ELSEVIER

Biophysical Chemistry 71 (1998) 9–20

Biophysical Chemistry

Effect of charge interactions on the carboxylate vibrational stretching frequency in *c*-type cytochromes investigated by continuum electrostatic calculations and FTIR spectroscopy

Monique Laberge *, Kim A. Sharp, Jane M. Vanderkooi

Johnson Research Foundation, Department of Biochemistry and Biophysics, School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Received 25 August 1997; revised 20 October 1997; accepted 21 October 1997

Abstract

The FTIR spectra of the asymmetric carboxylate absorption region of three *c*-type cytochromes—namely horse heart, yeast and bonito cytochromes *c*—as well as continuum electrostatic calculations performed on their respective protein matrices, show that these combined methods can target specific protein regions and yield pertinent protein charge information that correlates with the observed spectral data. Deconvolution of the IR carboxylate stretch frequency region (1525–1675 cm^{-1}) in the three cytochromes yield different ν^{a}_{oco} distributions. In the case of the bonito cytochrome *c* carboxylates, two ν^{a}_{oco} populations are clearly distinguishable in the deconvoluted spectra—which is not the case for the more complex ν^{a}_{oco} deconvolutions of the other two cytochromes. The frequency distributions of the calculated potentials are consistent with the experimental observations and we conclude that the IR carboxylate absorption in proteins can be modified by the electrostatic environment. © 1998 Elsevier Science B.V.

Keywords: Carboxylate; *c*-Type cytochromes; Electrostatic calculations

1. Introduction

The IR spectral parameters of amino acid residue absorption bands have been characterized well enough [1–3] to make ν^{a}_{oco} a useful protein conformational marker. We recently completed comprehensive IR surveys of the antisymmetric carboxylate absorption (ν^{a}_{oco}) in cytochrome *c* and model compounds, and showed that ν^{a}_{oco} was responsive to protein surface conditions [4]. Electrostatic interactions are now acknowledged to play a major role in the relationship between protein structure and function [5–9] and in the solvent interaction forces responsible for protein folding [10–12]. Cytochrome *c*, one of the best characterized heme proteins, plays a well-known role in biological electron transfer and the electrostatic charge distribution within the protein is assumed to be crucial in facilitating the electron transfer process [13,14]. In this work, we use the ν^{a}_{oco} marker and try to understand the observed spectroscopy in the context of protein matrix

Abbreviations: FDPB: Finite Difference Poisson–Boltzmann; ν^{a}_{oco}: Antisymmetric carboxylate stretch

* Corresponding author. E-mail: labergem@mail.med.upenn.edu

PII S0301-4622(97)00127-0

charge phenomena using a continuum approach to model the charge distribution of the protein [5].

For this purpose, we use the Finite Difference Poisson–Boltzmann (FDPB) method. It has been shown to work well for a wide variety of problems [15–24] and has the advantage of avoiding the computational cost of explicit solvent calculations. In classical electrostatics, the Poisson equation relates the spatial variation of the potential Φ at position $\vec{r}$ to the charge distribution ρ and the position-dependent dielectric permittivity ε as follows:

$$\nabla\varepsilon(\vec{r})\nabla\Phi(\vec{r}) = -4\pi\rho(\vec{r})/\mathrm{kT} \quad (1)$$

If mobile ions are present in the system, the Poisson equation can be combined with the Boltzmann expression for ion concentration, yielding:

$$\nabla\varepsilon(\vec{r})\nabla\Phi(\vec{r}) - \varepsilon\kappa^2\Phi(\vec{r}) = -4\pi\rho^{\mathrm{f}}(\vec{r})/\mathrm{kT} \quad (2)$$

where the linearized form of the Poisson–Boltzmann equation is sufficient for most protein applications. The term κ^2 is equal to $1/\lambda^2$ or $8\pi q^2 I/\mathrm{ekT}$ where λ is the Debye length, I is the ionic strength of the solution, and ρ^{f} is the fixed charge density. Φ, ε, κ and ρ are all functions of the vector $\vec{r}$. The second term in Eq. (2) describes the salt effect. Since water molecules are more polarized by an electric field than the protein, the use of two dielectric constants allows for consideration of the solvation effect experienced by polar molecules in an aqueous solution. The derivative of $\varepsilon(\vec{r})$ in Eq. (2) is nonzero only when $\varepsilon(\vec{r})$ varies. This is the 'dielectric discontinuity' region between the low dielectric solute (protein) and the high dielectric solvent or 'molecular surface'. The solute molecule—inside the molecular surface—is given a low dielectric of 2 [6] and the solvent is assigned a high dielectric (~ 80). The Poisson–Boltzmann equation solver implemented in the *Delphi* software package uses an algorithm which maps the solute/solvent variables on a three-dimensional grid; each grid lattice point is assigned a value for the charge, the dielectric constant and the ionic strength. Boundary conditions are set using either the Coulomb or Debye–Hückel equations [25] depending on whether the ionic strength is 0 or not, respectively. The Poisson–Boltzmann equation is replaced by a series of finite difference equations which are iteratively solved. Electrostatic free energies for charging the system are then calculated as:

$$\Delta G = 1/2\Sigma_i q_i \Phi_i \quad (3)$$

where q_i is the charge and Φ_i is the Poisson–Boltzmann calculated potential at each atom i. By summing over different groups of charges, Eq. (3) can be used to calculate the contributions of particular residues or groups as desired. The solute atomic charges, and the van der Waals radii parameters required to define the solute/solvent dielectric boundary are inputted into the program along with the solute and solvent dielectric constants.

Our interest here is to correlate the experimentally observed carboxylate stretching frequencies to the electrostatic potentials calculated in the protein matrices of the respective cytochromes.

2. Materials and methods

2.1. Materials and sample preparation

Horse heart, bonito and yeast cytochromes *c* were purchased from the Sigma Chemical (St. Louis, MO) as well as DCl and D_2O. Buffer water was deionized and glass-distilled. The protein was dissolved in deuterated phosphate buffer (20 mM), pD 7.0, and allowed to incubate for 24-h period, to ensure complete deuterium exchange of protein amide protons. Complete oxidation of the samples was monitored by UV–Vis spectroscopy in the Q-band region before and after FTIR data acquisition. Freshly prepared $K_3Fe(CN)_6$ (~ 5 μM) was added to samples which were not fully oxidized.

