$) is normally weak due to the lack of overlap between the different orbital systems. Figure 3 shows that the the CT_2 transition involves charge transfer from the nitrogen and the σ lone-pair of the oxygen of peptide group 1 to the π system of peptide group 2. Note that, in contrast to the CT_1 state, the π orbital of the oxygen of group 1 is not involved in the charge transfer.

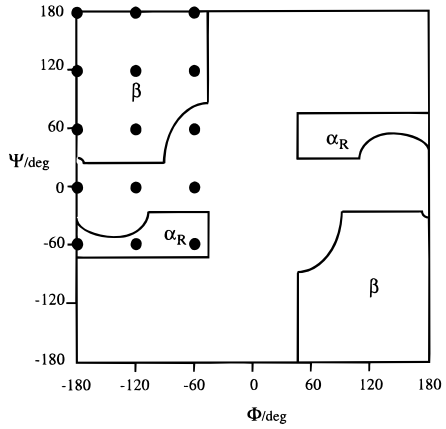


Figure 4. Simplified derivation diagram for the Ramachandran plot of the glycine residue. Black circles point out the employed geometries for the dipeptide. See text.

All of the following transitions have been computed with energies much higher than 9.0 eV and have not been included here. The 8^1A state can be characterized as a double excitation which involves simultaneous excitation of the lone-pair electron of each oxygen towards the π system of the peptide group. Thus, the 6^1A state appears to be the only suitable candidate, from the viewpoint of the energy as well as the intensity, to assign the 7.3−7.5-eV band in polypeptides.

3.1.2. Conformational Dependence of the Electronic Spectra. In this section we describe the electronic excitations computed for the dipeptide model at the different conformations resulting from rotating the peptide groups around the angles Φ ($N-C_\alpha$) and Ψ ($C-C_\alpha$) as defined in Figure 2. Evidently, the values of the rotation angles are constrained geometrically due to steric hindrance. It is common practice to indicate the permitted values of Φ and Ψ on a two-dimensional map of the $\Phi-\Psi$ plane, known as the Ramachandran plot.[28]

In the case of the present dipeptide model (which resembles the glycylglycine without terminal amine and carboxylic groups), the steric restrictions are less severe than in other polyamides. For instance, only 7.5% of the area of the Ramanchandran plot for polyalanine represents allowed conformations, and about 22.5% of the area defines allowed but energetically unfavorable conformations. These fractions increase to 45 and 61%, respectively, for the glycine residues. Figure 4 displays a simplified Ramachandran plot for glycine.[29,30] Two regions are highlighted, corresponding to the most common conformations, the β-sheets, in which the polypeptide is almost fully extended and the α helix, in which all of the CO bonds and NH bonds are oriented in the same direction, respectively.[29] The symmetry of the plot is due to the absence of a side chain in glycine.

We have computed the absorption spectra of the dipeptide at 15 conformations varying the torsion angles in steps of 60° in the following way: Φ ranges from -180 to $-60°$ and Ψ from 180 to $-60°$. Thus, the selected conformations cover the most common structural elements found in proteins. The geometries are shown in Figure 5.

(28) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283.
(29) Creighton, T. E. *Proteins. Structures and Molecular Properties.* W. E. Freeman and Co.: New York, 1993.
(30) Gould, I. R.; Cornell, W. D.; Hillier, I. H. *J. Am. Chem. Soc.* **1994**, *116*, 9250.

