CHIROPTICAL PROPERTIES AND CONFORMATION
OF N-ACETYL-L-AMINO ACIDS N'-METHYLAMIDES
WITH ALIPHATIC SIDE CHAINS*

Petr MALOŇ[a], Petr PANČOŠKA[b], Miloš BUDĚŠÍNSKÝ[a], Jan HLAVÁČEK[a], Jan POSPÍŠEK[a] and Karel BLÁHA[a]

[a] *Institute of Organic Chemistry and Biochemistry,*
Czechoslovak Academy of Sciences, 166 10 Prague 6, and
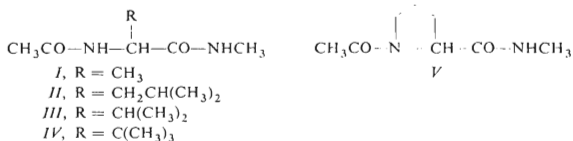[b] *Department of Chemical Physics, Charles University, 121 16 Prague 2*

Dedicated to Academician O. Wichterle on the occasion of his 70th birthday.

CD and ^{1}H NMR spectra of L-alanine (*I*), L-leucine (*II*), L-valine (*III*) and L-tert-leucine (*IV*) N-acetyl N'-methylamides were measured at various experimental conditions involving changes of temperature, concentration and solvent polarity. The least flexible tert-leucine derivative *IV* exists predominantly either in right handed α-helical (α_R) conformation (fluorinated alcohols, ethanol) or in extended (C_5) conformation (cyclohexane, acetonitrile). In this compound the Φ angle is constrained to about $-120°$, excepting aqueous solution. The flexible amides *I* and *II* exhibit more complex conformational equilibria involving probably the 3_{10}, α_R and C_5 conformations. The 3_{10} helical conformation is favoured in water, acetonitrile and non-fluorinated alcohols, while a higher participation of α_R conformation is observed in strongly polar fluorinated alcohols. Conformational distribution of the valine derivative *III* is similar to *IV* in fluorinated alcohols and to amides *I* and *II* in other solvents. The C_5^{eq} conformation is clearly detected only in cyclohexane solutions of non-associated forms of *I* and *II*.

Methylamides of N-acetyl-α-amino acids are models simulating fragments of a peptide chain which are composed of two amide groups and one complete amino acid residue (the so-called dipeptide unit). They play an important role in initial stages of conformational energy calculations of larger peptides and in their conformational investigation with the aid of spectroscopic and other methods. Despite these facts and numerous experimental and computational studies (*e.g.*[1-7]) the conformational arrangement of dipeptide units is still not clear in all details. This situation is caused primarily by the considerable flexibility, characteristic for the majority of residues of naturally occurring amino acids. Therefore it is of advantage to extend the series by a compound having more fixed conformational arrangement as a consequence of increased steric hindrance. In this communication we present a detailed conforma-

tional analysis of N′-methylamides of N-acetyl-L-alanine (*I*), N-acetyl-L-leucine (*II*), N-acetyl-L-valine (*III*) and N-acetyl-L-tert-leucine (*IV*). The series is chosen in such a way that a steric volume of the side chain is gradually increased up to the extreme case of tertiary butyl group. We performed also several measurements with N-acetyl-L-proline N′-methylamide (*V*) in order to make a proper comparison with previous results[8]. Two members of the set were already investigated by a wide range of physico-chemical methods by Ivanov and coworkers[1-3,9,10] (compounds *I* and *III*), Madison and Kopple[8] (compound *I*) and other authors[4,11-13]; recently also a preliminary report including *IV* has appeared[14]. The analysis is based on a combined use of CD spectra, ^{1}H NMR spectra and conformational energy calculation. The results are compared with our previous IR spectroscopic investigation[16]. For the analysis of CD data we use the modified factor analysis[16].

$$
\begin{array}{cc}
\overset{\displaystyle R}{\underset{|}{}} \\
CH_3CO-NH-CH-CO-NHCH_3 & \qquad CH_3CO-N \quad CH-CO-NHCH_3 \\
\textit{I},\ R = CH_3 & \textit{V} \\
\textit{II},\ R = CH_2CH(CH_3)_2 \\
\textit{III},\ R = CH(CH_3)_2 \\
\textit{IV},\ R = C(CH_3)_3
\end{array}
$$

The compounds *I*—*III* were synthesized from the corresponding L-amino acids by successive esterification, acetylation and reaction with methylamine (ref.[17]). N-Acetyl-L-tert-leucine N′-methylamide (*IV*) was prepared from L-tert-leucine N-carboxyanhydride[18] by a reaction with methylamine and subsequent acetylation.

EXPERIMENTAL

Melting points were determined on a Kofler block and were not corrected. Samples for elemental analysis were dried for several hours at room temperature over phosphorus pentoxide at 150 Pa. Optical rotations were determined on Perkin–Elmer 141 MCA polarimeter.

N-Acetyl-L-alanine methylamide (*I*), m.p. 182—183°C. $[\alpha]_D = -102 \cdot 8°$ (*c* 0·21, chloroform) and N-acetyl-L-leucine methylamide (*II*), m.p. 166—168°C, $[\alpha]_D = -82 \cdot 80$ (*c* 0·53, chloroform) were synthesized according to ref.[17] (reported m.p. 181—182°C, $[\alpha]_D = -51 \cdot 1°$ (*c* 2, ethanol) for the compound *I* and m.p. 165—167°C, $[\alpha]_D = -33 \cdot 90°$ (*c* 1, water) for the compound *II*). N-Acetyl-L-valine methylamide (*III*), m.p. 233—234°C, $[\alpha]_D = -41 \cdot 2°$ (*c* 0·49, water) was prepared by an analogous procedure (reported m.p. 235°C, $[\alpha]_D = -48°$ (*c* 1·7, water), see ref.[19]).

