Decomposition of Protein Tryptophan Fluorescence Spectra into Log-Normal Components. II. The Statistical Proof of Discreteness of Tryptophan Classes in Proteins

Yana K. Reshetnyak and Edward A. Burstein
Institute of Theoretical and Experimental Biophysics, Russia Academy of Sciences, Pushchino, Moscow Region, Russia 142290

ABSTRACT The physical causes for wide variation of Stokes shift values in emission spectra of tryptophan fluorophores in proteins have been proposed in the model of discrete states (Burstein, E. A., N. S. Vedenkina, and M. N. Ivkova. 1973. *Photochem. Photobiol.* 18:263–279; Burstein, E. A. 1977a. Intrinsic Protein Luminescence (The Nature and Application). *In* Advances in Science and Technology (Itogi Nauki i Tekhniki), Biophysics Vol. 7. VINITI, Moscow [In Russian]; Burstein, E. A. 1983. *Molecular Biology (Moscow)* 17:455–467 [In Russian; English translation]). It was assumed that the existence of the five most probable spectral classes of emitting tryptophan residues and differences among the classes were analyzed in terms of various combinations of specific and universal interactions of excited fluorophores with their environment. The development of stable algorithms of decomposition of tryptophan fluorescence spectra into log-normal components gave us an opportunity to apply two mathematically different algorithms, **SI**mple fitting with **M**ean-**S**quare criterion (SIMS) and **PH**ase-plot-based **RE**solving with **Q**uenchers (PHREQ) for the decomposition of a representative set of emission spectra of proteins. Here we present the results of decomposition of tryptophan emission spectra of >100 different proteins, some in various structural states (native and denatured, in complexes with ions or organic ligands, in various pH-induced conformations, etc.). Analysis of the histograms of occurrence of >300 spectral log-normal components with various maximum positions confirmed the statistical discreteness of several states of emitting tryptophan fluorophores in proteins.

INTRODUCTION

Tryptophan fluorescence is widely used to study the location, physical and dynamic properties of microenvironment of indole fluorophores, and the structural features and behavior of the protein molecule as a whole (Burstein, 1976, 1977a, 1983; Lakowicz, 1983; Demchenko, 1986). Depending on the environment of tryptophan residues in proteins, the maximum position (λ_m) and quantum yield (q) of tryptophan fluorescence could vary widely, from 308 to 353 nm and from 0.4 to immeasurably low, respectively.

In 1967 Konev put forward the hypothesis of the existence of two main classes of tryptophan residues in proteins, which possess discrete values of fluorescent parameters λ_m and q (Konev, 1967; Volotovski and Konev, 1967). One of the classes included tryptophan fluorophores inside the protein in a low-polar hydrophobic environment with a shorter-wavelength position of fluorescent maximum (λ_m of ~330 nm) and rather low quantum yield (0.04 to 0.07). The second class consisted of exposed tryptophan residues in a high-polar aqueous environment with long-wavelength position of spectra (λ_m of ~350 nm) and quantum yield equal or higher than that of free aqueous tryptophan (~0.13–0.17). This hypothesis had been based on the observation that protein spectrum shifts toward 350–353 nm upon denaturation by urea, and toward 330–332 nm upon addition of anionic detergents in acidic solutions. However, this model could not explain the existence of proteins with a high quantum yield (for example, 0.20–0.27 for serum albumins (Longworth, 1971)) and intermediate spectral maximum positions (341.5 nm) for some single-tryptophan-containing proteins, such as human serum albumin (Ivkova et al., 1971).

In 1973 one of us, with co-workers, revised and extended Konev's hypothesis of discrete classes of tryptophan residues in proteins and suggested a new model using some additional spectral parameters and approaches (Burstein et al., 1973). The spectral bandwidth at the half-maximal amplitude $\Delta\lambda$ was taken as an additional parameter. The linear relationship was found between values of $\Delta\lambda$ and maximum position λ_m for tryptophan and other C-3-substituted indole derivatives in solvents of various polarities. These spectra could be regarded as a series of "elementary" components representing the emission bands of individual tryptophan residues in proteins. Existence of a spectral shift accompanying the quenching of protein fluorescence by ionic solutes (NO_3^-, I^-, Cs^+) indicated the multicomponent character of a protein emission spectrum. It was demonstrated that the spectra of proteins, which are shifted upon quenching, possess $\Delta\lambda$ values exceeding those of "elementary" components. The analysis of the differences between initial spectra of such proteins and those after 20% quenching revealed that the best-quenched components have the longest-wavelength position at ~340 nm in native proteins, which is ~10 nm shorter than that postulated in the two-state model. However, the tryptophans in proteins denatured by urea or guanidinium chloride have a spectrum with λ_m of ~350 nm.

Received for publication 12 January 2001 and in final form 7 June 2001.
Address reprint requests to (present address) Dr. Yana K. Reshetnyak, Department of Molecular Biology and Immunology, Institute for Cancer Research, University of North Texas Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, TX 76107. Tel.: 817-735-5417; Fax: 817-735-2133; E-mail: yreshetn@hsc.unt.edu.

0006-3495/01/09/1710/25 $2.00

These results allowed us to develop an extended model of discrete states (classes) of tryptophan residues in proteins, which assumed the existence of five statistically most probable classes (Burstein et al., 1973; Burstein, 1977a, 1983). According to the model, the following discrete classes of tryptophan residues were predicted to be most probable in proteins (Burstein, 1977b, 1983):

1. Class A (λ_m = 308 nm, structured spectra); the fluorophores, which do not form hydrogen-bound complexes in the excited state (exciplexes) (Hershberger et al., 1981) with solvent or neighboring protein groups;
2. Class S (λ_m = 316 nm, structured spectra) includes the buried tryptophan residues that can form the exciplexes with 1:1 stoichiometry;
3. Class I (λ_m = 330–332 nm, $\Delta\lambda$ = 48–50 nm) represents the buried fluorophores that can form the exciplexes with 2:1 stoichiometry;
4. Class II (λ_m = 340–342 nm, $\Delta\lambda$ = 53–55 nm) represents the fluorophores exposed to the bound water possessing very long dipole relaxation time, which precludes completing the relaxation-induced spectral shift during the excited-state lifetime;
5. Class III (λ_m = 350–353 nm, $\Delta\lambda$ = 59–61 nm) contains rather fully exposed fluorophores surrounded by highly mobile water completely relaxing during the excitation lifetime, which makes their spectra almost coinciding with those of free aqueous tryptophan.

At that time, there was constructed and used the calibrated diagram $\Delta\lambda$ versus λ_m for estimating the contributions of three most frequent model classes (I, II, and III) of tryptophan residues in the fluorescence spectrum of a protein (Burstein et al., 1973). The application of this hypothesis was rather effective in the interpretation of protein tryptophan fluorescence data in various biophysical and molecular-biology studies. However, the idea of the existing of discrete classes of tryptophans in proteins was based on the analysis of very limited numbers of proteins. Now it seems reasonable to reappraise the hypothesis of discrete states applying more recent analytic and computing methods to the wide statistical database of protein spectra.

Such a reappraisal is possible now due to progress in protein preparation techniques and methods of component analysis of protein emission spectra. It was shown that the fluorescence spectra of tryptophan and its derivatives may be accurately described analytically by a log-normal function (Burstein, 1976), proposed initially for the absorption spectra by Siano and Metzler (1969). Thereupon, it was found that the shape of the emission spectra of tryptophan both in solutions and in a protein depends only on the spectral maximum position λ_m in a known manner (Burstein and Emelyanenko, 1996). This fact allowed us to develop effective algorithms of reliable decomposition of tryptophan fluorescence spectra of proteins into individual "elementary" log-normal components (Abornev and Burstein, 1992; Burstein et al., 2001).

Here we present the results of such decomposition of tryptophan emission spectra of >100 different proteins, some in various structural states (native and denatured, in complexes with ions or organic ligands, in various pH-induced conformations, etc.). The database of obtained log-normal components allowed us to check the hypothesis of the discrete classes of tryptophan residues on the basis of a representative set of proteins. Analysis of the histograms of occurrence of spectral components with various maximum positions confirmed the statistical discreteness of several states of emitting tryptophan fluorophores in proteins. The discrete classes observed here are similar to those proposed in the model of discrete states in 1973–1977.

MATERIALS AND METHODS

Fluorescence spectra

The large majority of protein fluorescence spectra used in this work and listed in Table 1 were taken from the archive of experimental data of the Laboratory of Functional Biophysics of Proteins (Institute of Theoretical and Experimental Biophysics of the Russia Academy of Sciences). The spectra have been published and/or analyzed in various special publications cited in the column "References." These publications also contain the data about protein preparations and techniques of fluorescence measurements. If the spectra were measured elsewhere and given to us, they were recorrected for instrument spectral sensitivity curves using the emission spectra of aqueous tryptophan solution as a standard. Such a recorrection was necessary because the component analysis algorithms were developed based on the model spectra measured with our lab-made instrument (Burstein et al., 1973; Bukolova-Orlova et al., 1974) with a carefully estimated spectral sensitivity curve (Burstein and Emelyanenko, 1996).

The preparations of wheat germ agglutinin, inorganic pyrophosphatase (pyrophosphate phosphohydrolase) from baker's yeast, monellin from *Dioscoreophyllum cummensii* fruits, *Escherichia coli* alkaline phosphatase, and subtilisin BPN′ produced by Sigma (St. Louis, MO) were kindly given to us by Dr. J. Borejdo (University of North Texas Health Science Center, Fort Worth, TX). Human and bovine carbonic anhydrases II, staphylococcal serine protease, and subtilisin Carlsberg (Sigma) were a kind gift from Dr. S. S. Lehrer (Boston Biomedical Research Institute, Boston, MA). The preparation of bovine carboxypeptidase A was from Reanal (Hungary). The yeast 3-phosphoglycerate kinase was from Sigma. Porcine pancreatic lipase was from Serva. Hen egg white ovalbumin was from Reakhim (Olaine, Latvia). Actin from dog heart was prepared by one of us (Ya. K. R.) according to Pardee and Spudich (1982).

Fluorescence measurements

Fluorescence spectra of the several proteins marked in Table 1 (in the "References" column) as "Present work" were measured by us during the last six years. Acrylamide, KCl, KI, CsCl, Tris, and ATP preparations used in these experiments were ultra-pure grade of Russian or Soviet production or Sigma.

Steady-state emission spectra were recorded on the lab-made spectrofluorimeter (Burstein et al., 1973; Bukolova-Orlova et al., 1974) with collection of emitted light from the cell front face, that allowed application of the strict correction function for screening and reabsorption inner filter effects (Burstein, 1968). The fluorescence was excited with the mercury

TABLE 1 The list of studied proteins, their conventional code name, content of tryptophan residues, and references containing the information about the fluorescence spectra used here for decomposition into log-normal components

No.	Protein Name and Conditions	Protein Codes	No. of Trps	References
1	Acetylcholine esterase (bovine brain); pH 4.9, pH 9.2	ACH	14	Burstein et al., 1973
2	F-actin (dog heart muscle); pH 7.5	ACD.F	4	Present work
3	G-actin (dog heart muscle); pH 7.4	ACD.G	4	Present work
4	F-actin, (rabbit skeletal muscle); pH 7.8	ACR.F	4	Burstein et al., 1973; Burstein, 1977a
5	*F-actin, (rabbit skeletal muscle);* 8 M urea	*ACR.F.UR*	4	Vedenkina et al., 1968
6	G-actin, (rabbit skeletal muscle); 2 mM ATP, 2 mM $CaCl_2$, pH 8.0	ACR.G	4	Vedenkina et al., 1968
7	Actomyosin (rabbit skeletal muscle); 0.4–0.5 M KCl	ACM	24	Burstein et al., 1973; Burstein, 1977a; Vedenkina et al., 1968
8	*Actomyosin (rabbit skeletal muscle);* 8 M urea	*ACM.UR*	24	Burstein et al., 1973; Burstein, 1977a; Vedenkina et al., 1968
9	Aminoacylase-I (porcine kidneys); pH 7.5	ACY	8	Abornev, 1993
10	Agglutinin (wheat germ); pH 7.3	AWG	3	Present work
11	α-1-Antitrypsin (human); pH 7.2	A1AT	2	Permyakov and Deikus, 1995
12	Serum albumin (bovine) F-form; pH 3.0	ASB.F	2	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1971
13	Serum albumin (bovine) N-form; pH 5.7–6.0	ASB.N	2	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1971
14	Serum albumin (bovine) N-form; 0.5 M isobutanol, pH 5.7	ASB.N.IB	2	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1971
15	Serum albumin (human) F-form; pH 3.1	ASH.F	1	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1971; Nyamaa et al., 1985
16	Serum albumin (human) N-form; pH 5.5	ASH.N	1	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1971; Nyamaa et al., 1985
17	Serum albumin (human) N-form; 0.5 M isobutanol, pH 5.5	ASH.N.IB	1	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1971
18	*Serum albumin (human);* 4.5 and 9 M urea	*ASH.UR*	1	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1971
19	Alcohol dehydrogenase (yeast); pH 7.5	ADY	2	Abornev, 1993
20	Aminoacyl-tRNA-synthase, tryptophanyl, A2 and K2 forms (*E. coli*)	ATRN.TRP A2-K2	2	Burstein et al., 1973; Vedenkina and Burstein, 1970
21	Aminoacyl-tRNA-synthase, tryptophanyl, E1+E2 form (*E. coli*)	ATRN.TRP E1+E2	2	Burstein et al., 1973; Vedenkina and Burstein, 1970
22	*Aminoacyl-tRNA-synthase, tryptophanyl, E1+E2 form (E. coli);* 8 M urea	*ATRN.TRP UR*	2	Burstein et al., 1973; Vedenkina and Burstein, 1970
23	Aminoacyl-tRNA-synthase, tyrosyl (bovine liver) dimer 2×40 kDa; pH 7.5	ATRN.TYR.40		Klimenko et al., 1991
24	Aminoacyl-tRNA-synthase.tyrosyl (bovine liver) dimer 2×59 kDa; pH 7.6	ATRN.TYR.59	2×3	Klimenko et al., 1991
25	α-Amylase (*Aspergillus oryzae*); pH 7.5	AMY	10	Abornev, 1993
26	Annexin VI* (human); pH 7.0	AX6	10	Abornev, 1993
27	L-Asparaginase (*E. coli* B); pH 4.3	ASP.AC	1	Burstein, 1977a, b; Burstein, 1983
28	L-Asparaginase; (*E. coli* B); pH 7.8–8.0, 293 K and 77 K	ASP.N	1	Burstein, 1977a, b; Burstein, 1983
29	Azurin, (*Pseudomonas aeruginosa*), holo-form; 77 K	AZU	1	Burstein et al., 1977
30	Calcium-binding protein† (*Astacus leptodactylus* sarcoplasmic reticulum); pH 6.0, pH 8.2	CBC	2	V. I. Emelyanenko (this lab)
31	Caldesmon (chicken gizzard); pH 7.4	CDS	5	Czurylo et al., 1991
32	*Caldesmon (chicken gizzard), fragment* 25 kDa	*CDS.25*		Czurylo et al., 1991
33	Caldesmon (chicken gizzard), fragment 25 kDa + calmodulin	CDS.25.CAM		Czurylo et al., 1991
34	Carbonic anhydrase-II (bovine); pH 7.3	CAB	7	Present work
35	Carbonic anhydrase II (human); pH 7.3	CAH	7	Present work
36	Carboxypeptidase A (bovine); 1 M KCl, pH 7.3	CPA	8	Present work
37	α-Chymotrypsin (bovine); in water; pH 7.3–7.8, pH 5.3	CHT	8	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1968; Reshetnyak and Burstein, 1997b
38	*α-Chymotrypsin (bovine) in 35% dioxane*	*CHT.D35*	8	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1968
39	*α-Chymotrypsin (bovine) in 60–80% dioxane*	*CHT.D60-80*	8	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1968

No.	Protein Name and Conditions	Protein Codes	No. of Trps	References
40	Chymotrypsinogen A (bovine); in water; pH 7.5 and pH 8.6	CHG	8	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1968; Reshetnyak and Burstein, 1997b
41	*Chymotrypsinogen A (bovine) in 35% dioxane*	*CHG.D35*	8	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1968
42	*Chymotrypsinogen A (bovine) in 60–70% dioxane*	*CHG.D60-70*	8	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1968
43	*Chymotrypsinogen A (bovine);* 8 M urea	*CHG.UR*	8	Burstein et al., 1973; Burstein, 1977a; Ivkova et al., 1968
44	Creatine phosphokinase (rabbit skeletal muscle); pH 6.0	CRK	2×4	Burstein et al., 1973; Burstein, 1977a
45	Elastase (human neutrophils)	ENH	3	Mely et al., 1997
46	Fibrinogen[‡] (bovine), fragment E; pH 7.5	FGE		V. I. Emelyanenko (this lab)
47	Fibronectin; pH 8.0	FBN	39	Koteliansky et al., 1981
48	Filamin (human); pH 7.8	FIL	17	Koteliansky et al., 1982
49	Glucagon	GLG	1	Ababou, 1998
50	Glyceraldehyde-3-phosphate dehydrogenase, apo- and holo-forms (porcine skeletal muscle); ATP, pH 7.5, pH 8.5	GPD	3	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970
51	Protein G (*Streptococcus* sp.)	GPS	1	Ababou, 1998
52	*Hemoglobin[§] (human), α-chains, reduced and carboxymethylated;* pH 7.5	*HBA.CM*	3	Abornev, 1993
53	Immunoglobulin G (bovine); pH 6.6	IGG		Burstein et al., 1973; Burstein, 1977a; Shanin et al., 1968
54	*Immunoglobulin G (bovine);* 8 M urea	*IGG.UR*		Burstein et al., 1973; Burstein, 1977a; Shanin et al., 1968
55	Inorganic pyrophosphatase (baker's yeast); pH 7.3	IOP	6	Present work
56	α-Lactalbumin (cow milk); 1 mM $CaCl_2$, pH 7.4	LAB	4	Permyakov et al., 1985
57	α-Lactalbumin (human milk); 1 mM $CaCl_2$, pH 7.4	LAH	3	Permyakov et al., 1991
58	β-Lactoglobulin (cow milk); pH 2.4, pH 4.6	LGB.AC	2	Burstein et al., 1973; Burstein, 1977a; Kaplanas et al., 1973, 1975
59	β-Lactoglobulin (cow milk); pH 6.4	LGB.N	2	Burstein et al., 1973; Burstein, 1977a; Kaplanas et al., 1973, 1975
60	β-Lactoglobulin (cow milk); pH 8.1	LGB.AL	2	Burstein et al., 1973; Burstein, 1977a; Kaplanas et al., 1975
61	β-Lactoglobulin[¶] (Bactrian camel milk); 0.01 M glycin, pH 2.0–3.0	LGC.AC	3	Abornev, 1993
62	β-Lactoglobulin[¶] (Bactrian camel milk); pH 7.0	LGC.N	3	Abornev, 1993
63	Lipase (porcine pancreas); pH 7.4	LPP	6	Present work
64	Lipase (wheat germ); pH 7.5	LWG		Abornev, 1993
65	Lysozyme (hen egg white); HCl, pH 2.2; glycin, pH 3.5	LZC.AC	6	Burstein et al., 1973; Burstein, 1977a
66	Lysozyme (hen egg white); pH 7.4–7.8	LZC.N	6	Burstein et al., 1973; Burstein, 1977a
67	*Lysozyme (hen egg white);* 8 M urea, pH 2.6, pH 8.0	*LZC.UR*	6	Burstein et al., 1973; Burstein, 1977a
68	Mastoparan (wasp venom)	MSP	1	Ababou, 1998
69	Melittin,[‖] monomer (*Apis mellifera* venom); water	MLT.MON	1	Permyakov and Deikus, 1995; Reshetnyak and Burstein, 1997a
70	Melittin,[‖] tetramer (*Apis mellifera* venom); 1.5–1.6 M KCl, pH 7.4	MLT.TETR	4×1	Permyakov and Deikus, 1995; Reshetnyak and Burstein, 1997a
71	Melittin,[‖] (*Apis mellifera* venom) + calmodulin + $CaCl_2$, pH 7.5	MLT.CAM	1 or 4×1	Permyakov and Deikus, 1995
72	Melittin, monomer (*Apis mellifera* venom) with liposomes; 2 M NaCl	MLT.LS	1	Ladokhin, 1990
73	H-meromyosin (rabbit smooth muscle)	HMM	11	Vishnevskaya et al., 1976
74	L-meromyosin, fraction 1, LMM59 and LMM77 (rabbit skeletal muscle)	LMM	2	Vishnevskaya et al., 1976; Zhou et al., 1998
75	Monellin (*Dioscoreophyllum cummensii*); pH 7.3	MON	1	Present work
76	Myoglobin** (beaver heart), apo-form; pH 7.5	MGB.APO	2	Abornev, 1993
77	Myosin (rabbit skeletal muscle); 0.5 M KCl	MSS	20	Burstein et al., 1973; Burstein, 1977a
78	*Myosin (rabbit skeletal muscle);* 8 M urea, pH 7.8	*MSS.UR*	20	Burstein et al., 1973; Burstein, 1977a
79	Myosin subfragment 1[††] (rabbit skeletal muscle); 50 mM KCl, 1 mM $MgCl_2$, pH 7.5	MS1	5	Reshetnyak et al., 2000
80	Myosin subfragment 1[††] (skeletal muscle, rabbit); 50 mM KCl, 1 mM $MgCl_2$, 0.2 mM ATP, pH 7.5	MS1.ATP	5	Reshetnyak et al., 2000