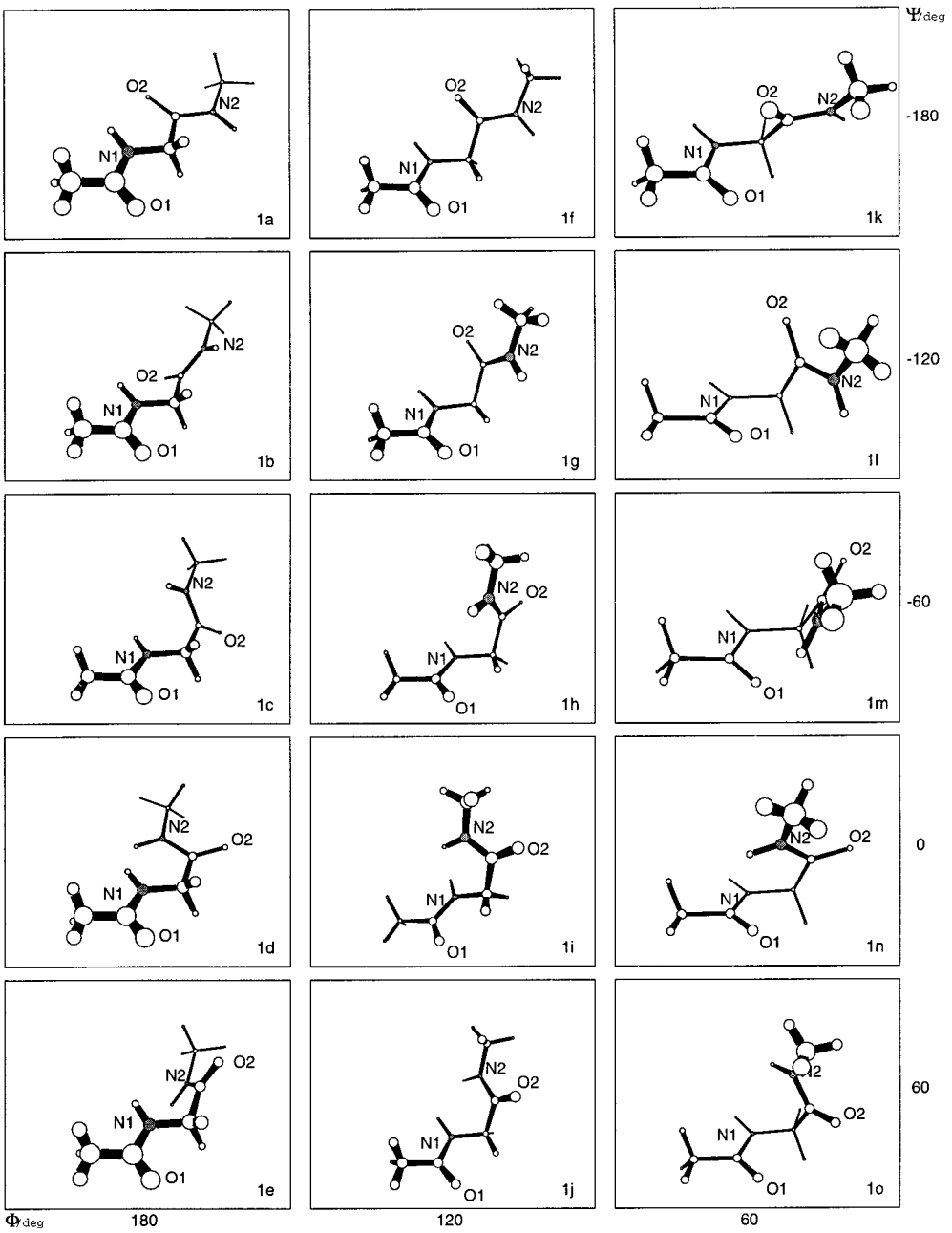


Figure 5. Structures of the different conformations computed for the model dipeptide by rotation of the Ramachandran angles (Φ, Ψ).

The results of the calculations are collected in Table 2. The computed excitation energies of the W bands range from 5.2−5.8 eV, and the oscillator strengths are for all conformations close to 0.001. Likewise, the NV_1 bands are predicted at energies ranging from 6.1 to 6.7 eV with oscillator strengths varying from 0.1 to 0.3. The character of these transitions can be clearly identified for all conformations, and the gap between the W and NV_1 bands is larger than 0.4 eV, throughout. Rotation around the angle Φ has an effect on the excitation energy significantly smaller (0.25 eV or less) than rotations around the Ψ angle (up to 0.5 eV). The effect of the conformation on the excitation energy is not consistent with a model of independent peptide units but is certainly overestimated due to the frozen fragment approximation which has been used to obtain the geometries.

