Quantitative IR Spectrophotometry of Peptide Compounds in Water (H_2O) Solutions. I. Spectral Parameters of Amino Acid Residue Absorption Bands

S. YU. VENYAMINOV and N. N. KALNIN

Institute of Protein Research, Academy of Sciences of the USSR, Puschino, Moscow region, USSR

SYNOPSIS

Infrared spectra of the amino acid residues in H_2O solution have been obtained in the 1800–1400-cm^{-1} region. It has been established that amino acid residues of arginine, asparagine, glutamine, aspartic and glutamic acids, lysine, tyrosine, histidine, and phenylalanine have intensive absorption in this spectral region. Infrared spectra for a set of model compounds have been measured. On the basis of these data, spectral parameters of amino acid residue absorption bands have been determined.

INTRODUCTION

Application of ir spectroscopy to analysis of protein secondary structure is based on the sensitivity of peptide group absorption bands to the polypeptide chain conformation. This was demonstrated by theoretical calculations and experimentally.[1-4] Later a method of estimation of protein and polypeptide secondary structure on the basis of contour shape analysis of amide I band was suggested.[5-7] The method includes quantitative measurements of intensities and other spectral parameters of absorption bands of polypeptides in α-, β-, and random conformations,[6,7] and of amino acid residue side chains.[5] The measurements were made for D_2O solutions.

Another approach to this problem was suggested by Byler and Susi.[8] They assigned particular absorption bands in ir spectra of proteins to different secondary structures using Fourier self-deconvolution procedure.

However, these methods have some disadvantages. The major one is the necessity of complete deuterium exchange of labile protons in protein. This is difficult to achieve for some proteins. Thus, estimation of the protein secondary structure for water solutions is still a problem.

CCC 0006-3525/90/13-141243-15 $04.00
Biopolymers, Vol. 30, 1243–1257 (1990)

In 1967 Susi, Timasheff, and Stevens published ir spectra of proteins in H_2O solution.[9] Later, some measurements of ir spectra of various compounds in H_2O were been reported.[10-12] It has been demonstrated that measurements of ir spectra of water solution are possible, but this is a rather complicated procedure. The major difficulty is to achieve fine compensation of water absorption using cells about 10 μ thick.[13]

We assumed that the ir spectrum of any protein is an additive sum of peptide and amino acid side chains absorptions. The first task was to estimate the contribution of the amino acid side-chain absorption for various proteins. For D_2O this contribution is known to achieve for some proteins 30% of amide I integral intensity.[5] We expected approximately the same value for H_2O solutions. The amino acid residue absorption can affect the protein secondary structure estimation. On the other hand, the knowledge of amino acid residues absorption is useful for the study of structural transitions in proteins upon variation of pH, temperature, or upon binding of ligands and substrates.

To obtain the parameters of amino acid residue absorption bands, we examined water (H_2O) solutions of 20 native amino acids. We picked out from them the amino acids that have intensive absorption in the amide I,II spectral region. In most amino acid spectra, substantial overlapping between bands related to $-COO^-$, $-NH_3^+$, and R group is observed.

Therefore, it is preferable to use simpler model compounds to transfer the band parameters to more complicated molecules. On the other hand, spectral parameters of side group absorption bands depend to a certain extent on the adjacent part of the molecule. Therefore, compounds containing the polypeptide chain fragment (-CONH-CHR-CONH-) should be chosen as models. The N_α-acetyl methyl ester of L-amino acid (CH_3-CONH-CHR-$COOCH_3$) is a good model for a polypeptide chain fragment.[5] When the peptide absorption band of this compound was strongly overlapped by the R-group absorption, other models were studied.

EXPERIMENTAL

The L-amino acids were obtained from Calbiochem (USA), except for cysteine and oxyproline, which were obtained from Reanal (Hungary). Methyl esters and N_α-acetyl methyl esters of amino acids were synthesized by Dr. K. H. Zikherman. Purity of the obtained preparations was checked by chromatography, and the melting temperature and C,N,H-content determined on a C,N,H-analyzer model 240B (Perkin Elmer, USA). All samples were in the crystalline form with the exception of N_α-acetyl methyl ester of glutamic acid. Acetic, propionic, butyric and valeric acids, amides of these acids, methylamine, butylamine, dodecylamine, cresol, imidazole, and toluene were recrystallized commercial preparations from Souzreaktiv (USSR).

The pH measurements were made with a pH meter model 262 (USSR) and combined electrode GK-2421C (Radiometer, Denmark). The pH of the solutions was adjusted by adding saturated KOH and HCl.

Concentrations of the low molecular weight compounds were determined by dry weight. Protein concentration was found by nitrogen analysis.[14] The concentrations were in the region of 5–30 mg/mL.

We recorded four spectra for every sample with two different concentrations and two cells. Thus, the resulting spectrum was the average of four curves.

Spectral measurements were made with the M-180 spectrophotometer (Perkin Elmer, USA) connected with the R-10 Videoton computer (Hungary). The spectral slit width was 3 cm^{-1}. Linearity of the ordinate scale and correspondence of the transmittance scale to the absorbance one were checked with revolving sector disks from LOMO (USSR). The precision of the spectrophotometer ordinate scale was better than 0.2%, and the wavenumber accuracy was 0.3 cm^{-1}. The monochromator and cell compartment were purged from H_2O vapors and CO_2.

The cells were thermostated during measurements at 25.0 ± 0.2°C.

Accurate compensation of water absorption was made for every measurement by repeated adjustment of the path length of the reference cell. Compensation was controlled in the 3700–3600-cm^{-1} region free of the sample absorption. All spectra were recorded in the region of 1800–1000 cm^{-1}. After baseline subtraction the absorption at 1800 cm^{-1} was admitted to be zero.

Demountable fluorite cells 6–12 μm thick were used.[13] The cell thickness was determined from interferentional patterns of empty cells and from the water absorption at 1643 cm^{-1}, using the value $A_{10\mu m}$ (1643 cm^{-1}) = 1.22. The accuracy was about 1%. Reproducibility of the cell thickness was also 1%.

Spectra were put into the computer memory with a data interval of 0.5 cm^{-1}. Further processing included transmittance-absorbance transformation, baseline correction, smoothing,[15] etc.

The complex spectrum was resolved into separate components with a specially designed seven-channel curve resolver similar to the Du Pont Curve Resolver 310. This device serves to generate and summarize seven different curves, each being the function $F(\nu) = f_G\, G(\nu) + (1 - f_G)\, L(\nu)$, where $G(\nu)$ is the Gaussian function, $L(\nu)$ is the Lorentzian function, ν is the wavenumber, and f_G is the shape parameter varying from 0 to 1 with the step 0.1. To solve the deconvolution problem correctly, we used the data for the analogous compounds in D_2O,[5] the second and the fourth derivative spectra.[16]

RESULTS AND DISCUSSION

We measured ir spectra of all 20 amino acids and extracted those with the R-group absorption intensity more than 40 L mol^{-1} cm^{-1} in the 1800–1480-cm^{-1} region. These were Asn, Gln, Asp, Glu, Arg, Lys, Tyr, Phe, and His. We shall call them "absorptive" amino acids.

The results are independent of the nature of counterion in the solution.[5] This was checked for Asn, Gln, and valeric acid by comparing their spectra recorded in the presence of different ions: Cl^-, SO_4^{2-}, PO_4^{3-}, K^+, Na^+, and Mg^{2+}. No change of the ir spectra was detected.

Spectral Characteristics of Hydration

Sample molecules in the solution deform the water H-bond network and change the water spectrum.[17]

A certain amount of bound water molecules are present in the sample cell, but absent from the reference one. Infrared spectra of NaCl solutions at 0.01, 0.1, and 1 *M* concentrations are represented in Figure 1a. To compensate, the H_2O absorption we took the reference cell with pure water. At NaCl concentration of 0.1 *M* and higher the characteristic curve appears in the region of 1700–1600 cm^{-1}, which represents the difference between the spectral parameters of the free and bound water molecules.