No.	Protein Name and Conditions	Protein Codes	No. of Trps	References
81	Neurotoxin I (cobra *Naja naja oxiana* venom); pH 4.6–6.0	NO1	2	Bukolova-Orlova et al., 1974; Bukolova-Orlova et al., 1976
82	Neurotoxin II (cobra *Naja naja oxiana* venom); pH 6.0	NO2	2	Bukolova-Orlova et al., 1974; Bukolova-Orlova et al., 1976
83	Nuclease (*Staphylococcus aureus*); pH 7.5	NST	1	Ababou, 1998
84	Nucleoside diphosphate kinase α[‡‡] (rat, recombinant); pH 7.7–9.3	NDPK.A.8	3Trp + $TyrO^-$	Orlov et al., 1997; Orlov et al., 1999
85	Nucleoside diphosphate kinase α[‡‡] (rat, recombinant); pH 4.0–5.5	NDPK.A.5.LIS	3	Orlov et al., 1997; Orlov et al., 1999
86	Nucleoside diphosphate kinase α[‡‡] (rat, recombinant); 0.5–1.0 M KCl, pH 4.0–5.5	NDPK.A.5.HIS	3	Orlov et al., 1997; Orlov et al., 1999
87	Nucleoside diphosphate kinase β[‡‡] (rat, recombinant); pH 7.5–9.5	NDPK.B.8	3	Orlov et al., 1997; Orlov et al., 1999
88	Nucleoside diphosphate kinase β[‡‡] (rat, recombinant); pH 4.5–5.2	NDPK.B.AC	3	Orlov et al., 1997; Orlov et al., 1999
89	Ovalbumin (hen egg white); low ionic strength, pH 7.4	OVH.LIS	3	Present work
90	Ovalbumin (hen egg white); high ionic strength, pH 7.4	OVH.HIS	3	Present work
91	Papain (*Carica papaya*); glycin, pH 3.1–4.0	PAP.AC	5	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970; Vedenkina et al., 1971
92	Papain (*Carica papaya*); water, pH 6.8–7.0	PAP.N	5	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970; Vedenkina et al., 1971
93	Papain (*Carica papaya*); glycin, pH 9.3–9.8	PAP.AL	5	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970; Vedenkina et al., 1971
94	*Papain (Carica papaya);* 8 M urea, pH 6.2	*PAP.UR*	5	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970; Vedenkina et al., 1971
95	Parvalbumin II[§§] (cod *Gadus morrhua*), acid form; pH 2.0	PAC.AC	1	Abornev, 1993
96	Parvalbumin II[§§] (cod *Gadus morrhua*), apo-form; pH 7.7	PAC.APO	1	Abornev, 1993
97	Parvalbumin II[§§] (cod *Gadus morrhua*), Mg-form; 1 mM $MgCl_2$, pH 7.7	PAC.MG	1	Abornev, 1993
98	Parvalbumin II[§§] (cod *Gadus morrhua*), Ca-form; 1 mM $CaCl_2$, pH 7.7	PAC.CA	1	Abornev, 1993
99	Parvalbumin (whiting *Gadus merlangus*), apo-form; pH 6.4	PAM.APO	1	Permyakov et al., 1980
100	Parvalbumin (whiting *Gadus merlangus*); μ 0.5 M, pH 8.8	PAM	1	Permyakov et al., 1980
101	Parvalbumin[§§§] (Alaska pollack, *Theragra chalcogramma*), apo-form; pH 8.0	PMT.APO	1	V. I. Emelyanenko (this lab)
102	Parvalbumin[§§§] (Alaska pollack, *Theragra chalcogramma*), Ca-form; 1 mM $CaCl_2$, pH 8.0	PMT.CA	1	V. I. Emelyanenko (this lab)
103	Parvalbumin[§§§] (Alaska pollack, *Theragra chalcogramma*), Mg-form; 1 mM $MgCl_2$, pH 8.0	PMT.MG	1	V. I. Emelyanenko (this lab)
104	Parvalbumin[§§§] (Pacific navaga, *Eleginus gracilis*), apo-form; pH 8.0	PAN.APO	1	V. I. Emelyanenko (this lab)
105	Parvalbumin[§§§] (Pacific navaga, *Eleginus gracilis*), Ca-form; 1 mM $CaCl_2$, pH 8.0	PAN.CA	1	V. I. Emelyanenko (this lab)
106	Parvalbumin[§§§] (Pacific navaga, *Eleginus gracilis*), Mg-form, 1 mM $MgCl_2$, pH 8.0	PAN.MG	1	V. I. Emelyanenko (this lab)
107	Pepsin (bovine); pH 1.3–2.8	PEB.1-3	5	Burstein et al., 1973; Burstein, 1977a
108	Pepsin (bovine); water or pH 4.0–5.2	PEB.4-5	5	Burstein et al., 1973; Burstein, 1977a
109	*Pepsin (bovine);* pH 5.2–9.0	*PEB.6-9*	5	Burstein et al., 1973; Burstein, 1977a
110	*Pepsin (bovine);* 8 M urea, pH 5.2	*PEB.UR*	5	Burstein et al., 1973; Burstein, 1977a
111	Phosphatase alkaline (*E. coli*); pH 7.1	PHA	2×3	Present work
112	3-Phosphoglycerate kinase (yeast); pH 7.4	PGK	2	Present work
113	Phospholipase A_2 (bovine pancreas); pH 6.0, pH 7.4	PLB	1	Permyakov and Deikus, 1995
114	Phospholipase A_2 (porcine pancreas); pH 7.4	PLS	1	Permyakov and Deikus, 1995

No.	Protein Name and Conditions	Protein Codes	No. of Trps	References
115	Phospholipase A_2 (cobra *Naja naja oxiana* venom); low ionic strength, pH 7.5	PAO.LIS	3	Orlova, 1979
116	Phospholipase A_2 (cobra *Naja naja oxiana* venom); high ionic strength, pH 7.5	PAO.HIS	3	Orlova, 1979
117	*Phospholipase A_2 (cobra Naja naja oxiana venom);* 6 M GdnHCl, pH 7.5	*PAO.GDN*	3	Orlova, 1979
118	Phospholipase C, phosphatidylcholine-specific, holo-form (*Bac. cereus*); pH 7.1	PLC	9	D. B. Veprintsev (this lab)
119	Prolactin (bovine); pH 8.5	PRL	2	Permyakov et al., 1997
120	Propionylcholine esterase (human); pH 7.5	PCE		Abornev, 1993
121	HIV-1 protease; low ionic strength, T=20−67°C, 0–50% glycerol, pH 8.0	PRH.LIS	2+2	A. I. Kornelyuk[¶¶]
122	HIV-1 protease; high ionic strength, pH 8.0	PRH.HIS	2+2	A. I. Kornelyuk[¶¶]
123	HIV-1 protease with pepstatin, pH 8.0	PRH.PST	2+2	A. I. Kornelyuk[¶¶]
124	Serine protease (*Staphylococcus aureus*); pH 7.5	PRS	2	Present work
125	Protease K (*Tritirachium album* Limber); pH 7.5	PRK	2	Reshetnyak and Burstein, 1997a
126	Pyruvate kinase (rabbit skeletal muscle); pH 7.5	PYK	3	Abornev, 1993
127	Pyruvate kinase (rabbit skeletal muscle) PEG conjugate	PYK.PEG	3	Abornev, 1993
128	Recoverin MP26, apo-form; pH 8.1	REC.APO	3	S. E. Permyakov (this lab)
129	Recoverin MP26 + $CaCl_2$, pH 8.1	REC.CA	3	S. E. Permyakov (this lab)
130	Red cells ghosts (human); pH 7.8	RCG		Orlova, 1979
131	Ribonuclease C_2 (*Aspergillus clavatus*); 0–20% glycerol, pH 4.0, pH 6.0, pH 8.0–8.4	RNC	1	Burstein, 1977a; Burstein, 1983; Grishchenko et al., 1976
132	Ribonuclease intracellular[‖‖] (*Aspergillus clavatus*); pH 8.0	RNI		V. M. Grishchenko (this lab)
133	RNAse T_1*** (*Aspergillus oryzae*); 20% glycerol, pH 7–8	RNT	1	Burstein, 1977a; Burstein, 1983; Grishchenko et al., 1976
134	Ribosomal protein S4[†††] (*E. coli*); pH 7.0	RS4	1	T. L. Bushueva (this lab)
135	Ribosomal protein S7[†††] (*E. coli*); pH 7.0	RS7	2	Burstein et al., 1973; Burstein, 1977a
136	Ribosome-inactivating protein[‡‡‡] (ricin agglutinin, *Ricinus communis*); pH 4.0–7.0	RIP		T. L. Bushueva[§§§]
137	Ricin chain A (*Ricinus communis*); pH 4.0–7.0	RCA	1	Bushueva and Tonevitsky, 1987a, b; Bushueva and Tonevitsky, 1988; Bushueva et al., 1990
138	Ricin chain B (*Ricinus communis*); pH 4.0	RCB.4	9	Bushueva and Tonevitsky, 1987a, b; Bushueva and Tonevitsky, 1988; Bushueva et al., 1990
139	Ricin chain B (*Ricinus communis*); pH 5.0–7.0	RCB.5-7	9	Bushueva and Tonevitsky, 1987a, b; Bushueva and Tonevitsky, 1988; Bushueva et al., 1990
140	Ricin chain B (*Ricinus communis*) + lactose; pH 4.0	RCB.4.LAC	9	Bushueva and Tonevitsky, 1987a, b; Bushueva and Tonevitsky, 1988; Bushueva et al., 1990
141	Ricin (*Ricinus communis*) (chain A + chain B) with and without lactose; pH 4.0–7.0	RCC	1+9	Bushueva and Tonevitsky, 1987a, b; Bushueva and Tonevitsky, 1988; Bushueva et al., 1990
142	Somatotropin, growth hormone (human); pH 8.2	STP	1	Permyakov and Deikus, 1995
143	Spectrin, α-chain and β-chain	SPC	42+42	T. L. Bushueva[§§§]
144	Subtilisin BPN′; pH 7.3	BPN	3	Present work
145	Subtilisin Carlsberg; pH 7.5	SLC	1	Present work
146	Telokin KRP, myosin light-chain kinase (chicken gizzard)	KRP	1	Bushueva et al., 1999
147	*Tobacco mosaic virus, protein A;* 8 M urea	*TMV.UR*	3	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970
148	Troponin T (rabbit skeletal muscle); pH 6.8	TNT	2	Morozova et al., 1988
149	Trypsin (bovine); water, pH 7.3, pH 8.7	TRY	4	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970; Ivkova et al., 1968; Reshetnyak and Burstein, 1997b

No.	Protein Name and Conditions	Protein Codes	No. of Trps	References
150	*Trypsin (bovine) in 35% dioxane*	*TRY.D35*	4	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970; Ivkova et al., 1968
151	*Trypsin (bovine) in 60–70% dioxane*	*TRY.D60-70*	4	Burstein et al., 1973; Burstein, 1977a; Vedenkina and Burstein, 1970; Ivkova et al., 1968
152	Trypsinogen (bovine); pH 7.4	TRG	4	Reshetnyak and Burstein, 1997b
153	Vinculin[¶¶¶] (human)	VIN	10	Bushueva[§§§]
154	*Vinculin[¶¶¶] (human);* 8 M urea	*VIN.UR*	10	Bushueva[§§§]
155	Vipoxin, protein A (*Vipera ammodytes ammodytes*); pH 3.8–4.8	VTA.AC	1	Bukolova-Orlova et al., 1979; Bukolova-Orlova et al., 1980; Tchorbanov et al., 1977
156	Vipoxin, protein A (*Vipera ammodytes ammodytes*); high ionic strength, pH 7.2–8.0	VTA.HIS	2×1	Bukolova-Orlova et al., 1979; Bukolova-Orlova et al., 1980; Tchorbanov et al., 1977
157	Vipoxin, protein A (*Vipera ammodytes ammodytes*); low ionic strength, pH 7.2–8.0	VTA.LIS	1	Bukolova-Orlova et al., 1979; Bukolova-Orlova et al., 1980; Tchorbanov et al., 1977
158	Vipoxin, protein B (*Vipera ammodytes ammodytes*); pH 3.6, pH 6.2–8.1	VTB	2	Bukolova-Orlova et al., 1979; Bukolova-Orlova et al., 1980; Tchorbanov et al., 1977
159	Vipoxin, complex A + B (*Vipera ammodytes ammodytes*); pH 4.8, pH 8.0	VTC	1+2	Bukolova-Orlova et al., 1979; Bukolova-Orlova et al., 1980; Tchorbanov et al., 1977
160	*Viscumin (lectin ML-1; misletoe Viscum album);* pH 4.0	*VSC.AC*	2+9	Bushueva and Tonevitsky, 1988; Bushueva et al., 1988
161	Viscumin (lectin ML-1; misletoe *Viscum album*); pH 5.0–7.5	VSC.5-7	2+9	Bushueva and Tonevitsky, 1988; Bushueva et al., 1988
162	*Viscumin (lectin ML-1; misletoe Viscum album), modified by ethylammonium nitrate;* pH 4.0, pH 7.0	*VSC.MOD*	2+9	Bushueva and Tonevitsky, 1988; Bushueva et al., 1988
163	*Viscumin (lectin ML-1; misletoe Viscum album), modified by ethylammonium nitrate;* 8 M urea, pH 4.0, pH 7.0	*VSC.UR*	2+9	Bushueva and Tonevitsky, 1988; Bushueva et al., 1988

Denatured protein codes appear in italics.

*Presented by Dr. Andrzej Sobota, N. Nencki Institute of Experimental Biology, Warsaw, Poland (Bandorowicz et al., 1992; Sobota et al., 1993).

[†]Prepared and presented by Dr. G. Benzonana (Benzonana et al., 1974; Cox et al., 1976).

[‡]Presented by Dr. L. V. Medved, Institute of Biochemistry, Ukraine Academy of Science, Kiev.

[§]Prepared and presented by Dr. L. V. Abaturov, Institute of Molecular Biology, RAS, Moscow.

[¶]Prepared by Dr. L. P. Kalinichenko (this lab) from camel milk presented by Dr. D. Nyamaa, Institute of Biotechnology, Mongolian Academy of Science, Ulan Baatar, Mongolia.

[‖]Presented by Dr. A. P. Demchenko, Institute of Biochemistry, National Academy of Science of Ukraine, Kiev.

**Presented by Dr. R. I. Artyukh and Dr. E. A. Shlyapnikova, Institute of Theoretical and Experimental Biophysics, RAS, Pushchino.

[††]Prepared by Dr. O. Andreev, University of North Texas, Fort Worth, Texas.

[‡‡]Prepared and presented by Dr. N. Ya. Orlov, this lab.

[§§]Prepared by Dr. L. P. Kalinichenko (this lab), using modified Haiech method (Haiech et al., 1979).

[¶¶]Presented by A. I. Kornelyuk, Institute of Molecular Biology and Genetics, National Academy of Science of Ukraine, Kiev.

[‖‖]Prepared by Dr. S. I. Bezborodova and co-workers, Institute of Biochemistry and Physiology of Microorganisms, Academy of Science USSR, Pushchino.

***Presented by Dr. C. A. Ghiron.

[†††]Prepared and presented by Dr. A. T. Gudkov, Institute of Protein Research, Academy of Science USSR, Pushchino.

[‡‡‡]Prepared by Dr. A. G. Tonevitsky (Tonevitsky et al., 1996).

[§§§]Spectra measured by Dr. T. L. Bushueva, Cardiology Research Center, Russian Academy of Medical Science, Moscow.

[¶¶¶]Prepared by Dr. V. E. Kotelyansky (Weller et al., 1990); spectra were measured by Dr. T. L. Bushueva, Cardiology Research Center, Russian Academy of Medical Science, Moscow.

line 296.7 nm from an SVD-120A mercury lamp isolated with a monochromator. The use of the mercury lamp with narrow spectral lines in emission provided the possibility of rather simple correction for the stray light presented in the registered emission spectra. The spectral width of slits for excitation and emission did not exceed 2 nm. After correction for the instrument spectral sensitivity (Burstein and Emelyanenko, 1996), the intensities were proportional to the number of photons emitted per unit wavelength interval during a unit time interval. Spectra of almost every protein were measured in the presence of varying concentrations of external quenchers (acrylamide, KI, CsCl, and/or $NaNO_3$). When using ionic quenchers (I^-, Cs^+, or NO_3^-) the total ionic strength was kept constant (from 0.4 to 0.6 M in various experiments) by addition of KCl. In experiments with acrylamide the intensities in fluorescence spectra were corrected for the screening inner filter effect of this quencher at the excitation wavelength (Burstein, 1968).

Analysis and representation of spectral data

The decomposition of tryptophan fluorescence spectra of proteins into log-normal components was performed for sets of spectra measured at varying concentrations of quenchers using two algorithms: SIMS (**SI**mple fitting with **M**ean-**S**quare criterion) (Abornev and Burstein, 1992; Burstein et al., 2001) and PHREQ (**PH**ase-plot-based **RE**solving with **Q**uenchers, for two-component decomposition) (Burstein et al., 2001). For the decomposition of a single spectrum (without quenchers) the modified SIMS algorithm was applied (Burstein et al., 2001). Using SIMS, the spectra were independently fitted by one, two, or three components. The criterion of attaining the solution (a minimal sufficient number of components describing the spectra) was the minimal root-mean-square differences (residuals) between theoretical and experimental spectra multiplied by the number of parameters under fitting. For several proteins two solutions with different numbers of components, which had the similar values of this criterion, were selected as the best ones.

The fact that the components of composite tryptophan emission spectra of proteins are wide (~30–60 nm at half-maximal level) compared with the differences between the values of their maximum positions (from 5 to ~50 nm), might considerably reduce the reliability of decomposition. Therefore, we present in Table 2 the averaged results (values of the maximum position, $\lambda_m(i)$ and the contribution into the area under spectrum, $S(i)$ and the standard deviations of averaging) obtained separately with the mathematically different algorithms: PHREQ and SIMS. The good agreement between the parameters that were obtained with these two methods demonstrate the reliability of a result. We can note that for the absolute majority of proteins the results obtained with these two methods differ less than by errors of means. It allowed us to present unified results of the component analysis for individual proteins independent of the decomposition method.

Moreover, the majority of protein spectra were measured in the presence of various quenchers, in most cases more than once, using preparations from different sources or varying solvent conditions. To extract unified values of the parameters for any given protein or its conformer, two averaging schemes were applied to the sets of decomposition results: "Averaging" and "Mean Best Fit" (Table 2). According to the first scheme, the parameters of components were calculated as averages of all solutions obtained for all individual experiments, independent of the quality of the decomposition. Within the second scheme, only the data obtained with minimal values of the decomposition functional for each experiment were averaged.

RESULTS

In Table 2 we present the results of decomposition of tryptophan fluorescence spectra of >100 different proteins, some in various structural states (native and denatured, in complexes with ions or organic ligands, and in various pH-induced states), obtained by the above-mentioned algorithms, SIMS and PHREQ, and different averaging schemes of the obtained solutions, averaging and mean best fit. The SIMS algorithm was used to the log-normal decomposition of 163 sets of tryptophan emission spectra. The PHREQ algorithm was applied to the 105 sets of fluorescence spectra because the other 58 spectra of proteins were measured without quenchers, and only the modified SIMS algorithm could be applied to decompose them. The results obtained with both SIMS and PHREQ methods for absolute majority of proteins differ less than errors of means presented in the $\lambda_m(i)$ and $S(i)$ columns. It allows presenting the unified results of the component analysis independent of the decomposition method. Besides the results of log-normal component analysis, Table 2 contains the maximum position values of whole, non-decomposed fluorescence spectra of proteins ("λ_m total," the third column). The denatured proteins are italicized in the tables.

The database of log-normal components of protein tryptophan fluorescence spectra shown in Table 2 may be regarded as a generalized representative database of fluorescence spectra of individual tryptophan residues in proteins. It allowed us to perform statistical testing of the hypothesis of discrete states, i.e., the existence of the set of several classes of tryptophan residues in proteins discretely differing in their fluorescence spectral maximum positions, which was put forward as early as 1973–1977 (Burstein et al., 1973; Burstein, 1977a, b, 1983). Such a test is performed here by constructing and analyzing the histograms of occurrence of log-normal components presented in the database of Table 2.

Before examining histograms of occurrence of individual log-normal components, it seems reasonable to consider how randomly are distributed spectral maximum positions of whole, non-decomposed emission spectra of proteins of our database. The histograms of occurrence of λ_m values of whole spectra are shown in Fig. 1, *A* and *B* for native and denatured proteins, respectively, constructed using the "λ_m total" values from Table 2. It is seen that the histogram of λ_m for native proteins is rather well fitted by normal distribution with the peak at 336.7 ± 0.4 nm and $\sigma = 6.0 \pm 0.9$ nm, which is a good confirmation of random character of the excerpt of native proteins. A worse fitting by the normal distribution was obtained for the maximum positions of fluorescence spectra of denatured proteins (the peak at 346.5 ± 0.6 nm; $\sigma = 4.1 \pm 1.3$ nm); however, it is expected due to the small number of proteins and, mainly, due to essentially different structural features of protein states formed under various denaturing effectors used (urea, dioxane, extreme pH, etc.).

The canonical histograms of occurrence of maximum position values of log-normal components (with the 2-nm steps) of emission spectra of native proteins are presented in Fig. 2. Panels *A* and *B* represent the distributions $\lambda_m(i)$

TABLE 2 The database of parameters of log-normal components obtained as result of decomposition of tryptophan fluorescence spectra of proteins using the SIMS and PHREQ algorithms