The dependency of the computed excitation energies of the CT_1 band on the conformation is similar to those of the W and NV_1 bands, i.e., the transitions occur at energies ranging from 7.1−7.7 eV and the changes are largest for rotation around the Ψ angle. In contrast, the variation of the computed oscillator strength is larger, 0.03−0.2. The properties of the CT_2 band expose the largest sensitivity to changes in geometry. This pattern can be well understood, recalling that the CT_1 state can be described in many cases as a $\pi \rightarrow \pi^*$ excitation, whereas transitions to the CT_2 state have a significant impact on the σ orbitals, especially the lone-pair orbital of the oxygen of group 1. Because the geometries of the molecules have not been fully optimized for each combination of Φ and Ψ angle, here too, the effect of the conformation on the excitation energies and oscillator strengths is certainly overestimated.

By inspection of Table 2, we also find that the results can be divided into groups. For some conformations, the CT_1 transition is computed at energies between 7.2 and 7.6 eV carrying most of the intensity, whereas the CT_2 transition is located at energies higher than 8.0 eV with low intensity. This situation is typical of the planar or quasiplanar geometries such as the structures **1a**, **1c**, **1d**, **1g**, and **1j** in Figure 5. In these cases, the CT_1 can be easily identified as $\pi \rightarrow \pi^*$ excitation between the two

peptide units. The CT_2 state, on the other hand, is described as the excitation from the oxygen O_1 lone-pair orbital to the π system of peptide group 2. Another important group of structures presents a charge-transfer CT_1 transition computed at 7.1−7.2 eV with oscillator strengths lower than 0.05. The Mulliken population analysis reveals that these transitions still can be described, in general, as the $\pi \rightarrow \pi^*$ excitations, but in most situations, the absence of overlap between the π systems leads to a weak band. In many cases, however, the CT_1 state has mixed character. The CT_2 excitation, on the other hand, is computed at 7.5−7.6 eV with oscillator strengths 0.1−0.2. In most cases, the classification by the character of the transition is hardly possible.

As shown in Figure 4, there is a rather large range of combinations of rotation angles allowed in β-sheets. The structures falling into this range are **1a**−**1d**, **1f**−**1h**, and **1k**−**1l**. Structure **1g** ($\Phi = -120°$, $\Phi = 120°$) is the most characteristic one for which the CT bands are well separated and easily identifiable. The intense CT_1 band is due to a $\pi \rightarrow \pi^*$ transition at 7.10 eV ($f = 0.13$) and is followed by the CT_2 band at 8.71 eV ($f = 0.01$) and clear $n \rightarrow \pi^*$ character. The excited-state properties of the structures **1a** and **1c** ressemble the situation found for structure **1g**. The opposite situation is, on the other hand, observed in structures **1b**, **1f**, **1h**, **1k**, and **1l**. In these cases, CT_2 is the most intense of the two charge-transfer transitions. Though eminently an $n \rightarrow \pi^*$ excitation, this transition is the most intense one because of the larger overlap between the orbital systems. McMillin et al.[10] reported that in polypeptides with β-sheet structure, two transitions were found around 6.4−6.5 eV and 7.2−7.6 eV, respectively, with similar intensities (see Table 1). Particularly in polyglycine, the transitions are located at 6.5 and 7.3 eV. The present calculations are in agreement with these observations and suggest that the second band could be assigned to the CT states.

Structure **1m** corresponds to a situation not found in proteins. Most of the intensity in the high energy region is carried by the CT_2 band for which the excitation energy rises to 8.76 eV. This state corresponds to the 9^1A state in the actual calculation. Because the N_2 atom and the C_1O_1 group are closer than at any other conformation, the charge transfer from the peptide group 2 to group 1 becomes lower in energy. The corresponding CT_1-(2→1) and CT_2(2→1) band were computed at 5.46 and 7.17 eV with oscillator strengths 0.01 and 0.08, respectively. A similar situation has been observed for structure **1i**.

Structures **1e**, **1j**, and **1o** correspond to the conformations found in α helices of which structure **1o** is the most characteristic ($\Phi = -60°$, $\Phi = 60°$). The two charge-transfer transitions are computed at CT_1 7.02 eV with $f = 0.21$ and CT_2 7.56 eV with $f = 0.04$. Moreover, the character of both CT bands is strongly mixed. It is also interesting to note that the CT_2 state is shifted more than 1 eV to lower energies as compared to structure **1g** which represents an idealized β-sheet conformation ($\Phi = -120°$, $\Phi = 120°$).