L-Tert-leucine Methylamide (*VI*)

Etheral solution (20 ml) of methylamine, liberated from methylamine hydroride (1 g, 15 mmol), was added to N-carboxy-L-tert-leucine anhydride (0·5 g, 3·2 mmol) in dichloromethane (10 ml) and the mixture was allowed to stand for 24 h at 0°C. The precipitate was filtered off, washed with ether and the solution was taken down. The residue was dissolved in minimum amount

of ether and the amide was precipitated by light petroleum. Yield 0·33 g (72%), m.p. 92—94°C. $[\alpha]_D = +59 \cdot 1°$ (*c* 0·30, dimethylformamide). For $C_7H_{16}N_2O$ (144·2) calculated: 58·29% C, 11·18% H, 19·43% N; found: 57·97% C, 10·98% H, 19·02% N.

N-Acetyl-L-tert-leucine Methylamide (*IV*)

A mixture of *VI* (0·3 g, 2·1 mmol) and acetanhydride (1 g, 9·8 mmol) was allowed to stand for 3 h at room temperature and taken down. The residue was dissolved in a minimum amount of 50%, aqueous methanol and filtered through columns of Dowex 50 (25 ml) and Amberlite IR 4C (10 ml). The eluate was taken down and the compound was crystallized from benzene–light petroleum. Yield 0·24 g (62%), m.p. 218—220°C, $[\alpha]_D = +41 \cdot 7°$ (*c* 0·52, dimethylformamide). For $C_9H_{18}N_2O_2$ (186·2) calculated: 58·03% C, 9·74% H, 15·04% N; found: 58·22% C, 9·75% H, 15·22% N.

Spectroscopic Measurements

CD Spectra were measured on Roussel–Jouan Dichrographe CD 185/II equipped with a cryostate. The measurements were performed in acetonitrile, ethanol, water, 2,2,2-trifluoroethanol, 1,1,1,3,3,3-hexafluoro-2-propanol at room temperature. The dependence on temperature was measured in cyclohexane (25—73°C), acetonitrile (−40 to +30°C) and methanol–ethanol (1 : 1) mixture (−80 to +40°C). The concentration of solutions was about $2 \cdot 10^{-3}$ mol l^{-1} excepting cyclohexane solutions ($\sim 2 \cdot 10^{-4}$ mol l^{-1}, limited solubility). The dependence on concentration was measured in acetonitrile. The measurements were performed in quartz cells with the optical path length of 0·01—2 cm. Spectral data are given as molar ellipticities $[[\Theta]$ (deg cm^2 dmol^{-1})] and are not corrected for solvent refractive index.

The modified factor analysis was applied to the set of twenty CD curves of compounds *I*—*V*. each of them measured in five solvents at room temperature. The procedure was carried out as described in our previous paper[16]. The necessary number of subspectra was determined using the eigenvalue drop criterion together with the graphical comparison of calculated and experimental curves. The formalism is described in the appendix.

^{1}H NMR measurements were carried out using Varian HA-100 (100 MHz, CW mode) and Varian XL-200 (200 MHz, FT mode) spectrometers. The spectra were recorded in deuteriochloroform, trideuterioacetonitrile and 1,1,1,3,3,3-hexafluoro-2-propanol (or 2,2,2-trifluoroethanol) at room temperature (22—27°C) at a concentration of about $2 \cdot 10^{-1}$ mol l^{-1}. The measurement in trideuterioacetonitrile was performed also at a concentration of $3—4 \cdot 10^{-3}$ mol l^{-1} (FT mode, 500 accumulations) in order to investigate the effect of dilution. The assignment of particular proton signals is based on homonuclear decoupling experiments. The chemical shifts and coupling constants were obtained by first order analysis.

Conformational energy calculations of the compound *IV* were carried out using potential functions and computational procedures described in ref.[20].

RESULTS

Chiroptical properties. CD spectra of compounds *I*—*V* are given in Tables I—IV and Figs 1—3. For the curves already reported (refs[3,8,14]) a reasonable agreement is achieved. The couples of amide chromophores give rise to distinct Cotton effects within the $n—\pi^*$ (205—260 nm) and $\pi—\pi^*$ (185—205 nm) region. The assignment of particular CD bands to these transitions is straightforward with the exception

TABLE I

CD Spectra of compounds $I—V$ in solvents of various polarity

Solvent[a]	$n—\pi^*$ region	$\pi—\pi^*$ region	
	$\lambda_{max}([\theta]_{max})^b$	$\lambda_{max}([\theta]_{max})^b$	end value[b]
		compound I	
ACN	218 (−5·8); 213 (−5·7) n	198·5 (−13·0)	190 (+1·9)
Ethanol	210 (−7·5) s	196·5 (−13·2)	185 (+6·6)
Water	c	195 (−20·1)	185 (−6·4)
TFE	209 (−10·9) s	204^d (−11·4)	185 (+5·9)
HFP	208 (−11·8)	193 (−7·3) s; 189 (−6·7) s	185 (−2·6)
		compound II	
ACN	222 (−3·1); 215 (−2·8) n	196·5 (−11·8)	185 (+8·4)
Ethanol	214 (−5·1) s	186 (−10·4)	185 (+5·8)
Water	215 (+1·1)	194 (−16·0)	185 (−8·0)
TFE	213 (−6·9)	195 (−3·9) s	187 (−0·2)
HFP	213 (−8·0)	195 (+1·5)	185 (−8·4)
		compound III	
ACN	216 (−5·9); 209 (−5·7) n	196 (−10·0)	187 (+7·7)
Ethanol	214 (−10·5)	c	185 (+10·8)
Water	205 (−7·2) s	196 (−10·5)	185 (−3·4)
TFE	211 (−11·9)	187 (+4·9)	185 (+4·7)
HFP	209 (−14·9)	192 (+5·0)	185 (+1·2)
		compound IV	
ACN	225 (−1·4) s; 215 (−2·2)	e	186 (+6·2)
Ethanol	215 (−7·5)	e	190 (+14·5)
Water	211 (−4·8)	192 (−0·6) n	185 (−3·7)
TFE	213 (−8·7)	e	188 (+14·3)
HFP	213 (−8·7)	194 (+15·5)	190 (+13·5)
		compound V	
TFE	210 (−18·7) s	200 (−23·4)	185 (0)
HFP	210 (−22·5) s	203 (−26·2)	185 (−10·5)

^{a} ACN acetonitrile, TFE 2,2,2-trifluoroethanol, HFP 1,1,1,3,3,3-hexafluoro-2-propanol; ^{b} λ_{max} wavelength of the maximum (nm), $[\theta]_{max}$ molar ellipticity of the maximum . 10^{-3} (deg cm^2 . dmol^{-1}), n negative minimum, s shoulder; ^{c} no distinct dichroic absorption in the $n—\pi^*$ transition region; ^{d} value influenced by superposition of closely lying bands with opposite signs; ^{e} maximum of the $\pi—\pi^*$ band outside the region accessible to measurement.