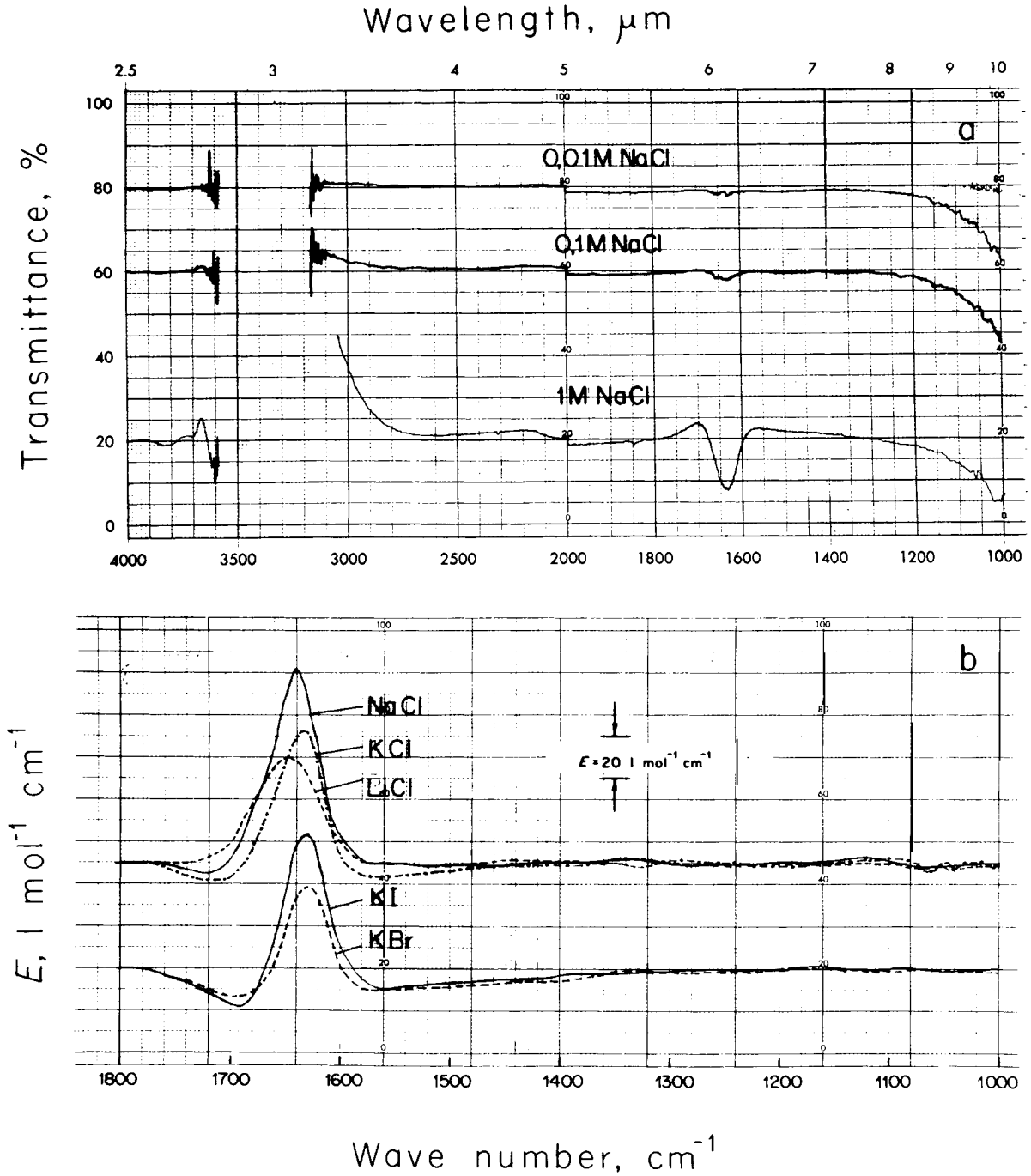

Figure 1. Bound water in the IR spectra. (a) Spectra of NaCl solutions at 0.01, 0.1, and 1 *M* concentrations. The curves are shifted along the transmittance scale. (b) Spectra of NaCl, KCl, LiCl, KI, and KBr in *E* (L mol^{-1} cm^{-1}) units. The spectra are shifted along the *E* scale.

Analogous curves for various salts are shown in Figure 1b. Similar curves were obtained for ir spectra of amino acids and model compounds.

Here we can describe the hydration effect in the spectra by an absorption band with the parameters

$$\nu_0 = 1652/E_0 = 60/\Delta\nu_{1/2} = 50/f_G = 0.5$$

Here and further the band parameters are represented in the following sequence: ν_0, band position, cm^{-1}; E_0, intensity at the maximum, l mol^{-1} cm^{-1}; $\Delta\nu_{1/2}$, band half width, cm^{-1}; f_G, band shape parameter.

α-Carboxyl and α-Amino Groups Absorption

The ionized carboxyl group in D_2O has a strong absorption band with the maximum at 1590 cm^{-1}.[5] In H_2O the ionized amino group has two bands corresponding to symmetrical and asymmetrical vibrations of the $-NH_3^+$ group in the 1800–1440-cm^{-1} region (Figure 2a). Carboxyl and amino groups of amino acids are separated by one C_α atom and affect each other through the $^-OOC\text{-}C_\alpha\text{-}NH_3^+$ chain. The band parameters of these groups differ from those for the carbonic acids and amines. Furthermore, titration of amino acids (Figures 2b and 2c) leads to a change of absorption bands of both $-COO^-$ and $-NH_3^+$ groups. This apparently results from electron density redistribution in the molecule, which leads to a change of the transition dipole of both absorptive groups.

As for the interaction of $-COO^-$ and $-NH_3^+$ groups with the side chains of amino acids, it depends on the charge of the R group and the molecule configuration. Below we shall demonstrate that intramolecular interaction is the strongest for asparagine and aspartic acid.

There are two additional absorption bands of the α-carboxyl group, which are absent in the figures and tables:

$C{=}O_{str}$ (COOH)

$$\nu_0 = 1260/E_0 = 170/\Delta\nu_{1/2} = 42/f_G = 0.5$$

COO^-_{sym}

$$\nu_0 = 1412/E_0 = 200/\Delta\nu_{1/2} = 24/f_G = 0.2$$

Absorption of Asparagine and Glutamine Residues

These amino acids have an amide group $-CONH_2$, which gives two absorption bands in the considered spectral region. Amides of acetic, propionic, butyric, and valeric acids were chosen as models of R groups of Asn and Gln. The spectra of these compounds (Figures 3a and 3b) have the absorption bands

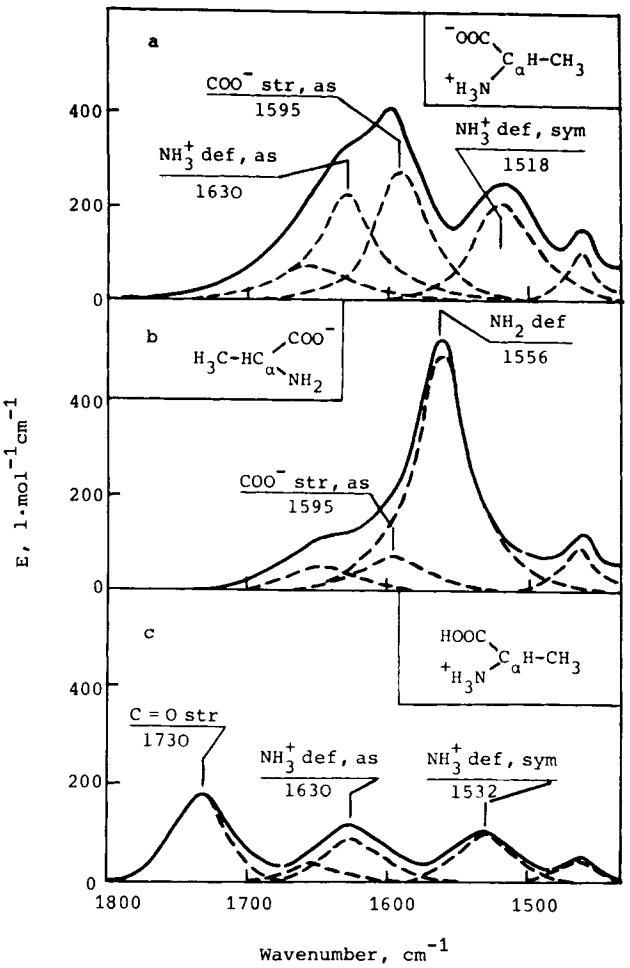

Figure 2. Absorption of α-carboxyl and α-amino group of Ala at various pH. (a) Ala, pH 6.1. (b) Ala, pH 11.5. (c) Ala, pH 1.0.