The present calculations predict that the CT_1 band is polarized in a direction perpendicular to the axis of the chain and parallel to the C_1-O_1 bond. In planar systems, this transition is obviously polarized in the plane of the polypeptide chain. For instance, in structure **1a**, the polarization angle of the CT_1 excitation is +65°, polarized with respect to the inertial axes of the system. On the other hand, the CT_2 bands are polarized in an out-of-plane direction in accord with the $n \rightarrow \pi^*$ character of these transitions. For α-helical structures, the CT_1 transition is polarized along the axis of the helix. As mentioned in the introduction, these directions are compatible with experimental information[9] and led Robin[6] to the assumption, that the 7.5-eV band observed in electronic spectra of proteins is due to an $n \rightarrow \sigma^*$ transition. However, Robin's assignment is unlikely to explain the intensity of the 7.5-eV band, whereas the CT bands carry a substantial fraction of intensity and the calculated polarization directions are in agreement with experimental values as well. For these reasons, we prefer the tentative assignment of the 7.5-eV band to the CT states.

Table 3. Computed CASPT2 Excitation Energies (ΔE, eV) and Oscillator Strengths (f) for the Model Tripeptide and Model β Dipeptide at the all-trans Planar Conformation[a]

	tripeptide		β dipeptide	
state	ΔE	f	ΔE	f
W(2)	5.61	.001	5.10	.001
W(1)	5.74	.001	5.40	.001
W(3)	5.91	.001		
NV_1(2, 3)	6.14	.421		
NV_1(2, 3)	6.16	.211	6.79[b]	.353
NV_1(1)	6.55	.256	6.58	.414
CT_1(1 → 2)	7.01	.169	7.99	.063
CT_1(2 → 3)	7.39	.105		
CT_1(1 → 3)	8.74	.017		
CT_2(1 → 2)	8.12	.000	9.13	.000
CT_2(2 → 3)	8.33	.000		
CT_2(1 → 3)	9.30	.000		

[a] See Figure 1 for structures. [b] NV_1(2) for β dipeptide.

3.2. Electronic Spectra of the Tripeptide and β-Dipeptide. To support the interpretation of the electronic spectra of polypeptides given in the previous section, we also computed the spectra of a model of a tripeptide and a β peptide in an all-trans planar conformation (cf. Figure 1). The results are collected in Table 3.

In accord with the localized nature of the W state, we computed for the tripeptide three W type ($n \rightarrow \pi^*$) transitions with low intensity and three intense NV_1 states. An analysis of the orbitals and configurations involved, as well as the Mulliken population analysis, shows that the two lowest transitions computed at 6.14 and 6.16 eV correspond to a combination of the NV_1(2) and NV_1(3) excitations. Here (2) and (3) denote the peptide group 2 ($C_2O_2N_2H$) and peptide group 3 ($C_3O_3N_3H$), respectively. The excitation NV_1(1) is, however, computed at 6.55 eV. This effect is due to the different environment of the carbonyl carbon. C_2 and C_3 both suffer from charge withdrawing caused by the nitrogens of the neighboring peptide group. In contrast, C_1 has more than 0.1 e excess charge. The effect is to lower the excitation energy of groups 2 and 3 as compared to group 1. Thus, in long polypeptides, one may expect the NV_1 bands to shift to lower energies.

The three following excitations are computed to be the CT_1 type of charge-transfer transitions, that is, excitations from the peptide π system to the neighboring peptide π system. Two transitions computed at 7.01 and 7.39 eV with medium intensities correspond to the CT_1(1→2) and CT_1(2→3) excitations, respectively. These are transitions between adjacent peptide groups. A third one-electron promotion CT_1(1→3) is computed at much higher energy, 8.74 eV, with low oscillator strength, 0.017, which is a transition between peptide groups that are not adjacent neighbors. The CT_2 type transitions ($n \rightarrow \pi^*$ promotions between different peptide groups) are computed at much higher energies and with negligible intensities.