TABLE II

Temperature dependence of CD spectra of compounds $I-V$ in acetonitrile[a]

Compound	Temperature °C	$[\theta]_{211}$	$[\theta]_{220}$	$[\theta]_{230}$
I	+ 40	− 6 200	− 6 900	− 4 900
	− 40	− 3 900	− 1 900	− 1 100
II	+ 40	− 2 800	− 3 100	− 2 600
	− 40	− 300	+ 1 100	+ 300
III	+ 40	− 4 500	− 4 800	− 3 300
	− 40	− 4 500	− 3 400	− 1 600
IV	+ 40	− 1 300	− 1 600	− 1 100
	− 40	− 1 900	− 1 600	− 500
V	+ 40	− 17 000	− 18 500	− 14 500
	− 40	− 20 600	− 16 600	− 10 300

[a] For units see notes to Table I. $[\theta]$ molar ellipticity (deg cm^2 dmol^{-1}).

TABLE III

Temperature dependence of CD spectra of compounds $I-V$ in methanol–ethanol (1 : 1) mixture[a]

Compound	Temperature °C	$[\theta]_{200}$	$[\theta]_{210}$	$[\theta]_{215}$	$[\theta]_{220}$
I	+ 40	− 12 600	− 7 400	− 6 300	− 5 400
	− 80	− 18 600	+ 1 400	+ 3 400	+ 3 100
II	+ 40	− 9 300	− 4 400	− 4 200	− 3 800
	− 80	− 15 300	− 4 700	+ 6 300	+ 4 800
III	+ 40	− 4 900	− 9 300	− 9 700	− 8 300
	− 80	− 9 000	− 5 900	− 4 700	− 3 000
IV	+ 40	+ 3 200	− 5 800	− 7 400	− 6 600
	− 80	+ 7 200	− 6 900	− 8 000	− 6 000
V	+ 40	− 20 300	− 17 400	− 13 400	− 9 800
	− 80	− 32 700	− 19 500	− 12 600	− 6 600

[a] For units see notes to Tables I and II.

of cyclohexane solutions. The short wavelength branch of the $\pi-\pi^*$ couplet lies in most cases outside the region accessible to measurement, since the $\pi-\pi^*$ absorption maximum of secondary diamides is located at about $185-188$ nm[3]. Simultaneously with the decreasing flexibility of compounds $I-IV$ their CD spectra reflect a gradual decrease of sensitivity to changes of solvent polarity (Table I, Figs 1, 2) and temperature of the solution (Tables II, III, Fig. 3). Within the whole set of spectra we can trace three basic spectral types according to the signs of $n-\pi^*$ and $\pi-\pi^*$ bands. The first spectral type (A) is represented by a combination of the negative $n-\pi^*$ band and the positive $\pi-\pi^*$ band (e.g. IV in hexafluoropropanol, Fig. 2b). Both these bands exhibit negative signs for the second spectral type (B) (e.g. leucine derivative II in acetonitrile, Fig. 1b). In this case the $n-\pi^*$ dichroic band is usually detected as a wide shoulder. The third spectral type (C) is approximately enantiomorphous to the first type (positive $n-\pi^*$ band and negative $\pi-\pi^*$ band (e.g. II in water, Fig. 1b).

Analysing first the dependence of CD spectra on solvent polarity at room temperature (Table I, Fig. 1, 2) we can differentiate the spectra of compounds $I-III$ from those of the compound IV. The spectra of flexible compounds $I-III$ exhibit sign alterations of both the $n-\pi^*$ and $\pi-\pi^*$ bands and their dependence on solvent polarity displays characteristic isodichroic points at $200-205$ nm and, less clearly, at about 225 nm. On the contrary, the spectra of compound IV are less sensitive. The latter curves belonging invariably to the type A are almost uniformly shaped in ethanol, hexafluoropropanol and trifluoroethanol, while a pronounced intensity decrease is observed in acetonitrile and water. The spectra of compounds I and II

TABLE IV

Temperature dependence of CD spectra of compounds $I-IV$ in cyclohexane[a]

Compound	Temperature °C	$[\theta]_{200}$	$[\theta]_{215}$	$[\theta]_{230}$
I	25	$+\ \ \ 400$	$-\ 3\ 200$	$-\ 3\ 000$
	72	$+\ 1\ 700$	$-\ 5\ 800$	$-\ 7\ 800$
II	25	$-\ 1\ 900$	$-11\ 100$	$-13\ 200$
	72	$+\ 1\ 500$	$-\ 9\ 000$	$-14\ 700$
III	25	[b]	$-\ 2\ 600$	$-\ 2\ 600$
	72	[b]	$-\ 2\ 000$	$-\ 2\ 000$
IV	25	$+21\ 200$	$-\ 2\ 500$	$-\ 3\ 300$
	72	$+\ 3\ 000$	$-\ \ \ 600$	$-\ 4\ 700$

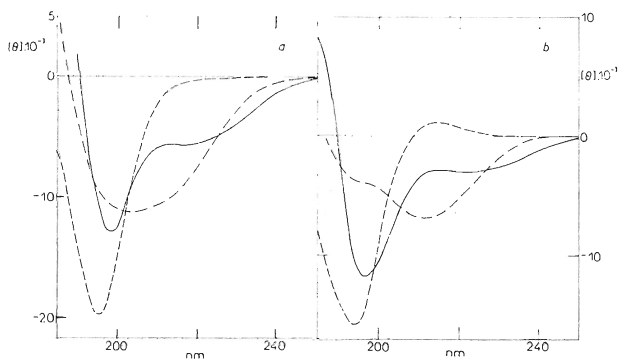[a] For units see notes to Tables I and II; [b] not measured.