2.2. FTIR spectroscopy

Spectral measurements were performed at 22°C on a Bruker IFS 66 FTIR (Bruker Billerica, MA) using a Globar source, a KBr beam splitter and MCT detector. The samples were placed in a ZnSe ATR crystal (Graseby-Specac, Fairfield, CT) cell. The D_2O phosphate buffer was used as reference. A total of

128 interferograms were averaged per recorded spectrum at a resolution of 0.5 cm^{-1}. Acquisition time was 2 min per spectrum.

2.3. Calculation of electric potentials

2.3.1. Atomic coordinates

The three-dimensional structures were obtained from the Protein Data Bank [26]: pdb1cyc.ent for bonito cytochrome *c* [27], pdb1hrc.ent for horse heart cytochrome *c* [28] and pdb1ycc.ent for yeast cytochrome *c* [29]. The X-ray waters were not retained and the structures were prepared for the electrostatic calculations using the *Insight II* software package (Molecular Simulations, San Diego, CA) on a Silicon Graphics IRIS Indigo workstation. The hemes of all three cytochromes were similarly oriented along the *x–y* axes and the three cytochromes superimposed to ensure good backbone alignment.

2.3.2. Finite Difference Poisson–Boltzmann calculations

All electrostatic potential fields were calculated using the finite difference solutions to the Poisson–Boltzmann equation [25] as implemented in the *Delphi* [30,31] and *Grasp* [32] software packages. Visualization of the potentials and fields was achieved with *Grasp*.

Two sets of calculations were performed: (i) first, to account for the contributions of ionizable residues, protein matrix charges were assigned, as most likely to apply at neutral pH: -0.5 to the od1 and od2 Asp oxygens and to the oe1 and oe2 Glu oxygens, -0.5 to the oxygens of the two heme propionates, $+0.5$ to the nh1 and nh2 Arg nitrogens and $+1$ to the Lys nz; the potentials were then calculated at the oxygens atoms of all Glu, Asp and propionic acid carboxylate groups; (ii) a second calculation was performed to correct for the contribution of the charge on the oxygen where the potential was being calculated. This was done by setting both protein and solvent dielectric to 2, charging the oxygen in question and calculating the potential which was then subtracted from the value obtained in the first calculation. The effect of salt in the solvent was taken into account by performing the calculations at the experimental ionic strength (20 mM). The cytochromes were placed on a $65 \times 65 \times 65$ grid and scaled so that the longest dimension of the molecule as a percent of grid length was 70%. The probe radius, used to define the molecular surface, was set at 1.4 Å. Calculations were performed at different scales and lattice mappings to ensure that the grid mapping procedure did not affect computational results.

2.4. Statistical analysis

The calculated potentials were organized as frequency distributions for each cytochrome. The class intervals were arbitrarily fixed and three different sets of class boundaries were tested. It was found that all three types of frequency distributions provided the same statistical conclusions and the distribution type finally selected for illustrative purposes was chosen because it simplified the discussion. For each frequency distribution, the expected value of the mean (x) was calculated, the standard deviation of the sample (s) as well as the standard deviation of the mean (σ_x). Additionally, we calculated confidence levels using the *t* distribution as required for small samples at 99% confidence limits.

3. Results

3.1. Observed carboxylate stretches

Fig. 1 presents the infrared ν^{a}_{oco} stretch frequency region, observed between 1525 and 1675 cm^{-1} for all three cytochromes, and the best fits obtained in deconvoluting the spectra. The absorptions at ~ 1650 cm^{-1} and 1518 cm^{-1} are assigned to amide I′ and tyrosine vibrations, respectively, and are of no relevance in this context. Focusing on the ν^{a}_{oco} absorption—centered at ~ 1570 cm^{-1}, we note that the deconvolutions yield different distributions of ν^{a}_{oco} populations in the different cytochromes. The difference is especially evident when comparing the bonito spectrum to that of the horse and the yeast cytochromes. We have shown recently that the protein charge distribution was an important factor affecting the stretch frequency of the small CO heme ligand in carbonmonoxy cytochromes *c* [21,22,33] and the question immediately arises as to whether the respective protein matrices of the three cytochromes under