$$C{=}O_{str} - \nu_0 = 1655/E_0 = 320/\Delta\nu_{1/2} = 32/f_G = 0.7$$

$$NH_{2def} - \nu_0 = 1600/E_0 = 200/\Delta\nu_{1/2} = 43/f_G = 0.0$$

The spectra of Asn and Gln (Figures 3c and 3d) are very complex. Despite the strong band overlapping, it can be seen from the differential spectra (Asn) − (Gln) (Figure 3e) that the $C{=}O_{str}$ band of Asn is shifted from that of Gln by about 10 cm^{-1}. The frequency shift appears due to intramolecular interaction in Asn. At the same time the experimental spectrum of glutamine can be satisfactorily

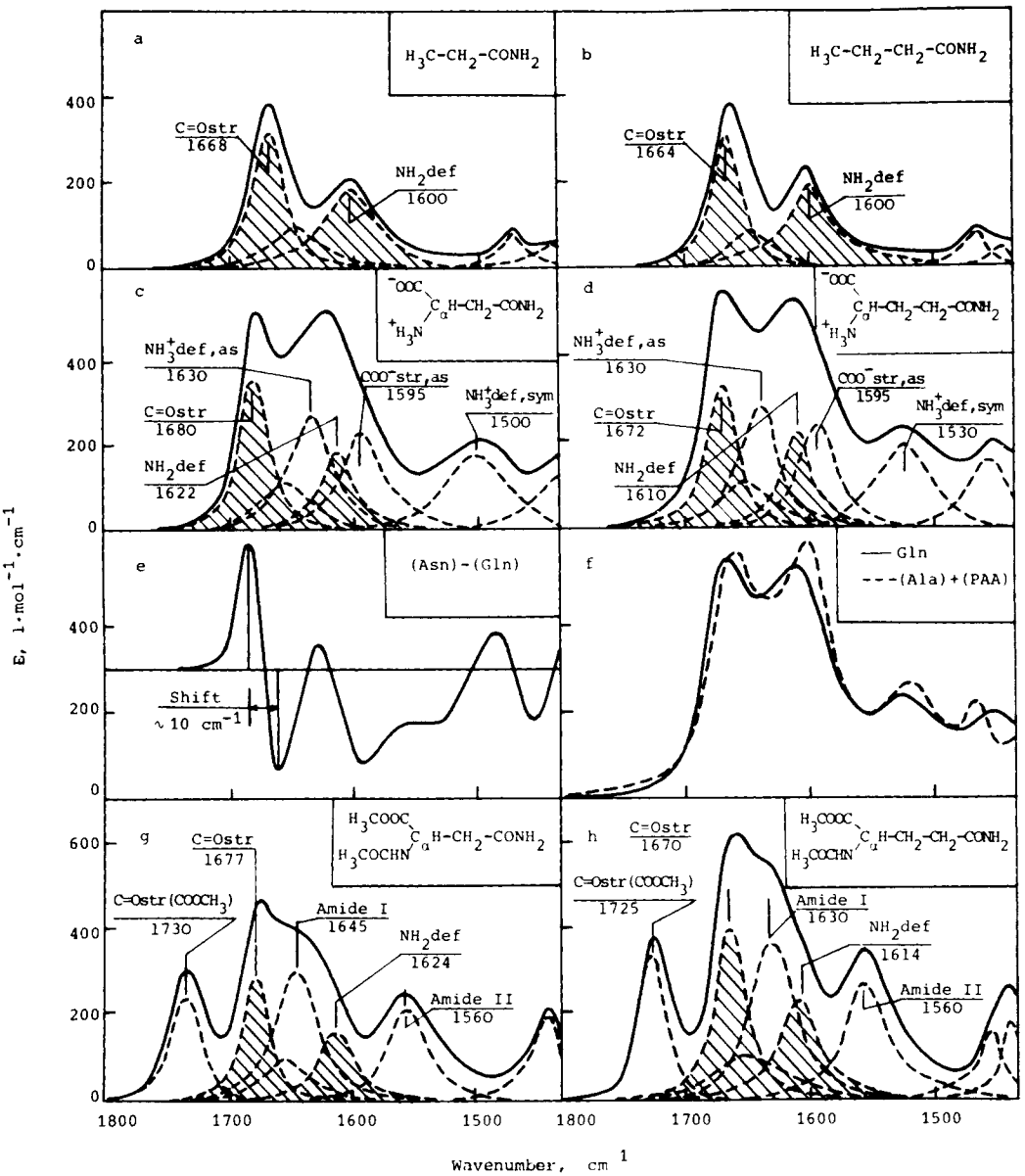

Figure 3. Absorption of Asn and Gln residues (dashed area). (a) Propionic acid amide, pH 6.5. (b) Butyric acid amide, pH 6.5. (c) Asn, pH 6.5. (d) Gln, pH 6.5. (e) Difference (Asp) − (Gln), pH 6.5. (f) Gln at pH 6.5 (solid line) and the sum (Ala) + (propionic acid amide), pH 6.5. (g) N_α-acetyl methyl ester of Asn, pH 6.5. (h) N_α-acetyl methyl ester of Gln, pH 6.5.

fitted by the sum (Ala) + (propionic acid amide) (Figure 3f). It seems that the presence of two $-CH_2-$ groups between the amide group of Gln and the C_α atom is quite enough to localize the R-group vibrations.

N_α-acetyl methyl esters of Asn and Gln display strong overlapping of the main- and side-chain amide bands (Figures 3g and 3h). We took the parameters of the peptide absorption band from the spectrum of N_α-acetyl methyl ester of Ala. Besides the model compounds mentioned above, we measured the spectra of Asn and Gln at acidic and alcaline pH (spectra are not shown). It can be seen from Figures 3c and 3d that upon deionization of the $-COO^-$ group the band at 1595 cm^{-1} disappears and the spectrum becomes simpler. In the other case deionization of the $-NH_3^+$ group also leads to simplification of the spectra.

Absorption of Aspartic and Glutamic Acid Residues

Side chains of these amino acids at neutral pH have an absorption band belonging to the $-COO^-_{str,as}$ vibrational mode of the carboxyl group. The band position differs from that of the α-carboxyl group discussed above. Acetic, propionic, butyric, and valeric acids were chosen as the simplest models (Figures 4a and 4b). According to the data for D_2O,[5] the band position of Asp is shifted by 15–20 cm^{-1} to higher wave numbers as compared with Glu. For H_2O we obtained the same result (Figures 4c and 4d). In addition, Asp has an anomalous position of the