Fig. 1

CD Spectra of compounds *I* (*a*) and *II* (*b*) measured in acetonitrile (————·), water (– – – –), trifluoroethanol (–·–·–·), hexafluoropropanol (······)
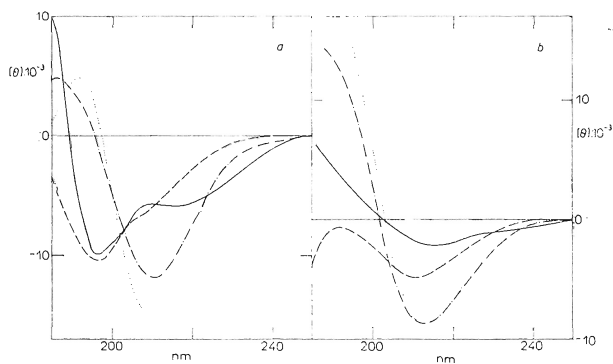

Fig. 2

CD Spectra of compounds *III* (*a*) and *IV* (*b*) measured in acetonitrile (————), water (– – – –), trifluoroethanol (–·–·–·), hexafluoropropanol (······)

possess a gradually changing shape from the type B towards type A when the solvent polarity increases. The B spectra are characteristic for compounds I and II in acetonitrile and ethanol and for the compound III in acetonitrile and water. In strongly polar solvents the contribution of A spectra is gradually lowered with decreasing bulkiness of the side chain (Tle < Val < Leu < Ala). The transformation of the spectral type B to A with increasing solvent polarity is rather smooth with the exception of spectra measured in water. Although in our solvent polarity scale (based on solvent-induced shifts of the amide $n-\pi^*$ dichroic band in rigid polycyclic lactams, see ref.[21,22]) water stands between ethanol and trifluoroethanol, the spectra indicate a significant contribution of the spectral type C at least for the flexible compounds I and II. For II the type C is quite clearly discernible due to the presence of a positive $n-\pi^*$ band, while in the case of I there is no distinct dichroic absorption within the $n-\pi^*$ transition region, probably as a consequence of compensation of oppositely signed bands. The contribution of A spectra is suppressed in aqueous solution, since it is qualitatively preserved only in the spectrum of the least flexible IV. Even in the case of III the spectrum is of the type B despite the clear appearance of the type A in less polar ethanol. Generally, the spectra of compounds I–IV show distinct hypsochromic shifts of the $n-\pi^*$ transition with increasing solvent polarity, but it is difficult to evaluate its magnitude due to overlapping of bands. In order to quantify the above observations this fundamental set of CD curves has been further
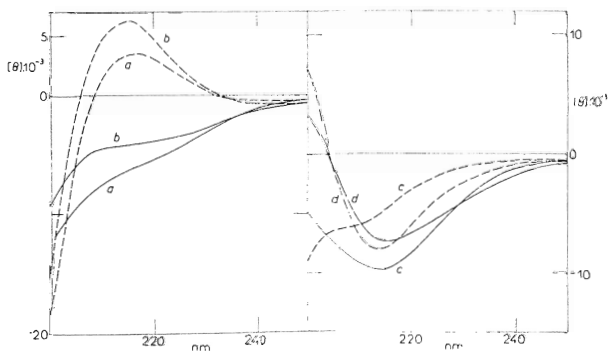

FIG. 3

Temperature dependence of CD spectra of compounds I (a), II (b), III (c), IV (d) measured in methanol-ethanol (1 : 1) mixture. (———) $t = +40°C$, (-----) $t = -80°C$

evaluated by the modified factor analysis. It has been proved by dilution studies that these spectra are not concentration dependent (range $10^{-3}-10^{-4}$ mol l^{-1}). The dependence of CD spectra on temperature was measured in acetonitrile (Table II) and methanol–ethanol mixture (Table III, Fig. 3, see also[23]). The parallel behaviour of compounds $I-IV$ was observed in both solvent systems. Again, we can differentiate the spectra of tert-leucine derivative IV from other members of the set. The spectra of IV do not exhibit any substantial changes when the temperature is decreased. In both solvent systems the spectra can be ascribed to the type A although band intensities in acetonitrile are markedly lower. On the contrary, CD spectra of compounds $I-III$ exhibit a gradual change from spectral type B to C when the temperature is lowered. The change is largest for compounds I and II in the spectra of which the sign reversal of the $n-\pi^*$ band is observed. The spectra of III exhibit significant decrease of intensity of the negative $n-\pi^*$ band.

Since N-acetyl-L-proline N'-methylamide (V) serves as a model[8] in which the ϕ angle is constrained to about $-60°C$ we performed CD measurements in trifluoroethanol and hexafluoropropanol (Table I) as well as low temperature studies in acetonitrile (Table II) and methanol–ethanol mixture (Table III). Both CD spectra recorded in fluorinated alcohols are of the type B with the $n-\pi^*$ dichroic absorption intensity comparable to that found in acetonitrile[8]. Lowering of the negative $n-\pi^*$ dichroic band with decreasing temperature is observed in both low temperature studies with the simultaneous increase of the negative $\pi-\pi^*$ band. This process corresponds qualitatively to the change of spectral type B to C as observed with compounds I and II but in this case no positive $n-\pi^*$ dichroic band can be detected.

In cyclohexane the spectra of compounds $I-IV$ exhibit invariably a negative band of the $n-\pi^*$ transition (Table IV). The maximum of the positive $\pi-\pi^*$ transition has been detected only for IV at elevated temperature. In addition, the third small band of negative sign is observed in spectra of all compounds at about 215 to 220 nm. Its intensity decreases with increasing temperature while the intensity of the $n-\pi^*$ band increases. This process is clearly observed with compounds I and IV. The positive $\pi-\pi^*$ band of the compound IV decreases markedly with increasing temperature. The observed spectral changes correlate with the gradual decrease of self-association of these compounds which has been observed under similar conditions by IR spectroscopy[15]. The intensity of the negative $n-\pi^*$ band, when recorded at the highest accessible temperature, decreases in the series Leu < < Ala < Tle < Val, but these data can be interpreted only semiquantitatively due to poor solubility of these diamides, especially of the compound III, in cyclohexane.