			PHREQ			SIMS			Average			Mean Best Fit		
No.	Protein Code	λ_m total (nm)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)
1	ACH	335.4 ± 0.5	1	324.8 ± 0.2	39.2 ± 3.5	1	321.6 ± 6.0	46.7 ± 18.3	1	322.6 ± 4.9	36.6 ± 4.0	1	322.6 ± 3.7	36.8 ± 6.9
			2	344.8 ± 0.2	60.8 ± 3.5	2	342.5 ± 0.7	53.3 ± 18.3	2	342.9 ± 1.0	63.4 ± 4.0	2	343.0 ± 1.6	63.2 ± 6.9
2	ACD.F	329.7 ± 0.6	1	317.9 ± 2.5	48.4 ± 0.1	1	326.0 ± 5.0	78.4 ± 24.4	1	325.0 ± 5.4	79.8 ± 21.7	1	327.8 ± 1.6	88.4 ± 7.0
			2	335.8 ± 2.7	51.6 ± 0.1	2	342.7 ± 5.8	21.6 ± 24.4	2	341.9 ± 5.9	20.2 ± 21.7	2	342.5 ± 2.0	11.6 ± 7.0
3	ACD.G	329.9 ± 0.9	1	328.3 ± 1.2	100 ± 0	1	329.9 ± 1.0	100 ± 0	1	329.4 ± 1.3	100 ± 0	1	329.2 ± 1.0	100 ± 0
4	ACR.F	330.5 ± 0.5	1	329.8 ± 4.5	100 ± 0	1	330.2 ± 0.5	100 ± 0	1	330.3 ± 0.4	100 ± 0	1	330.5 ± 0.5	100 ± 0
5	*ACR.F.UR*	*350.0 ± 0.5*	*1*	*348.5 ± 3.2*	*100 ± 0*	*1*	*349.9 ± 0.4*	*100 ± 0*	*1*	*349.6 ± 1.0*	*100 ± 0*	*1*	*350.0 ± 0.5*	*100 ± 0*
6	ACR.G	329.6 ± 0.5	1	321.0 ± 5.3	47.9 ± 1.9	1	320.9 ± 0.5	35.1 ± 2.0	1	321.0 ± 0.1	41.5 ± 9.1	1	320.9 ± 0.5	35.1 ± 2.0
			2	335.5 ± 4.5	52.1 ± 1.6	2	332.8 ± 0.5	64.9 ± 4.0	2	334.0 ± 1.0	58.5 ± 9.1	2	332.8 ± 0.5	64.9 ± 4.0
7	ACM	334.7 ± 1.2	1	325.6 ± 6.4	51.2 ± 14.9	1	325.2 ± 5.6	53.8 ± 10.7	1	325.4 ± 5.6	50.7 ± 11.0	1	326.0 ± 6.2	52.2 ± 8.2
			2	343.2 ± 1.7	48.8 ± 14.9	2	343.0 ± 2.5	46.2 ± 10.7	2	343.4 ± 2.1	49.3 ± 11.0	2	344.4 ± 2.6	47.8 ± 8.2
8	*ACM.UR*	*347.8 ± 0.9*	*1*	*333.4 ± 10.3*	*15.7 ± 2.1*	*1*	*339.6 ± 9.6*	*37.2 ± 9.9*	*1*	*336.5 ± 8.9*	*26.5 ± 13.7*	*1*	*337.7 ± 10.5*	*25.2 ± 16.5*
			2	*351.3 ± 0.2*	*84.4 ± 2.1*	*2*	*353.7 ± 1.4*	*62.8 ± 9.9*	*2*	*352.5 ± 1.6*	*73.5 ± 13.7*	*2*	*351.7 ± 0.9*	*74.8 ± 16.5*
9	ACY	335.5 ± 0.5	1	330.0 ± 2.9	60.6 ± 0.1	1	332.9 ± 0.6	84.2 ± 5.2	1	331.3 ± 1.8	70.6 ± 14.1	1	331.8 ± 2.9	70.6 ± 14.1
			2	345.6 ± 3.1	39.4 ± 0.2	2	353.3 ± 2.9	15.8 ± 5.2	2	348.4 ± 4.0	29.4 ± 14.1	2	348.4 ± 3.1	29.4 ± 14.1
10	AWG (1-comp.)	349.0 ± 0.3	1	349.1 ± 0.2	100 ± 0	1	349.1 ± 0.3	100 ± 0	1	349.1 ± 0.2	100 ± 0	1	349.1 ± 0.2	100 ± 0
	(2-comp.)		1	333.4 ± 2.2	5.5 ± 0.8	1	330.6 ± 6.4	8.6 ± 7.4	1	331.4 ± 5.5	7.3 ± 6.3	1	330.4 ± 7.1	4.3 ± 1.4
			2	350.0 ± 0.3	94.5 ± 0.8	2	350.6 ± 1.0	91.4 ± 7.4	2	350.4 ± 0.8	92.4 ± 6.3	2	349.9 ± 0.1	95.7 ± 1.4
11	A1AT	334.8 ± 0.5		—	—	1	324.4 ± 0.1	38.9 ± 0.2	1	324.4 ± 0.1	38.9 ± 0.2	1	324.4 ± 0.1	38.9 ± 0.2
						2	340.1 ± 0.1	61.1 ± 0.2	2	340.1 ± 0.1	61.1 ± 0.2	2	340.1 ± 0.1	61.1 ± 0.2
12	ASB.F	333.0 ± 0.5	1	314.7 ± 2.8	15.3 ± 0.4	1	319.9 ± 0.5	20.7 ± 1.0	1	317.3 ± 3.7	18.0 ± 3.8	1	314.7 ± 2.8	15.3 ± 0.4
			2	335.1 ± 3.0	84.7 ± 0.1	2	335.6 ± 0.5	79.3 ± 3.0	2	335.4 ± 0.4	82.0 ± 3.8	2	335.1 ± 3.0	84.7 ± 0.1
13	ASB.N	342.8 ± 0.2	1	342.3 ± 0.5	100 ± 0	1	342.9 ± 0.2	100 ± 0	1	342.7 ± 0.4	100 ± 0	1	342.8 ± 0.2	100 ± 0
14	ASB.N.IB	338.6 ± 0.5	1	338.5 ± 3.6	100 ± 0	1	336.6 ± 1.8	100 ± 0	1	337.6 ± 1.7	100 ± 0	1	338.5 ± 3.6	100 ± 0
15	ASH.F	334.8 ± 0.5	1	323.8 ± 3.7	42.6 ± 0.6	1	328.6 ± 0.5	31.3 ± 2.0	1	326.2 ± 3.4	55.7 ± 18.5	1	323.8 ± 3.7	42.6 ± 0.6
			2	342.5 ± 3.6	57.4 ± 0.5	2	348.5 ± 0.5	68.7 ± 2.0	2	345.5 ± 4.2	44.3 ± 18.5	2	342.5 ± 3.6	57.4 ± 0.5
16	ASH.N	341.3 ± 0.5	1	328.7 ± 2.9	20.0 ± 0.1	1	325.7 ± 5.4	19.4 ± 15.3	1	326.7 ± 4.2	25.1 ± 7.2	1	328.7 ± 2.9	20.0 ± 0.1
			2	344.8 ± 3.1	80.0 ± 0.1	2	345.1 ± 1.8	80.6 ± 15.3	2	345.0 ± 1.3	74.9 ± 7.2	2	344.8 ± 3.1	80.0 ± 0.1
17	ASH.N.IB	338.5 ± 0.5	1	326.2 ± 3.4	40.6 ± 0.4	1	327.1 ± 0.5	46.1 ± 1.0	1	326.7 ± 0.6	43.4 ± 3.9	1	326.2 ± 3.4	40.6 ± 0.4
			2	346.8 ± 3.2	59.4 ± 0.3	2	347.3 ± 0.5	53.9 ± 2.0	2	347.1 ± 0.4	56.6 ± 3.9	2	346.8 ± 3.2	59.4 ± 0.3
18	*ASH.UR*	*345.6 ± 3.8*	*1*	*323.1 ± 4.9*	*21.5 ± 6.9*	*1*	*325.4 ± 0.5*	*35.2 ± 1.0*	*1*	*323.8 ± 3.7*	*26.0 ± 9.3*	*1*	*326.0 ± 0.8*	*25.9 ± 13.2*
			2	*350.5 ± 4.9*	*78.5 ± 6.9*	*2*	*348.4 ± 0.5*	*64.8 ± 1.0*	*2*	*349.8 ± 3.7*	*74.0 ± 9.3*	*2*	*352.1 ± 4.0*	*74.1 ± 13.2*
19	ADY	344.1 ± 0.5	1	331.1 ± 2.9	55.7 ± 0.5	1	330.8 ± 1.6	35.7 ± 6.4	1	330.9 ± 1.1	42.4 ± 12.4	1	331.1 ± 2.9	55.7 ± 0.5
			2	356.2 ± 3.2	44.3 ± 0.5	2	356.6 ± 1.2	64.3 ± 6.4	2	356.4 ± 0.9	57.6 ± 12.4	2	356.2 ± 3.2	44.3 ± 0.5
20	ATRN.TRP.	338.1 ± 0.1	1	330.0 ± 1.0	47.5 ± 7.9	1	333.0 ± 3.8	57.0 ± 30.2	1	332.3 ± 3.0	55.7 ± 29.7	1	332.3 ± 3.0	55.7 ± 29.7
	A2-K2 (2-comp.)		2	344.7 ± 1.5	52.5 ± 7.9	2	345.7 ± 3.8	43.0 ± 30.2	2	344.4 ± 3.3	44.3 ± 29.7	2	344.4 ± 3.3	44.3 ± 29.7
	(3-comp.)			—	—	1	318.4 ± 0.5	5.0 ± 3.0	1	317.7 ± 1.1	6.0 ± 1.4	1	318.4 ± 0.5	5.0 ± 3.0
						2	336.3 ± 0.5	52.0 ± 6.0	2	337.0 ± 0.9	66.1 ± 19.9	2	336.3 ± 0.5	52.0 ± 6.0
						3	341.9 ± 0.5	43.0 ± 10.0	3	345.7 ± 5.3	27.9 ± 21.4	3	341.9 ± 0.5	43.0 ± 10.0
21	ATRN.TRP.	338.4 ± 0.5	1	330.3 ± 3.3	41.0 ± 0.7	1	331.7 ± 4.0	47.4 ± 33.8	1	331.2 ± 2.9	45.1 ± 24.4	1	330.3 ± 3.3	41.0 ± 0.7
	E1+E2		2	343.9 ± 3.2	59.0 ± 0.7	2	344.7 ± 4.9	52.6 ± 33.8	2	344.4 ± 3.5	54.9 ± 24.4	2	343.9 ± 3.2	59.0 ± 0.7
22	*ATRN.TRP.UR*	*350.3 ± 0.5*	*1*	*316.5 ± 3.1*	*0.4 ± 9.4*	*1*	*340.0 ± 0.5*	*14.7 ± 7.0*	*1*	*332.4 ± 13.8*	*8.9 ± 12.0*	*1*	*316.5 ± 3.1*	*0.4 ± 9.4*
			2	*350.4 ± 3.1*	*99.6 ± 0.1*	*2*	*352.6 ± 1.5*	*85.3 ± 7.0*	*2*	*351.9 ± 1.3*	*91.1 ± 12.0*	*2*	*350.4 ± 3.1*	99.6 ± 0.1

			PHREQ			SIMS			Average			Mean Best Fit		
No.	Protein Code	λ_m total (nm)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)
23	ATRN.TYR.40	338.1 ± 1.0	1	326.5 ± 3.2	35.8 ± 3.5	1	328.8 ± 3.5	41.6 ± 4.3	1	328.0 ± 3.5	38.7 ± 4.8	1	326.7 ± 3.0	37.3 ± 5.2
			2	345.0 ± 1.4	64.2 ± 3.5	2	346.4 ± 2.5	58.4 ± 4.3	2	345.9 ± 2.2	61.3 ± 4.8	2	345.1 ± 1.3	62.7 ± 5.2
24	ATRN.TYR.59	338.4 ± 0.3	1	327.7 ± 2.6	27.0 ± 8.5	1	328.3 ± 4.7	31.7 ± 19.1	1	328.0 ± 3.5	29.3 ± 13.9	1	324.7 ± 7.6	29.6 ± 17.1
			2	341.9 ± 0.6	73.0 ± 8.5	2	344.3 ± 2.4	68.3 ± 19.1	2	342.7 ± 1.9	70.7 ± 13.9	2	342.4 ± 3.4	70.4 ± 17.1
25	AMY	333.4 ± 1.4	1	327.9 ± 1.2	63.0 ± 13.6	1	330.5 ± 1.7	79.1 ± 5.4	1	329.6 ± 1.9	72.5 ± 13.6	1	327.9 ± 1.2	63.0 ± 13.6
			2	350.4 ± 6.9	37.0 ± 13.6	2	354.3 ± 4.3	20.9 ± 5.4	2	352.8 ± 5.3	27.5 ± 13.6	2	350.4 ± 6.9	37.0 ± 13.6
26	AX6 (2-comp.)	331.3 ± 0.5	1	325.5 ± 2.9	67.7 ± 0.1	1	326.0 ± 0.5	74.0 ± 1.0	1	325.8 ± 0.4	70.9 ± 4.5	1	325.5 ± 2.9	67.7 ± 0.1
			2	346.5 ± 3.1	32.3 ± 0.1	2	347.7 ± 0.5	26.0 ± 2.0	2	347.1 ± 0.8	29.1 ± 4.5	2	346.5 ± 3.1	32.3 ± 0.1
	(3-comp.)			—	—	1	317.2 ± 0.5	10.3 ± 3.0	1	317.2 ± 0.5	10.3 ± 3.0	1	317.2 ± 0.5	10.3 ± 3.0
						2	329.4 ± 0.5	75.6 ± 3.0	2	329.4 ± 0.5	75.6 ± 3.0	2	329.4 ± 0.5	75.6 ± 3.0
						3	356.3 ± 0.5	14.2 ± 2.0	3	356.3 ± 0.5	14.2 ± 2.0	3	356.3 ± 0.5	14.2 ± 2.0
27	ASP.AC	329.4 ± 1.0		—	—	1	327.2 ± 1.4	91.7 ± 7.8	1	327.2 ± 1.4	91.7 ± 7.8	1	327.2 ± 1.4	91.7 ± 7.8
						2	355.4 ± 3.9	8.3 ± 7.8	2	355.4 ± 3.9	8.3 ± 7.8	2	355.4 ± 3.9	8.3 ± 7.8
28	ASP.N	324.7 ± 1.6	1	325.2 ± 1.2	100 ± 0	1	324.5 ± 1.8	100 ± 0	1	324.5 ± 1.7	100 ± 0	1	323.9 ± 1.8	100 ± 0
29	AZU.77K	307.9 ± 3.6		—	—	1	307.9 ± 4.4	100 ± 0	1	307.9 ± 4.4	100 ± 0	1	307.9 ± 4.4	100 ± 0
30	CBC	338.3 ± 0.2	1	327.8 ± 3.9	39.0 ± 11.4	1	323.3 ± 1.6	27.3 ± 3.9	1	325.6 ± 3.6	33.1 ± 10.0	1	323.3 ± 1.6	27.3 ± 3.9
	(2-comp.)		2	344.4 ± 1.7	61.0 ± 11.4	2	342.3 ± 0.6	72.7 ± 3.9	2	343.3 ± 1.7	66.7 ± 10.0	2	342.3 ± 0.6	72.7 ± 3.9
	(3-comp.)			—	—	1	315.5 ± 5.5	14.9 ± 4.7	1	315.5 ± 5.5	14.9 ± 4.7	1	315.5 ± 5.5	14.9 ± 4.7
						2	337.2 ± 1.9	61.0 ± 3.2	2	337.2 ± 1.9	61.0 ± 3.2	2	337.2 ± 1.9	61.0 ± 3.2
						3	347.8 ± 1.1	24.1 ± 7.1	3	347.8 ± 1.1	24.1 ± 7.1	3	347.8 ± 1.1	24.1 ± 7.1
31	CDS	349.3 ± 0.5	1	348.6 ± 3.2	100 ± 0	1	349.4 ± 0.1	100 ± 0	1	349.1 ± 0.5	100 ± 0	1	349.3 ± 0.5	100 ± 0
32	*CDS.25*	*351.2 ± 0.5*		—	—	*1*	*351.2 ± 1.0*	*100 ± 0*	*1*	*351.2 ± 1.0*	*100 ± 0*	*1*	*351.2 ± 1.0*	*100 ± 0*
33	CDS.25.CAM	342.3 ± 3.7			—	—1	326.9 ± 4.0	40.5 ± 19.1	1	326.9 ± 4.0	40.5 ± 19.1	1	328.5 ± 6.5	42.7 ± 19.3
						2	348.1 ± 2.8	59.5 ± 19.1	2	348.1 ± 2.8	59.5 ± 19.1	2	348.0 ± 3.5	57.3 ± 19.3
34	CAB	339.5 ± 0.6	1	339.8 ± 0.6	100 ± 0	1	339.8 ± 0.7	100 ± 0	1	339.7 ± 0.6	100 ± 0	1	339.8 ± 0.1	100 ± 0
35	CAH (2-comp.)	332.8 ± 0.4	1	325.9 ± 3.3	49.8 ± 24.9	1	320.5 ± 4.6	26.4 ± 15.5	1	322.3 ± 4.9	43.3 ± 19.2	1	325.9 ± 3.3	49.8 ± 24.9
			2	338.6 ± 3.1	50.2 ± 24.9	2	335.9 ± 1.9	73.6 ± 15.5	2	336.8 ± 2.6	56.8 ± 19.2	2	338.6 ± 3.1	50.2 ± 24.9
	(3-comp.)			—	—	1	317.1 ± 3.8	15.3 ± 4.7	1	317.1 ± 3.8	15.3 ± 4.7	1	317.1 ± 3.8	15.3 ± 4.7
						2	333.1 ± 1.4	60.0 ± 18.2	2	333.1 ± 1.4	60.0 ± 18.2	2	333.1 ± 1.4	60.0 ± 18.2
						3	340.3 ± 4.7	24.7 ± 13.5	3	340.3 ± 4.7	24.7 ± 13.5	3	340.3 ± 4.7	24.7 ± 13.5
36	CPA (2-comp.)	331.0 ± 0.3	1	327.7 ± 1.6	77.5 ± 8.4	1	325.7 ± 1.6	66.5 ± 12.8	1	326.4 ± 1.8	70.2 ± 12.0	1	326.2 ± 1.5	69.1 ± 12.4
			2	347.4 ± 4.4	22.5 ± 8.4	2	341.7 ± 3.0	33.5 ± 12.8	2	343.6 ± 4.2	29.8 ± 12.0	2	342.6 ± 3.1	30.9 ± 12.4
	(3-comp.)			—	—	1	304.4 ± 11.9	18.6 ± 11.3	1	304.4 ± 11.9	18.6 ± 11.3	1	304.4 ± 11.9	18.6 ± 11.3
						2	330.7 ± 1.6	73.2 ± 12.7	2	330.7 ± 1.6	73.2 ± 12.7	2	330.7 ± 1.6	73.2 ± 12.7
						3	351.5 ± 2.9	8.2 ± 1.7	3	351.5 ± 2.9	8.2 ± 1.7	3	351.5 ± 2.9	8.2 ± 1.7
37	CHT	333.7 ± 1.3	1	329.3 ± 2.2	57.6 ± 14.6	1	330.3 ± 2.5	64.6 ± 21.2	1	330.6 ± 3.4	64.8 ± 17.6	1	328.2 ± 2.7	54.9 ± 17.0
			2	341.6 ± 3.3	42.4 ± 14.6	2	342.1 ± 3.5	35.4 ± 21.2	2	342.4 ± 3.6	35.2 ± 17.6	2	341.0 ± 2.9	45.1 ± 17.0
38	*CHT.D35*	*339.2 ± 2.4*	*1*	*328.8 ± 0.1*	*26.1 ± 26.4*	*1*	*324.2 ± 12.8*	*34.8 ± 20.6*	*1*	*326.1 ± 9.6*	*32.3 ± 23.2*	*1*	*324.2 ± 12.8*	*34.8 ± 20.6*
			2	*342.5 ± 1.3*	*73.9 ± 26.4*	*2*	*340.7 ± 2.5*	*65.2 ± 20.6*	*2*	*341.6 ± 2.0*	*67.7 ± 23.2*	*2*	*340.7 ± 2.5*	*65.2 ± 20.6*
39	*CHT.D60-80*	*339.8 ± 0.1*	*1*	*327.8 ± 1.3*	*34.0 ± 1.9*	*1*	*328.9 ± 2.0*	*44.0 ± 10.3*	*1*	*328.5 ± 1.8*	*35.3 ± 1.9*	*1*	*327.8 ± 1.3*	*34.0 ± 1.9*
			2	*343.7 ± 1.6*	*66.0 ± 1.9*	*2*	*343.9 ± 1.1*	*56.0 ± 10.3*	*2*	*344.3 ± 1.4*	*64.7 ± 1.9*	*2*	*343.7 ± 1.6*	*66.0 ± 1.9*
40	CHG (1-comp.)	331.7 ± 0.2	1	332.0 ± 0.5	100 ± 0	1	331.9 ± 0.4	100 ± 0	1	331.9 ± 0.4	100 ± 0	1	332.2 ± 0.7	100 ± 0
	(2-comp.)		1	327.5 ± 1.5	61.5 ± 13.9	1	327.9 ± 3.1	50.0 ± 29.4	1	327.7 ± 2.3	55.7 ± 22.8	1	327.5 ± 1.3	58.5 ± 19.2
			2	339.1 ± 1.9	38.5 ± 13.9	2	338.7 ± 6.5	50.0 ± 29.4	2	338.9 ± 4.6	44.3 ± 22.8	2	339.4 ± 4.0	41.5 ± 19.2
41	*CHG.D35*	*333.2 ± 0.5*	*1*	*322.7 ± 2.9*	*33.3 ± 0.2*	*1*	*316.7 ± 4.5*	*35.9 ± 11.1*	*1*	*318.7 ± 4.7*	*30.7 ± 3.7*	*1*	*322.7 ± 2.9*	*33.3–0.2*
			2	*336.2 ± 3.0*	*66.7 ± 0.1*	*2*	*335.6 ± 0.1*	*64.1 ± 11.1*	*2*	*335.8 ± 0.4*	*69.3 ± 3.7*	*2*	*336.2 ± 3.0*	*66.7 ± 0.1*