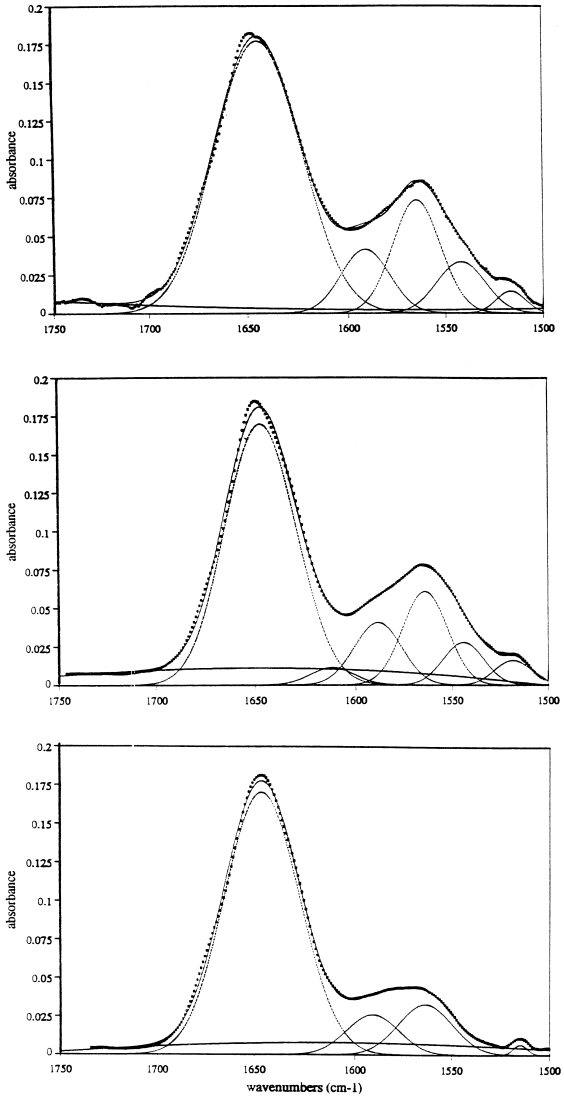

Fig. 1. IR carboxylate stretch frequency region for cytochrome *c* at neutral pH: (top) yeast, (middle) horse heart, (bottom) bonito. Deconvolutions and best fits obtained using gaussian fitting procedures. Fitting parameters: *yeast*: peak @ 1542.2 cm^{-1} (HW = 12.3 cm^{-1}); peak @ 1565.0 cm^{-1} (HW = 12.3 cm^{-1}); peak @ 1591.5 cm^{-1} (HW = 12.3 cm^{-1}); *horse*: peak @ 1544.1 cm^{-1} (HW = 12.7 cm^{-1}); peak @ 1564.2 cm^{-1} (HW = 12.2 cm^{-1}); peak @ 1588.6 cm^{-1} (HW = 12.3 cm^{-1}); peak @ 1611.8 cm^{-1} (HW = 11.9 cm^{-1}); *tuna*: peak @ 1563.7 cm^{-1} (HW = 12.6 cm^{-1}); peak @ 1590.9 cm^{-1} (HW = 12.3 cm^{-1}):

consideration—and their different charge distributions—could be responsible for the different ν^{a}_{oco} distribution pattern observed here.

3.2. Electrostatic potentials at the carboxylate groups

The potentials calculated for the carboxylate groups of all Glu and Asp residues and for the heme propionate side chains are listed in Table 1 for the respective cytochromes. Fig. 2 compares the fre-

Table 1
Summary of $\Sigma q\Phi$ per COO^- in horse heart, yeast and bonito cytochromes *c*

	$\Sigma q\Phi$ (kT)	Statistics
Yeast cyt c		
Glu 4	−142.63	
Glu 21	−127.47	$x = -121.45$ kT
Glu 44	−159.02	
Asp 50	−135.48	$s = 28.21$
Asp 60	−94.06	
Glu 61	−135.61	$\sigma_x = 7.82$
Glu 66	−127.91	
Glu 88	−141.06	Confidence at $\alpha = 0.99$
Asp 90	−122.54	$= 0.0981$
Asp 93	−111.95	
Glu 103	−143.04	
Prop (old)	−73.05	
Prop (ola)	−65.05	
Horse cyt c		
Asp 2	−115.80	$x = -127.22$ kT
Glu 4	−153.37	
Glu 21	−124.99	$s = 24.01$
Asp 50	−135.65	
Glu 61	−173.42	$\sigma_x = 6.42$
Glu 62	−135.21	
Glu 66	−136.73	Confidence at $\alpha = 0.99$
Glu 69	−144.75	$= 0.0804$
Glu 90	−125.11	
Glu 92	−116.56	
Asp 93	−118.78	
Glu 104	−132.31	
Prop (old)	−77.65	
Prop (ola)	−90.58	
Bonito cyt c		
Asp 2	−119.34	$x = -115.43$ kT
Glu 21	−133.65	
Glu 44	−127.25	$s = 14.73$
Asp 50	−107.84	
Glu 61	−110.03	$\sigma_x = 4.44$
Glu 66	−134.33	
Glu 69	−112.87	Confidence at $\alpha = 0.99$
Glu 90	−131.43	$= 0.0557$
Asp 93	−100.14	
Prop (old)	−93.31	
Prop (ola)	−99.52	

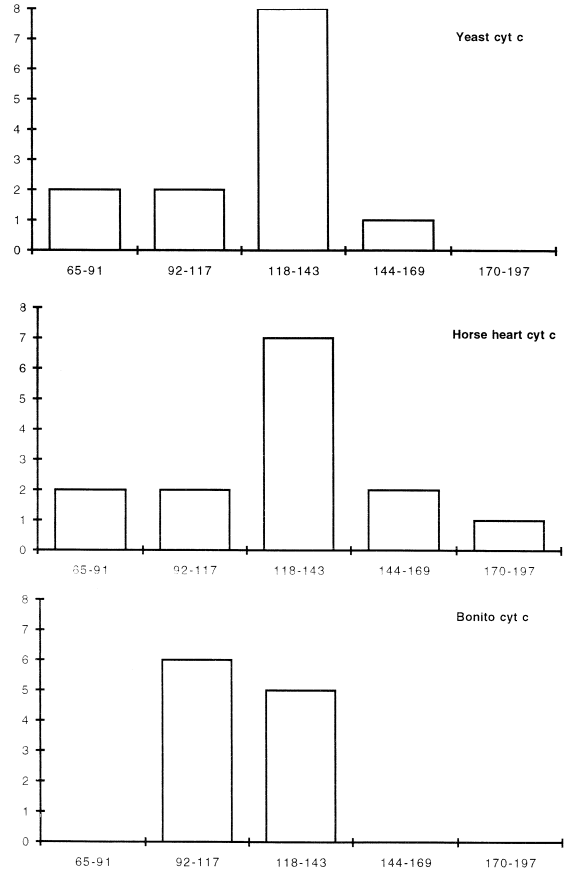

Fig. 2. Frequency distributions of the ΣqΦ calculated at all Glu, Asp residues and propionates in all three cytochromes: (top) yeast, (middle) horse heart, (bottom), bonito. Statistical parameters: yeast: $x = -121.45$ kT, $\sigma_x = 7.82$; horse: $x = -127.22$ kT, $\sigma_x = 6.42$; bonito; $x = -115.43$ kT, $\sigma_x = 4.44$. All potentials are negative.

quency distributions of the potentials calculated for the three cytochromes. Table 1 also presents a few statistical parameters (v. supra, Section 2.2). In the case of bonito cytochrome *c*, the ν^{a}_{oco} potentials can be seen to clearly distribute in a more 'polarized' pattern when compared to the frequency distributions of the horse and yeast $\sum q\Phi$ values. Two distributions are distinguishable: one with average $\sum q\Phi_{avg}$ equal to -129.2 kT and another with $\sum q\Phi_{avg} = -103.9$ kT. Statistically, no such grouping can be achieved with the frequency distributions of the other two cytochromes (horse and yeast). Their electrostatic potential distributions are harder to distribute into a specific number of populations, but we can