Modified factor analysis. Table V summarizes numerical criteria used for the determination of the number of subspectra which are necessary for the description of the given set. The sudden drop of significance given by the order of magnitude

difference between eigenvalues of the fourth and fifth subspectrum is consistent with the graphical comparison of calculated and experimental curves. The deviations become smaller than the experimental error when four subspectra are used in the linear combination (Table V). The first subspectrum adopting exclusively positive weight values participates most significantly in the spectra of *I* measured in polar solvents (trifluoroethanol, hexafluoropropanol). The lowest contribution of the first subspectrum is found in the spectra of *IV* excepting aqueous solution. The first subspectrum can be identified with the spectrum of *I* in trifluoroethanol on the basis of comparison with experimental curves (Fig. 4). The second subspectrum adopts positive as well as negative weights in the range -50 to $+76\%$ (Fig. 5a). In all solvents the weights of the second subspectrum increase with increasing bulkiness of the side chain (the exception exhibited by acetonitrile solutions of *II* and *III* is not significant due to minimal absolute values of the corresponding weights). There is a marked differentiation of compound *IV* having maximal positive weight and of the compound *I* exhibiting always negative weights. The magnitude of contributions of second subspectrum depends also on the character of solvent. There is a tendency to increase weights in the series water–acetonitrile–ethanol–trifluoroethanol–hexafluoropropanol with the exception of compound *I*. Apparently, the specific character of water is emphasized. The second subspectrum can be identified with the experimental curve of *IV* in hexafluoropropanol, *i.e.* the compound having the most bulky side chain in the solvent of highest polarity. The third subspectrum, possessing no direct experimental equivalent, adopts weights in the range -44 to $+32\%$. The magnitude of its contribution is markedly solvent dependent. The weights of the

TABLE V

Criteria for the determination of the number of necessary subspectra

Sub-spectrum i	Eigenvalue λ_i	% Description of experimental data	$\bar{\Delta}^a$ [deg cm^2 dmol^{-1}]
1	12·752	63·76	338
2	6·548	96·45	80
3	0·298	97·94	44
4	0·102	97·94	18
5	0·009	98·49	15
6	0·005	98·52	11
7	0·002	98·54	8

^{a} Average mean quadratic deviation $\bar{\Delta} = 1/p \sum_i \Delta_i$, p denotes number of experimental CD curves.

third subspectrum are rather similar for all compounds measured in the same solvent and exhibit an increase with increasing solvent polarity (Fig. 5b). For compounds *III* and *IV* the highest weights are exhibited by aqueous solutions while in the case of compounds *I* and *II* the weights of third subspectrum increase in the series aceto-nitrile–ethanol–water–trifluoroethanol–hexafluoropropanol. The fourth subspectrum exhibits significant contributions only to spectra of *II* and *III*. There is also no direct experimental equivalent of this subspectrum.

^{1}H NMR *spectra*. The values of coupling constants $^3J_{NH,CaH}$ (Table VI) increase gradually with increasing bulkiness of the side chain up to the extreme case of *IV*. In accord with the results of Delaney and coworkers[14], it can be derived from these high values that the ϕ angle in *IV* is constrained to about $-120°$ with the exception of aqueous solutions. In aqueous solution, a systematic decrease of coupling constants by about $1 - 1.5$ Hz is observed. In other solvents the coupling constants are not very sensitive to solvent polarity, the values found in trideuterioacetonitrile and hexa-fluoropropanol are almost identical. The dependence on concentration measured in trideuterioacetonitrile is negligible.
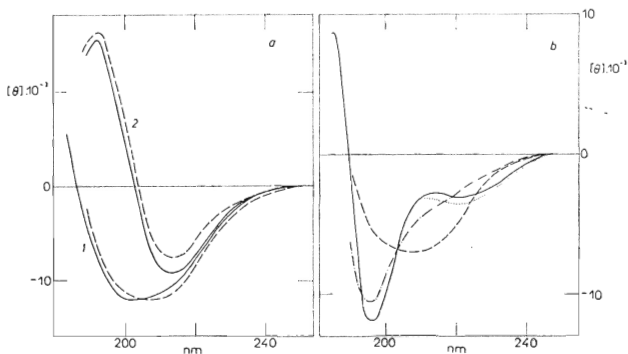

Fig. 4

Modified factor analysis. *a*) Comparison of the first subspectrum (*1*, ———) with the experimental curve of compound *I* in trifluoroethanol (-----) and the second subspectrum (*2*, ———) with the curve of *IV* in hexafluoropropanol (-----). *b*) Decomposition of the curve of compound *II* in acetonitrile into partial linear combinations of subspectra. (-----) first subspectrum, (-·-·-·) first + second subspectrum, (······) first + second + third subspectrum, (———) combination of first four subspectra which is indistinguishable from the experimental curve

Conformational energy calculations. The map for the tert-leucine derivative *IV* exhibits global minimum at $\phi = -130°$, $\psi = +134°$, *i.e.* in the region of extend (C_5) conformation. The second minimum, the relative energy of which is about 17 kJ . . mol^{-1}, is located at $\phi = -105°$, $\psi = -27°$ (*i.e.* in the region of distorted α-helical conformation). The position of global minimum is not in contrast to the position calculated by Delaney and coworkers ($\psi = +100°$)[14].

DISCUSSION

The first decision that has to be made is the determination of the minimal number of significantly populated conformers which are capable to describe chiroptical properties of the given set as a whole. The base for this decision can be extracted from the results of modified factor analysis. In agreement with our previous results[16] the contribution of the first subspectrum represents a basic structural property which is more or less present in all members of the set. In our set this property can be identified with the appearance of averaged mixture of conformers the actual composition of which is adapted to particular experimental situation mainly by the contribution of second subspectrum. Since the shape of second subspectrum can be associated
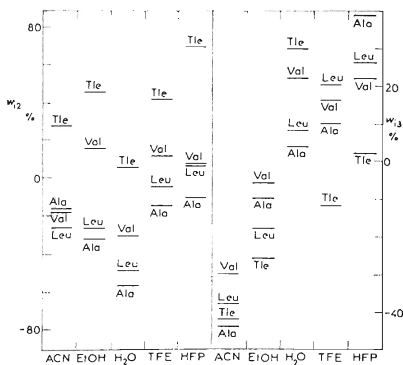