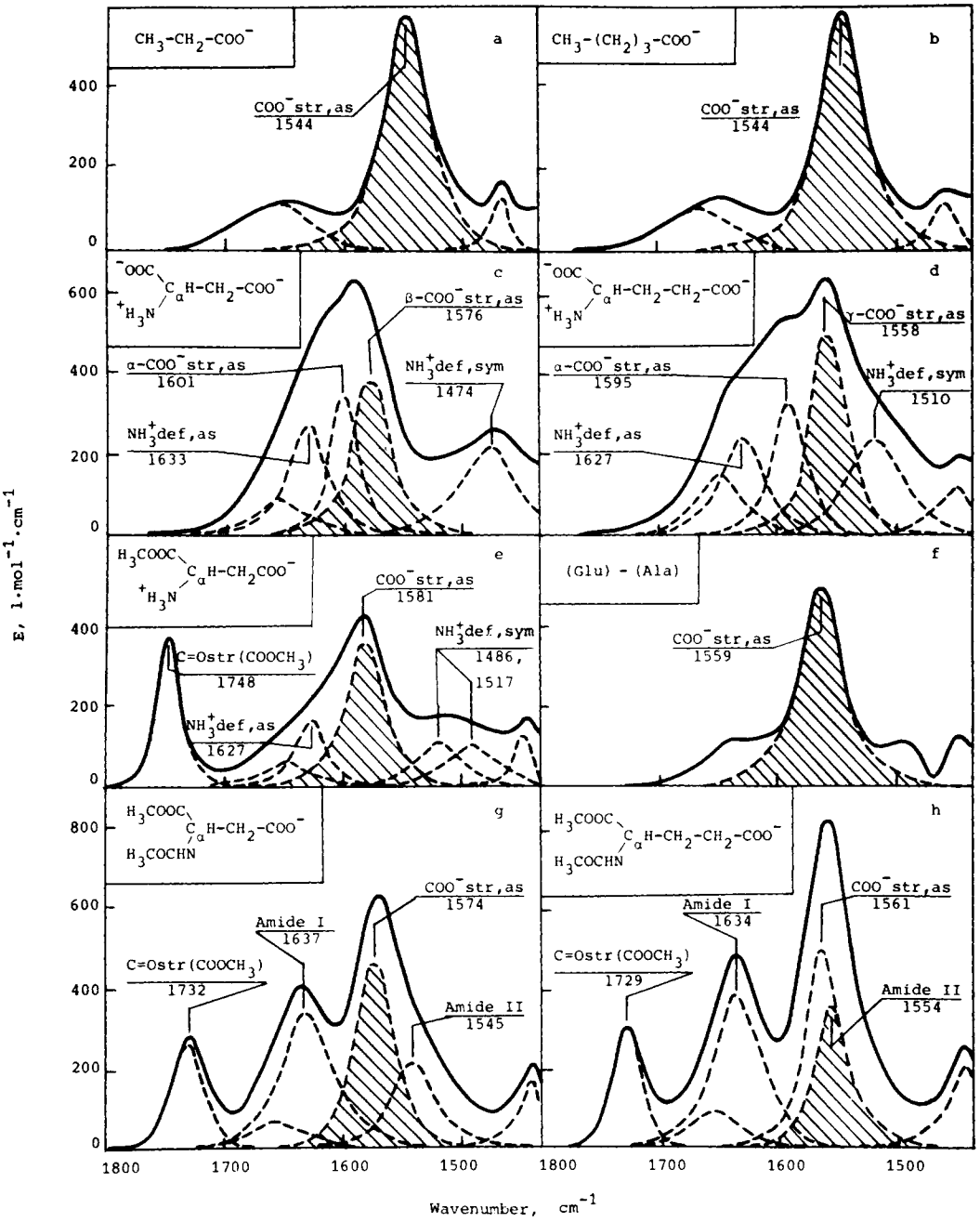

Figure 4. Absorption of Asp and Glu residues in ionized form (dashed area). (a) Propionic acid, pH 6.5. (b) Valeric acid, pH 6.5. (c) Asp, pH 6.5. (d) Glu, pH 7.0. (e) Asp-OMe, pH 5.5. (f) Difference (Glu) − (Ala), pH 7.0. (g) N_α-acet-Asp-OMe, pH 7.0. (h) N_α-acet-Glu-OMe, pH 6.1.

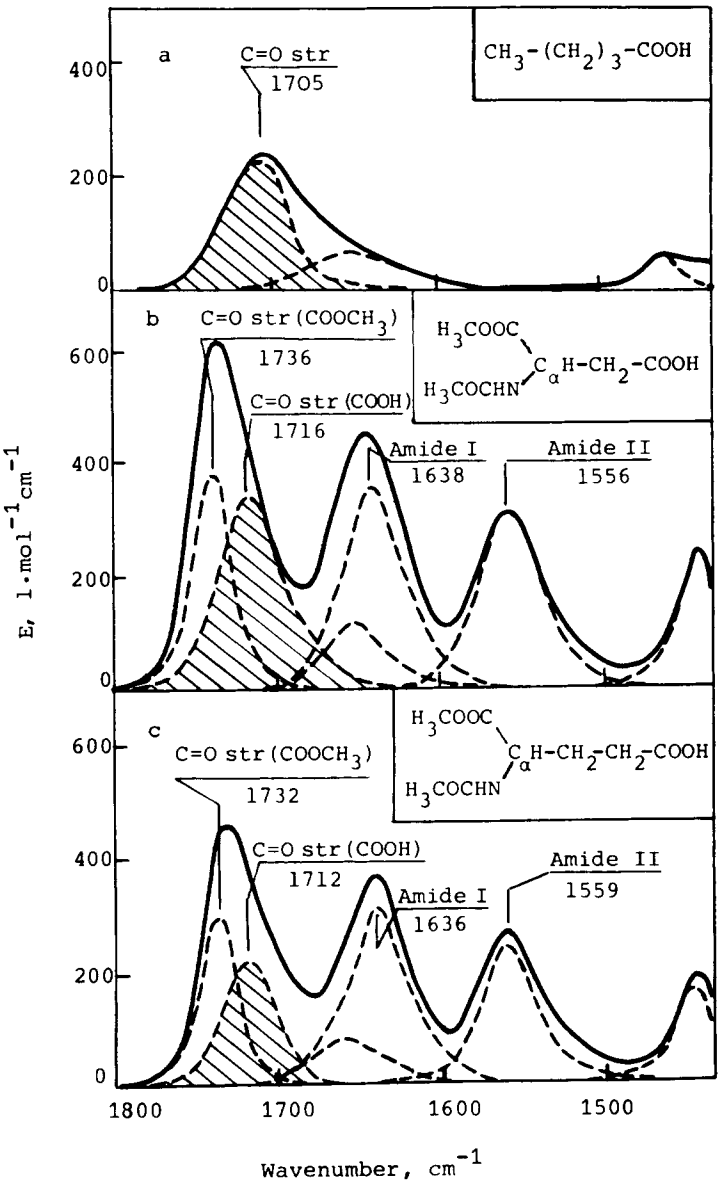

Figure 5. Absorption of Asp and Glu residues in deionized form (dashed area). (a) Valeric acid, pH 2.8. (b) N_α-acet-Asp-OMe, pH 2.5. (c) N_α-acet-Glu-OMe, pH 2.7.

$NH^+_{3\,def,sym}$ band: $\nu_0 = 1474\ cm^{-1}$ (normal $\nu_0 = 1500$–$1530\ cm^{-1}$). Apparently the β-COO^- group of Asp undergoes strong intramolecular interaction in contrast with Glu.

To eliminate overlapping between α-$COO^-_{str,as}$ ($\nu_0 = 1601\ cm^{-1}$) and β-$COO^-_{str,as}$ ($\nu_0 = 1576\ cm^{-1}$), we measured the spectra of methyl ester and N_α-acetyl methyl ester of Asp (Figures 4e and 4g). The resulting spectral parameters of the Asp R-group absorption were derived from the latter.

Vibrations of the γ-carboxyl group of Glu are localized from the end-group vibrations of the amino acid. Therefore we used the difference (Glu) − (Ala) (Figure 4f). Due to strong overlapping between $COO^-_{str,as}$ and amide II bands in the spectrum of N_α-acetyl methyl ester of Glu (Figure 4h), we derived the R-group absorption band parameters from the (Glu) − (Ala) (Figure 4f) and Glu-methyl ester (not shown) spectra. Spectrum 4h was used only to check additivity.

In addition to the absorption bands mentioned above, the R groups of Asp and Glu in ionized form have the bands corresponding to $-COO^-_{str,sym}$ vibrational mode with the parameters