The β-dipeptide model contains a chain of two alkyl groups between the two peptide units and thus mimics the situation in Nylon 3, a polyamide formed by the unit $(\text{-CONH-CH2-CH2-})_n$. Increasing the separation between the peptide groups has a small effect on the W and NV_1 type transitions. It induces, however,

a noticeable blue-shift of the excitation energy of $NV_1(2)$ with respect to the corresponding transition in the dipeptide and tripeptide models. This increase can be related to the withdrawing of charge from C_2 due to the longer alkyl chain. The alkyl group destabilizes the excited state and results in a partial transfer of charge from the nitrogen to the carbonyl group. This effect was also observed in paper I.

The most important changes observed in the excited-state properties of the β-dipeptide with respect to the dipeptide and tripeptide models are the increase of the energies and decrease of the intensities of the interpeptide charge-transfer transitions. The $CT_1(1\rightarrow2)$ transition is computed here at 7.99 eV with an oscillator strength 0.063. The excitation energy is shifted by almost 1 eV to higher energies as compared to the dipeptide, and the intensity is less than $^1/_2$ the previous value. The same effect is observed for CT_2. It is therefore clear that the increase in the peptide−peptide distance has a decisive role in the effect on the properties of the charge-transfer states and explains the absence of the 7.3−7.5-eV band observed in many polyamides.[6,31] β-Peptides have recently been found to fold into helical structures similar to those of the natural peptides.[32] It can be predicted that the 7.3−7.5-eV band observed in the α-peptides will be absent in the β-peptides or that it will at least decrease in intensity.

4. Summary and Conclusions

The present study reports calculations on electronically excited states of simple model compounds aiming to determine qualitatively the nature and characteristics of electronic spectra of polyamides. The CASSCF/CASPT2 method has been employed to compute the excited-state properties. Three model systems have been used (cf. Figure 1), a dipeptide, a tripeptide, and a β-dipeptide of which the planar geometries were optimized at the MP2/6-31G* level of calculation. In addition, the spectrum of the dipeptide has been studied at 15 different conformations. The selected structures cover most of the common regular structures of polypeptides (cf. Figures 2, 3, and 5).

The present calculations complement previous studies (papers I and II) which analyzed the electronic spectra of isolated amides and amino acids. It has been shown that knowledge of the electronic spectra of isolated molecules was sufficient to explain the origin of the W and NV_1 absorption bands of polypeptides at 5.6 and 6.5 eV, respectively. These transitions can be correlated with intrapeptide $n\rightarrow\pi^*$ and $\pi\rightarrow\pi^*$ excited states, respectively. However, there was no convincing evidence for the assignment of the 7.3−7.5 eV found in some polypeptides. Therefore, a proposal was made that this band could be assigned to a charge-transfer excitation between neighboring peptide units.[2]

The present results support the hypothesis that charge-transfer transitions between neighboring peptide groups in polypeptides are the origin of the band observed near 60 000 cm^{-1} (7.4 eV) in many polypeptides. For each peptide group, we were able to identify two charge-transfer transitions named CT_1 and CT_2. The former is due to a $\pi\rightarrow\pi^*$ excitation involving neighboring peptide π systems, and the latter is an $n\rightarrow\pi^*$ excitation. Unlike the W and NV_1 bands, the energy and intensity of the CT bands depend strongly upon the conformation of the system. The computed intensities, however, suggest that the CT bands are intense enough to be observed in polypeptides such as polyalanine[6] in solution or solid phases. It also has been shown that polypeptides in a β-sheet conformation tend to show similar oscillator strengths for the NV_1 and CT_1-CT_2 bands, while for a α-helical conformation the charge-transfer transitions tend to increase their relative intensities. In addition, the character of the CT states in peptides with an α-helical structure is strongly mixed, i.e., these states can not be classified as a transitions of $\pi\rightarrow\pi^*$ or $n\rightarrow\pi^*$ character.

Good agreement with experimental values is also found for the polarization of the bands. The NV_1 transitions are polarized pairwise in perpendicular directions, one parallel and another perpendicular to the axis of the α-helix. The CT_1 excitation has been found to have a transition dipole moment almost perpendicular to the CN bond of the peptide group. Thus, in the α-helical structures, the CT band contributes a polarization parallel to all CO bonds in agreement with the observed directions.[9]

The calculation on the electronically excited states of the tripeptide model confirm the conclusions derived from the calculations on the dipeptides. For each of the units, we found a W and an NV_1 band localized at the peptide group. The substitution effect is to enhance the intensity of the lower energy NV_1 and CT_1 bands. At high energies new, very weak CT bands were found.