Fɪɢ. 5

Weights of the second (w_{i2}) and third (w_{i3}) subspectrum. The symbols Ala, Leu, Val and Tle denote the respective compounds *I–IV*. ACN acetonitrile, TFE trifluoroethanol, HFP hexafluoropropanol

with the spectral type *A* (when operating with positive weight coefficient) or with type *C* (with negative weight coefficient) and since both these spectral types can be identified with distinct members of the set we can infer that the contribution represented by second subspectrum reflects the weighted participation of two distinct conformers characterized by nearly enantiomorphous CD curves. This assignment is supported by the fact that we can locate experimental conditions under which the *A* and *C* curves behave like spectra of pure conformers (independence on temperature of the spectrum of compound *IV* in methanol–ethanol mixture and the general trend to change *B* spectra of compounds *I* and *II* towards *C* when the temperature is lowered). As further follows from the orthogonality of subspectra at least one additional conformer (or family of closely related conformers) has to be present in the basic mixture of conformers (represented by the first subspectrum). Chiroptical manifestation of this additional chiral entity must differ from the above two conformers at least in relative (generally low) intensities of $n-\pi^*$ and $\pi-\pi^*$ bands or possibly even in relative sign pattern. In accord with weights of the first subspectrum, which are always positive, the additional conformer should be present at least to a minor extent in all members of the set. It seems sufficient to consider only the contributions of first two subspectra since the third and fourth subspectra describe the respective effect of solvent and potential side chain asymmetry within the given conformational type. In conclusion, at least three different conformers play a significant role in modelling CD spectra of compounds *I−IV* in solvents ranging from acetonitrile to hexafluoropropanol.

The coupling constants $^3J_{NH,C^\alpha H}$ exhibit reasonable agreement with the above analysis. The systematic decrease of about $1-1·5$ Hz when going from acetonitrile to water indicates that the tendency to form CD spectra of the type *C* involves a conformational change characterized by a shift in Φ angle values. On the other hand, the almost non-changing values of $^3J_{NH,C^\alpha H}$ in acetonitrile and hexafluoro-

TABLE VI
Solvent and concentration (mol l^{-1}) dependence of $^3J_{NH,C^\alpha H}$ (Hz) in compounds *I−IV*

Compound	C^2HCl_3 0·2	C^2H_3CN 0·2	C^2H_3CN 0·004	H_2O	$(CF_3)_2CHOH$ 0·2
I	7·85	7·1	6·9	6·0[a]	7·2[b]
II	8·5	8·2	8·2	6·4	7·9
III	9·0	8·6	8·6	7·5[c]	8·4
IV	9·6	9·1	9·0	7·9[c]	9·5

[a] Ref.[8]; [b] value determined in CF_3CH_2OH; [c] ref.[14].

propanol suggest that the conformer characterized by the A spectrum and the third conformer the chiroptical manifestation of which is hidden in the shape of the first subspectrum possess similar ϕ angles. Consequently, their individuality is concentrated to the changes of ψ.

In the next step we can try to define those particular conformational states which are responsible for the A and C CD spectra of compounds $I-IV$. Since the spectral type A appears in strongly polar protic solvents it should refer to a conformation having carbonyl oxygen atoms accessible to the attack of solvent. In the case of the most typical example of the tert-leucine derivative IV in hexafluoropropanol, trifluoroethanol or ethanol, the ϕ angle value is constrained to about $-120°$ (see very high values of $^3J_{NH,C^{\alpha}H}$ in hexafluoropropanol, Table VI, and methanol, ref.[14]). There are two allowed conformations (conformational energy calculations of the compound IV) characterized by this value — extended C_5 and distorted α-helical α_R conformation. From these only the latter can exhibit, according to calculations of optical activity[24,25], the observed pattern of rather intense CD bands. The loss in energy which is caused by the fact that the α_R conformation lies about 17 kJ mol^{-1} above global minimum (for the compound IV) can be compensated by good accessibility of both oxygen atoms to hydrogen bonding with even those solvents whose steric demands are considerable (hexafluoropropanol). In fact, steric bulkiness when present either in the dipeptide (tert-leucine derivative) or in the solvent (hexafluoropropanol) supports strongly the appearance of the spectral type A. Hence, this spectral type can be ascribed to α_R conformation. In principle, our assignment is in agreement with results of the very detailed spectroscopic and computational study of alanine (I) and proline (V) dipeptides by Madison and Kopple[8]. However, they did not observe pure type A spectra since population of the α_R conformation in acetonitrile solution of I cannot be large. The proline dipeptide V, which was used as model in the mentioned study, does not exhibit A spectra at any conditions (see Tables $I-III$ and ref.[8]) confirming thus the importance of a change of the ϕ angle.

The assignments reported up to this date which can be related to the spectral type C operate with polyproline II conformation (P_I: $\phi \sim -80°$, $\psi \sim +150°$)[8,11,18]. It has been argued, on the basis of optical activity calculations, that P_{II} regions is the only one which can exhibit the positive $n-\pi^*$ Cotton effect if we exclude left handed α-helical conformation. According to Madison and Kopple[8], P_{II} conformation is present in solutions of compounds I and VI in acetonitrile, dimethyl sulfoxide and water. The argumentation is based on close parallelity of CD spectra of both compounds together with ^{13}C chemical shifts and ^1H nuclear Overhauser effects of compound V although NMR data do not exclude other explanations. Detailed analysis of our CD data reveals that the assignment of P_{II} conformation to spectral type C is not fully satisfactory. The positive $n-\pi^*$ band, when appearing in real spectra of compounds $I-IV$, is invariably accompanied by the negative $\pi-\pi^*$ band of much higher intensity. In several situations (e.g. compounds I in water) the positive $n-\pi^*$