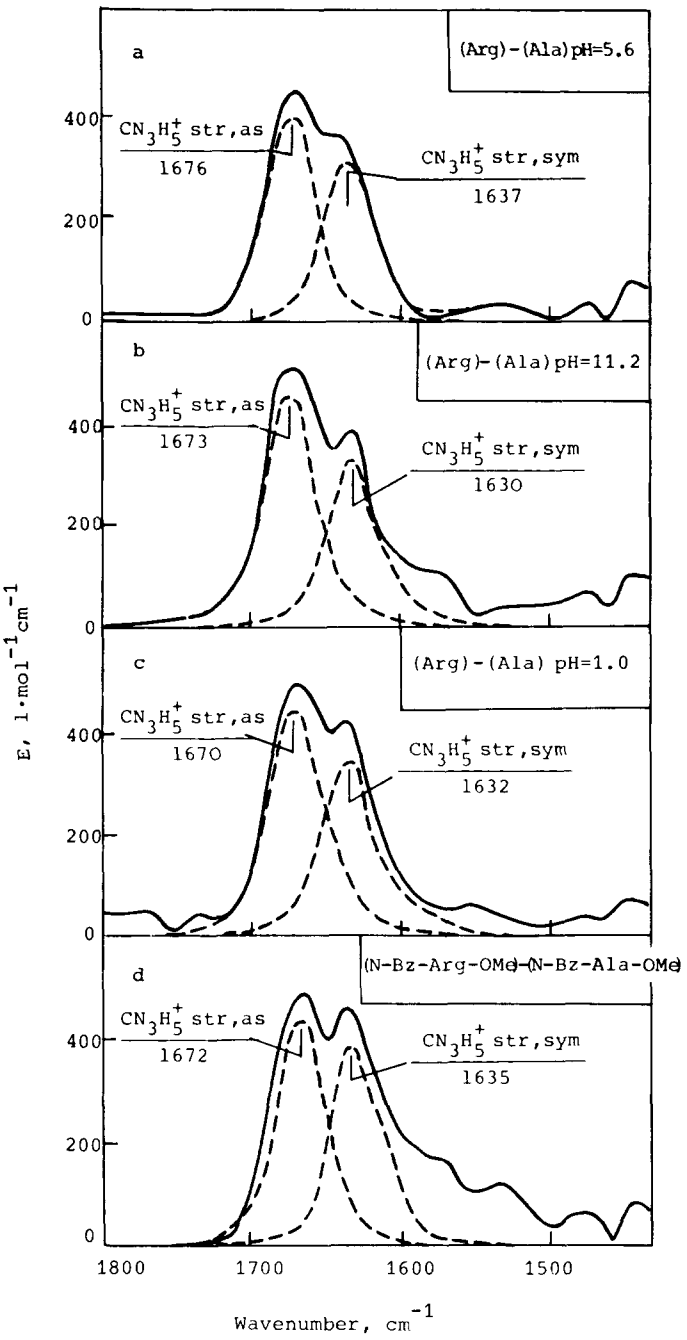

Figure 6. Absorption of Arg residue. Differential spectra: (a) (Arg) − (Ala), pH 5.6; (b) (Arg) − (Ala), pH 11.2; (c) (Arg) − (Ala), pH 1.0; (d) (N_α-Bz-Arg-OMe) − (N_α-Bz-Ala-OMe), pH 7.5.

Asp $\nu_0 = 1402/E_0 = 256/\Delta\nu_{1/2} = 28/f_G = 0.3$

Glu $\nu_0 = 1404/E_0 = 316/\Delta\nu_{1/2} = 28/f_G = 0.4$

Asp and Glu in deionized form exhibit an absorption band pertaining to $C{=}O_{str}$ vibrations of the -COOH group. We measured at acidic pH the spectra of carbonic acids (an example is presented in Figure 5a), of amino acids (not shown), and of N_α-acetyl methyl esters of Asp and Glu (Figure 5b and 5c). It should be noted that this band has very constant parameters, which are practically the same for H_2O and D_2O solutions.

The R groups of Asp and Glu at acidic pH have

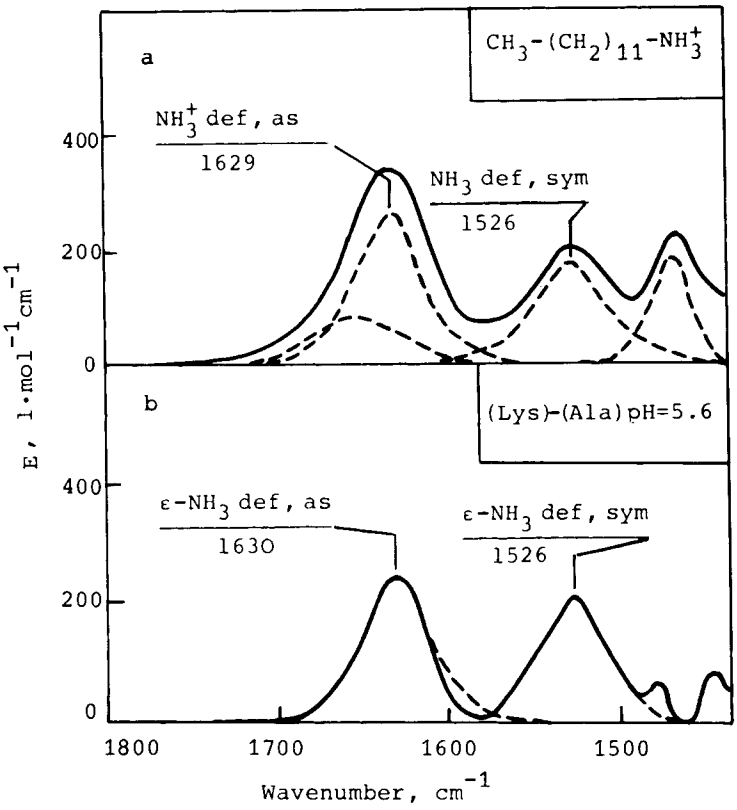

Figure 7. Absorption of Lys residue in ionized form. (a) Dodecylamine, pH 5.0. (b) Difference (Lys) − (Ala), pH 5.6.

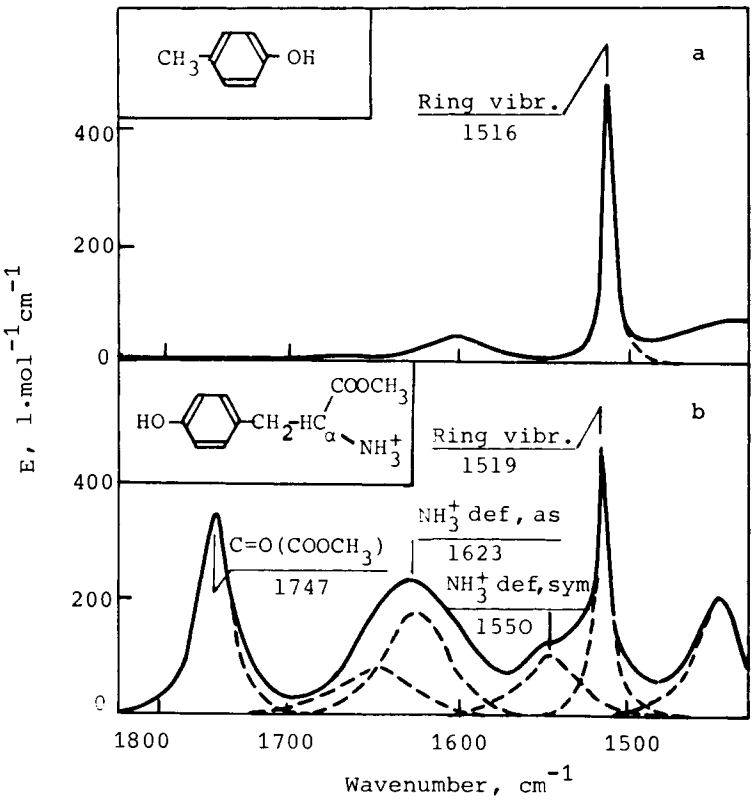

Figure 8. Absorption of Tyr residue in deionized form. (a) Cresol, pH 5.8. (b) Tyr-OMe, pH 6.2.

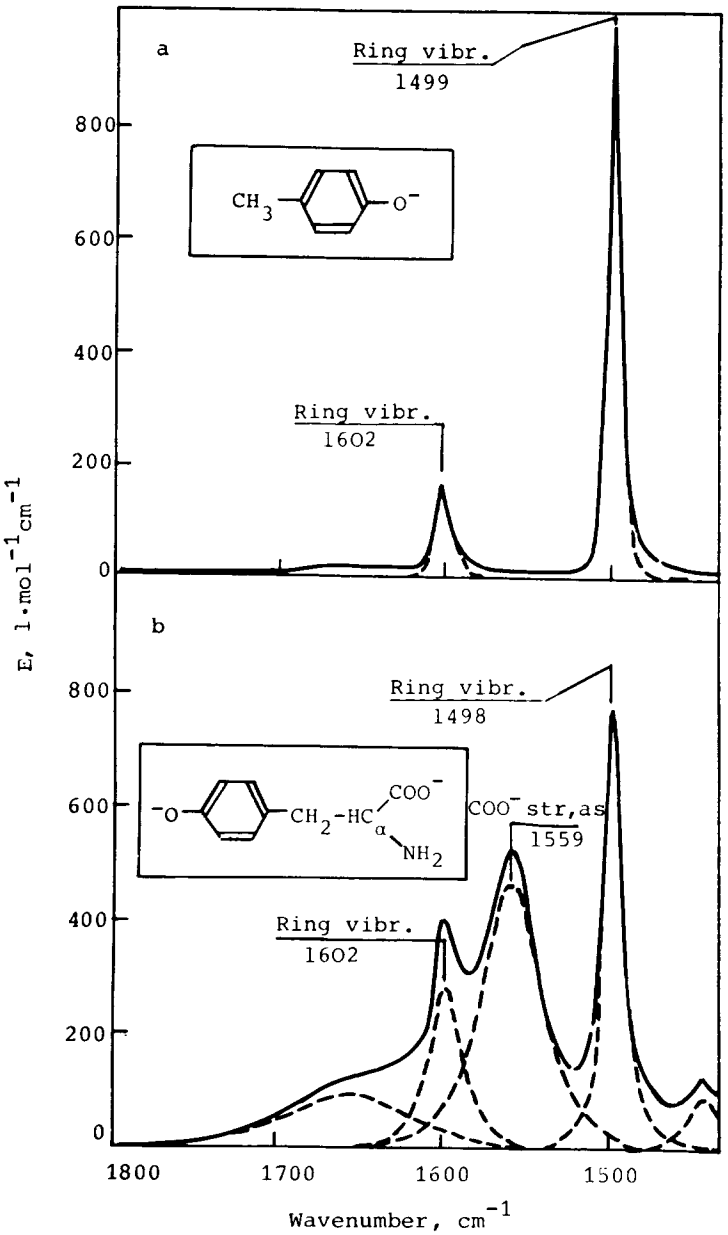

Figure 9. Absorption of Tyr residue in ionized form. (a) Cresol, pH 12.2. (b) Tyr, pH 12.2.