			PHREQ			SIMS			Average			Mean Best Fit		
No.	Protein Code	λ_m total (nm)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)
42	*CHG.D60-70*	*339.2 ± 0.3*	*1*	*336.8 ± 0.9*	*69.6 ± 9.3*	*1*	*336.0 ± 1.8*	*61.0 ± 28.4*	*1*	*336.2 ± 1.5*	*64.9 ± 21.0*	*1*	*336.8 ± 0.9*	*69.6 ± 9.3*
			2	*347.5 ± 3.8*	*30.4 ± 9.3*	*2*	*345.9 ± 5.1*	*39.0 ± 28.4*	*2*	*346.5 ± 4.3*	*35.1 ± 21.0*	*2*	*347.5 ± 3.8*	*30.4 ± 9.3*
43	*CHG.UR*	*349.7 ± 0.5*	*1*	*348.4 ± 3.1*	*100 ± 0*	*1*	*350.0 ± 0.4*	*100 ± 0*	*1*	*349.4 ± 0.5*	*100 ± 0*	*1*	*350.0 ± 3.2*	*100 ± 0*
44	CPK	333.5 ± 0.4	1	323.9 ± 2.8	34.1 ± 6.4	1	321.3 ± 2.9	34.9 ± 3.7	1	321.9 ± 2.8	34.0 ± 4.4	1	320.9 ± 3.9	34.8 ± 5.4
			2	338.9 ± 0.9	65.9 ± 6.4	2	338.7 ± 0.4	65.1 ± 3.7	2	338.7 ± 0.5	66.0 ± 4.4	2	338.9 ± 0.7	65.2 ± 5.4
45	ENH (1-comp.)	332.6 ± 0.1	1	332.3 ± 0.1	100 ± 0	1	332.5 ± 0.1	100 ± 0	1	332.4 ± 0.2	100 ± 0	1	332.2 ± 0.2	100 ± 0
	(2-comp.)		1	327.7 ± 0.8	47.6 ± 13.0	1	329.6 ± 1.0	53.0 ± 17.8	1	329.0 ± 1.3	44.1 ± 11.9	1	327.7 ± 0.8	47.6 ± 13.0
			2	336.6 ± 1.4	52.4 ± 13.0	2	334.6 ± 0.3	47.0 ± 17.8	2	335.2 ± 1.3	55.9 ± 11.9	2	336.6 ± 1.4	52.4 ± 13.0
46	FGE	345.8 ± 0.8	1	330.7 ± 7.6	23.5 ± 14.3	1	333.0 ± 2.6	26.7 ± 7.9	1	332.2 ± 4.2	27.4 ± 9.9	1	330.7 ± 7.6	23.5 ± 14.3
			2	351.2 ± 3.7	76.5 ± 14.3	2	351.3 ± 2.3	73.3 ± 7.9	2	351.3 ± 2.4	72.6 ± 9.9	2	351.2 ± 3.7	76.5 ± 14.3
47	FBN	331.8 ± 0.6	1	321.7 ± 1.3	45.5 ± 5.8	1	321.6 ± 2.5	60.7 ± 12.0	1	321.6 ± 2.1	55.4 ± 11.8	1	321.6 ± 2.5	60.7 ± 12.0
			2	340.0 ± 1.7	64.5 ± 5.8	2	345.0 ± 3.4	39.3 ± 12.0	2	343.3 ± 3.8	44.6 ± 11.8	2	345.0 ± 3.4	39.3 ± 12.0
48	FIL	337.4 ± 0.7	1	326.8 ± 3.7	26.8 ± 6.5	1	320.6 ± 2.8	31.7 ± 3.5	1	322.6 ± 4.2	29.5 ± 5.0	1	321.0 ± 2.6	32.0 ± 1.3
			2	341.8 ± 3.3	73.2 ± 6.5	2	342.8 ± 1.7	68.3 ± 3.5	2	342.5 ± 3.2	70.5 ± 5.0	2	343.0 ± 1.7	68.0 ± 1.3
49	GLG	350.8 ± 0.1	1	349.6 ± 0.4	100 ± 0	1	350.8 ± 1.0	100 ± 0	1	350.4 ± 0.7	100 ± 0	1	350.8 ± 0.1	100 ± 0
50	GPD	334.9 ± 0.8	1	327.4 ± 1.7	57.9 ± 13.6	1	328.7 ± 3.0	63.9 ± 17.0	1	328.6 ± 2.7	59.4 ± 14.3	1	329.3 ± 2.1	62.8 ± 15.2
			2	345.5 ± 3.2	42.1 ± 13.6	2	345.2 ± 3.2	36.1 ± 17.0	2	345.8 ± 4.0	40.6 ± 14.3	2	346.3 ± 3.3	37.2 ± 15.2
51	GPS	342.5 ± 0.1	1	341.8 ± 0.2	100 ± 0	1	342.5 ± 1.0	100 ± 0	1	342.2 ± 1.0	100 ± 0	1	342.5 ± 0.1	100 ± 0
52	*HBA.CM*	*349.4 ± 0.5*	*1*	*324.6 ± 0.8*	*3.4 ± 3.1*	*1*	*330.3 ± 3.4*	*7.9 ± 1.9*	*1*	*327.5 ± 3.8*	*5.6 ± 3.4*	*1*	*324.6 ± 0.8*	*3.4 ± 3.1*
			2	*350.7 ± 0.3*	*96.6 ± 3.1*	*2*	*351.4 ± 0.2*	*92.1 ± 1.9*	*2*	*351.0 ± 0.5*	*94.4 ± 3.4*	*2*	*350.7 ± 0.3*	*96.6 ± 3.1*
53	IGG	334.7 ± 0.3	1	328.4 ± 1.3	53.1 ± 3.7	1	325.4 ± 4.0	29.4 ± 6.4	1	326.9 ± 3.0	41.3 ± 14.3	1	328.3 ± 0.9	46.7 ± 11.4
			2	339.4 ± 0.5	46.9 ± 3.7	2	336.6 ± 0.4	70.6 ± 6.4	2	338.0 ± 1.7	58.7 ± 14.3	2	338.3 ± 1.8	53.3 ± 11.4
54	*IGG.UR*	*348.3 ± 0.5*	*1*	*344.0 ± 3.2*	*51.8 ± 0.6*	*1*	*346.6 ± 0.5*	55.8 ± 10.0	*1*	*345.3 ± 1.8*	*53.8 ± 2.8*	*1*	*346.6 ± 0.5*	*55.8 ± 10.0*
			2	*354.6 ± 3.5*	*48.2 ± 0.6*	*2*	*351.2 ± 0.5*	*44.2 ± 9.0*	*2*	*352.9 ± 2.4*	*46.2 ± 2.8*	*2*	*351.2 ± 0.5*	*44.2 ± 9.0*
55	IOP	334.6 ± 0.6	1	323.1 ± 2.8	47.2 ± 7.2	1	327.9 ± 0.5	71.1 ± 1.3	1	326.3 ± 2.8	63.1 ± 12.7	1	327.9 ± 0.5	71.1 ± 1.3
			2	347.7 ± 2.3	52.8 ± 7.2	2	359.5 ± 1.5	29.0 ± 1.3	2	355.6 ± 6.3	36.9 ± 12.7	2	359.5 ± 1.5	29.0 ± 1.3
56	LAB	325.9 ± 0.2	1	322.4 ± 1.1	74.9 ± 3.4	1	321.9 ± 3.9	70.1 ± 18.1	1	322.1 ± 2.8	76.2 ± 6.4	1	322.4 ± 1.1	74.9 ± 3.4
			2	338.3 ± 3.8	25.1 ± 3.4	2	338.0 ± 3.0	29.9 ± 18.1	2	338.1 ± 6.8	23.8 ± 6.4	2	338.3 ± 3.8	25.1 ± 3.4
57	LAH	327.3 ± 0.3	1	322.4 ± 2.2	67.2 ± 4.9	1	324.5 ± 1.0	82.5 ± 5.5	1	323.6 ± 1.7	74.5 ± 9.5	1	322.4 ± 2.2	67.2 ± 4.9
			2	341.9 ± 2.1	32.8 ± 4.9	2	350.1 ± 5.8	17.5 ± 5.5	2	346.6 ± 6.1	25.5 ± 9.5	2	341.9 ± 2.1	32.8 ± 4.9
58	LGB.AC	323.3 ± 0.3		—	—	1	332.3 ± 0.3	100 ± 0	1	332.3 ± 0.3	100 ± 0	1	332.3 ± 0.4	100 ± 0
59	LGB.N	332.1 ± 0.5		—	—	1	332.2 ± 0.1	100 ± 0	1	332.2 ± 0.1	100 ± 0	1	332.1 ± 0.5	100 ± 0
60	LGB.AL	333.3 ± 0.5		—	—	1	333.3 ± 0.3	100 ± 0	1	333.3 ± 0.2	100 ± 0	1	333.3 ± 0.5	100 ± 0
61	LGC.AC	337.5 ± 0.2	1	323.2 ± 6.9	29.0 ± 28.1	1	326.1 ± 1.1	30.5 ± 5.3	1	325.4 ± 3.1	29.6 ± 18.3	1	326.1 ± 1.1	30.5 ± 5.3
			2	347.5 ± 7.7	71.0 ± 28.1	2	343.6 ± 0.3	69.5 ± 5.3	2	343.3 ± 0.7	70.4 ± 18.3	2	343.6 ± 0.3	69.5 ± 5.3
62	LGC.N	329.1 ± 0.8	1	322.9 ± 1.1	55.8 ± 2.0	1	324.1 ± 0.4	64.3 ± 0.1	1	323.8 ± 1.1	60.1 ± 5.0	1	324.1 ± 0.4	64.3 ± 0.1
	(2-comp.)		2	347.3 ± 0.4	44.2 ± 2.0	2	349.0 ± 1.1	35.7 ± 0.1	2	348.4 ± 1.3	39.9 ± 5.0	2	349.0 ± 1.1	35.7 ± 0.1
	(3-comp.)			—	—	1	323.7 ± 0.4	54.6 ± 6.6	1	323.7 ± 0.4	54.6 ± 6.6	1	323.7 ± 0.4	54.6 ± 6.6
						2	335.5 ± 8.6	21.4 ± 2.8	2	335.5 ± 8.6	21.4 ± 2.8	2	335.5 ± 8.6	21.4 ± 2.8
						3	353.0 ± 4.0	24.2 ± 9.3	3	353.0 ± 4.0	24.2 ± 9.3	3	353.0 ± 4.0	24.2 ± 9.3
63	LPP (2-comp.)	344.0 ± 0.3	1	326.3 ± 3.4	22.1 ± 7.1	1	326.7 ± 2.4	22.1 ± 8.4	1	327.6 ± 4.7	24.7 ± 6.0	1	326.3 ± 3.4	22.1 ± 7.1
			2	351.6 ± 1.9	77.9 ± 7.1	2	351.0 ± 1.9	77.9 ± 8.4	2	351.2 ± 1.8	75.3 ± 6.0	2	351.6 ± 1.9	77.9 ± 7.1
	(3-comp.)			—	—	1	325.5 ± 2.7	17.0 ± 8.9	1	325.5 ± 2.7	17.0 ± 8.9	1	325.5 ± 2.7	17.0 ± 8.9
						2	339.2 ± 7.2	24.4 ± 8.7	2	339.2 ± 7.2	24.4 ± 8.7	2	339.2 ± 7.2	24.4 ± 8.7
						3	353.8 ± 1.2	58.6 ± 13.5	3	353.8 ± 1.2	58.6 ± 13.5	3	353.8 ± 1.2	58.6 ± 13.5
64	LWG	335.8 ± 0.4	1	325.3 ± 0.4	38.3 ± 2.5	1	329.8 ± 0.6	55.5 ± 5.0	1	328.3 ± 2.4	48.0 ± 11.3	1	329.7 ± 0.3	57.6 ± 0.8

			PHREQ			SIMS			Average			Mean Best Fit		
No.	Protein Code	λ_m total (nm)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)
			2	343.8 ± 0.2	61.7 ± 2.5	2	345.9 ± 1.3	44.5 ± 5.0	2	345.2 ± 1.5	52.0 ± 11.3	2	346.4 ± 0.6	42.4 ± 0.8
65	LZC.AC	337.2 ± 0.2	1	330.0 ± 1.2	44.5 ± 5.7	1	329.9 ± 1.4	45.8 ± 7.7	1	329.9 ± 1.2	42.9 ± 6.8	1	330.0 ± 1.2	44.5 ± 5.7
	(2-comp.)		2	346.4 ± 0.8	55.5 ± 5.7	2	345.6 ± 12.2	54.2 ± 7.7	2	345.8 ± 1.8	57.1 ± 6.8	2	346.4 ± 0.8	55.5 ± 5.7
	(3-comp.)			—	—	1	317.4 ± 0.3	8.7 ± 2.8	1	317.4 ± 0.3	8.7 ± 2.8	1	317.4 ± 0.3	8.7 ± 2.8
						2	333.6 ± 1.7	41.4 ± 0.3	2	333.6 ± 1.7	41.4 ± 0.3	2	333.6 ± 1.7	41.4 ± 0.3
						3	345.7 ± 1.9	49.9 ± 2.4	3	345.7 ± 1.9	49.9 ± 2.4	3	345.7 ± 1.9	49.9 ± 2.4
66	LZC.N	338.3 ± 0.2		—	—	1	313.9 ± 0.5	5.0 ± 1.0	1	313.9 ± 0.5	5.0 ± 1.0	1	313.9 ± 0.5	5.0 ± 1.0
						2	334.8 ± 0.5	67.5 ± 5.0	2	334.8 ± 0.5	67.5 ± 5.0	2	334.8 ± 0.5	67.5 ± 5.0
						3	343.0 ± 0.5	27.5 ± 8.0	3	343.0 ± 0.5	27.5 ± 8.0	3	343.0 ± 0.5	27.5 ± 8.0
67	*LZC.UR*	*345.2 ± 3.7*	*1*	*333.2 ± 4.5*	*30.2 ± 34.1*	*1*	*340.6 ± 6.6*	*50.6 ± 2.6*	*1*	*337.6 ± 6.6*	*40.0 ± 22.7*	*1*	*342.3 ± 8.3*	*53.1 ± 1.8*
			2	*350.1 ± 2.1*	*69.8 ± 34.1*	*2*	*349.0 ± 1.1*	*49.4 ± 2.6*	*2*	*349.4 ± 1.3*	*60.0 ± 22.7*	*2*	*350.7 ± 1.3*	*46.9 ± 1.8*
68	MSP	349.8 ± 0.2	1	349.6 ± 0.1	100 ± 0	1	350.0 ± 0.2	100 ± 0	1	349.0 ± 0.2	100 ± 0	1	349.9 ± 0.2	100 ± 0
69	MLT.MON	349.8 ± 0.5		—	—	1	349.2 ± 0.5	100 ± 0	1	349.2 ± 0.5	100 ± 0	1	349.8 ± 0.5	100 ± 0
70	MLT.TETR	336.9 ± 1.7	1	331.3 ± 4.9	56.6 ± 21.9	1	329.4 ± 4.1	38.5 ± 11.1	1	330.0 ± 4.1	47.3 ± 17.9	1	329.5 ± 4.5	48.7 ± 13.9
			2	348.0 ± 6.5	43.4 ± 21.9	2	342.3 ± 1.8	61.5 ± 11.1	2	344.2 ± 4.5	52.7 ± 17.9	2	343.9 ± 3.3	51.3 ± 13.9
71	MLT.CAM	338.3 ± 0.5		—	—	1	332.3 ± 1.9	71.5 ± 12.2	1	332.3 ± 1.9	71.5 ± 12.2	1	330.9 ± 0.5	62.8 ± 2.0
						2	352.8 ± 3.3	28.5 ± 12.2	2	352.8 ± 3.3	28.5 ± 12.2	2	350.5 ± 0.5	37.2 ± 2.0
72	MLT.LS	332.2 ± 0.5	1	327.3 ± 2.9	65.2 ± 0.1	1	323.2 ± 3.0	50.0 ± 24.7	1	324.5 ± 3.2	42.5 ± 31.2	1	324.5 ± 3.2	42.5 ± 31.2
			2	343.5 ± 3.0	34.8 ± 0.1	2	335.7 ± 0.2	50.0 ± 24.7	2	338.3 ± 4.5	57.5 ± 31.2	2	338.3 ± 4.5	57.5 ± 31.2
73	HMM	337.5 ± 0.5	1	324.0 ± 5.2	17.6 ± 2.8	1	323.6 ± 0.5	13.0 ± 2.0	1	323.8 ± 0.3	15.2 ± 3.1	1	323.6 ± 0.5	13.0 ± 2.0
			2	340.1 ± 3.1	82.4 ± 0.6	2	339.3 ± 0.5	87.0 ± 3.0	2	341.1 ± 2.5	84.8 ± 3.1	2	339.3 ± 0.5	87.0 ± 3.0
74	LMM	339.4 ± 0.5	1	337.4 ± 3.1	100 ± 0	1	339.4 ± 0.6	100 ± 0	1	339.1 ± 0.9	100 ± 0	1	339.4 ± 0.6	100 ± 0
75	MON (1-comp.)	340.7 ± 0.4	1	341.1 ± 0.5	100 ± 0	1	341.1 ± 0.5	100 ± 0	1	340.2 ± 1.3	100 ± 0	1	340.3 ± 1.5	100 ± 0
	(2-comp.)		1	331.0 ± 3.4	28.3 ± 10.9	1	323.9 ± 7.1	22.7 ± 8.2	1	325.6 ± 7.2	22.9 ± 8.8	1	327.5 ± 6.6	25.9 ± 8.6
			2	344.8 ± 1.1	71.7 ± 10.9	2	344.1 ± 1.3	77.3 ± 8.2	2	344.3 ± 1.3	77.1 ± 8.8	2	344.1 ± 1.5	74.1 ± 8.6
76	MGB.APO	336.4 ± 0.5	1	319.1 ± 7.8	22.5 ± 19.2	1	325.6 ± 3.2	28.3 ± 8.6	1	323.4 ± 5.4	24.4 ± 13.0	1	325.3 ± 3.7	22.5 ± 19.2
			2	340.5 ± 2.7	77.5 ± 19.2	2	340.2 ± 0.6	71.7 ± 8.6	2	340.3 ± 1.3	75.6 ± 13.0	2	340.3 ± 0.7	77.5 ± 19.2
77	MSS	336.9 ± 0.4	1	327.0 ± 9.6	41.1 ± 28.2	1	328.6 ± 4.8	24.4 ± 26.2	1	328.0 ± 6.3	28.0 ± 23.1	1	323.4 ± 6.6	12.3 ± 2.7
			2	342.2 ± 5.5	59.9 ± 28.2	2	339.4 ± 1.7	75.6 ± 26.2	2	340.5 ± 2.7	72.0 ± 23.1	2	338.6 ± 0.6	87.7 ± 2.7
78	*MSS.UR*	*345.1 ± 0.5*	*1*	*338.2 ± 3.3*	*31.9 ± 1.1*	*1*	*340.3 ± 0.5*	*38.7 ± 6.0*	*1*	*340.5 ± 2.5*	*35.3 ± 4.8*	*1*	*338.2 ± 3.3*	*31.9 ± 1.1*
			2	*348.7 ± 3.3*	*68.1 ± 0.6*	*2*	*348.4 ± 0.1*	*61.3 ± 5.0*	*2*	*348.5 ± 0.2*	*64.8 ± 4.8*	*2*	*348.7 ± 3.3*	*68.1 ± 0.6*
79	MS1	335.7 ± 0.4		—	—	1	318.8 ± 1.3	20.7 ± 3.0	1	318.8 ± 1.3	20.7 ± 3.0	1	318.8 ± 1.3	20.7 ± 3.0
						2	331.8 ± 0.9	23.0 ± 0.5	2	331.8 ± 0.9	23.0 ± 0.5	2	331.8 ± 0.9	23.0 ± 0.5
						3	342.5 ± 0.5	56.3 ± 2.5	3	342.5 ± 0.5	56.3 ± 2.5	3	342.5 ± 0.5	56.3 ± 2.5
80	MS1.ATP	335.5 ± 0.3		—	—	1	317.4 ± 1.8	13.8 ± 3.0	1	317.4 ± 1.8	13.8 ± 3.0	1	317.4 ± 1.8	13.8 ± 3.0
						2	330.3 ± 2.4	28.8 ± 5.7	2	330.3 ± 2.4	28.8 ± 5.7	2	330.3 ± 2.4	28.8 ± 5.7
						3	339.2 ± 0.5	57.5 ± 8.7	3	339.2 ± 0.5	57.5 ± 8.7	3	339.2 ± 0.5	57.5 ± 8.7
81	NO1 (1-comp.)	346.1 ± 0.2	1	346.6 ± 0.4	100 ± 0	1	346.5 ± 0.6	100 ± 0	1	346.5 ± 0.6	100 ± 0	1	346.6 ± 0.4	100 ± 0
	(2-comp.)		1	329.1 ± 4.4	4.9 ± 1.3	1	335.6 ± 6.8	28.1 ± 20.1	1	332.8 ± 6.5	19.5 ± 20.3	1	329.1 ± 4.4	4.9 ± 1.3
			2	347.4 ± 0.3	95.1 ± 1.3	2	350.1 ± 2.3	71.9 ± 20.1	2	348.9 ± 2.2	80.5 ± 20.3	2	347.4 ± 0.3	95.1 ± 1.3
82	NO2	344.5 ± 0.2	1	344.1 ± 3.1	100 ± 0	1	344.5 ± 0.3	100 ± 0	1	344.4 ± 0.3	100 ± 0	1	344.4 ± 0.2	100 ± 0
83	NST	336.0 ± 0.1	1	330.9 ± 0.8	45.7 ± 5.7	1	327.0 ± 7.2	41.7 ± 44.1	1	328.3 ± 6.1	33.3 ± 14.2	1	330.9 ± 0.8	45.7 ± 5.7
			2	340.7 ± 0.4	54.9 ± 5.7	2	339.4 ± 1.0	58.3 ± 44.1	2	339.8 ± 1.0	66.7 ± 14.2	2	340.7 ± 0.4	54.9 ± 5.7
84	NDPK.A.8	340.5 ± 0.4	1	340.2 ± 0.4	100 ± 0	1	340.4 ± 0.4	100 ± 0	1	340.4 ± 0.4	100 ± 0	1	340.4 ± 0.3	100 ± 0
85	NDPK.A.5.LIS	337.3 ± 0.8		—	—	1	326.4 ± 4.7	25.9 ± 22.6	1	326.4 ± 4.7	25.9 ± 22.6	1	327.0 ± 5.0	17.9 ± 11.7
						2	340.0 ± 2.1	74.1 ± 22.6	2	340.0 ± 2.1	74.1 ± 22.6	2	340.0 ± 2.3	82.1 ± 11.7