Table 2
H-bonding in bonito cytochrome *c*

Donor	Acceptor	Distance (Å)	Angle (deg)
2:HN	**2:OD1**	2.02	122.69
6:HN	2:O	2.28	148.06
7:HN	3:O	2.33	161.99
7:HN	4:O	2.08	127.12
8:HN	4:O	2.35	160.11
9:HN	5:O	1.86	178.51
10:HN	6:O	1.35	164.44
11:HN	7:O	1.38	140.59
12:HN	8:O	2.17	159.10
13:HN	9:O	2.10	156.94
14:HN	10:O	2.48	137.26
15:HN	10:O	2.01	125.11
16:HE22	11:O	2.22	126.17
17:HN	14:O	1.66	170.36
16:HE21	16:N	1.84	125.04
16:HN	16:NE2	1.77	132.32
27:HZ1	17:O	1.62	151.63
31:HD21	19:OG1	2.15	132.61
24:HN	21:O	2.46	169.10
23:HN	31:OD1	2.29	122.92
31:HN	26:ND1	1.61	141.21
26:HE2	44:O	2.42	143.59
33:HN	31:O	2.33	140.10
59:HN	38:O	1.53	134.32
52:HD21	40:OG1	1.73	160.21
46:H:H	1:H:O2A	2.30	141.28
47:HG	47:O	2.36	133.55
79:HN	48:O:H	1.73	125.19
52:HD22	2::HO	2.15	121.50
56:HN	53:O	1.54	153.76
1:H:H2A	59:NE1	2.49	123.22
59:HE1	1:H:O2A	2.39	132.02
61:HN	60:OD1	1.79	165.67
66:HN	62:O	1.98	163.03
67:HN	63:O	1.55	169.51
68:HN	64:O	1.78	142.90
69:HN	65:O	1.70	173.73
2:H:HO1	67:O:H	2.00	151.97
67:H:H	2:H:O	2.04	145.85
72:HN	70:ND2	2.04	134.89
70:HD21	71:N	2.48	142.03
70:HD21	72:N	2.15	124.48
91:HN	87:O	2.03	151.32
92:HN	88:O	2.00	123.55
93:HN	89:O	1.95	161.82
94:HN	90:O	1.65	172.42
95:HN	91:O	2.35	146.30
96:HN	92:O	1.92	157.38
97:HN	93:O	1.72	161.29
98:HN	94:O	2.23	160.07
100:HN	96:O	2.15	173.21
101:HN	97:O	1.75	170.90
102:HN	98:O	2.22	131.51

Table 3
H-bonding in HH cytochrome *c*

Donor	Acceptor	Distance (Å)	Angle (deg)
6:HN	2:O	1.90	144.53
4:HN	**2:OD2**	2.19	137.80
5:HN	**2:OD2**	2.28	148.34
7:HN	3:O	2.23	146.74
9:HN	5:O	2.36	163.43
10:HN	6:O	2.08	162.95
11:HN	7:O	1.97	173.66
12:HN	8:O	2.10	159.64
13:HN	9:O	1.78	141.41
13:HN	10:O	2.49	126.76
14:HN	10:O	1.76	157.48
15:HN	10:O	2.06	155.29
17:HN	14:O	1.99	166.60
29:HN	17:O	2.02	131.00
32:HN	19:O	1.65	153.68
21:HN	**21:OE1**	1.96	147.39
24:HN	21:O	1.81	150.72
31:HD22	21:O	1.72	154.94
31:HN	26:ND1	1.90	163.87
26:HE2	44:O	2.07	165.22
27:HN	29:O	1.94	141.65
33:HN	31:OD1	2.01	143.52
35:HN	32:O	2.40	133.79
38:HH21	33:O	2.46	132.34
34:HN	102:O	1.75	168.70
38:HN	35:O	2.24	149.51
37:HN	59:O	1.81	166.36
59:HN	38:O	1.85	177.93
39:HZ1	56:O	2.45	146.99
56:HN	40:OG1	2.24	138.63
40:HG1	52:O	2.24	141.00
41:HN	**105H:O2A**	2.16	167.40
43:HN	48:OH	2.03	141.49
46:HN	43:O	2.48	164.58
48:HH	**105H:O1A**	1.94	126.85
49:HN	**105H:O1D**	1.96	167.55
53:HN	49:O	2.15	153.57
52:HN	49:OG1	2.28	148.39
54:HN	50:O	1.95	159.72
52:HD21	52:O	1.65	159.00
105H:H2A	52:ND2	2.33	157.84
52:HD22	**105H:O2A**	2.37	151.10
55:HZ1	74:O	1.90	140.35
59:HE1	**105H:O2A**	1.71	152.88
60:HN	63:OG1	2.14	148.52
64:HN	60:O	2.13	163.90
65:HN	61:O	1.85	148.00
99:HZ1	**61:OE2**	1.86	143.89
67:HN	63:O	2.08	164.67
68:HN	64:O	1.95	164.40
91:HH12	65:O	1.75	161.34
69:HN	66:O	2.40	138.03
73:HN	70:O	2.17	152.10

Table 3 (continued)

Donor	Acceptor	Distance (Å)	Angle (deg)
74:HN	70:O	2.19	131.45
73:HN	70:OD1	2.38	156.12
75:HN	71:O	1.89	135.92
82:HN	72:NZ	1.77	156.29
72:HZ2	80:O	2.36	162.46
78:HN	75:O	1.89	136.09
80:HN	78:OG1	2.17	139.56
79:HN	**105H:O2D**	1.74	164.84
91:HH22	86:O	2.37	159.33
91:HN	87:O	2.31	158.13
92:HN	88:O	1.90	151.69
93:HN	89:O	1.87	172.82
94:HN	90:O	1.93	159.36
95:HN	91:O	1.98	159.36
96:HN	92:O	2.05	160.20
97:HN	93:O	2.04	158.28
98:HN	94:O	1.85	171.79
99:HN	95:O	1.96	164.66
101:HN	97:O	1.84	178.26
102:HN	98:O	1.72	154.20
103:HD21	100:O	2.10	125.45
103:HD21	103:N	2.46	125.42

confidently say that they do distribute into more populations than the bonito, and that, as such, they reflect the ν_{oco}^{a} spectral deconvolution patterns of the respective cytochromes.