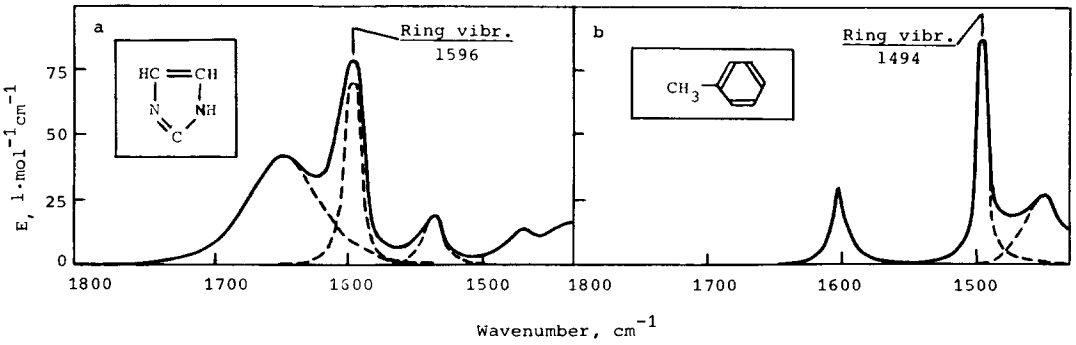

Figure 10. (a) Spectrum of imidazole, pH 6.8. (b) Spectrum of toluene, pH 6.8.

Table I IR Absorption Band Spectral Parameters of Amino Acid Residue Side-Chain Groups in H_2O for the 1800–1440-cm^{-1} Range

	Arg		Lys		Asn		Gln		Tyr (Deionized)	Tyr (Ionized)	
Band Parameters[a]	$-CN_3H_5^+$ str, as	$-CN_3H_5^+$ str, sym	$-NH_3^+$ def, as	$-NH_3^+$ def, sym	-C=O str	$-NH_2$ def	-C=O str	$-NH_2$ def	Ring Vibration	Ring Vibration	Ring Vibration
ν_0, cm^{-1}	1673 ± 3	1633 ± 3	1629 ± 1	1526 ± 3	1678 ± 3	1622 ± 2	1670 ± 4	1610 ± 4	1518 ± 1	1602 ± 2	1498 ± 1
E_0, L mol^{-1} cm^{-1}	420 ± 40	300 ± 20	130 ± 10	100 ± 10	310 ± 20	160 ± 15	360 ± 20	220 ± 20	430 ± 20	160 ± 20	700 ± 10
$\Delta\nu_{1/2}$, cm^{-1}	40 ± 2	40 ± 4	46 ± 2	48 ± 2	32 ± 2	44 ± 2	32 ± 1	44 ± 2	8 ± 1	14 ± 2	10 ± 1
f_G	0.9	0.5	0.5	0.7	0.8	0.0	0.8	0.0	0.5	0.4	0.0
B 10^{-4}, L mol^{-1} cm^{-2}	4.3 ± 0.6	3.6 ± 0.6	1.8 ± 0.2	1.3 ± 0.3	2.7 ± 0.3	2.5 ± 0.3	3.1 ± 0.3	3.5 ± 0.3	1.0 ± 0.1	0.7 ± 0.1	2.5 ± 0.2

	Asp		Glu		His	Phe	$-COO^-$	-COOH	$-NH_3^+$		$-NH_2$
	$-COO^-$ str, as	-COOH str	$-COO^-$ str, as	-COOH str	Ring Vibration	Ring Vibration	$-COO^-$ str, as	-COOH str	$-NH_3^+$ def, as	$-NH_3^+$ def, sym	$-NH_2$ def
ν_0, cm^{-1}	1574 ± 2	1716 ± 2	1560 ± 3	1712 ± 2	1596 ± 1	1494 ± 1	1598 ± 2	1740 ± 5	1631 ± 3	1515 ± 5	1560 ± 2
E_0, L mol^{-1} cm^{-1}	380 ± 20	280 ± 20	470 ± 30	220 ± 10	70 ± 10	80 ± 10	240 ± 15	170 ± 15	210 ± 10	200 ± 30	450 ± 30
$\Delta\nu_{1/2}$, cm^{-1}	44 ± 2	50 ± 2	48 ± 2	56 ± 2	14 ± 1	6 ± 1	47 ± 5	50 ± 5	54 ± 7	60 ± 3	46 ± 2
f_G	0.3	0.6	0.4	0.6	0.4	0.2	0.4	1.0	0.2	0.0	0.0
B 10^{-4}, L mol^{-1} cm^{-2}	5.5 ± 0.5	4.1 ± 0.3	7.1 ± 0.7	3.6 ± 0.4	0.3 ± 0.1	0.2 ± 0.1	3.5 ± 0.4	2.1 ± 0.4	3.8 ± 0.7	4.3 ± 0.6	7.5 ± 0.9

[a] ν_0: Band position; E_0: absorption at the maximum; $\Delta\nu_{1/2}$: band half width; f_G: contour shape parameter; B: integral intensity.

Table II Numerical Values of Molar Extinction Coefficients of Amino Acid Residue Side Groups in H_2O, in E (L mol^{-1} cm^{-1}) Units