			PHREQ			SIMS			Average			Mean Best Fit		
No.	Protein Code	λ_m total (nm)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)
86	NDPK.A.5.HIS	336.5 ± 0.3	1	329.1 ± 0.4	39.4 ± 1.7	1	323.8 ± 4.5	20.0 ± 10.9	1	325.5 ± 4.4	26.5 ± 13.0	1	325.1 ± 4.5	24.3 ± 12.6
			2	342.2 ± 0.6	60.6 ± 1.7	2	339.0 ± 1.3	80.0 ± 10.9	2	340.1 ± 1.9	73.5 ± 13.0	2	339.0 ± 1.8	75.7 ± 12.6
87	NDPK.B.8	335.5 ± 0.3	1	335.3 ± 0.8	100 ± 0	1	335.5 ± 0.3	100 ± 0	1	335.4 ± 0.3	100 ± 0	1	335.4 ± 0.5	100 ± 0
88	NDPK.B.AC	334.2 ± 0.8		—	—	1	334.7 ± 0.7	100 ± 0	1	334.7 ± 0.7	100 ± 0	1	334.4 ± 1.1	100 ± 0
	(1-comp.)													
	(2-comp.)			—	—	1	330.0 ± 0.7	78.9 ± 6.1	1	330.0 ± 0.7	78.9 ± 6.1	1	330.0 ± 0.7	78.9 ± 6.1
			2	353.9 ± 4.9	21.1 ± 6.1	2	353.9 ± 4.9	21.1 ± 6.1	2	353.9 ± 4.9	21.1 ± 6.1			
89	OVH.LIS	334.4 ± 0.5	1	325.8 ± 2.6	17.6 ± 0.2	1	327.5 ± 0.2	23.0 ± 0.4	1	326.9 ± 1.0	20.4 ± 4.0	1	325.8 ± 2.6	17.6 ± 0.2
			2	336.9 ± 2.7	82.4 ± 0.1	2	337.0 ± 0.4	77.0 ± 0.4	2	336.9 ± 0.3	79.6 ± 4.0	2	336.9 ± 2.7	82.4 ± 0.1
90	OVH.HIS	334.9 ± 0.1	1	335.1 ± 0.2	100 ± 0	1	335.2 ± 0.2	100 ± 0	1	335.1 ± 0.2	100 ± 0	1	334.9 ± 0.1	100 ± 0
91	PAP.AC	340.2 ± 1.9	1	326.8 ± 8.3	43.1 ± 16.7	1	323.5 ± 5.4	44.9 ± 17.8	1	324.6 ± 6.4	39.7 ± 15.3	1	323.8 ± 7.3	37.7 ± 13.4
	(2-comp.)		2	349.7 ± 3.2	56.9 ± 16.7	2	349.8 ± 3.7	55.1 ± 17.8	2	349.8 ± 3.4	60.3 ± 15.3	2	348.5 ± 3.2	62.3 ± 13.4
	(3-comp.)			—	—	1	304.2 ± 6.6	14.0 ± 5.4	1	304.2 ± 6.6	14.0 ± 5.4	1	304.2 ± 6.6	14.0 ± 5.4
						2	331.9 ± 5.5	39.5 ± 13.5	2	331.9 ± 5.5	39.5 ± 13.5	2	331.9 ± 5.5	39.5 ± 13.5
						3	351.1 ± 3.7	46.5 ± 17.5	3	351.1 ± 3.7	46.5 ± 17.5	3	351.1 ± 3.7	46.5 ± 17.5
92	PAP.N	343.3 ± 0.8	1	324.8 ± 4.0	19.3 ± 8.3	1	327.4 ± 4.4	29.5 ± 13.7	1	326.5 ± 4.2	20.3 ± 6.4	1	324.8 ± 4.0	19.3 ± 8.3
	(2-comp.)		2	348.5 ± 1.3	80.7 ± 8.3	2	349.1 ± 0.9	70.5 ± 13.7	2	348.9 ± 1.0	79.7 ± 6.4	2	348.5 ± 1.3	80.7 ± 8.3
	(3-comp.)			—	—	1	317.7 ± 6.9	17.5 ± 5.9	1	317.7 ± 6.9	17.5 ± 5.9	1	317.7 ± 6.9	17.5 ± 5.9
						2	341.0 ± 3.9	20.1 ± 13.0	2	341.0 ± 3.9	20.1 ± 13.0	2	341.0 ± 3.9	20.1 ± 13.0
						3	349.6 ± 4.0	62.4 ± 15.7	3	349.6 ± 4.0	62.4 ± 15.7	3	349.6 ± 4.0	62.4 ± 15.7
93	PAP.AL	348.2 ± 0.6	1	349.0 ± 1.2	100 ± 0	1	348.7 ± 1.0	100 ± 0	1	348.4 ± 1.0	100 ± 0	1	348.5 ± 1.2	100 ± 0
	(1-comp.)													
	(2-comp.)		1	333.6 ± 8.8	10.3 ± 8.5	1	337.7 ± 6.8	23.2 ± 28.6	1	336.1 ± 7.3	21.7 ± 26.3	1	333.6 ± 8.8	10.3 ± 8.5
			2	349.6 ± 0.4	89.7 ± 8.5	2	350.6 ± 0.6	76.8 ± 28.6	2	350.2 ± 0.7	78.3 ± 26.3	2	349.6 ± 0.4	89.7 ± 8.5
94	*PAP.UR*	*344.0 ± 0.5*	*1*	*316.8 ± 4.3*	*18.4 ± 0.7*	*1*	*320.7 ± 3.5*	*21.8 ± 5.8*	*1*	*319.4 ± 3.3*	*18.1 ± 0.5*	*1*	*318.2 ± 0.5*	*17.7 ± 1.0*
			2	*346.5 ± 3.1*	*81.6 ± 0.7*	*2*	*346.1 ± 0.1*	*78.2 ± 5.8*	*2*	*346.2 ± 0.3*	*81.9 ± 0.5*	*2*	*346.0 ± 0.5*	*82.3 ± 1.0*
95	PAC.AC	341.5 ± 0.5		—	—	1	310.1 ± 0.6	8.6 ± 3.5	1	310.1 ± 0.6	8.6 ± 3.5	1	309.6 ± 2.0	6.1 ± 3.0
	(2-comp.)					2	342.6 ± 0.4	91.4 ± 3.5	2	342.6 ± 0.4	91.4 ± 3.5	2	342.3 ± 2.0	93.9 ± 3.0
	(3-comp.)			—	—	1	309.6 ± 2.0	6.1 ± 3.0	1	309.6 ± 2.0	6.1 ± 3.0	1	309.6 ± 2.0	6.1 ± 3.0
						2	338.8 ± 0.5	66.7 ± 3.0	2	338.8 ± 0.5	66.7 ± 3.0	2	338.8 ± 0.5	66.7 ± 3.0
						3	350.8 ± 0.5	27.2 ± 4.0	3	350.8 ± 0.5	27.2 ± 4.0	3	350.8 ± 0.5	27.2 ± 4.0
96	PAC.APO	348.0 ± 0.5		—	—	1	308.9 ± 12.0	17.0 ± 6.2	1	308.9 ± 12.0	17.0 ± 6.2	1	300.4 ± 4.0	22.1 ± 7.8
						2	349.4 ± 0.6	83.0 ± 6.2	2	349.4 ± 0.6	83.0 ± 6.2	2	349.0 ± 0.5	77.8 ± 1.0
97	PAC.MG	331.1 ± 0.5		—	—	1	318.1 ± 0.5	55.3 ± 2.0	1	318.1 ± 0.5	55.3 ± 2.0	1	318.1 ± 0.5	55.3 ± 2.0
						2	341.1 ± 0.5	44.7 ± 2.0	2	341.1 ± 0.5	44.7 ± 2.0	2	341.1 ± 0.5	44.7 ± 2.0
98	PAC.CA	325.8 ± 0.5		—	—	1	326.5 ± 0.7	100 ± 0	1	326.5 ± 0.7	100 ± 0	1	326.5 ± 0.7	100 ± 0
99	PAM.APO	336.6 ± 0.5	1	326.9 ± 2.9	43.9 ± 0.1	1	327.7 ± 2.6	54.9 ± 17.5	1	327.4 ± 1.9	43.2 ± 1.0	1	326.9 ± 2.9	43.9 ± 0.1
	(2-comp.)		2	345.7 ± 3.1	56.1 ± 0.1	2	341.7 ± 3.4	45.1 ± 17.5	2	346.6 ± 2.5	56.8 ± 1.0	2	345.7 ± 3.1	56.1 ± 0.1
	(3-comp.)			—	—	1	310.6 ± 0.5	8.8 ± 1.0	1	310.6 ± 0.5	8.8 ± 1.0	1	310.6 ± 0.5	8.8 ± 1.0
						2	332.4 ± 0.5	58.4 ± 1.0	2	332.4 ± 0.5	58.4 ± 1.0	2	332.4 ± 0.5	58.4 ± 1.0
						3	349.5 ± 0.5	32.8 ± 3.0	3	349.5 ± 0.5	32.8 ± 3.0	3	349.5 ± 0.5	32.8 ± 3.0
100	PAM (2-comp.)	336.4 ± 0.5	1	328.5 ± 4.1	30.1 ± 1.3	1	318.9 ± 1.5	36.3 ± 3.3	1	322.1 ± 5.6	32.1 ± 2.8	1	317.9 ± 0.5	34.0 ± 2.0
			2	340.4 ± 3.4	69.9 ± 0.6	2	341.7 ± 0.2	63.7 ± 3.3	2	341.2 ± 0.7	67.9 ± 2.8	2	341.5 ± 0.5	66.0 ± 2.0
	(3-comp.)			—	—	1	317.2 ± 0.5	32.1 ± 3.0	1	317.2 ± 0.5	32.1 ± 3.0	1	317.2 ± 0.5	32.1 ± 3.0
						2	333.2 ± 0.5	6.5 ± 2.0	2	333.2 ± 0.5	6.5 ± 2.0	2	333.2 ± 0.5	6.5 ± 2.0
						3	341.8 ± 0.5	61.4 ± 7.0	3	341.8 ± 0.5	61.4 ± 7.0	3	341.8 ± 0.5	61.4 ± 7.0

			PHREQ			SIMS			Average			Mean Best Fit		
No.	Protein Code	λ_m total (nm)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)
101	PMT.APO	349.9 ± 0.5		—	—	1	349.8 ± 0.1	100 ± 0	1	349.8 ± 0.1	100 ± 0	1	349.8 ± 0.5	100 ± 0
	(1-comp.)													
	(2-comp.)			—	—	1	331.2 ± 0.5	9.5 ± 2.0	1	331.2 ± 0.5	9.5 ± 2.0	1	331.2 ± 0.5	9.5 ± 2.0
						2	351.7 ± 0.5	90.5 ± 2.0	2	351.7 ± 0.5	90.5 ± 2.0	2	351.7 ± 0.5	90.5 ± 2.0
102	PMT.CA	324.8 ± 0.5		—	—	1	324.4 ± 1.1	100 ± 0	1	324.5 ± 0.8	100 ± 0	1	324.8 ± 0.5	100 ± 0
	(1-comp.)													
				—	—	1	313.0 ± 1.5	18.3 ± 2.0	1	313.0 ± 1.5	18.3 ± 2.0	1	313.0 ± 1.5	18.3 ± 2.0
	(2-comp.)					2	326.0 ± 0.5	81.7 ± 5.0	2	326.0 ± 0.5	81.7 ± 5.0	2	326.0 ± 0.5	81.7 ± 5.0
103	PMT.MG	32.2 ± 0.5		—	—	1	332.0 ± 0.2	100 ± 0	1	332.0 ± 0.2	100 ± 0	1	331.8 ± 0.5	100 ± 0
	(1-comp.)													
	(2-comp.)			—	—	1	325.1 ± 0.5	74.1 ± 1.0	1	325.1 ± 0.5	74.1 ± 1.0	1	325.1 ± 0.5	74.1 ± 1.0
						2	351.3 ± 0.5	25.9 ± 1.0	2	351.3 ± 0.5	25.9 ± 1.0	2	351.3 ± 0.5	25.9 ± 1.0
104	PAN.APO	348.6 ± 0.5		—	—	1	321.4 ± 2.5	6.5 ± 0.8	1	321.4 ± 2.5	6.5 ± 0.8	1	323.1 ± 0.5	7.1 ± 1.0
						2	349.9 ± 0.2	93.5 ± 0.8	2	349.9 ± 0.2	93.5 ± 0.8	2	350.0 ± 0.5	92.9 ± 1.0
105	PAN.CA	324.7 ± 0.5		—	—	1	324.7 ± 0.1	100 ± 0	1	324.7 ± 0.1	100 ± 0	1	324.7 ± 0.5	100 ± 0
106	PAN.MG	332.6 ± 0.5		—	—	1	317.0 ± 1.8	42.4 ± 3.5	1	317.0 ± 1.8	42.4 ± 3.5	1	318.2 ± 0.5	44.8 ± 2.0
						2	339.2 ± 0.8	57.6 ± 3.5	2	339.2 ± 0.8	57.6 ± 3.5	2	339.7 ± 0.8	55.2 ± 2.0
107	PEB.1-3	340.2 ± 0.4	1	334.4 ± 1.4	52.0 ± 13.8	1	332.0 ± 4.9	44.0 ± 21.9	1	332.8 ± 4.1	38.3 ± 17.5	1	330.8 ± 6.2	41.9 ± 23.0
	(2-comp.)		2	347.4 ± 2.4	48.0 ± 13.8	2	343.9 ± 1.5	56.0 ± 21.9	2	345.1 ± 2.4	61.7 ± 17.5	2	345.9 ± 3.4	58.1 ± 23.0
	(3-comp.)			—	—	1	319.4 ± 4.2	12.1 ± 6.8	1	319.4 ± 4.2	12.1 ± 6.8	1	319.4 ± 4.2	12.1 ± 6.8
						2	340.7 ± 0.8	51.4 ± 12.6	2	340.7 ± 0.8	51.4 ± 12.6	2	340.7 ± 0.8	51.4 ± 12.6
						3	344.3 ± 2.2	36.5 ± 6.6	3	344.3 ± 2.2	36.5 ± 6.6	3	344.3 ± 2.2	36.5 ± 6.6
108	PEB.4-5	342.5 ± 0.8	1	327.5 ± 7.0	22.0 ± 15.6	1	327.4 ± 4.6	20.6 ± 6.5	1	327.4 ± 5.2	19.0 ± 11.0	1	325.5 ± 6.0	15.7 ± 7.2
	(2-comp.)		2	347.1 ± 4.8	78.0 ± 15.6	2	345.7 ± 2.1	79.4 ± 6.5	2	346.2 ± 3.2	81.0 ± 11.0	2	345.1 ± 1.6	84.3 ± 7.2
	(3-comp.)			—	—	1	317.5 ± 4.1	12.0 ± 6.0	1	317.5 ± 4.1	12.0 ± 6.0	1	317.5 ± 4.1	12.0 ± 6.0
						2	338.7 ± 4.6	14.8 ± 6.0	2	338.7 ± 4.6	14.8 ± 6.0	2	338.7 ± 4.6	14.8 ± 6.0
						3	345.9 ± 2.9	73.2 ± 2.1	3	345.9 ± 2.9	73.2 ± 2.1	3	345.9 ± 2.9	73.2 ± 2.1
109	*PEB.6-9*	*342.5 ± 0.5*	*1*	*343.4 ± 0.4*	*100 ± 0*	*1*	*342.9 ± 0.6*	*100 ± 0*	*1*	*343.1 ± 0.6*	*100 ± 0*	*1*	*342.3 ± 0.4*	*100 ± 0*
	(1-comp.)													
	(2-comp.)		*1*	*337.2 ± 2.3*	*44.2 ± 10.4*	*1*	*335.2 ± 3.3*	*36.2 ± 11.4*	*1*	*336.2 ± 2.7*	*42.4 ± 10.3*	*1*	*336.5 ± 2.8*	*48.5 ± 10.2*
			2	*348.6 ± 2.0*	*55.8 ± 10.4*	*2*	*346.9 ± 0.6*	*63.8 ± 11.4*	*2*	*347.6 ± 1.5*	*57.6 ± 10.3*	*2*	*349.7 ± 1.0*	*51.5 ± 10.2*
110	*PEB.UR*	*348.1 ± 0.5*	*1*	*326.2 ± 3.1*	*10.3 ± 0.5*	*1*	*327.6 ± 7.5*	*12.5 ± 6.9*	*1*	*327.1 ± 5.4*	*13.9 ± 5.0*	*1*	*326.2 ± 3.1*	*10.3 ± 0.5*
			2	*350.0 ± 3.1*	*89.7 ± 0.1*	*2*	*349.9 ± 1.1*	*87.5 ± 6.9*	*2*	*349.9 ± 0.8*	*86.1 ± 5.0*	*2*	*350.0 ± 3.1*	*89.7 ± 0.1*
111	PHA	327.3 ± 0.5	1	320.2 ± 0.7	61.0 ± 5.2	1	323.3 ± 1.5	81.1 ± 4.7	1	322.5 ± 2.0	74.8 ± 11.7	1	322.8 ± 2.1	76.2 ± 13.5
			2	337.5 ± 2.3	39.0 ± 5.2	2	346.5 ± 2.5	18.9 ± 4.7	2	345.2 ± 7.1	25.2 ± 11.7	2	345.8 ± 9.9	23.8 ± 13.5
112	PGK	335.8 ± 0.2	1	323.7 ± 3.2	41.0 ± 3.1	1	323.7 ± 0.4	53.8 ± 0.1	1	323.7 ± 2.0	47.2 ± 7.5	1	323.0 ± 0.8	48.7 ± 8.7
			2	344.6 ± 1.7	59.0 ± 3.1	2	348.1 ± 1.5	46.2 ± 0.1	2	346.3 ± 2.4	52.8 ± 7.5	2	346.6 ± 1.1	51.3 ± 8.7
113	PLB	344.3 ± 0.5	1	331.8 ± 2.4	21.6 ± 17.7	1	330.0 ± 4.8	35.5 ± 4.7	1	330.5 ± 4.1	30.4 ± 12.2	1	331.7 ± 4.5	36.7 ± 2.7
			2	349.0 ± 3.7	79.4 ± 17.7	2	351.5 ± 1.5	65.5 ± 4.7	2	350.8 ± 2.3	69.6 ± 12.2	2	352.3 ± 1.5	63.3 ± 2.7
114	PLS	343.5 ± 0.1	1	323.7 ± 13.7	28.7 ± 22.6	1	325.5 ± 4.7	21.2 ± 8.9	1	324.9 ± 7.2	24.7 ± 14.2	1	323.7 ± 8.5	27.3 ± 16.1
			2	350.6 ± 5.6	71.3 ± 22.6	2	348.3 ± 1.5	78.8 ± 8.9	2	349.0 ± 3.0	75.3 ± 14.2	2	349.4 ± 3.6	72.7 ± 16.1
115	PAO.LIS	337.1 ± 0.4		—	—	1	320.1 ± 4.7	35.3 ± 10.6	1	320.1 ± 4.7	35.3 ± 10.6	1	322.4 ± 2.5	40.9 ± 7.9
						2	343.4 ± 2.9	64.7 ± 10.6	2	343.4 ± 2.9	64.7 ± 10.6	2	344.7 ± 2.8	59.1 ± 7.9
116	PAO.HIS	344.0 ± 0.2	1	329.0 ± 2.3	21.9 ± 6.9	1	330.3 ± 2.8	28.3 ± 11.9	1	329.9 ± 2.5	29.7 ± 9.8	1	331.4 ± 1.1	26.9 ± 10.0
			2	349.9 ± 1.2	78.1 ± 6.9	2	350.8 ± 2.3	71.8 ± 11.9	2	350.5 ± 1.9	70.3 ± 9.8	2	351.6 ± 1.3	73.1 ± 10.0
117	*PAO.GDN*	*346.5 ± 0.5*		—	—	*1*	*346.6 ± 0.2*	*100 ± 0*	*1*	*346.6 ± 0.2*	*100 ± 0*	*1*	*346.5 ± 0.5*	*100 ± 0*

			PHREQ			SIMS			Average			Mean Best Fit		
No.	Protein Code	λ_m total (nm)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)
118	PLC (1-comp.)	343.3 ± 0.3		—	—	1	342.9 ± 0.6	100 ± 0	1	342.9 ± 0.6	100 ± 0	1	343.0 ± 0.2	100 ± 0
	(2-comp.)			—	—	1	327.6 ± 2.6	8.4 ± 0.4	1	327.6 ± 2.6	8.4 ± 0.4	1	327.6 ± 2.6	8.4 ± 0.4
						2	344.3 ± 0.1	91.6 ± 0.4	2	344.3 ± 0.1	91.6 ± 0.4	2	344.3 ± 0.1	91.6 ± 0.4
119	PRL* (1-comp.)	336.9 ± 0.3		—	—	1	337.3 ± 0.4	100 ± 0	1	337.3 ± 0.4	100 ± 0	1	336.9 ± 0.4	100 ± 0
	(2-comp.)			—	—	1	334.0 ± 1.2	79.7 ± 9.2	1	334.0 ± 1.2	79.7 ± 9.2	1	333.6 ± 0.5	83.2 ± 2.0
						2	352.2 ± 4.2	20.2 ± 9.2	2	352.2 ± 4.2	20.2 ± 9.2	2	354.8 ± 1.5	16.8 ± 4.0
120	PCE	335.3 ± 0.1	1	323.5 ± 1.4	46.3 ± 2.2	1	329.9 ± 0.0	75.2 ± 0.7	1	326.7 ± 3.8	60.7 ± 16.8	1	327.4 ± 4.3	65.0 ± 17.6
			2	348.4 ± 0.5	53.7 ± 2.2	2	355.6 ± 0.1	24.8 ± 0.7	2	357.0 ± 9.9	39.3 ± 16.8	2	359.7 ± 10.1	35.0 ± 17.7
121	PRH.LIS	345.0 ± 1.0	1	344.9 ± 0.5	100 ± 0	1	345.2 ± 0.9	100 ± 0	1	345.0 ± 0.9	100 ± 0	1	345.4 ± 0.7	100 ± 0
122	PRH.HIS	342.8 ± 0.7	1	328.8 ± 3.5	23.3 ± 3.2	1	332.2 ± 1.7	39.1 ± 3.7	1	331.1 ± 2.7	30.8 ± 9.0	1	328.8 ± 3.5	23.3 ± 3.2
			2	348.5 ± 1.9	76.7 ± 3.2	2	350.7 ± 1.6	60.9 ± 3.7	2	349.9 ± 1.9	69.2 ± 9.0	2	348.5 ± 1.9	76.7 ± 3.2
123	PRH.PST	344.3 ± 1.1	1	344.4 ± 0.5	100 ± 0	1	344.6 ± 1.1	100 ± 0	1	344.6, 1.0	100 ± 0	1	344.9 ± 0.9	100 ± 0
124	PRS	333.9 ± 0.1	1	333.0 ± 0.9	100 ± 0	1	333.9 ± 0.2	100 ± 0	1	333.6 ± 0.6	100 ± 0	1	333.9 ± 0.1	100 ± 0
125	PRK	332.9 ± 0.6	1	323.7 ± 2.9	33.9 ± 8.4	1	328.0 ± 1.6	72.5 ± 9.8	1	326.4 ± 3.0	58.4 ± 21.4	1	326.9 ± 2.9	59.3 ± 22.2
			2	337.8 ± 1.4	66.1 ± 8.4	2	347.1 ± 3.7	27.5 ± 9.8	2	343.7 ± 5.5	41.6 ± 21.4	2	344.2 ± 6.5	40.7 ± 22.2
126	PYK	338.7 ± 1.3	1	328.2 ± 5.4	32.0 ± 12.7	1	331.2 ± 3.8	44.9 ± 14.4	1	330.5 ± 4.5	33.0 ± 10.9	1	328.2 ± 5.4	32.0 ± 12.7
			2	344.5 ± 4.1	68.0 ± 12.7	2	346.0 ± 2.3	55.1 ± 14.4	2	345.5 ± 3.0	67.0 ± 10.9	2	344.5 ± 4.1	68.0 ± 12.7
127	PYK.PEG	339.9 ± 0.5	1	321.1 ± 2.9	25.2 ± 0.1	1	325.6 ± 1.8	45.3 ± 8.6	1	324.1 ± 2.9	32.2 ± 9.9	1	321.1 ± 2.9	25.2 ± 0.1
			2	346.2 ± 3.1	74.8 ± 0.0	2	350.0 ± 1.7	54.7 ± 8.6	2	348.7 ± 2.5	67.8 ± 9.9	2	346.2 ± 3.1	74.8 ± 0.0
128	REC.APO*	329.9 ± 0.5		—	—	1	319.4 ± 0.5	51.3 ± 1.0	1	319.4 ± 0.5	51.3 ± 1.0	1	319.4 ± 0.5	51.3 ± 1.0
						2	336.7 ± 0.5	48.7 ± 3.0	2	336.7 ± 0.5	48.7 ± 3.0	2	336.7 ± 0.5	48.7 ± 3.0
129	REC.CA*	336.6 ± 0.5		—	—	1	326.0 ± 0.5	50.6 ± 1.0	1	326.0 ± 0.5	50.6 ± 1.0	1	326.0 ± 0.5	50.6 ± 1.0
						2	345.2 ± 0.5	49.4 ± 2.0	2	345.2 ± 0.5	49.4 ± 2.0	2	345.2 ± 0.5	49.4 ± 2.0
130	RCG	339.1 ± 0.7	1	325.6 ± 3.3	33.5 ± 6.1	1	332.3 ± 2.5	64.2 ± 16.5	1	330.0 ± 4.2	53.8 ± 22.6	1	330.5 ± 3.8	65.0 ± 18.3
			2	346.8 ± 2.2	66.5 ± 6.1	2	351.8 ± 1.1	35.8 ± 16.5	2	348.8 ± 3.2	46.2 ± 22.6	2	359.2 ± 12.9	35.0 ± 18.3
131	RNC (1-comp.)	327.4 ± 0.4	1	327.5 ± 0.6	100 ± 0	1	327.5 ± 0.7	100 ± 0	1	327.4 ± 0.6	100 ± 0	1	327.4 ± 0.6	100 ± 0
	(2-comp.)		1	324.5 ± 1.6	69.9 ± 16.2	1	325.8 ± 1.6	74.7 ± 18.5	1	325.3 ± 1.7	64.5 ± 21.9	1	323.7 ± 3.1	73.4 ± 16.1
			2	335.1 ± 3.5	30.1 ± 16.2	2	334.0 ± 4.3	25.3 ± 18.5	2	334.6 ± 3.9	35.5 ± 21.9	2	335.1 ± 3.6	26.6 ± 16.1
132	RNI	334.6 ± 0.5	1	318.6 ± 2.9	28.8 ± 0.1	1	328.1 ± 0.6	67.1 ± 6.3	1	324.9 ± 5.5	50.2 ± 30.2	1	328.5 ± 0.5	71.5 ± 1.0
			2	342.0 ± 3.0	71.2 ± 0.1	2	357.6 ± 1.7	32.9 ± 6.3	2	352.4 ± 9.1	49.8 ± 30.2	2	358.8 ± 0.5	28.5 ± 1.0
133	RNT	325.6 ± 0.6	1	325.1 ± 1.0	100 ± 0	1	324.6 ± 1.9	100 ± 0	1	324.5 ± 1.8	100 ± 0	1	325.1 ± 0.8	100 ± 0
134	RS4 (2-comp.)	336.1 ± 0.1	1	330.0 ± 1.8	57.1 ± 4.2	1	329.6 ± 0.1	59.1 ± 4.6	1	329.8 ± 1.0	58.1 ± 3.8	1	330.0 ± 1.8	57.1 ± 4.2
			2	345.1 ± 0.1	42.9 ± 4.2	2	345.2 ± 1.6	40.9 ± 4.6	2	345.1 ± 0.9	41.9 ± 3.8	2	345.1 ± 0.1	42.9 ± 4.2
	(3-comp.)			—	—	1	324.6 ± 0.7	18.5 ± 2.5	1	324.6 ± 0.7	18.5 ± 2.5	1	324.6 ± 0.7	18.5 ± 2.5
						2	335.0 ± 0.8	61.6 ± 11.1	2	335.0 ± 0.8	61.6 ± 11.1	2	335.0 ± 0.8	61.6 ± 11.1
						3	350.8 ± 3.9	19.9 ± 8.7	3	350.8 ± 3.9	19.9 ± 8.7	3	350.8 ± 3.9	19.9 ± 8.7
135	RS7	341.9 ± 0.7	1	329.6 ± 3.4	33.1 ± 11.9	1	324.9 ± 5.3	26.6 ± 7.4	1	326.5 ± 5.1	31.2 ± 8.0	1	329.6 ± 3.4	31.8 ± 10.1
			2	351.1 ± 2.8	66.9 ± 11.9	2	349.0 ± 1.6	73.4 ± 7.4	2	349.7 ± 2.1	69.8 ± 8.0	2	351.1 ± 2.8	68.2 ± 10.1
136	RIP (1-comp.)	340.6 ± 0.5	1	341.8 ± 0.8	100 ± 0	1	342.7 ± 1.9	100 ± 0	1	342.0 ± 1.5	100 ± 0	1	341.9 ± 1.2	100 ± 0
137	RCA (2-comp.)	330.5 ± 1.4	1	325.8 ± 3.0	56.4 ± 20.8	1	326.9 ± 1.1	83.1 ± 6.0	1	326.2 ± 1.7	67.6 ± 18.6	1	324.9 ± 1.8	62.7 ± 21.2
			2	343.7 ± 4.8	43.6 ± 20.8	2	353.4 ± 6.4	16.9 ± 6.0	2	351.2 ± 7.1	32.4 ± 18.6	2	345.5 ± 4.9	37.3 ± 21.2
	(3-comp.)			—	—	1	324.0 ± 3.3	44.7 ± 25.9	1	324.0 ± 3.3	44.7 ± 25.9	1	324.0 ± 3.3	44.7 ± 25.9
						2	330.3 ± 1.7	43.9 ± 23.7	2	330.3 ± 1.7	43.9 ± 23.7	2	330.3 ± 1.7	43.9 ± 23.7
						3	359.0 ± 3.8	11.4 ± 3.9	3	359.0 ± 3.8	11.4 ± 3.9	3	359.0 ± 3.8	11.4 ± 3.9
138	RCB.4	336.1 ± 1.3	1	331.7 ± 1.9	48.9 ± 12.7	1	331.5 ± 1.4	63.7 ± 11.0	1	331.6 ± 1.4	60.3 ± 14.8	1	330.1 ± 1.7	61.0 ± 12.0
			2	345.9 ± 3.2	51.1 ± 12.7	2	351.0 ± 4.5	36.3 ± 11.0	2	348.3 ± 6.5	39.7 ± 14.8	2	349.7 ± 5.0	39.0 ± 12.0