4. Discussion

We are interested here in trying to elucidate what could possibly give rise to the different carboxylate populations as inferred from the ν_{oco}^{a} deconvolutions and the electrostatic calculations. Three possibilities can account for the observed and calculated carboxylate distributions: (i) they could reflect different H-bonding patterns, either intramolecular or solvent mediated, (ii) the carboxylates are attached to Glu, Asp and heme propionate residues and this could conceivably cause them to absorb at different frequencies; and (iii) the observed distributions could be representative of different electrostatic environments for the vibrational oscillators which could then be Stark-shifted with respect to one another due to the presence of a nearby charge.

The Glu and Asp carboxylates of all three cytochromes all have their carboxyl groups fully solvated on the protein surface. Vibrational frequency shifts and intensity changes have been associated before with changes in the degree of H-bonded structure and attributed to the different enthalpies of different types of hydrogen bonds [34]. In view of recent IR work [35] showing that amide region spectra can distinguish between two populations of amide bonds undergoing proton–deuterium exchange at different rates—depending on whether the population is solvent-exposed or not—we similarly attempted to distinguish between different H-bonded carboxylate populations. We determined the H-bonds patterns for all three cytochromes using their respective X-ray structures (cf. Tables 2–4: the intramolecular H-bonds involving the IR-monitored residues, i.e., Asp, Glu and propionates, are boldfaced). H-bonding was considered to occur when the distance between the proton on the donor atom and the heavy atom acceptor was less than 2.5 Å for H–N and H–O and less than 3.0 Å for N–O and when the angle between the proton acceptor, the proton and the proton donor was greater than 120°. The results show that the carbox-

Table 4
H-bonding in yeast cytochrome *c*

Donor	Acceptor	Distance (Å)	Angle (deg)
−3:HN	−5:OG1	2.25	129.33
−5:HG1	61:OE1	2.30	137.28
96:HG1	1:N	2.29	131.07
1:HN	96:OG1	2.12	152.34
2:HN	93:OD1	1.75	174.27
6:HN	2:O	1.89	150.70
5:HN	2:OG	1.98	165.11
7:HN	3:O	2.04	140.05
8:HN	4:O	2.45	153.15
9:HN	5:O	2.09	142.32
10:HN	6:O	1.91	159.90
11:HN	7:O	2.11	159.10
12:HN	8:O	2.00	168.30
13:HN	9:O	1.97	142.61
14:HN	10.O	1.83	153.54
15:HN	10:O	2.04	149.18
17:HN	14:O	2.05	158.59
18:HN	14:O	2.30	133.62
27:HZ1	15:O	1.91	164.17
29:HN	17:O	1.83	138.08
18:HE2	1H:NO	2.15	121.88
32:HN	19:O	1.93	172.35
31:HD21	19:OG1	1.78	173.12
19:HG1	25:O	2.15	122.90
20:HN	**21:OE1**	2.40	140.28
21:HN	**21:OE1**	1.94	139.34
24:HN	21:O	2.08	144.07
31:HD22	21:O	2.09	141.71
31:HN	26:ND1	1.81	168.20
26:HE2	44:O	1.74	167.65
27:HN	29:O	1.71	161.66
46:HH	28:O	1.49	168.88
33:HN	31:OD1	2.13	139.79
35:HN	32:O	2.22	138.87
38:HH11	33:O	2.30	131.15
38:HH21	33:O	2.16	133.88
34:HN	102:O	1.88	153.72
38:HN	35:O	2.20	157.15
37:HN	59:O	1.69	175.76
59:HN	38:O	2.02	173.28
40:HN	57:O	1.96	137.08
57:HN	40:OG	2.26	124.98
41:HN	1H:O2A	2.26	153.91
43:HN	48:OH	2.33	147.85
79:HZ1	47:O	1.78	133.41
49:HN	1H:O1D	1.96	158.19
53:HN	49:O	2.50	175.52
52:HN	49:OG1	2.14	159.38
49:HG1	1H:O2D	1.62	172.33
54:HN	50:O	2.46	145.79
55:HN	52:O	2.49	135.77
52:HD22	1H:O2A	2.34	165.75
55:HZ3	74:O	1.92	172.97
59:HE1	1H:O2A	2.14	152.25
64:HN	60:O	2.19	159.10
63:HN	**60:OD1**	1.72	158.02
63:HD21	**60:OD2**	2.40	174.01
67:HN	63:O	1.72	161.83
68:HN	64:O	1.93	171.03
69:HN	65:O	2.44	126.35
91:HE	65:OG	1.90	173.77
69:HN	66:O	2.37	135.29
70:HN	67:O	2.10	140.21
85:HN	68:O	1.79	158.23
86:HZ3	69:O	1.36	174.14
73:HN	70:O	2.36	128.37
74:HN	70:O	2.04	154.23
73:HN	70:OD1	2.50	136.53
75:HN	71:O	1.74	165.25
78:HN	75:O	2.00	150.10
80:HN	78:OG1	2.31	167.77
1H:H2D	79:N	2.22	152.63
79:HN	1H:O2D	2.19	157.95
82:HN	80:O	2.43	137.42
91:HH22	85:O	2.10	165.73
91:HN	87:O	2.11	175.57
92:HN	88:O	1.84	156.66
93:HN	89:O	2.01	171.28
94:HN	90:O	1.85	167.37
95:HN	91:O	1.81	173.50
96:HN	92:O	2.05	143.83
97:HN	93:O	2.22	158.96
98:HN	94:O	1.77	157.68
99:HN	95:O	1.66	167.89
100:HN	96:O	2.49	152.92
101:HN	97:O	1.81	157.76
102:HN	98:O	1.91	152.46

ylates in the bonito species are mostly only H-bonded to solvent molecules on the surface (cf. Table 2): only one carboxylate residue (Asp 2) is intramolecularly H-bonded. In contrast, the H-bonding pattern in the horse cytochrome shows that four of its 13 carboxylates (Asp 2, Glu 21, Glu 61 and propionates) can also H-bond to the nitrogens of neighboring residues (boldface residues in Table 3). The same pattern is observed for the yeast cytochrome (Table 4) which shows possible H-bonding to neighboring residues for Glu 21 and Asp 60. Such strong H-bonds would be expected to lower the carboxylate frequency with respect to that of weaker H-bonded carboxylates, the difference in frequency then attributable to the different H-bond strengths or to