Wave-number (cm^{-1})	Asn	Gln	His	Phe	Arg	Lys	Tyr	Asp	Glu	-COO^-	-NH_3^+	Asp	Glu	-COOH	Tyr	-NH_2	Wave-number (cm^{-1})
								(Ionized Form)						(Deionized Form)			
1800	2	2	0	0	2	0	0	1	1	0	5	8	7	2	0	3	1800
1795	2	3	0	0	2	0	0	1	1	1	5	9	8	4	0	3	1795
1790	2	3	0	0	2	0	0	1	2	1	5	10	9	9	0	3	1790
1785	2	3	0	0	2	0	0	1	2	1	6	12	11	16	0	3	1785
1780	3	3	0	0	2	0	0	1	2	1	6	15	13	27	0	3	1780
1775	3	3	0	0	2	0	0	2	2	1	6	19	17	42	0	4	1775
1770	3	4	0	0	3	0	0	2	2	1	7	25	22	61	0	4	1770
1765	4	4	0	0	3	0	0	2	2	1	7	33	29	80	0	4	1765
1760	4	5	0	0	4	0	0	2	3	1	8	45	38	108	0	4	1760
1755	5	5	0	0	4	1	0	2	3	2	9	62	50	131	0	5	1755
1750	5	6	0	0	5	1	0	3	3	2	10	84	66	151	0	5	1750
1745	6	7	0	0	6	1	0	3	3	2	10	112	86	164	0	5	1745
1740	7	8	0	0	7	1	0	3	3	2	11	145	106	169	0	6	1740
1735	8	9	0	0	8	1	0	3	4	3	12	182	134	164	0	6	1735
1730	10	10	0	0	10	2	0	4	4	3	13	219	160	151	0	7	1730
1725	12	11	0	0	13	2	0	4	4	3	15	251	185	131	0	7	1725
1720	15	14	0	0	20	2	0	4	5	4	17	273	205	108	0	8	1720
1715	22	17	0	0	32	3	0	5	5	4	18	278	216	80	0	8	1715
1710	35	22	0	0	52	3	0	5	6	5	20	266	218	61	0	9	1710
1705	59	32	0	0	84	4	0	6	6	5	22	239	208	42	0	10	1705
1700	98	52	0	0	132	5	0	6	7	6	26	204	189	27	0	10	1700
1695	154	85	0	0	195	6	0	7	7	6	30	167	165	16	0	11	1695
1690	211	139	0	0	268	7	0	8	8	7	34	132	139	9	0	12	1690
1685	285	210	0	0	344	9	0	9	9	8	41	101	113	4	0	13	1685
1680	325	289	0	0	409	12	0	9	9	9	48	75	90	2	0	14	1680
1675	323	355	0	0	449	16	0	11	10	11	57	55	70	0	0	16	1675
1670	284	385	0	0	459	21	0	12	11	13	70	40	53	0	0	17	1670
1665	226	363	0	0	440	29	0	13	13	15	84	30	40	0	0	19	1665
1660	169	305	0	0	405	40	0	15	14	17	103	22	30	0	0	21	1660
1655	127	235	0	0	369	55	0	17	15	21	125	17	23	0	0	23	1655
1650	104	175	0	0	345	71	0	19	17	26	150	14	18	0	0	26	1650
1645	99	134	0	0	337	90	1	22	19	33	176	11	14	0	0	29	1645
1640	108	116	0	0	338	108	2	25	22	43	199	10	11	0	0	33	1640
1635	126	117	0	0	336	122	3	30	25	57	215	8	9	0	0	37	1635
1630	146	132	0	0	315	129	4	35	29	76	220	7	8	0	0	42	1630
1625	161	158	1	0	275	126	7	43	33	99	212	6	7	0	0	49	1625
1620	161	188	2	0	224	114	12	54	40	128	195	6	6	0	0	56	1620
1615	147	213	4	0	172	98	25	68	49	160	172	5	5	0	0	65	1615
1610	125	223	9	0	127	80	61	89	61	193	149	4	5	0	0	77	1610
1605	101	212	23	0	91	62	130	116	77	221	128	4	4	0	0	92	1605
1600	81	184	52	0	64	47	153	153	101	237	110	3	4	0	0	110	1660
1595	64	152	67	0	45	36	86	198	132	235	95	3	3	0	0	134	1595
1590	51	121	40	0	33	28	36	251	172	216	85	3	3	0	0	165	1590
1585	41	97	16	0	25	23	17	307	222	186	78	2	3	0	0	205	1585
1580	34	77	6	0	19	20	10	354	281	153	74	2	2	0	0	255	1580
1575	28	62	3	0	16	19	7	378	345	121	72	2	2	0	0	314	1575
1570	24	51	1	0	13	21	6	367	405	94	72	2	2	0	0	377	1570
1565	20	42	1	0	11	25	5	328	451	71	75	1	2	0	0	428	1565
1560	17	35	0	0	10	32	5	274	468	54	81	1	1	0	0	449	1560
1555	15	30	0	0	9	41	6	219	451	41	90	1	1	0	1	428	1555
1550	12	26	0	0	8	52	6	170	405	32	99	1	1	0	1	377	1550

Table II (*Continued from the previous page.*)

Wave-number (cm^{-1})	Asn	Gln	His	Phe	Arg	Lys	Tyr	Asp	Glu	-COO⁻	$-NH_3^+$	Asp	Glu	-COOH	Tyr	$-NH_2$	Wave-number (cm^{-1})
								(Ionized Form)						(Deionized Form)			
1545	11	22	0	0	7	66	7	130	345	25	114	1	1	0	2	314	1545
1540	9	19	0	0	5	80	9	99	281	20	130	1	1	0	4	255	1540
1535	8	16	0	0	4	91	12	76	222	17	149	1	1	0	7	205	1535
1530	7	14	0	0	4	99	15	59	172	14	170	0	1	0	14	165	1530
1525	6	12	0	0	3	101	21	47	132	12	189	0	0	0	35	134	1525
1520	6	11	0	0	3	96	31	38	101	11	202	0	0	0	183	110	1520
1515	5	10	0	0	3	86	51	32	77	9	207	0	0	0	423	92	1515
1510	4	9	0	1	2	73	95	27	61	8	200	0	0	0	131	77	1510
1505	4	8	0	3	2	59	214	23	49	7	186	0	0	0	28	65	1505
1500	4	7	0	12	2	45	559	20	40	6	165	0	0	0	12	56	1500
1495	3	6	0	71	2	33	559	18	33	6	143	0	0	0	6	49	1495
1490	3	6	0	26	1	23	214	16	29	5	122	0	0	0	4	42	1490
1485	3	5	0	5	1	16	95	14	25	4	104	0	0	0	2	37	1485
1480	10	5	0	1	1	11	51	12	22	4	89	0	0	0	1	32	1480
1475	2	4	0	0	1	8	31	11	19	4	75	0	0	0	1	29	1475
1470	2	4	0	0	1	6	21	10	17	3	64	0	0	0	0	26	1470
1465	2	3	0	0	1	4	15	9	15	3	55	0	0	0	0	23	1465
1460	1	3	0	0	0	3	11	8	14	3	48	0	0	0	0	21	1460
1455	1	3	0	0	0	3	8	7	13	2	42	0	0	0	0	19	1455
1450	1	3	0	0	0	2	6	7	11	2	37	0	0	0	0	17	1450
1445	1	2	0	0	0	2	5	6	10	2	33	0	0	0	0	16	1445
1440	1	2	0	0	0	1	4	5	9	2	29	0	0	0	0	14	1440

one more absorption band at 1250 cm^{-1} with E_0 = 100–200 L mol^{-1} cm^{-1}. Very strong band overlapping does not permit accurate determination of the band parameters.

Absorption of Arginine Residue

In the spectral region considered, the arginine residue has two absorption bands belonging to sym-

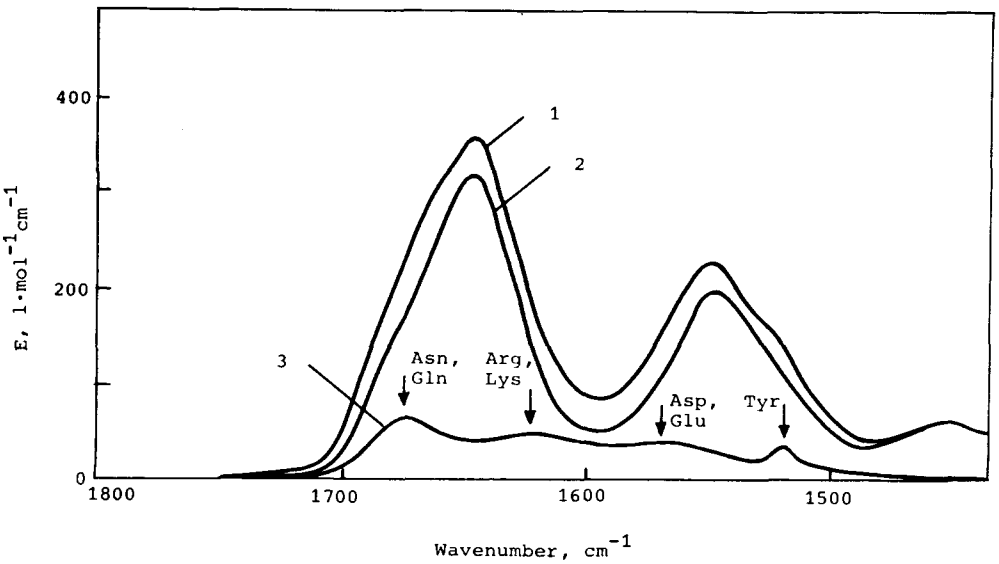

Figure 11. Elimination of side-chain group absorption from bovine ribonuclease A spectrum. (1) Ribonuclease A in 0.2 *M* phosphate buffer, pH 6.8, experimental curve. (2) Peptide absorption. (3) Amino acid side-group absorption.

metrical and antisymmetrical vibrations of the guanidine group $-CN_3H_5^+$. We studied only the ionized form of Arg due to the extremely high pK value (12.5).

The guanidine group of Arg is separated from the C_α atom by three $-CH_2-$ groups. Thus, guanidine vibrations are completely localized. Despite strong band overlapping, analysis of the spectra was rather simple. We used the differential spectra (Arg) − (Ala) at various pH (Figure 6). All spectra were identical. They were averaged to obtain the parameters of the arginine residue absorption bands.

Absorption of Lysine Residue

The ionized ϵ-amino group of Lys has two absorption bands corresponding to antisymmetrical and symmetrical vibrations (Figure 7). Following the procedure used for Arg, we calculated differential spectra (Lys) − (Ala) to obtain Lys side-chain absorption bands. The spectra of methylamine, butylamine, and dodecylamine in the 1800–1500-cm^{-1} region are essentially identical to the differential spectra mentioned above.