			PHREQ			SIMS			Average			Mean Best Fit		
No.	Protein Code	λ_m total (nm)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)	No.	λ_m (i) (nm)	$S(i)$ (%)
139	RCB.4.LAC	330.9 ± 0.9	1	326.1 ± 0.4	54.9 ± 10.5	1	326.3 ± 0.8	74.7 ± 6.2	1	326.2 ± 0.5	64.8 ± 13.4	1	326.0 ± 0.4	77.8 ± 1.7
			2	340.6 ± 7.3	45.1 ± 10.5	2	349.7 ± 1.3	25.3 ± 6.2	2	345.1 ± 6.8	39.2 ± 13.4	2	348.2 ± 3.5	22.2 ± 1.7
140	RCB.5-7	331.4 ± 0.9	1	322.8 ± 3.7	41.7 ± 18.3	1	326.6 ± 0.9	62.5 ± 10.4	1	325.5 ± 2.6	55.3 ± 19.1	1	324.2 ± 3.2	58.7 ± 15.2
	(2-comp.)		2	343.1 ± 6.2	58.3 ± 18.3	2	344.9 ± 3.2	37.5 ± 10.4	2	343.8 ± 3.8	44.7 ± 19.1	2	344.5 ± 4.6	41.3 ± 15.2
	(3-comp.)			—	—	1	327.0 ± 1.0	54.9 ± 11.2	1	327.0 ± 1.0	54.9 ± 11.2	1	327.0 ± 1.0	54.9 ± 11.2
						2	333.7 ± 1.7	29.0 ± 11.1	2	333.7 ± 1.7	29.0 ± 11.1	2	333.7 ± 1.7	29.0 ± 11.1
						3	356.0 ± 2.7	16.2 ± 1.9	3	356.0 ± 2.7	16.2 ± 1.9	3	356.0 ± 2.7	16.2 ± 1.9
141	RCC	331.4 ± 1.2	1	320.5 ± 2.9	40.6 ± 11.4	1	327.2 ± 1.1	67.8 ± 12.3	1	324.9 ± 4.0	58.3 ± 20.1	1	324.1 ± 3.7	58.4 ± 20.5
			2	339.8 ± 3.3	59.4 ± 11.4	2	346.0 ± 5.0	32.2 ± 12.3	2	342.8 ± 7.5	41.7 ± 20.1	2	345.6 ± 6.8	41.6 ± 20.5
142	STP	333.5 ± 0.5		—	—	1	335.6 ± 1.0	100 ± 0	1	335.6 ± 1.0	100 ± 0	1	335.0 ± 0.5	100 ± 0
143	SPC (2-comp.)	341.7 ± 0.1	1	324.9 ± 0.4	13.4 ± 3.0	1	326.3 ± 7.7	33.3 ± 26.9	1	326.2 ± 6.0	18.3 ± 8.8	1	325.6 ± 1.3	18.3 ± 8.8
			2	344.7 ± 0.1	86.6 ± 3.0	2	345.5 ± 2.0	66.7 ± 26.9	2	345.2 ± 1.5	81.7 ± 8.8	2	345.4 ± 1.2	81.7 ± 8.8
	(3-comp.)			—	—	1	322.0 ± 6.0	352.2 ± 2.8	1	322.0 ± 6.0	352.2 ± 2.8	1	322.0 ± 6.0	352.2 ± 2.8
						2	342.1 ± 3.6	58.6 ± 14.7	2	342.1 ± 3.6	58.6 ± 14.7	2	342.1 ± 3.6	58.6 ± 14.7
						3	352.2 ± 2.8	24.4 ± 0.9	3	352.2 ± 2.8	24.4 ± 0.9	3	352.2 ± 2.8	24.4 ± 0.9
144	BPN	345.7 ± 0.7	1	346.2 ± 0.6	100 ± 0	1	345.7 ± 0.7	100 ± 0	1	346.2 ± 0.7	100 ± 0	1	346.2 ± 0.6	100 ± 0
145	SLC	350.0 ± 2.0	1	315.8 ± 8.6	17.5 ± 5.8	1	319.8 ± 4.9	22.0 ± 4.6	1	318.4 ± 6.5	20.4 ± 5.8	1	315.8 ± 8.6	17.5 ± 5.8
			2	353.0 ± 1.0	82.5 ± 5.8	2	354.5 ± 1.7	78.0 ± 4.6	2	354.0 ± 1.6	79.6 ± 5.8	2	353.0 ± 1.0	82.5 ± 5.8
146	KRP	332.4 ± 0.5	1	332.4 ± 3.1	100 ± 0	1	332.9 ± 0.5	100 ± 0	1	332.6 ± 0.4	100 ± 0	1	332.4 ± 0.5	100 ± 0
147	*TMV.UR*	*352.2 ± 3.3*	1	*352.2 ± 3.3*	*100 ± 0*	*1*	*352.2 ± 0.1*	*100 ± 0*	*1*	*352.2 ± 0.1*	*100 ± 0*	*1*	*352.5 ± 0.5*	*100 ± 0*
148	TNT	341.8 ± 0.5	1	343.5 ± 3.1	100 ± 0	1	343.5 ± 2.4	100 ± 0	1	343.2 ± 1.8	100 ± 0	1	343.9 ± 1.9	100 ± 0
149	TRY	334.5 ± 1.3	1	329.5 ± 2.4	55.8 ± 22.7	1	330.7 ± 2.1	72.6 ± 16.4	1	330.3 ± 2.2	63.1 ± 19.9	1	328.8 ± 3.0	59.9 ± 20.4
			2	346.1 ± 4.6	44.2 ± 22.7	2	347.0 ± 4.3	23.4 ± 16.4	2	346.6 ± 4.3	36.9 ± 19.9	2	346.8 ± 4.8	40.1 ± 20.4
150	*TRY.D35*	*339.0 ± 0.5*	*1*	*333.8 ± 3.0*	*65.8 ± 0.1*	*1*	*331.2 ± 4.2*	*54.9 ± 25.9*	*1*	*332.1 ± 3.4*	*51.2 ± 20.6*	*1*	*333.8 ± 3.0*	*65.8 ± 0.1*
			2	*349.9 ± 3.1*	*34.2 ± 0.1*	*2*	*347.8 ± 4.5*	*45.1 ± 25.9*	*2*	*348.5 ± 3.4*	*48.8 ± 20.6*	*2*	*349.9 ± 3.1*	*34.2 ± 0.1*
151	*TRY.D60-70*	*338.0 ± 0.9*	*1*	*329.0 ± 4.2*	*32.9 ± 19.4*	*1*	*332.1 ± 3.0*	*31.5 ± 23.6*	*1*	*330.8 ± 3.6*	*37.5 ± 22.9*	*1*	*329.0 ± 4.2*	*32.9 ± 19.4*
			2	*342.1 ± 3.0*	*67.1 ± 19.4*	*2*	*341.7 ± 3.6*	*68.5 ± 23.6*	*2*	*341.8 ± 3.3*	*62.6 ± 22.9*	*2*	*342.1 ± 3.0*	*67.1 ± 19.4*
152	TRG (1-comp.)	332.6 ± 0.4	1	332.0 ± 0.9	100 ± 0	1	332.5 ± 0.4	100 ± 0	1	332.3 ± 0.6	100 ± 0	1	332.3 ± 0.5	100 ± 0
	(2-comp.)		1	327.5 ± 4.2	44.8 ± 29.7	1	327.6 ± 1.4	48.0 ± 18.9	1	327.6 ± 2.1	48.5 ± 24.2	1	327.0 ± 3.1	49.4 ± 24.9
			2	338.9 ± 3.9	55.2 ± 29.7	2	337.2 ± 2.8	52.0 ± 18.9	2	337.3 ± 2.8	51.5 ± 24.2	2	337.7 ± 3.3	50.6 ± 24.9
153	VIN	338.4 ± 0.5	1	333.8 ± 1.6	55.7 ± 11.5	1	333.3 ± 3.4	61.0 ± 3.2	1	333.5 ± 2.5	63.7 ± 15.1	1	333.8 ± 1.6	55.7 ± 11.5
			2	346.1 ± 1.4	44.3 ± 11.2	2	345.6 ± 0.5	39.0 ± 3.2	2	345.8 ± 0.8	36.3 ± 15.1	2	346.1 ± 1.4	44.3 ± 11.2
154	*VIN.UR*	*345.3 ± 0.6*	*1*	*341.8 ± 2.1*	*51.2 ± 26.7*	*1*	*343.3 ± 0.1*	*57.9 ± 13.4*	*1*	*342.5 ± 1.7*	*54.5 ± 17.7*	*1*	*341.8 ± 2.1*	*51.2 ± 26.7*
			2	*351.7 ± 2.6*	*48.8 ± 26.7*	*2*	*350.6 ± 1.1*	*42.1 ± 13.4*	*2*	*351.1 ± 1.8*	*45.5 ± 17.7*	*2*	*351.7 ± 2.6*	*48.8 ± 26.7*
155	VTA.AC	339.4 ± 1.7		—	—	1	333.9 ± 0.9	74.3 ± 6.8	1	333.9 ± 0.9	74.3 ± 6.8	1	333.7 ± 0.7	75.5 ± 8.7
						2	357.3 ± 8.3	25.7 ± 6.8	2	357.3 ± 8.3	25.7 ± 6.8	2	353.7 ± 8.4	24.5 ± 8.7
156	VTA.HIS	339.9 ± 2.6	1	321.1 ± 1.9	13.8 ± 4.2	1	321.6 ± 4.6	25.3 ± 12.7	1	321.5 ± 3.9	23.1 ± 11.5	1	323.6 ± 3.4	30.6 ± 17.6
			2	345.6 ± 1.8	86.2 ± 4.2	2	345.3 ± 3.2	74.7 ± 12.7	2	345.3 ± 2.8	76.9 ± 11.5	2	346.2 ± 2.1	69.4 ± 17.6
157	VTA.LIS	348.2 ± 0.8	1	349.1 ± 4.0	100 ± 0	1	348.9 ± 0.5	100 ± 0	1	349.3 ± 0.7	100 ± 0	1	349.1 ± 0.7	100 ± 0
158	VTB (1-comp.)	343.2 ± 0.8	1	343.8 ± 3.2	100 ± 0	1	343.8 ± 1.0	100 ± 0	1	343.7 ± 0.9	100 ± 0	1	344.0 ± 0.7	100 ± 0
	(2-comp.)		1	331.7 ± 3.2	23.2 ± 0.3	1	336.4 ± 4.9	45.4 ± 38.0	1	335.9 ± 4.8	47.2 ± 32.8	1	336.2 ± 4.3	48.4 ± 33.4
			2	347.5 ± 3.1	76.8 ± 0.2	2	353.5 ± 10.9	54.6 ± 38.0	2	352.9 ± 10.5	52.8 ± 32.8	2	351.2 ± 9.0	51.6 ± 33.4
159	VTC	334.3 ± 0.3		—	—	1	334.3 ± 0.3	100 ± 0	1	334.3 ± 0.3	100 ± 0	1	334.2 ± 0.3	100 ± 0
160	*VSC.AC*	*344.4 ± 0.5*	*1*	*336.2 ± 5.8*	*40.3 ± 4.1*	*1*	*329.6 ± 3.3*	*38.3 ± 13.5*	*1*	*331.8 ± 4.5*	*39.0 ± 9.7*	*1*	*331.9 ± 0.5*	*47.9 ± 1.0*
			2	*350.3 ± 6.4*	*59.7 ± 3.2*	*2*	*358.6 ± 3.3*	*41.7 ± 13.6*	*2*	*355.8 ± 5.3*	*61.0 ± 9.7*	*2*	*360.9 ± 0.5*	*52.1 ± 1.0*
161	VSC.N	332.5 ± 1.9	1	324.6 ± 2.9	45.1 ± 16.6	1	325.1 ± 1.5	59.5 ± 10.7	1	324.9 ± 2.0	54.6 ± 16.5	1	325.2 ± 1.6	64.0 ± 10.6

No.	Protein Code	λ_m total (nm)	PHREQ			SIMS			Average			Mean Best Fit		
			No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)	No.	λ_m (*i*) (nm)	*S*(*i*) (%)
	(2-comp.)		2	341.9 ± 7.3	54.9 ± 16.6	2	345.9 ± 4.4	40.5 ± 10.7	2	344.6 ± 5.6	45.4 ± 16.5	2	347.6 ± 4.2	36.0 ± 10.6
	(3-comp.)			—	—	1	324.1 ± 2.1	43.2 ± 21.4	1	324.1 ± 2.1	43.2 ± 21.4	1	324.1 ± 2.1	43.2 ± 21.4
						2	333.7 ± 6.4	34.8 ± 15.2	2	333.7 ± 6.4	34.8 ± 15.2	2	333.7 ± 6.4	34.8 ± 15.2
						3	350.8 ± 3.6	22.0 ± 7.6	3	350.8 ± 3.6	22.0 ± 7.6	3	350.8 ± 3.6	22.0 ± 7.6
162a	VSC.MOD	341.7 ± 0.2	1	328.2 ± 2.3	25.7 ± 6.9	1	330.1 ± 1.2	32.2 ± 6.1	1	329.5 ± 1.7	31.4 ± 7.8	1	328.2 ± 2.3	25.7 ± 6.9
	(2-comp.)		2	350.7 ± 0.5	74.3 ± 6.9	2	355.1 ± 2.9	67.8 ± 6.1	2	353.3 ± 3.2	68.6 ± 7.8	2	350.7 ± 0.5	74.3 ± 6.9
162b	*VSC.MOD*			—	—	*1*	*329.7 ± 1.6*	*27.3 ± 3.0*	*1*	*329.7 ± 1.6*	*27.3 ± 3.0*	*1*	*329.7 ± 1.6*	*27.3 ± 3.0*
	(3-comp.)					*2*	*344.7 ± 1.3*	*42.1 ± 5.2*	*2*	*344.7 ± 1.3*	*42.1 ± 5.2*	*2*	*344.7 ± 1.3*	*42.1 ± 5.2*
						3	*361.1 ± 4.0*	*30.6 ± 2.2*	*3*	*361.1 ± 4.0*	*30.6 ± 2.2*	*3*	*361.1 ± 4.0*	*30.6 ± 2.2*
163	*VSC.UR*	*349.5 ± 0.5*	*1*	*331.2 ± 3.2*	*9.6 ± 0.1*	*1*	*338.6 ± 0.9*	*19.6 ± 4.5*	*1*	*336.1 ± 4.4*	*16.2 ± 9.2*	*1*	*331.2 ± 3.2*	*9.6 ± 0.1*
			2	*353.5 ± 3.4*	*90.4 ± 0.1*	*2*	*354.8 ± 0.1*	*80.4 ± 4.5*	*2*	*354.3 ± 0.7*	*83.8 ± 9.2*	*2*	*353.5 ± 3.4*	*90.4 ± 0.1*

Information on denatured proteins appears in italics.
*Fluorescence was excited at 280 nm.

obtained with two decomposition algorithms, SIMS (375 components) and PHREQ (249 components), respectively. Panels *C* and *D* reflect the results obtained using two schemes of averaging ("Average" and "Mean Best Fit" in Table 2), respectively. All four histograms have the well-expressed deep minimum at 335–337 nm as a common feature. This feature seems to clearly demonstrate the existence of a statistical discreteness of at least two large spectral classes of tryptophan residues in native proteins. Moreover, panels *A–D* demonstrate reproducible dips at 329 ± 1 and 347 ± 1 nm, which may also mark the frontiers between overlapping discrete classes of tryptophans in proteins.

However, the canonical form of histogram, constructed from columns, cannot reflect the accuracy of calculated maximum position values and of relative contributions of components in the whole emission spectrum. Therefore, we constructed a new kind of histogram accounting for these factors (Figs. 3-5). In such histograms, each element (the individual *i*th spectral component) is represented by the little Gauss distribution $y(\lambda_m(i))$ with maximum at $\lambda_m(i)$, $\sigma(i)$ equal to the standard (root-mean-square one) error of the mean $\lambda_m(i)$ value (see Table 2), and the maximal amplitude of Gauss curves proportional to the relative contribution $S(i)$ of the *i*th component to the total protein emission. Such a histogram is a sum of these elements:

$$y(\lambda_m) = \alpha(i)\sum_i \left\{ \frac{S(i)}{\sigma(i)} \exp\left[\frac{\lambda - \lambda(i)}{\sigma(i)} \right]^2 \right\}$$

Here, $S(i)$ is expressed in decimal fractions and $\alpha(i)$ is a multiplier equal to 1 if only one decomposition result is selected for a protein spectrum, or to 0.5 in the cases when two solutions with similar functional values are taken.

In such a representation, the histograms of Fig. 2 acquire the view seen in Fig. 3 (curves N). In this case, the $S(i)$ values are taken to be 1, and $\sigma = 1.0$ for every spectral component. One can see that the main features of the canonical histograms are retained in the new, Gauss-curve-based representation. Besides the histograms for components of spectra of native proteins, Fig. 3 also contains the curves obtained for proteins denatured under various influences (see Table 2). The shorter-wavelength "continua" reflect the fact that denaturation is not completed in some cases and the emission spectra possess components belonging to the retained populations of native and/or partly unfolded molecules. However, the highest peak at ~350 nm and peaks or shoulders at 340–347 nm are due to the components belonging to tryptophans in proteins denatured by urea or guanidinium chloride and by dioxane, pH and other less unfolding denaturing factors. It is interesting that the histograms for denatured proteins also have the minima at 335–340 nm, which position coincides with the deep minima of histograms for native proteins.

We constructed two more Gauss-curve-based histograms with values of $\sigma(i)$ and $S(i)$ varying according to

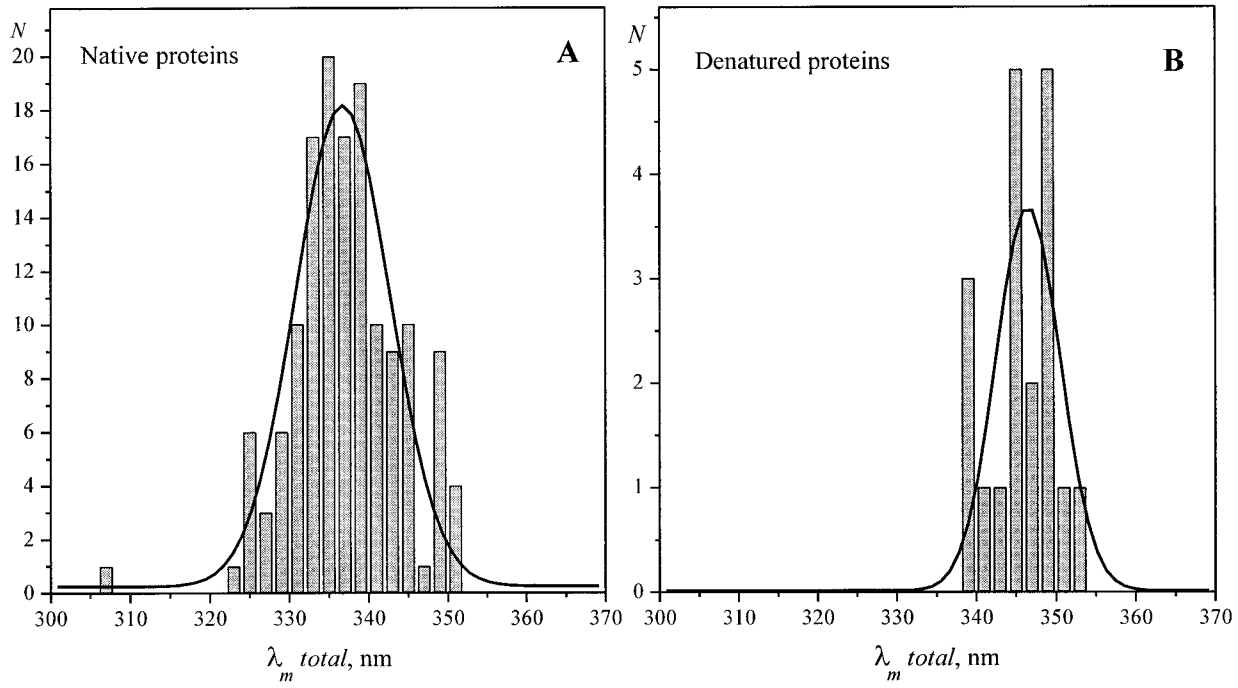

FIGURE 1 Canonical distribution of occurrence of maximum positions of total, non-decomposed emission spectra for native (*A*) and denatured (*B*) proteins with 2-nm step. Solid lines represent the fitting by Gauss function.

the calculation data for native (N) and denatured (D) proteins. The histograms accounting only for the differences in contributions of individual components in whole spectra are presented in Fig. 4 ($\sigma(i)$ = 1.0 nm; $S(i)$ = var). The resulting reduction in contributions of "small" components ($S(i) < 0.01$) in the histogram essentially improved the resolution of individual peaks in it. The "small" components may not only have worse determined λ_m values, but also may be erroneous in principle in the cases of bad quality of a spectrum under decomposition. The fact that the improved resolution of peaks in such histograms at ~326, 331, 345, and 350 nm allowed us to assume that they reflect the existence of discrete classes of tryptophan emitters in native proteins. The reduction of contributions of minor components essentially decreased the relative amplitude of a shorter-wavelength wing in the denatured-proteins' histogram. The remaining shorter-wavelength peaks (instead of a "continuum") roughly correspond to the peaks in the native-proteins' histogram. However, the 350-nm peak and the shoulder at 340–347 nm became much better-expressed.