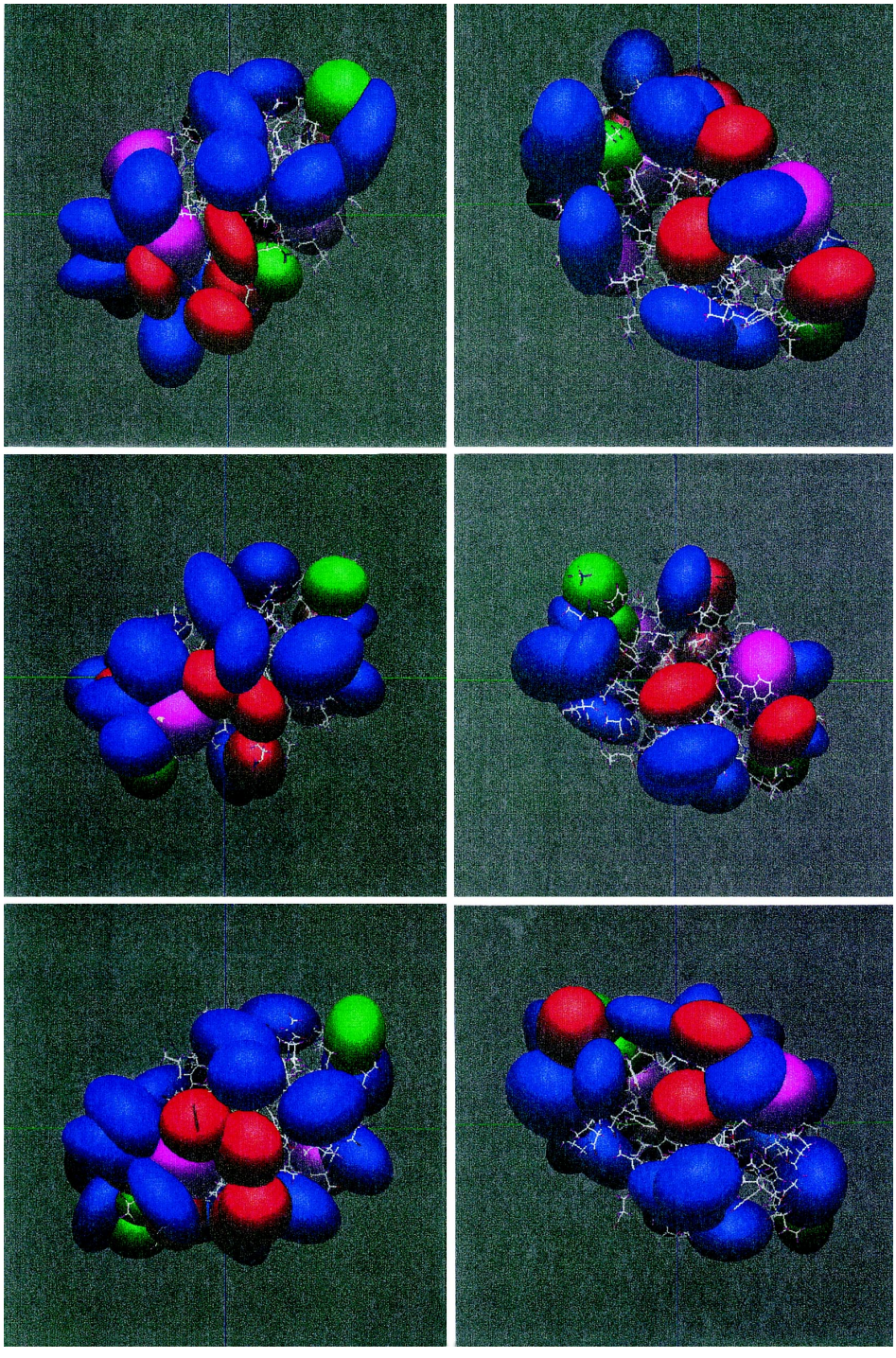

Fig. 3. Distribution of residues with charged R-groups at neutral pH in the cytochromes. Color scheme: Glu (red). Asp (green), Arg (pink), Lys (blue). Top: Yeast; middle: Bonito; bottom: Horse heart. The three cytochromes are oriented in the same plane with superimposed backbones. RHS/LHS shows 180°-rotation about the y-axis.

differences between proton acceptor and proton donor torsion angles. We propose that different H-bonding patterns cannot be invoked to account for the two carboxylate populations observed in the bonito cytochrome—since all carboxylates exhibit the same type of H-bonding, i.e., to solvent—but that they may play a role in the more complex horse and yeast carboxylate distributions and account for some of the additional populations inferred from the deconvolutions, especially those seen at lower frequency.

We can immediately dismiss the second possibility proposed above, that is that the populations distribute according to species, i.e., that the Glu, Asp and propionate ν^{a}_{oco} bands absorb at different frequencies. The deconvolutions do not support such an interpretation, since the number of bands can not be correlated to the number of Glu/Asp residues present in the respective cytochromes. In a recent ν^{a}_{oco} pH-dependence study, we also address this question and the pK_a titration curves are unambiguous: Glu and Asp ν^{a}_{oco} stretches can not be localized within the ν^{a}_{oco} spectral envelope [4].

A more likely interpretation is that the populations of vibrational oscillators distribute as a function of the charge environment to which they are exposed. The effect of electric fields on vibrational spectra is directly related to changes in the potential energy surface resulting from the interaction of the vibrational oscillator with the imposed electric field. Such Stark shifts have been observed in spectra as line shifts and changes in intensity [36] both in the visible [37–39] and in the IR regions [21,22,33,40,41]. Our deconvoluted ν^{a}_{oco} bands (cf. Fig. 1 show two different oscillators for the bonito cytochrome (1563 and 1564 cm^{-1}), three for yeast (1542, 1564 and 1591 cm^{-1}) and four for horse heart (1544, 1564, 1588 and 1611 cm^{-1}). We propose that they represent carboxylate residues exposed to different electric fields imposed by different protein matrix environments.

Fig. 3 illustrates the distribution of residues with charged R-groups at neutral pH in all three cytochromes. In Table 5, we list the distances separating the cytochrome carboxylates from charged residues belonging to the R-groups of positively and negatively charged amino acids in two cytochromes (horse and bonito) at neutral pH.