Under titration of the ϵ-amino group (pK 8.9), intensities of absorption bands pertaining to $-NH_3^+$ group vibrations disappeared but no new bands appeared in the spectrum. Thus, we draw the conclusion that the intensities of the absorption bands corresponding to ϵ-NH_2 group vibrations are less than 40 L mol^{-1} cm^{-1}. We do not take into account the contribution of these bands to protein spectra.

Absorption of Tyrosine Residue

The tyrosine residue absorption band is well known from the spectra of protein films and D_2O solutions.[5,18,19] It has a characteristic band near 1515 cm^{-1}. The band can be seen in the protein spectra due to a very small band width.

The spectrum of cresol solution at pH 5.8 (Figure 8a) has an intensive absorption band at 1516 cm^{-1} corresponding to phenol group vibration. The same band can be seen in the spectrum of methyl ester of Tyr at neutral pH (Figure 8b).

In addition to the deionized form, we studied the ionized R group of Tyr (Figure 9). Tyr in H_2O solution has a pK of 10.1. At extremely high pH the phenol group absorption band shifts to 1499 cm^{-1} and the intensity increases up to 700 L mol^{-1} cm^{-1}. Another absorption band appears at 1602 cm^{-1}.

There is one more absorption band in the Tyr spectrum that is not described in tables and figures:

pH < 9

$$\nu_0 = 1245/E_0 = 200/\Delta\nu_{1/2} = 32/f_G = 0.5$$

pH > 11

$$\nu_0 = 1269/E_0 = 580/\Delta\nu_{1/2} = 27/f_G = 0.6$$

We assign this band to the phenol C−O group valent vibration.

Absorption of Histidine and Phenylalanine Residues

The His and Phe residues have the most weak absorption in the set of absorptive amino acids.

Imidazole spectrum is represented in Figure 10a. It is the model of the histidine residue. We considered only the band at 1596 cm^{-1}. The band parameters are independent on the deionization of His (pK 6.1). The His spectrum, as well as the imidazole one, do not change at pH variation.

We chose toluene as a model of phenylalanine residue (Figure 10b). There is an absorption band at 1494 cm^{-1} in the spectrum. The band corresponds to benzene ring vibration.

Quantitative Estimation of Amino Acid Side-Chain Group Absorption in the Protein Spectra

As mentioned above, the contribution of amino acid side-chain group absorption to the protein spectra depends on the protein amino acid composition. The shape of the side-chain group absorption curve differs from one protein to another. This is very important for analysis of the amide band shape to derive the protein secondary structure. Otherwise, the presence of shoulders in the amide bands can be incorrectly interpreted as the display of some particular conformation of the polypeptide chain.

The spectral parameters of the absorption bands of amino acid side-chain groups and the numerical values of the extinction coefficients are represented in Tables I and II.

The data obtained concern side-chain group vibrations in a water environment. It is well known that amino acids in proteins are subjected to strong influence of the protein intramolecular environment. In most cases charged groups are exposed into the solvent, and either interact with the water molecules or form "salt bonds" with other groups. To check the effect of the salt bond on the ir spectrum, we compared the spectrum of an equimolar mixture of dodecylamine and butyric acid at a concentration of 0.2 *M* at neutral pH with the sum of individual spec-

tra. We found about a 3 cm^{-1} shift of the $-COO^-_{str,as}$ band and no change of intensity. Thus the presented data should be considered only as a good approximation for globular proteins. Application of such data to D_2O solutions of many proteins[20-26] demonstrates validity of this approach.

An example of elimination of side-chain group absorption from bovine native ribonuclease A spectrum is represented in Figure 11. Synthesis of curve 3 was made according to the formula

$$E(\nu) = \frac{\sum_k n_k E_k(\nu)}{n}$$

where $E_k(\nu)$ is the molar extinction coefficient of the kth amino acid residue at the wave number ν, n_k is the amount of kth amino acid in the protein,[27] and n is the total number of amino acids in the protein (in our case $n = 124$). N- and C-end groups should be included in this formula as amino acid residues. We obtain pure peptide absorption of ribonuclease (curve 2) by subtracting synthesized side-chain absorption (curve 3) from the experimental protein spectrum (curve 1). The contribution of amino acid side-chain group absorption to the protein spectrum is 19% of total integral intensity of amide I and II.

The authors express their thanks to Dr. K. H. Zicherman for synthesis of the most important compounds, to Mr. A. I. Umnov for making fluorite cells, to Mrs. N. B. Ilyna for technical assistance, and to Prof. Yu. N. Chirgadze and Dr. O. V. Fedorov for very helpful discussions.

REFERENCES

1. Miyazava, T. (1960) *J. Chem. Phys.* **32,** 1647–1652.
2. Chirgadze, Yu. N. & Nevskaya, N. A. (1976) *Biopolymers* **15,** 607–625.
3. Chirgadze, Yu. N. & Nevskaya, N. A. (1976) *Biopolymers* **15,** 627–636.
4. Chirgadze, Tu. N. & Nevskaya, N. A. (1976) *Biopolymers* **15,** 637–648.
5. Chirgadze, Yu. N., Fedorov, O. V. & Trushina, N. P. (1975) *Biopolymers* **14,** 679–694.
6. Chirgadze, Yu. N., Shestopalov, B. V. & Venyaminov, S. Yu. (1973) *Biopolymers* **12,** 1337–1351.
7. Chirgadze, Yu. N. & Brazhnikov, E. V. (1974) *Biopolymers* **13,** 1701–1712.
8. Byler, D. M. & Susi, H. (1986) *Biopolymers* **25,** 469–487.
9. Susi, H., Timasheff, S. N. & Stevens, L. (1967) *J. Biol. Chem.* **242,** 5460–5466.
10. Hannah, R. W. & Patacini, S. C. (1971) *Perkin Elmer IR Bull.* No. 2.
11. Atwood, M. R. & Hannah, R. W. (1972) *Perkin Elmer IR Bull.* No. 10.
12. Lee, D. C., Hayward, J. A., Restall, C. J. & Chapman, D. (1984) *Biochemistry* **24,** 4364–4373.
13. Kalnin, N. N. & Venyaminov, S. Yu. (1988) *J. Prikladnoy Spectrosc.* **49,** 592–597.
14. Jaenicke, L. (1974) *Anal. Biochem.* **61,** 623–627.
15. Savitzky, A. & Golay, M. J. E. (1964) *Anal. Chem.* **36,** 1627–1638.
16. Butler, W. L. & Hopkins, D. W. (1970) *Photochem. Photobiol.* **12,** 439–450.
17. Zundel, G. (1969) *Hydration and Intermolecular Interaction, Academic Press,* New York and London, pp. 40–132.
18. Bellamy, L. J. (1963) *The Infrared Spectra of Complex Molecules,* Methuen & Co. Ltd., London, and John Wiley & Sons, New York, pp. 355–373.
19. Bendit, E. B. (1967) in *Symposium on Fibrous Proteins,* Buttersworth, Australia, pp. 386–396.
20. Shestopalov, B. V. & Chirgadze, Yu. N. (1976) *Eur. J. Biochem.* **67,** 123–128.
21. Hegyi, G. & Venyaminov, S. Yu. (1980) *FEBS Lett.* **109,** 134–136.
22. Venyaminov, S. Yu. & Gogiya, Z. V. (1981) *Dokl. Akad. Nauk SSSR* **257,** 737–741.
23. Koteliansky, V. E., Glukhova, M. A., Bejanian, M. V., Smirnov, V. N., Filimonov, V. V., Zalite, O. M. & Venyaminov, S. Yu. (1981) *Eur. J. Biochem.* **119,** 619–624.
24. Koteliansky, V. E., Glukhova, M. A., Shirinsky, V. P., Smirnov, V. N., Bushueva, T. L., Filimonov, V. V. & Venyaminov, S. Yu. (1982) *Eur. J. Biochem.* **121,** 553–559.
25. Venyaminov, S. Yu., Metsis, M. L., Chernousov, M. A. & Koteliansky, V. E. (1983) *Eur. J. Biochem.* **135,** 485–493.
26. Fedorov, O. V. (1978) in *Proceedings of the 7th European Conference on Muscle and Motility,* Warsaw, p. 65.
27. Wlodawer, A., Bott, R. & Sjolin, L. (1982) *J. Biol. Chem.* **257,** 1325–1333.

Received June 21, 1989
Accepted June 6, 1990