As expected, the calculations on the β-dipeptide model (two CH_2 groups between peptide groups) do not reveal new features. The properties of the W and NV_1 bands remain almost unchanged. In contrast, the CT bands are predicted to shift at much higher energy and the intensity to be lowered. These results explains the absence of the 7.3−7.5-eV band in polyamides such as nylons[6,31] which are formed by peptide groups separated by large alkyl chains.

Traditionally, the 7.3−7.5-eV band observed in electronic spectra of polypeptides was considered to correspond to the NV_2 band of isolated amides.[10,26] In view of the calculations presented in papers I and II, we could not confirm the assignment. The NV_2 band is predicted at energies higher than 9.0 eV for all of the molecules included in papers I and II. Rejecting the abundance of Rydberg states in condensed media, it was found that none of the computed $n\rightarrow\pi^*$ or $n\rightarrow\sigma^*$ transitions was characterized by an oscillator strength strong enough to be observed. In contrast, the assignment of the band to the charge-transfer transitions between neighboring peptide groups explains the observed experimental facts and fulfills the additivity conditions required to explain such an intense band.

Evidently, the present calculations are hampered by a number of limitations of technical and principle nature. On the one hand, we applied only small basis sets and simple models. Despite these shortcomings the excited-state properties of formamide computed within this framework reproduce the results presented in paper I with an accuracy of about 0.3 eV. On the other hand, there are more severe limitations which hardly can be solved by ab initio quantum mechanical methods such as the conformational constraints imposed on a pair of peptides by the ensemble of a protein, hydrogen bonding to the nearest neighbors, and other solvation effects. For these reasons, we concentrated on the dominant bands and did not dare to assign the weak feature observed near 7.0 eV in the electronic spectrum of myristamide.[4] In order to converge to a consistent picture, more detailed experimental information on di- and tripeptides in needed.

If one is to speculate about the general implications of the present results, one may note that the peptide backbone is known to have properties of an electric conductor.[33−36] For instance, the photolytic deamination of peptides[37] has been explained by

(31) Onari, S. *Jpn. J. Appl. Phys.* **1970**, *9*, 227.
(32) Iverson, B. L. *Nature* **1997**, *385*, 113.

a mechanism involving electron transfer from the carboxylate anion through the pepide bonds. In view of the present results, such pathways appear resonable because the additional electron will enhance drastically the probability of charge transfer. Additional studies aiming to extend the size and complexity of the systems studied in the present work are in progress.

(33) Wasielewski, M. R. Distance dependencies of electron transfer reactions. In *Photoinduced Electron Transfer*; Fox, M. A., Chanon, M., Eds.; Amsterdam: Elsevier, 1988; Vol. A.

(34) Hill, R. R.; Coyle, J. D.; Birch, D.; Dawe, E.; Jeffs, G. E.; Randall, D.; Stec, I.; Stevenson, T. M. *J. Am. Chem. Soc.* **1991**, *113*, 1805.

(35) Hill, R. R.; Jeffs, G. E.; Banaghan, F.; McNally, T.; Werninck, A. R. *J. Chem. Soc., Perkin Trans. II* **1996**, p 1595.

(36) Isied, S. S.; Vassilian, A.; Magnuson, R.; Schwarz, H. A. *J. Am. Chem. Soc.* **1985**, *107*, 7432.

(37) Birch, D.; Coyle, J. D.; Hill, R. R.; Jeffs, G. E.; Stec, I.; Stevensson, T. M. *J. Chem. Soc., Chem. Commun.* **1991**, 507.

Acknowledgment. The research reported in this paper has been supported by a grant from the Swedish Natural Science Research Council (NFR), by projects PB94-0986 and PB97-1377 of the Dirección General de Investigación Científica y Técnica of Spain, and by the European Commission through the TMR network FMRX-CT96-0079.

JA981148+