Fig. 4 helps to visualize their distribution in two

Table 5
Distance of carboxylates to nearest charged residues in bonito and horse heart cytochromes *c*

υ^{a}_{oco} residue	Distance (Å)	To charged residues
Bonito cyt c		
Asp 2	4.14	asp93(od2)
	8.71	lys5(nz)
Glu 21	7.41	lys25(nz)
	10.95	lyz27(nz)
Glu 44	> 13	to all R-charges
Asp 50	9.73	lys53(nz)
Glu 61	4.36	lys 99
Glu 66	10.50	lys73(nz)
	7.15	glu69(oc2)
	4.90	lys55(nz)
Glu 69	7.15	glu66(oc2)
	7.38	lys73(nz)
Glu 90	11.79	lys86(nz)
Asp 93	3.02	lys5
	9.16	lys88(nz)
	9.82	lys8(nz)
HH cyt c		
Asp 2	4.49	glu4(odl)
	4.27	lys5(nz)
	9.95	lys100(nz)
	6.15	asp93(odl)
Glu 4	4.97	asp2(od2)
	7.57	lys100(nz)
	8.88	lys7(nz)
Glu 21	8.89	lys25(nz)
Asp 50	4.90	lys53(nz)
Glu 61	2.75	lys99(nz)
	8.16	glu92(ocl)
	8.43	glu62(ocl)
Glu 62	7.12	lys88(nz)
	8.43	glu61(ocl)
Glu 66	6.88	glu62(ocl)
	7.41	lys55(nz)
Glu 69	9.27	lys88(nz)
	4.14	srg91(nhl)
	9.75	lys73(nz)
Glu 90	3.31	lys13(nz)
	5.57	lys86(nz)
Glu 92	8.16	glu61(ocl)
	4.63	lys88(nz)
Asp 93	5.38	lys5(nz)
	6.15	asp2(odl)
	7.65	glu90(oc2)
Glu 104	7.13	lys100(nz)
	5.38	lys22(nz)

od1,od2,oe2,oel: Carboxyl oxygen.
nz, nhl: R-group nitrogens.

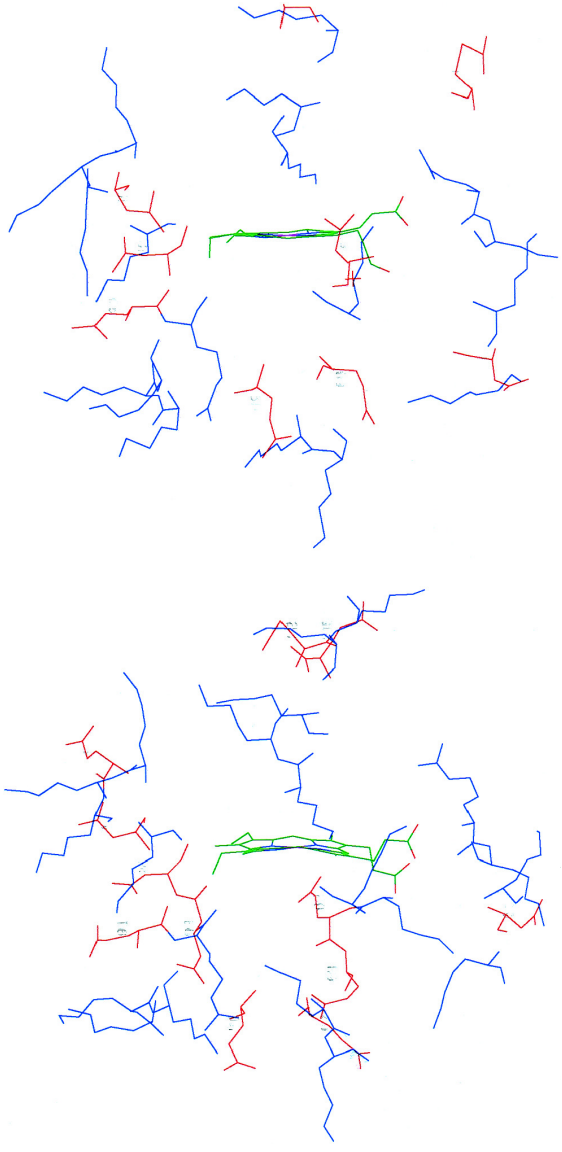

Fig. 4. Stereo view of the carboxylate-containing residues (Glu, Asp) shown in red and numbered according to sequence number and of positively charged residues (Lys, Arg) at neutral pH in blue. The heme is shown in green. RHS: Horse heart cytochrome *c*. LHS: Bonito cytochrome *c*.

of the studied cytochromes, i.e., in horse heart (RHS) and bonito (LHS). Striking is the higher level of charge interactions present in horse heart cytochrome *c*: the majority of its carboxylates are within strong interaction distance from a charged R-group (cf. Table 5). This is not the case for the bonito cytochrome: in this case, only two carboxylates—belonging to Glu 61 and Asp 93—are exposed to nearby R-group charges (4.36 and 3.02 Å, respectively). We do not wish to argue that these specific carboxylates should be associated with one of the two deconvoluted experimental ν^{a}_{oco} populations but simply wish to point out that we are able to identify carboxylates that are subject to monopolic interaction in the bonito species and that they would be likely to shift spectrally as a result of this interaction.

This interpretation—assigning different ν^{a}_{oco} populations to carboxylates experiencing a different charge environment—is supported both by theory [42–44]—describing the effects of uniform electric fields on the infrared spectra of carbonyls—and by experiment, such as recent FTIR work [45] showing that a Glu residue exposed to charge within a protein will have an abnormally high pK_a as a result of electrostatic repulsion when compared to the pK_a of a carboxylate not exposed to charge interactions. Caution must be exercised however, in that the field exerted by a protein matrix is non-uniform in nature. But this factor was fully addressed by Augspurger et al. [46] who studied the effect of external perturbing electrical potentials on the vibrational frequency of the CO ligand in carbonmonoxy heme proteins, comparing uniform electric fields, gradient-type fields and point charges arranged as dipoles. They concluded that the electrical perturbation arising from almost any molecular charge distribution has a uniform field component which is predominantly responsible for any vibrational frequency perturbation.

5. Conclusions

The deconvoluted IR carboxylate stretches and the electrostatic calculations presented here provide evidence for the existence of a correlation between charge phenomena and observable vibrational marker bands. This finding represents a promising approach for the conformational investigation of protein regions characterized by different charge environments. In our work, this correlation was more obvious for one cytochrome (bonito) and, although much more complex in the case of the two other cytochromes under consideration (horse heart and yeast), a case could still be made for the occurrence

of different carboxylate populations subjected to different electrostatic influence.