To demonstrate that the resolution of peaks obtained in previous histograms (and, thus, the observed spectral discrete classes) are beyond the error of the estimation of the maximum position of spectral components, we constructed the histograms (Fig. 5) accounting for the precision of $\lambda_m(i)$ estimation ($\sigma(i)$ = var) and the relative contributions of components ($S(i)$ = var). These histograms demonstrate the presence of discreteness of emission of tryptophan residues in proteins in the same manner as Figs. 3 and 4. There are seen the global minimum at ~337 nm and four maxima at ~326, 333, 344, and 350 nm for native proteins. The PHREQ algorithm well-resolves only two classes, because it was applied only to the limited number of protein spectra, possessing one or two components. For denatured proteins, accounting for $\sigma(i)$ made peaks even more expressive. The main maxima for denatured proteins are at ~346 and 350 nm; however, the peak at ~334 nm also exists.

DISCUSSION

In 1973–1977 Burstein and co-workers postulated the hypothesis of the existence of discrete classes of tryptophan residues in proteins (Burstein et al., 1973; Burstein, 1977a, 1983). Later, algorithms of stable and reliable decomposition of composite tryptophan fluorescence spectra into log-normal components were developed (Abornev and Burstein, 1992; Burstein and Emelyanenko, 1996; Burstein et al., 2001) that allowed us to verify and revise the hypothesis. Here, we used mathematically different algorithms of decomposition of emission spectra to examine the stability of solutions. The fitting algorithm (SIMS) and the analytical ones (PHREQ) gave, in general, very similar results: obtained maximum positions $\lambda_m(i)$ differed within an error of ±2.8 nm, and contributions $S(i)$ of components differed within the range of ±8.6%.

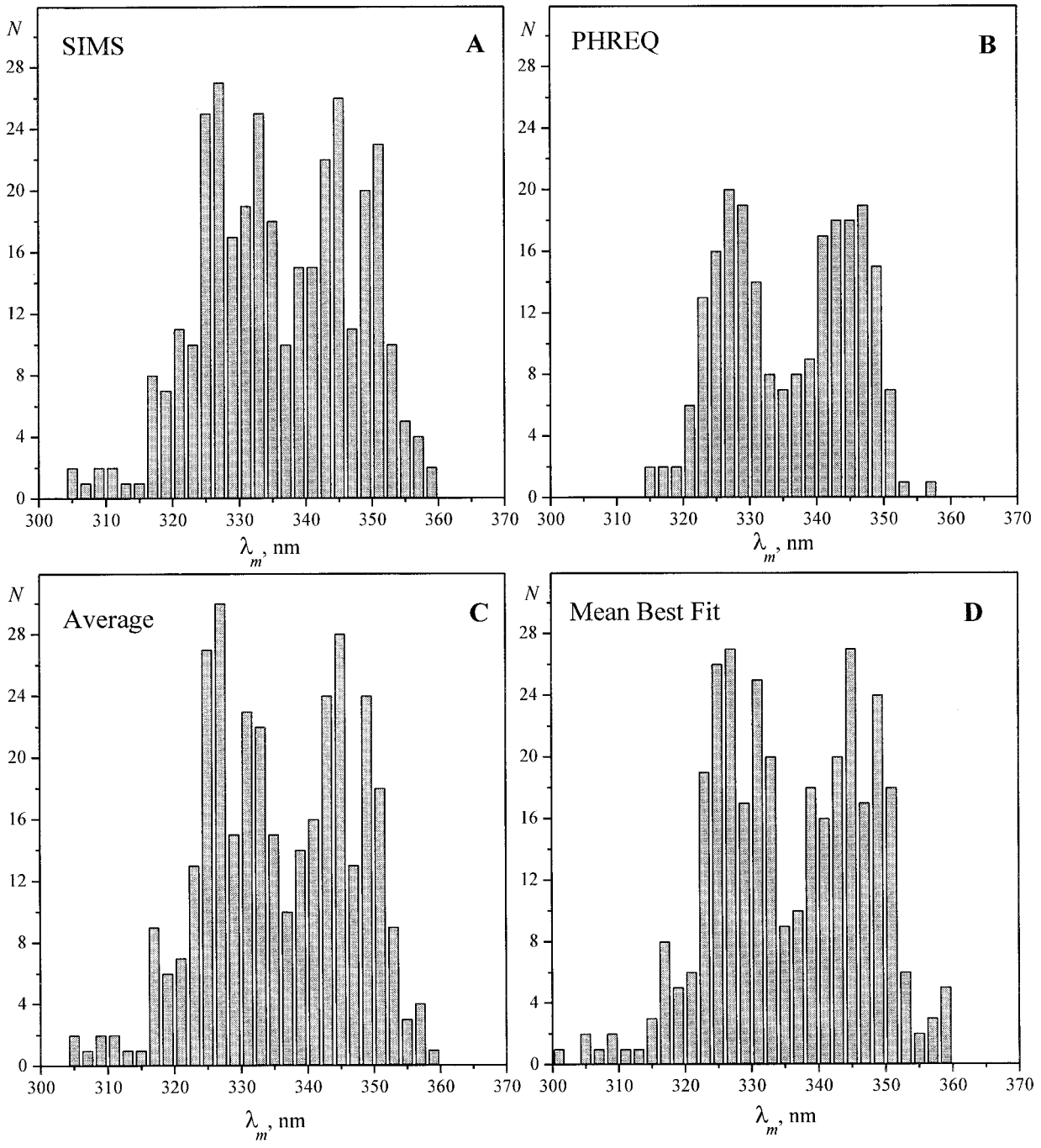

FIGURE 2 Canonical distribution of occurrence of maximum positions of log-normal components obtained by applying SIMS (*A*), PHREQ (*B*) algorithms, and different schemes of averaging of the results: "Average" (*C*) and "Mean Best Fit" (*D*) for native proteins with 2-nm step.

First of all, we showed that the distribution of maximum positions of total, nondecomposed spectra ("λ_m total" in Table 2) for native proteins is rather well-fitted by a Gauss equation with the maximum at ~337 nm and dispersion of ~12 nm (Fig. 1 *A*), which demonstrated the randomness and, thus, representativeness of the set of the proteins under study. Then, we analyzed the distributions of the maximum position of log-normal components taking into consideration the experimental deviations of maximum position and the contributions of components to spectra of native and denatured proteins (Figs. 2–5). All obtained data confirmed the existence of several statistically discrete classes of emitting tryptophan fluorophores in proteins. The discrete classes observed here are in approximate agreement with those proposed in 1973–1977 in the model of discrete states. A very important point is that there was no a priori assump-

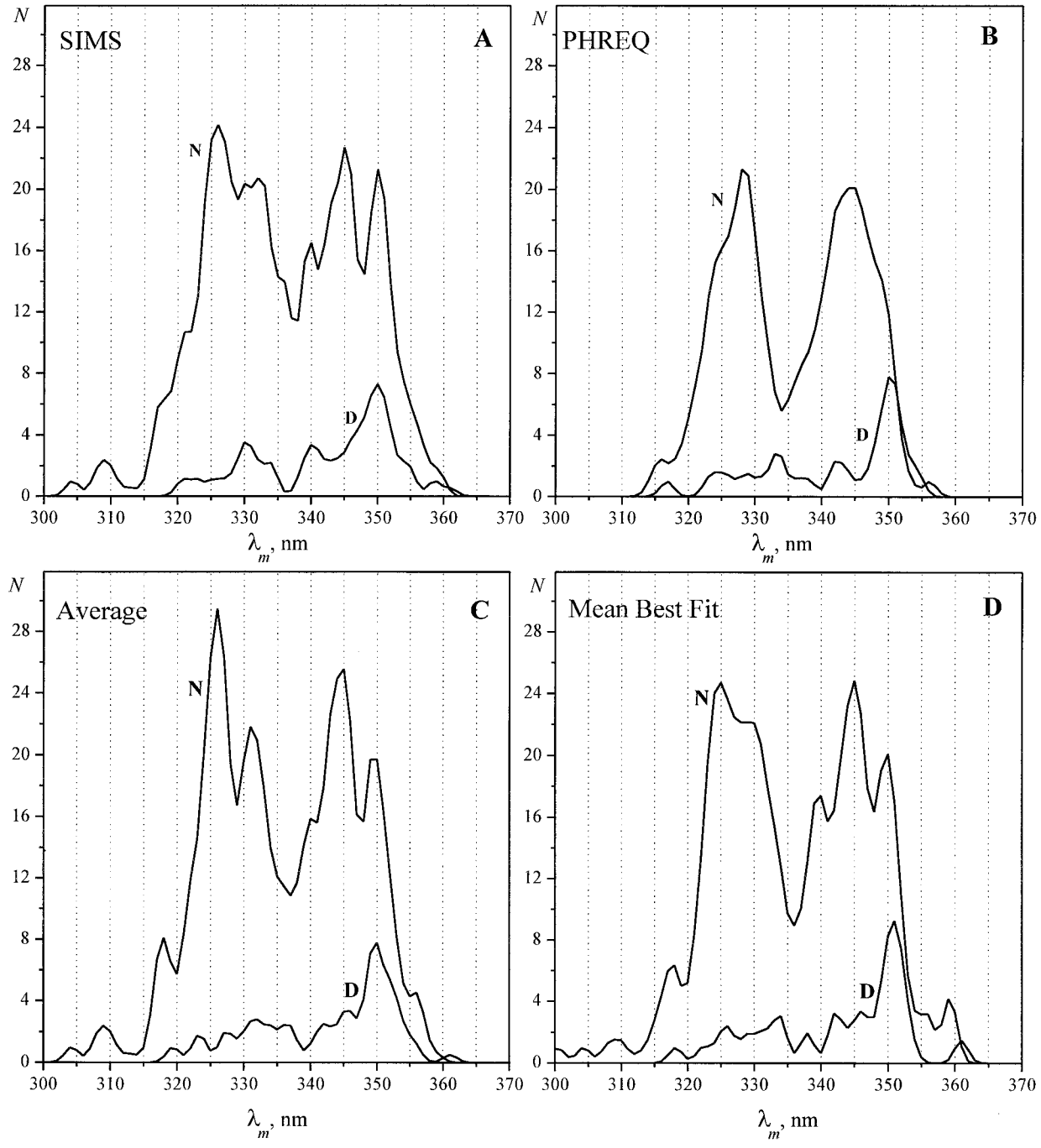

FIGURE 3 Distribution of occurrence of maximum positions of log-normal components obtained by applying SIMS (*A*), PHREQ (*B*) algorithms, and different schemes of averaging of the results: "Average" (*C*) and "Mean Best Fit" (*D*) with 1-nm step. Each curve is a sum of Gauss function (see explanation in the text) for each spectral component of native (N) and denatured (D) proteins with $\sigma(i) = 1.0$ and $S(i) = 1.0$.

tion about the discreteness in any algorithm used for component analysis of fluorescence spectra (Burstein et al., 2001).

It is evident that to find the causes for so widely differing and discrete Stokes shifts in the emission spectra of tryptophan fluorophores in proteins one has to look in a variety of combinations of interactions of individual fluorophores with their environment, occurring mainly during the lifetime of the fluorescent excited state. The physical causes for the differences among the spectroscopic classes can be analyzed in terms of various combinations of Bakhsiev's theory of specific and universal interactions of the excited fluorophores with their environment (Bakhshiev, 1972). Both kinds of interactions are induced by essential redistribution of electronic density on the atoms and chemical bonds of

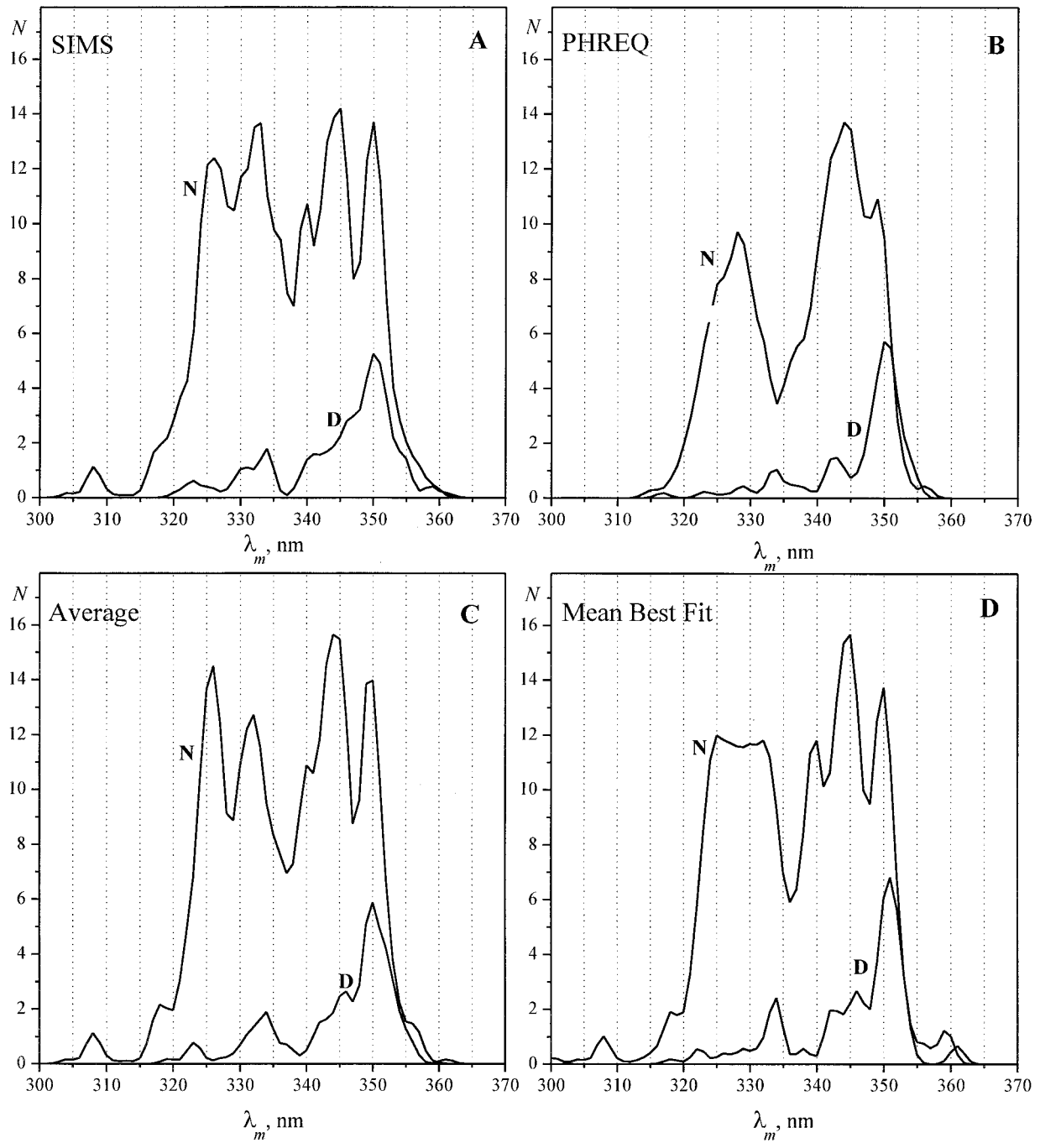

FIGURE 4 Distribution of occurrence of maximum positions of log-normal components obtained by applying SIMS (*A*), PHREQ (*B*) algorithms, and different schemes of averaging of the results: "Average" (*C*) and "Mean Best Fit" (*D*) with 1-nm step. Each curve is a sum of Gauss function (see explanation in the text) for each spectral component of native (N) and denatured (D) proteins with $\sigma(i) = 1.0$ and $S(i)$ equal the relative contribution of the component in the total protein emission (see Table 2).

fluorophore after its electronic excitation. The specific interactions include changes in the nearest range interactions with neighboring groups and/or solvent molecules, i.e., resolvation, hydrogen bond formation, or other noncovalent complexing (Bakhshiev, 1972; Mataga et al., 1955). The universal interactions occur due to the dipole relaxation of surrounding dielectric continuum in response to the changes in direction and magnitude of fluorophore dipole moment at the excitation (Bakhshiev, 1972; Bilot and Kawski, 1962; Lippert, 1957; Liptay, 1965;. Mataga et al., 1956).

The formation of stoichiometric complexes (most probably hydrogen-bonded ones) of excited indole fluorophore

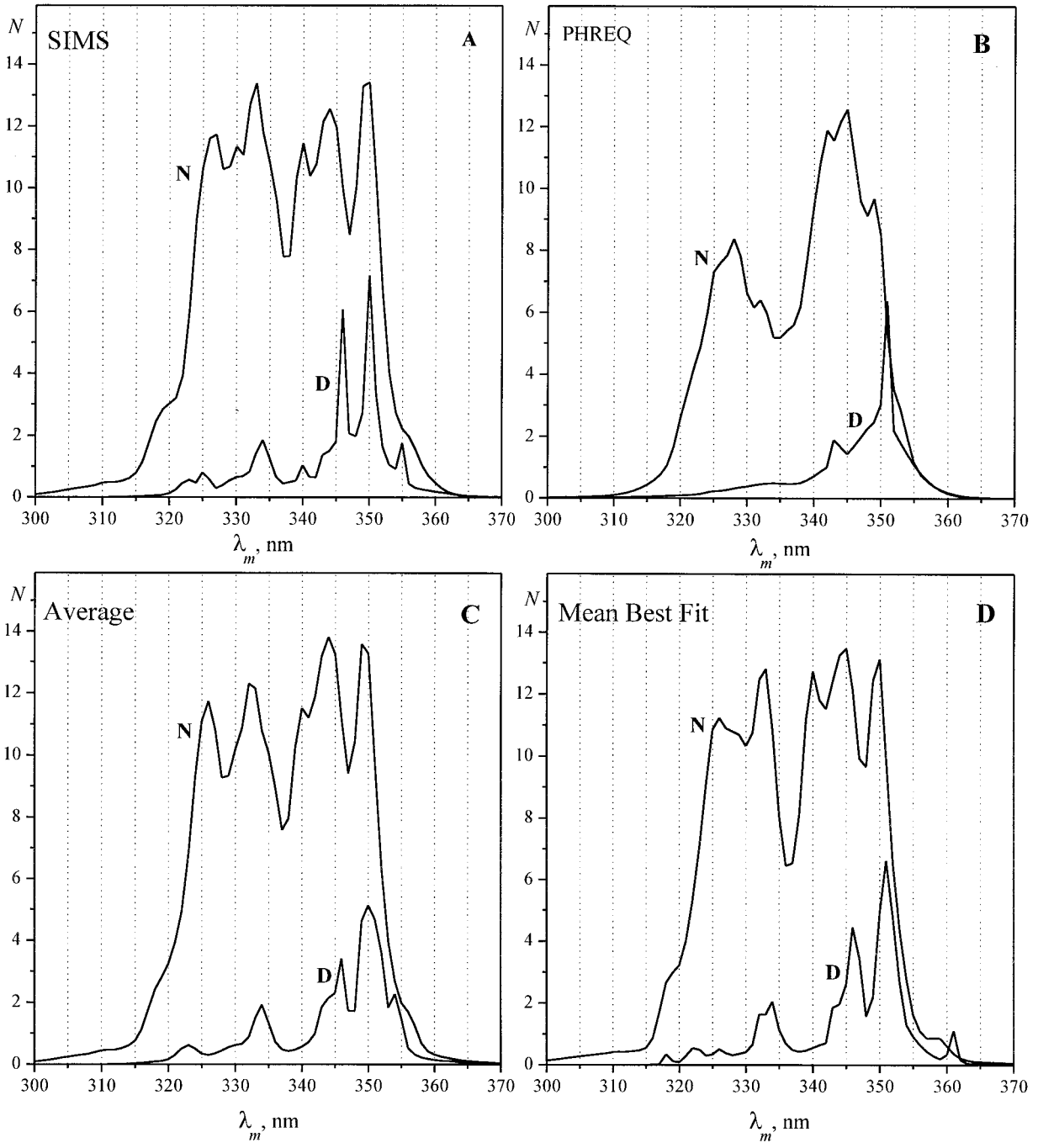

FIGURE 5 Distribution of occurrence of maximum positions of log-normal components obtained by applying SIMS (*A*), PHREQ (*B*) algorithms, and different schemes of averaging of the results: "Average" (*C*) and "Mean Best Fit" (*D*) with 1-nm step. Each curve is a sum of Gauss function (see explanation in the text) for each spectral component of native (N) and denatured (D) proteins with $\sigma(i)$ equal to the standard (root-mean-square one) error of the mean $\lambda_m(i)$ value and $S(i)$ equal to the relative contribution of the component in the total protein emission (see Table 2).

with alcohols was demonstrated experimentally (Walker et al., 1967; Lumry and Hershberger, 1978; Hershberger et al., 1981). Such complexes with alcohols in apolar cycloheptane had discrete positions of fluorescence spectral maxima dependent on the stoichiometric ratios of alcohol/fluorophore: the maximum was at 316 nm for the 1:1 ratio and at ~330 nm at the 2:1 ratio compared with the central peak at 307 nm in the structured spectrum obtained in the absence of any polar co-solvents. The authors interpreted these excited-state complexes (exciplexes) as a result of H-π-bonding of alcohol hydroxyls with indolic N- and Cγ-atoms as those possessing maximal electronic density in the excited

state. Such a location of H-bonds may be revised in the light of recent quantum-mechanical calculations of electronic density distribution in the excited indole and tryptophan fluorophores, which revealed that maximal densities are located at the Cϵ3, Cζ2, and Cδ2 atoms in the main fluorescent ^1L_a state of the indolic ring (Callis, 1997).

In proteins, the structured fluorescence bands coinciding with those of nonexciplexed indole fluorophores occur in azurin (Burstein et al., 1977) and in bacteriorhodopsin purple membranes suspended in 2 M CsCl (Permyakov and Shnyrov, 1983). Such tryptophan residues were attributed to class A in the hypothesis of discrete states. The structured spectra similar to those of the exciplexes 1:1 occurring in several single-tryptophan-containing proteins were assigned to class S (Burstein, 1977a, b, 1983). Emission spectra of classes A and S do not undergo any spectral shift under freezing the protein solutions down to −196°C, i.e., the dipole moments of these fluorophores practically do not change at the excitation. In the histograms obtained in this work, classes A and S are reflected by peaks and/or shoulders at 305–308 and 316–318 nm, respectively (see Figs. 3–5).