band is absent or exhibits very low intensity but the intense $\pi-\pi^*$ band remains present. Consequently, the negative sign of $\pi-\pi^*$ band seems highly significant for the relevant conformation. This observation contradicts current assignment, since optical activity calculations predict the positive sign of $\pi-\pi^*$ transition for the P_{II} conformation region. Such sign pattern was never observed either with compounds $I-IV$ or with compound V. Although semiempirical calculations of optical activity are probably not very accurate, the reversal of sign in the middle of a rather well defined zone (see Fig. 9c in ref.[25]) can hardly be expected. An alternative explanation can be deduced from the inspection of calculated maps of optical rotatory strengths[24,25]. There is a small enclave at $\phi \sim -50°$ to $-60°$ and $\psi \sim -20°-0°$ for which the desired sign pattern (negative $\pi-\pi^*$ rotatory strength and small or positive $n-\pi^*$ rotatory strength) was calculated. This conformation, corresponding to 3_{10}-helical arrangement of polypeptide chains, lies at the edge of conformational space which is allowed for Ala and Leu residues. It can be accessible for flexible compounds I and II in solvents of considerable solvation ability whose molecules are not bulky (acetonitrile, water). The 3_{10} conformation cannot be distinguished from P_{II} conformation on the basis of $C^\beta-C^\gamma$ chemical shift differences exhibited by the proline derivative V due to dual meaning of these values (cf. Fig. 4c in ref.[8]). The proposed ϕ value of $-50°$ to $-60°$ corresponds well with lower values of $J_{NH,C^\alpha H}$ observed with solutions of compounds $I-IV$ in water. It is known that larger segments of 3_{10} helices are rather rare due to unfavourable side chain packing, but shorter segments appear at N- and C-terminals of polypeptide chains[26,27]. Thus it is possible that a molecule of dipeptide can adopt a conformation of 3_{10} type in situations for which the C spectrum is displayed. Notably, this conformation has been calculated as the energetically most favourable for α-aminoisobutyric acid residue if the arrangement of bonds on the C^α atom was non-symmetrical[28].

Assignment of the third conformational type can be regarded as an indication only. The assumed low intensity (almost featureless) CD spectra might conform well to a conformation of extended type (C_5, ϕ $-120°$ to $-150°$, ψ $+150°$), since this region of allowed conformational space is crossed by nodal planes for both the $n-\pi^*$ and $\pi-\pi^*$ transitions[24,25]. In accord, the low intensity CD spectrum of IV in acetonitrile is not dependent on temperature and the associated value of $J_{NH,C^\alpha H}$ correspond to ϕ of about $-120°$. Similar CD spectra are exhibited by compounds III and IV in cyclohexane. Moreover, the lowest energy conformation of IV is of extended type and, as follows from calculations of Lewis and coworkers[29], C_5 conformation is considerably stabilized by librational entropy, at least in flexible dipeptides I and II. Of course, we cannot exclude minor but significant population of C_7^{eq} conformer ($\phi \sim -80°$, $\psi \sim +80°$) in acetonitrile solutions of I (see[8]) and II.

The conformational distribution of compounds $I-IV$ in cyclohexane seems slightly different due to very low polarity of this solvent. Our conclusions on conformation of compounds I and II do not differ significantly from the results achieved previously

with compound I (ref.[3,8]). The lowest energy C_7^{eq} conformer dominates the shape of CD spectra if self-association is avoided. The assignment is based on the pattern of CD bands (intense negative $n - \pi^*$ band together with the almost negligible $\pi - \pi^*$ band), as calculated by Bayley and coworkers[24]. Further support is provided by IR spectroscopic results achieved with compounds I and II in tetrachloromethane and tetrachloroethylene[15]. Low intensity CD spectra, which are characteristic for the less flexible compounds III and IV, are ascribed to a different type of backbone conformation. According to conformational energy calculations on the compound Ie and in agreement with IR spectroscopic investigation[15] we can suggest that th$V$ arevailing conformer is C_5. The observed rather plain CD spectra conform well to the proposal. The fact that C_7^{eq} and C_5 conformations persist in cyclohexane solutions at elevated temperature is in accord with their supposed stabilization by entropy.

The following survey of conformations can be suggested. In low polarity solvents (cyclohexane, and also n-heptane, tetrachloromethane, chloroform, 1,4-dioxane, see refs[3,8]) C_7^{eq} conformer is characteristic for compounds I, II and V while the C_5 conformer dominates in solutions of compounds III and IV. The population of C_7^{eq} is rapidly diminished with increasing solvent polarity and conformers of α_R and 3_{10} type become important. The participation of C_5 conformation decreases less rapidly. The ratio of α_R and 3_{10} conformations is dependent on solvent and nature of the amino acid side chain. It cannot be related simply to solvent polarity since specific solute–solvent interactions, including hydrogen bonding, are important. The general effect of solvent polarity seems to be observable separately in the third subspectrum (Fig. 5b) the weights of which follow the solvent polarity scale deduced from solvent dependent CD spectra of rigid lactams[21,22]. The population of α_R conformation increases in strongly polar solvents interacting predominantly with carbonyl oxygen atoms and steric hindrance has a marked positive effect on the population of this conformer. Hence, the increased population of α_R conformation is dictated by a need of free accessibility of carbonyl oxygen atoms to solvation, in accord with findings of Madison and Kopple[8]. On the other hand, 3_{10} conformation is favourized in solutions of flexible dipeptides in solvents interacting either selectively with N—H grouping (acetonitrile) or with both solvation sites (water). The very high population of this conformer in aqueous solutions of I and II gives evidence on the exceptional character of water as a solvent, even for bulky tert-leucine derivative IV a tendency to assume 3_{10} conformation (a decrease of $J_{NH,C^\alpha H}$) is observed. The highest population found for aqueous solution of II can be caused by hydrophobic nature of leucine side chain. The conformational distribution of the diamide III approaches that of IV in fluorinated alcohols and that of diamides I and II in other solvents. The above characterization of population of α_R and 3_{10} conformers is quantified approximately by weights of the second subspectrum (Fig. 5a). High positive weights denote increased population of α_R conformation and, similarly, high negative weights can be associated with increased population of 3_{10} conformation.