Acknowledgements

This research was supported by NIH (National Institute of Health) Grant No. RO1 GM55004-01 and by NIGMS (National Institutes of General Medical Science) Grants No. RO1-GM48130 and GM54105.

References

[1] S.Y. Venyaminov, N.N. Kalnin, Biopolymers 30 (1990) 1243.
[2] E. Goormaghtigh, V. Cabiaux, J.-M. Ruysschaert, Subcellular Biochem. 23 (1994) 329.
[3] J.L.R. Arrondo, A. Muga, J. Castresana, F.M. Goni, Prog. Biophys. Mol. Biol. 59 (1993) 23.
[4] W.W. Wright, M. Laberge, J.M. Vanderkooi, Biochemistry 36 (1997) 14724.
[5] K. Sharp, B. Honig, Ann. Rev. Biophys. Chem. 19 (1990) 301.
[6] B. Honig, K.A. Sharp, A.-S. Yang, J. Phys. Chem. 97 (1993) 1101.
[7] J.B. Matthew, Ann. Rev. Biophys. Biophys. Chem. 14 (1985) 387.
[8] S.C. Harvey, Proteins Struct. Funct. Genet. 5 (1989) 78.
[9] A. Warshel, J. Aqvist, Ann. Rev. Biophys. Biophys. Chem. 20 (1991) 267.
[10] J. Novotny, R. Bruccoleri, M. Karplus, J. Molec. Biol. 177 (1984) 787.
[11] C. Tanford, The Hydrophobic Effect, 2nd edn., Wiley, New York, 1980.
[12] D. Eisenberg, A.D. McLachlan, Nature 319 (1986) 199.
[13] S.H. Northrup, J.O. Boles, J.C.L. Reynolds, J. Phys. Chem. 91 (1987) 5991.
[14] S. Northrup, T.G. Wensel, C.F. Meares, J.J. Wendoloski, J.B. Matthew, Proc. Natl. Acad. Sci., USA 87 (1990) 9503.
[15] D. Bashford, K. Gerwert, Biochemistry 29 (1992) 10219.
[16] T. Takahashi, H. Nakamura, A. Wada, Biopolymers 23 (1992) 897.
[17] A.-S. Yang, M.R. Gunner, R. Sampogna, K. Sharp, B. Honig, Proteins 15 (1993) 252.
[18] D. Sitkoff, K.A. Sharp, B. Honig, J. Phys. Chem. 98 (1994) 1978.
[19] D. Sitkoff, University of Pennsylvania, Philadelphia, PA, 1995.
[20] H. Anni, J.M. Vanderkooi, K.A. Sharp, T. Yonetani, S.C. Hopkins, L. Herenyi, J. Fidy, Biochemistry 33 (1994) 3475.
[21] M. Laberge, J.M. Vanderkooi, K.A. Sharp, J. Phys. Chem. 100 (1996) 10793.
[22] M. Laberge, K. Sharp, J. Vanderkooi, in: B. Koppenhofer, U. Epperlein (Eds.), Chemistry and Computing: Terra Incognita or Land of Unlimited Opportunities? Eberhard-Karls U., Tuebingen, 1996.
[23] C.R. Lancaster, H. Michel, B. Honig, M.R. Gunner, Biophys. J. 70 (1996) 2469.
[24] E.G. Alexov, M.R. Gunner, Biophys. J. 74 (1997) 2075.
[25] M. Gilson, K. Sharp, B. Honig, J. Comp. Chem. 9 (1988) 327.
[26] F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, Eur. J. Biochem. 80 (1977) 319.
[27] N. Tanaka, T. Yamane, T. Tsukihara, T. Ashida, M. Kakudo, J. Biochem. (Tokyo) 77 (1975) 145.
[28] G.W. Bushnell, G.V. Louie, G.D. Brayer, J. Mol. Biol. 214 (1990) 585.
[29] G.V. Louie, G.D. Brayer, J. Mol. Biol. 214 (1990) 527.
[30] B. Jayaram, K.A. Sharp, B. Honig, Biopolymers 28 (1989) 975.
[31] A. Nicholls, B. Honig, J. Comp. Chem. 12 (1991) 435.
[32] A. Nicholls, K.A. Sharp, B. Honig, Proteins 11 (1991) 281.
[33] M. Laberge, K.A. Sharp, J.M. Vanderkooi, J. Phys. Chem. 101 (1997) 7364.
[34] D. Hecht, L. Tadesse, L. Walters, J. Am. Chem. Soc. 115 (1993) 3336.
[35] H.H.J. de Jongh, E. Goormaghtigh, J.-M. Ruysschaert, Biochemistry 34 (1995) 172.
[36] W. Liptay, in: E.C. Lim (Ed.), Excited States, Vol. 1, Academic Press, New York, 1974, p. 129.
[37] D.J. Lockhart, S.G. Boxer, Proc. Natl. Acad. Sci., USA 85 (1988) 107.
[38] T.R. Middendorf, L.T. Mazzola, K. Lao, A. Steffen, S.G. Boxer, Biochim. Biophys. Acta 1143 (1993) 223.
[39] M.A. Steffen, K. Lao, S.G. Boxer, Science 264 (1994) 810.
[40] A. Chattopadyay, S.G. Boxer, J. Am. Chem. Soc. 117 (1995) 1449.
[41] S. Maiti, B.R. Walker, R. Cowen, R. Pippenger, C.C. Moser, P.L. Dutton, R.M. Hochstrasser, Proc. Natl. Acad. Sci., USA 91 (1994) 10360.
[42] K. Ashley, S. Pons, Chem. Rev. 88 (1988) 673.
[43] D.K. Lambert, J. Phys. Chem. 89 (1988) 3847.
[44] J. Marti, A. Liedos, J. Bertran, M. Duran, J. Comput. Chem. 13 (1992) 821.
[45] J. Davoodi, W.W. Wakarchuk, R.L. Campbell, P.R. Carey, W.K. Surewicz, Eur. J. Biochem. 232 (1995) 839.
[46] J.D. Augspurger, C.E. Dykstra, E. Oldfield, J. Am. Chem. Soc. 113 (1991) 2447.