In the model experiments (Walker et al., 1967; Lumry and Hershberger, 1978; Hershberger et al., 1981), the maximum position of the exciplex 2:1 spectrum was not constant under increasing alcohol concentration, which reflected the rise of Stokes shift induced by the solvent dipole relaxation in response to the change of fluorophore dipole moment in the excited state of the exciplex. Therefore, the florescence of protein tryptophan residues possessing the spectra with maxima at ~330 nm and longer were assigned to the emission of exciplexes with stoichiometry not <2:1 in the model of discrete classes (Burstein, 1977a, b, 1983). The model provided three discrete classes (I–III) for such exciplexes with stoichiometry ≥2:1, differed in their accessibility to extrinsic quenchers (vanishing accessibility for class I, λ_m at ~330 nm, and the accessibility of >30% of that of free tryptophan for classes II and III). The difference between two accessible classes was assumedly interpreted as a result of different dipole relaxation rates in their environments.

The histograms (Figs. 3–5) contain peaks reflecting all three discrete classes. However, the class I band in the histograms splits up into two peaks, at ~325 and 332 nm. Almost all histograms contain the global minimum at 336–337 nm and two evident maxima at 344–345 and 349–351 nm, which correspond to the emission of exposed tryptophan residues of classes II and III, respectively. In the model of discrete classes, it was suggested that exposed tryptophan residues in native proteins should emit at 340–342 nm, and emission at 350–353 nm is due to tryptophan residues of denatured proteins (Burstein et al., 1973). However, the extended set of proteins investigated here indicates that tryptophans in native proteins also can have spectra with maxima at ~350 nm.

Tryptophan residues of denatured proteins (italicized in Tables 1 and 2) have the main peak positions at 340–345 and 351 nm. The longest-wavelength peak corresponds to the completely unfolded proteins in urea or guanidinium chloride solutions. However, the peak at 340–345 nm belongs mainly to the fluorophores in proteins denatured by extreme pH values (e.g., pepsin at pH 6–9; see Table 2) or high dioxane concentrations (chymotrypsin, chymotrypsinogen, and trypsin; see Table 2*)*. It is possible that such maximum positions of fluorescence spectra are characteristic of tryptophan residues in such intermediate unfolding states of proteins as molten globule (Ptitsyn, 1995). The shorter-wavelength peaks (λ_m < 337 nm) in histograms for denatured proteins belong, most probably, to the fluorophores in molecules retained native or partly unfolded in denaturing conditions.

The model of discrete states reflects the existence in proteins of the five most manifested classes of tryptophan residues. The discreteness of fluorescence parameters has to have a probabilistic nature. The histograms obtained in this work reflected all these classes and can be regarded as a statistical confirmation of the hypothesis of discrete classes of tryptophan residues in proteins. It could be assumed that such a situation might be realized because tryptophan residues are located in a few kinds of physically preferred environments in protein structures that provide realization of different combinations of specific and universal interactions of excited fluorophore with its environment.

There were a lot of papers published during the last several last years where it was demonstrated that tryptophan residues could have various fluorescence parameters depending on their environment in proteins (Callis, 1997; Callis and Burgess, 1997; Reshetnyak and Burstein, 1997a, b; Chen and Barkley, 1998; Kuznetsova and Turoverov, 1998; Meagher et al., 1998). Because highly resolved x-ray and NMR structures of many proteins are now available, we developed a system of describing physical and structural characteristics of the microenvironment of tryptophan residues and examined the existence of discreteness of structural parameters of the microenvironment of tryptophan residues in proteins. Moreover, that work allowed us to elucidate the main physical and structural factors determining the class of an individual tryptophan residue. The next paper of this series contains the results of that study.

The authors are thankful to all colleagues from the Laboratory of Protein Functional Biophysics of the Institute of Theoretical and Experimental Biophysics (Pushchino) for everyday help and discussions. We thank Drs. K. K. Turoverov and R. V. Polozov for fruitful discussions. We are especially grateful to Drs. V. M. Grishchenko, L. P. Kalinichenko, and T. G. Orlova for the help in experimental work and to Drs. D. B. Veprintsev and D. S. Rykunov for valuable consultations and maintaining the function of the computers and nets. We are much obliged to all colleagues who placed protein preparations and/or fluorescence spectra at our disposal.

This work was supported in part by Grants 95-04-12935, 97-04-49449, and 00-04-48127 from the Russia Foundation of Basic Research.

REFERENCES

Ababou, A. 1998. Contribution to the study of temporal fluorescence: study of the fluorescence decay heterogeneity of the single tryptophan containing proteins. Ph.D. thesis, Louis Pasteur University, Strasbourg, France. 238 pp.

Abornev, S. M. 1993. Resolution of the composed tryptophan fluorescence spectra of proteins. Phys. and Math. Sci. Candidate thesis, Institute of Theoretical and Experimental Biophysics, RAS, Pushchino. 122 pp.

Abornev, S. M., and E. A. Burstein. 1992. Resolution of protein tryptophan fluorescence spectra into elementary components. *Molecular Biology (Moscow).* 26:890–987.

Bakhshiev, N. G. 1972. Spectroscopy of Intermolecular Interactions. Nauka, Leningrad [in Russian].

Bandorowicz, J., S. Pikula, and A. Sobota. 1992. Annexins IV (p32) and VI (p68) interact with erythrocyte membrane in a calcium-dependent manner. *Biochim. Biophys. Acta.* 1105:201–206.

Benzonana, G., L. Kohler, and E. A. Stein. 1974. Regulatory proteins of crayfish tail muscle. *Biochim. Biophys. Acta.* 368:247–258.

Bilot, L., and A. Kawski. 1962. Zur Theorie des Einfl 252 sses von L 246 sungsmitteln auf die Elektronenspektren der Molek 252 le. *Z. Naturforsch.* 17a:621–627.

Bukolova-Orlova, T. G., E. A. Burstein, B. P. Tchorbanov, B. V. Alexiev, and B. P. Atanasov. 1979. Study of some physico-chemical properties of the neurotoxic complex and its components from the venom of Bulgarian sand viper, *Vipera ammodytes ammodytes.* I. Luminescence of tryptophan residues. *Biochim. Biophys. Acta.* 577:44–51.

Bukolova-Orlova, T. G., E. A. Burstein, and L. Ya. Yukelson. 1974. Fluorescence of neurotoxins from Middle-Asian cobra venom. *Biochim. Biophys. Acta.* 342:275–280.

Bukolova-Orlova, T. G., N. Y. Orlov, E. A. Burstein, B. P. Tchorbanov, and B. V. Alexiev. 1980. Study of the neurotoxic complex and its components from the venom of Bulgarian sand viper, *Vipera ammodytes ammodytes. Arch. Biochem. Biophys.* 200:216–222.

Bukolova-Orlova, T. G., E. A. Permyakov, E. A. Burstein, and L. Ya. Yukelson. 1976. Reinterpretation of luminescence properties of neurotoxins from the venom of Middle-Asian cobra, *Naja oxiana. Biochim. Biophys. Acta.* 439:426–431.

Burstein, E. A. 1968. Quenching of protein fluorescence. I. Principles of the method. Solutions of tryptophan, tyrosine and denatured proteins. *Biophysics (Moscow).* 13:433–442 [in Russian; English translation].

Burstein, E. A. 1976. Luminescence of protein chromophores (model studies). *In* Advances in Science and Technology (Itogi Nauki i Tekhniki), ser. Biophysics, vol. 6. VINITI, Moscow [In Russian].

Burstein, E. A. 1977a. Intrinsic protein luminescence (the nature and application). *In* Advances in Science and Technology (Itogi Nauki i Tekhniki) ser. Biophysics, vol. 7. VINITI, Moscow [In Russian].

Burstein, E. A., editor. 1977b. The study of fast dynamics of protein structure using the intrinsic fluorescence methods. *In* Equilibrium Dynamics of Native Protein Structure. Center of Biological Research, RAS, Pushchino. 60–83 [In Russian].

Burstein, E. A. 1983. The intrinsic luminescence of proteins is a method for studies of the fast structural dynamics. *Molecular Biology (Moscow).* 17:455–467 [In Russian; English translation].

Burstein, E. A., S. M. Abornev, and Ya. K. Reshetnyak. 2001. Decomposition of protein tryptophan fluorescence spectra into log-normal components. I. Decomposition algorithms. *Biophys. J.* 81:1699–1709.

Burstein, E. A., and V. I. Emelyanenko. 1996. Log-normal description of fluorescence spectra of organic fluorophores. *Photochem. Photobiol.* 64:316–320.

Burstein, E. A., E. A. Permyakov, V. A. Yashin, S. A. Burkhanov, and A. Finazzi Agrò. 1977. The fine structure of luminescence spectra of azurin. *Biochim. Biophys. Acta.* 491:155–159.

Burstein, E. A., N. S. Vedenkina, and M. N. Ivkova. 1973. Fluorescence and the location of tryptophan residues in protein molecules. *Photochem. Photobiol.* 18:263–279.

Bushueva, T. L., M. V. Teplova, V. N. Bushuev, D. S. Kudryashov, A. V. Vorotnikov, and V. P. Shirinsky. 1999. The structure stability of KPR (Kinase Related Protein). *Molecular Biology (Moscow).* 33:227–236.

Bushueva, T. L., and A. G. Tonevitsky. 1987a. The effect of pH on the conformation and stability of the structure of plant toxin ricin. *Molecular Biology (Moscow).* 21:414–420 [In Russian; English translation].

Bushueva, T. L., and A. G. Tonevitsky. 1987b. The effect of pH on the conformation and stability of the structure of plant toxin-ricin. *FEBS Lett.* 215:155–159.

Bushueva, T. L., and A. G. Tonevitsky. 1988. Similarity of protein conformation at low pH and high temperature observed for B-chains of two plant toxins: ricin and mistletoe lectin 1. *FEBS Lett.* 229:119–122.

Bushueva, T. L., A. G. Tonevitsky, and E. A. Burstein. 1990. The structure of ricin: the study by the fluorescence quenching method. *Molecular Biology (Moscow).* 24:614–620 [In Russian; English translation].

Bushueva, T. L., A. G. Tonevitsky, A. Kindt, and H. Franz. 1988. The structure of toxic protein: the mistletoe lectin at different pH. The study by intrinsic fluorescence methods. *Molecular Biology (Moscow).* 22: 628–634 [In Russian; English translation].

Callis, P. R. 1997. ^1L_a and ^1L_b transitions of tryptophan: applications of theory and experimental observations to fluorescence of proteins. *Methods Enzymol.* 278:113–151.

Callis, P. R., and B. K. Burgess. 1997. Tryptophan fluorescence shifts in proteins from hybrid simulations: an electrostatic approach. *J. Phys. Chem. B.* 101:9429–9432.

Chen, Yu, and M. D. Barkley. 1998. Toward understanding tryptophan fluorescence in proteins. *Biochemistry.* 37:9976–9982.

Cox, J. A., W. Wnuk, and E. A. Stein. 1976. Isolation and properties of a sarcoplasmic calcium-binding protein from crayfish. *Biochemistry.* 15: 2613–2617.

Czurylo, E. A., V. I. Emelyanenko, E. A. Permyakov, and R. Dabrowska. 1991. Spectrofluorimetric studies on C-terminal 34 kDa fragment of caldesmon. *Biophys. Chem.* 40:181–188.

Demchenko, A. P. 1986. Ultraviolet Spectroscopy of Proteins. Springer-Verlag, Berlin.

Grishchenko, V. M., S. I. Bezborodova, V. I. Emelyanenko, M. N. Ivkova, and E. A. Burstein. 1976. Fluorescence studies of extracellular guanyl-specific ribonuclease C2 from *Aspergillus clavatus. Bioorg. Chem. (Moscow).* 2:207–216 [In Russian; English translation].

Haiech, J., J. Derancourt, J.-F. Pechère, and J.G. Demaille. 1979. A new large-scale purification procedure for muscular parvalbumins. *Biochimie.* 61:583–587.

Hershberger, M. V., R. Lumry, and R. Verrall. 1981. The 3-methylindole/*n*-butanol exciplexes evidence for two exciplex sites in indole compounds. *Photochem. Photobiol.* 33:609–617.

Ivkova, M. N., V. V. Mosolov, and E. A. Burstein. 1968. Fluorescence of some proteins in water-dioxane solutions. *Molecular Biology (Moscow).* 2:839–850. [In Russian; English translation].

Ivkova, M. N., N. S. Vedenkina, and E. A. Burstein. 1971. Tryptophan fluorescence of serum albumins. *Molecular Biology (Moscow).* 5:214–224. [In Russian; English translation].

Kaplanas, R. I., T. G. Bukolova, and E. A. Burstein. 1973. Fluorescence of β-lactalbumin AB under various physico-chemical conditions. I. Influences of pH and binding of sodium dodecyl sulfate. *Molecular Biology (Moscow).* 7:753–759 [In Russian; English translation].

Kaplanas, R. I., T. G. Bukolova, and E. A. Burstein. 1975. Fluorescence of β-lactalbumin AB under various physico-chemical conditions. II. Denaturation by urea and various organic solvents. *Molecular Biology (Moscow).* 9:795–804 [In Russian; English translation].

Klimenko, I. V., T. O. Gushcha, and A. I. Korneliuk. 1991. Tryptophan fluorescence properties of two forms of the tyrosyl-aminoacyl-tRNA-synthase (bovine liver). *Biopolymers and Cells (Kiev).* 7:83–88 [In Russian; English translation].

Konev, S. V. 1967. Fluorescence and Phosphorescence of Proteins and Nucleic Acids. Plenum Press, New York.

Koteliansky, V. E., T. L. Bushueva, M. V. Betanyan, M. A. Glukhova, and V. N. Smirnov. 1981. The intrinsic fluorescence study of the structure of

fibronectin. *Dokl. AN SSSR (Proc. USSR Acad. Sci; Moscow).* 258: 1485–1488. [In Russian; English translation].

Koteliansky, V. E., M. A. Glukhova, V. P. Shirinsky, V. N. Smirnov, T. L. Bushueva, V. K. Filimonov, and S. Y. Venyaminov. 1982. A structural study of filamin, a high-molecular-weight actin-binding protein from chicken gizzard. *Eur. J. Biochem.* 121:553–559.

Kuznetsova, I. M., and K. K. Turoverov. 1998. What determines the characteristics of the intrinsic UV-fluorescence of proteins? Analysis of the properties of the microenvironment and features of the localization of their tryptophan residues. *Cytology (St. Petersburg).* 40:747–62. [In Russian].

Ladokhin, A. S. 1990. Equilibrium dynamics of proteins. Intrinsic fluorescence study of melittin. *Biopolymers and Cells (Kiev).* 6:84–91. [In Russian].

Lakowicz, J. R. 1983. Principles of Fluorescence Spectroscopy. Plenum Press, New York.

Lippert, E. 1957. Spektroskopische Bestimmungen des Dipolmomentes aromatischer Verbindungen im ersten angeregten Singulettzustand. *Z. Elektrochem.* 61:962.

Liptay, W. 1965. Die Lösungsmittelabhägigkeit der Wellenzahl von Elektronenbanden und chemisch-physikalischen Grundlagen. *Z. Naturforsch.* 20a:1441–1471

Longworth, J. W. 1971. Excited States of Proteins and Nucleic Acids. R. F. Steiner and I. Weinryb, editors. Plenum Press, New York.

Lumry, R., and M. Hershberger. 1978. Status of indole photochemistry with special reference to biological application. *Photochem. Photobiol.* 27:819–840.

Mataga, N., Y. Kaifu, and M. Koizumi. 1955. The solvent effect on fluorescence spectrum. Change of solute-solvent interaction during the lifetime of excited solute molecule. *Bull. Chem. Soc. Japan.* 28:690–691.

Mataga, N., Y. Kaifu, and M. Koizumi. 1956. Solvent effects upon fluorescence spectra and the dipole-moments of excited molecules. *Bull. Chem. Soc. Japan.* 29:465–470.

Meagher, J. L., J. M. Beechem, S. T. Olson, and P. G. Gettins. 1998. Deconvolution of the fluorescence emission spectrum of human antithrombin and identification of the tryptophan residues that are responsive to heparin binding. *J. Biol. Chem.* 273:23283–23289.

Mely, Y., M. Cadene, I. Sylte, and J. G. Breth. 1997. Mapping of suramin-binding sites of human neutrophil elastase: investigation by fluorescence energy transfer and molecular modeling. *Biochemistry*. 36: 15624–15631.

Morozova, L. A., N. B. Gusev, V. L. Shnyrov, and E. A. Permyakov. 1988. Study of the physico-chemical properties of troponins I and T from the heart and skeletal muscles using protein fluorescence and calorimetric methods. *Biochemistry (Moscow).* 53:531–540 [In Russian; English translation].

Nyamaa, D., O. Bat-Erdene, and E. A. Burstein. 1985. Solvent effects on the functional and structural properties of serum albumins. III. Temperature and ionic strength dependencies of human serum albumin N-F_1, F_1-F_2 and F_2-E transitions. *Molecular Biology (Moscow).* 19:833–840.

Orlov, N. Y., T. G. Orlova, Ya. K. Reshetnyak, E. A. Burstein, and N. Kimura. 1997. Interaction of recombinant rat nucleoside diphosphate kinase alpha with bleached bovine retinal rod outer segment membranes: a possible mode of pH and salt effects. *Biochem. Mol. Biol. Int.* 41:189–198.

Orlov, N. Ya., T. G. Orlova, Ya. K. Reshetnyak, E. A. Burstein, and N. Kimura. 1999. Comparative study of recombinant rat nucleoside diphosphate kinases alpha and beta by intrinsic protein fluorescence. *J. Biomol. Struct. Dyn.* 16:955–968.

Orlova, T. G. 1979. The study of toxic proteins of snake venom by intrinsic protein fluorescence techniques. Biol. Science Candidate thesis, Institute of Photobiology, Byelorus Acad. Sci., Minsk [In Russian].

Pardee, J. D., and J. A. Spudich. 1982. Purification of muscle actin. *Methods Enzymol.* 85(Pt. B):164–181.

Permyakov, E. A., and G. Y. Deikus. 1995. Analytical description of electron-vibrational spectra of proteins. *Molecular Biology (Moscow).* 29:159–167. [In Russian].

Permyakov, E. A., V. M. Grishchenko, L. P. Kalinichenko, N. Y. Orlov, K. Kuwajima, and S. Sugai. 1991. Calcium-regulated interactions of human alpha-lactalbumin with bee venom melittin. *Biophys. Chem.* 39: 111–117.

Permyakov, E. A., L. A. Morozova, and E. A. Burstein. 1985. Cation binding effects on the pH, thermal and urea denaturation transitions in alpha-lactalbumin. *Biophys. Chem.* 21:21–23.

Permyakov, E. A., and V. L. Shnyrov. 1983. A spectrofluorimetric study of the environment of tryptophans in bacteriorhodopsin. *Biophys. Chem.* 18:145–152.

Permyakov, E. A., D. B. Veprintsev, G. Y. Deikus, S. E. Permyakov, L. P. Kalinichenko, V. M. Grishchenko, and C. L. Brooks. 1997. pH-induced transition and Zn^{2+}-binding properties of bovine prolactin. *FEBS Lett.* 405:273–276.

Permyakov, E. A., V. V. Yarmolenko, V. I. Emelyanenko, E. A. Burstein, J. Closset, and C. Gerday. 1980. Fluorescence studies of the calcium binding to whiting (*Gadus merlangus*) parvalbumin. *Eur. J. Biochem.* 109:307–315.

Ptitsyn, O. B. 1995. Molten globule and protein folding. *Adv. Protein Chem.* 47:83–229.

Reshetnyak, Ya. K., O. A. Andreev, J. Borejdo, D. D. Toptygin, L. Brand, and E. A. Burstein. 2000. The identification of tryptophan residues responsible for ATP-induced increase in intrinsic fluorescence of myosin subfragment 1. *J. Biomol. Struct. Dyn.* 18:113–125.

Reshetnyak, Ya. K., and E. A. Burstein. 1997a. Assignment of log-normal components of protein fluorescence spectra to individual tryptophan residues using their microenvironment properties in three-dimensional structure. *Biophysics (Moscow).* 42:293–300 [In Russian; English translation].

Reshetnyak, Ya. K., and E. A. Burstein. 1997b. Assignment of components of fluorescence spectra of serine proteases to clusters of tryptophan residues. *Biophysics (Moscow).* 42:785–795 [In Russian; English translation].

Shanin, S. S., I. A. Tarkhanova, A. Y. Kulberg, and E. A. Burstein. 1968. Conformational changes of γG-globulin under dissociation on peptides. *Molecular Biology (Moscow).* 2:587–593 [In Russian; English translation].

Siano, D. B., and D. E. Metzler. 1969. Band shapes of the electronic spectra of complex molecules. *J. Chem. Phys.* 51:1856–1861.

Sobota, A., J. Bandorowicz, A. Jezierski, and A. F. Sikorski. 1993. The effect of annexins IV and VI on the fluidity of phosphatidylserine/phosphatidylcholine bilayers studied with the use of 5-deoxylstearate spin label. *FEBS Lett.* 315:178–182.

Tchorbanov, B. P., B. V. Alexiev, T. G. Bukolova, E. A. Burstein, and B. P. Atanasov. 1977. Subfractionation and recombination of a neurotoxic complex of Bulgarian viper (*Vipera ammodytes ammodytes*). *FEBS Lett.* 76:266–268.

Tonevitsky, A. G., D. E. Temyakov, E. C. Sweeney, N. V. Konareva, A. N. Kornev, R. A. Palmer, and A. M. Mikhailov. 1996. Structural features of ricin agglutinin from *Ricinus communis*. *Crystallography (Moscow).* 41:1018–1023 [In Russian; English translation].

Vedenkina, N. S., T. G. Bukolova, M. N. Ivkova, and E. A. Burstein. 1971. Fluorescence and states of tryptophan in papain. *Molecular Biology (Moscow).* 5:809–816 [In Russian; English translation].

Vedenkina, N. S., and E. A. Burstein. 1970. Tryptophan fluorescence of proteins in solutions. The spectral maxima positions. *Molecular Biology (Moscow).* 4:743–748 [In Russian; English translation].

Vedenkina, N. S., M. N. Ivkova, V. N. Leonova, and E. A. Burstein. 1968. The fluorescence of actin and actomyosin in various structural states. *Biophysics (Moscow).* 13:847–852 [In Russian; English translation].

Vishnevskaya, Z. I., N. S. Vedenkina, and M. V. Georgadze. 1976. Fluorescence study of myosin and its fragments, heavy and light meromyosins, in concentrated solutions of neutral salts. *Molecular Biology (Moscow).* 10:799–805 [In Russian; English translation].

Volotovski, I. D., and S. V. Konev. 1967. The relation between conformation and UV luminescence of proteins. *Biophysics (Moscow).* 12: 200–205 [In Russian; English translation].

Walker, M. S., T. W. Bednar, and R. Lumry. 1967. Exciplex studies. II. Indole and indole derivatives. *J. Chem. Phys.* 47:1020–1028.

Weller, P. A., E. P. Ogryzko, E. B. Corben, N. I. Zhidkova, B. Patel, G. J. Price, N. K. Spurr, V. E. Koteliansky, and D. R. Gritchley. 1990. Complete sequence of human vinculin and assignment of the gene to chromosome 10. *Proc. Natl. Acad. Sci. U.S.A.* 87:5667–5671.

Zhou, X., Y. Maeda, K. Mabuchi, and S. S. Lehrer. 1998. Unfolding domains and tryptophan accessibility of a 59 kDa coiled-coil light meromyosin. *J. Mol. Biol.* 276:829–838.