APPENDIX

The modified factor analysis is based on the fact that CD spectra of a set of analogous compounds form the finite dimensional space H_p of single variable functions. The algorithm finds the projection operator which creates the basis of orthogonal CD curves (subspectra ϑ_i) from the set of experimental CD curves representing vectors in the H_p space. Simultaneously, the matrix of coefficients $[\alpha_{ij}]$ is formed which enables to express each experimental CD curve Y_j of the given set as a linear combination of subspectra ϑ_i

$$Y_j = \sum_{i=1}^{p} \alpha_{ij}\vartheta_i , \qquad (1)$$

where p is the dimension of input data set. The coefficients α_{ij} are normalized according to the relation

$$w_{ij} = \frac{\alpha_{ij}}{\sum\limits_{i=1}^{p} |\alpha_{ij}|} \cdot 100 \qquad (2)$$

affording weights w_{ij}. The eigenvalue λ_i specifies the amount of information described by the respective subspectrum ϑ_i. The subspectrum characterized by the largest eigenvalue (λ_1) reflects basic characteristics which are common to all CD curves within the set. The number of subspectra necessary for reproducing the input CD curves within experimental error is usually much lower than p and the drop of eigenvalues λ_i together with graphical comparison of calculated and experimental curves are used for its determination.

The interpretation in structural terms is carried out using the notion of chiral entity. Chiral entity describes a situation in which a certain spatial arrangement of atoms connected by chemical interactions influences chiroptical properties. Hence, chiral entity can be regarded as structural equivalent of physical forces determining chiroptical properties of the investigated molecule or species. Two enantiomorphous arrangements are represented by one chiral entity. We assume that, in the given set, there is always only a limited number of chiral entities. Properties of chiral entities can be identified with vectors from the space of physical properties. It is always possible to find such linear combination of properties of chiral entities which corresponds to a vector from the orthogonal subspaces. Hence, we assume that each subspectrum ϑ_i corresponds to one independent combination of properties of chiral entities. If chiroptical properties of some chiral entity are linearly independent on other chiral entities then we can identify the respective subspectrum with the given chiral entity. In other cases it is possible to estimate participation of other chiral entities in the corresponding linear combination.

The identification of chiral entities is performed by a comparison of calculated subspectra ϑ_i with experimental curves and by a correlation of weights w_{ij} with variable structural parameters of the set (e.g. structural modifications, changes of solvent etc.). We try to find the best correlation between weights and structural parameters of the particular members of the set. The sign of w_{ij} indicates either presence of chiral entity in one or both enantiomorphous forms or higher or lower participation of the given factor as compared with the average situation described by the preceeding subspectra.

REFERENCES

1. Efremov D. S., Senyavina L. B., Zheltova V. N., Ivanova A. G., Kostetskii P. V., Ivanov V. T., Popov E. M., Ovchinnikov Yu. A.: Khim. Prir. Soedin. 9, 322 (1973).

2. Ivanov V. T., Kostetskii P. V., Balashova T. A., Portnova S. L., Efremov E. S., Ovchinnikov Yu. A.: Khim. Prir. Soedin. 9, 339 (1973).
3. Ivanov V. T., Kostetskii P. V., Meshcheryakova E. A., Efremov E. S., Popov E. M., Ovchinnikov Yu. A.: Khim. Prir. Soedin. 9, 363 (1973).
4. Crippen G. M., Yang J. T.: J. Phys. Chem. 78, 1127 (1973).
5. Pullman B., Pullman A.: Advan. Protein Chem. 28, 347 (1974).
6. Zimmerman S. S., Pottle M. S., Némethy G., Scheraga H. A.: Macromolecules 10, 1 (1977).
7. Hodes Z. I., Némethy G., Scheraga H. A.: Biopolymers 18, 1565 (1979).
8. Madison V., Kopple K. D.: J. Amer. Chem. Soc. 102, 4855 (1980).
9. Efremov E. S., Kostetskii P. V., Ivanov V. T., Popov E. M., Ovchinnikov Yu. A.: Khim. Prir. Soedin. 9, 348 (1973).
10. Efremov E. S., Kostetskii P. V., Ivanov V. T., Popov E. M., Ovchinnikov Yu. A.: Khim. Prir. Soedin. 9, 354 (1973).
11. Burgess A. W., Scheraga H. A.: Biopolymers 12, 2177 (1973).
12. Johnson W. C., jr, Tinoco I., jr: J. Amer. Chem. Soc. 94, 4389 (1972).
13. Nielsen E. B., Schellman J. A.: Biopolymers 10, 1559 (1971).
14. Delaney N. G., Lasky J., Madison V. in the book: Peptides. Synthesis-Structure-Function. Proceedings of the Seventh American Peptide Symposium. (D. H. Rich, E. Gross, Eds), p. 303. Pierce Chemical Co., Rockford 1981.
15. Smoliková J., Pospíšek J., Bláha K.: This Journal 46, 772 (1981).
16. Pančoška P., Frič I., Bláha K.: This Journal 44, 1296 (1979).
17. Applewhite T. H., Niemann C.: J. Amer. Chem. Soc. 81, 2208 (1959).
18. Pospíšek J., Bláha K.: This Journal 42, 1069 (1977).
19. Kostetskii P. V., Tsetlin V. I., Ivanov V. T., Ovchinnikov Yu. A.: Khim. Prir. Soedin. 9, 319 (1973).
20. Nikiforovich G. V., Galaktionov S. G., Chipens G. I.: Izv. Akad. Nauk Latvian SSR 1979, 94.
21. Frič I., Maloň P., Tichý M., Bláha K.: This Journal 42, 678 (1977).
22. Bláha K., Maloň P., Tichý M., Frič I., Usha R., Ramakumar S., Venkatesan K.: This Journal 43, 3241 (1978).
23. Ivanov V. T. in the book: Peptides. Proceedings of the Fifth American Peptide Symposium, (M. Goodman, J. Meienhofer, Eds), p. 307. Halsted Press, New York 1977.
24. Bayley P. M., Nielsen E. B., Schellman J. A.: J. Phys. Chem. 73, 228 (1969).
25. Bayley P. M.: Progr. Biophys. Mol. Biol. 27, 1 (1973).
26. Donohue J.: Proc. Nat. Acad. Sci. U.S.A. 39, 470 (1953).
27. Hendrickson W. A., Love W. E., Karle I.: J. Mol. Biol. 74, 331 (1973).
28. Paterson Y., Rumsey S. M., Benedetti E., Némethy G., Scheraga H. A.: J. Amer. Chem. Soc. 103, 2947 (1981).
29. Lewis P. N., Momany F. A., Scheraga H. A.: Isr. J. Chem. 11, 121 (1973).

Translated by the author (P. M.).