Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.
Estell, D. A., Graycar, T. P., & Wells, J. A. (1985) *J. Biol. Chem. 260*, 6518–6521.
Frauenfelder, H., Petsko, G. A., & Tsernoglou, D. C. (1979) *Nature (London) 280*, 558–563.
Freedman, R. B., & Hillson, D. A. (1980) in *The Enzymology of Post-Translational Modification of Proteins* (Freedman, R. B., & Hawkins, H. C., Eds.) Vol. 1, pp 167–212, Academic, New York.
Greene, R. F., & Pace, C. N. (1974) *J. Biol. Chem. 249*, 5388–5393.
Haber, E., & Anfinsen, C. B. (1962) *J. Biol. Chem. 237*, 1839–1844.
Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramaseshan, S., & Venkatesan, K., Eds.) pp 13.1–13.23, Indian Institute of Science, Bangalore, India.
Howell, E. E., Villafranca, J. E., Warren, M. S., Oatley, S. J., & Kraut, J. (1986) *Science (Washington, D.C.) 231*, 1123–1128.
Huennekens, F. M., Vitols, K. S., Whiteley, J. M., & Neef, V. G. (1976) in *Methods in Cancer Research* (Busch, H., Ed.) pp 199–223, Academic, New York.
Itzhaki, R. F., & Gill, D. M. (1964) *Anal. Biochem. 9*, 401–410.
Michaelis, S., Inouye, H., Oliver, D., & Beckwith, J.(1983) *J. Bacteriol. 154*, 366–374.
Pace, C. N. (1975) *CRC Crit. Rev. Biochem. 3*, 1–43.
Perry, L. J., & Wetzel, R. (1986) *Biochemistry 25*, 733–739.
Poland, D. C., & Scheraga, H. A. (1965) *Biopolymers 3*, 379–399.
Pullman, B., & Pullman, A.(1974) *Adv. Protein Chem. 28*, 347–526.
Richards, F. M. (1977) *Annu. Rev. Biophys. Biochem. 6*, 151–175.
Richardson, J. (1981) *Adv. Protein Chem. 34*, 167–339.
Saxena, V. P., & Wetlaufer, D. (1970) *Biochemistry 19*, 5015–5022.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1–95.
Thornton, J. M. (1981) *J. Mol. Biol. 151*, 261–287.
Villafranca, J E., Howell, E. E., Voet, D. H., Strobel, M. S., Ogden, R. C., Abelson, J. N., & Kraut, J. (1983) *Science (Washington, D.C.) 222*, 782–788.
Wells, J. A., & Powers, D. B. (1986) *J. Biol. Chem. 261*, 6564–6570.
Xuong, N.-h., Freer, S. T., Hamlin, R., Nielson, C., & Vernon, W. (1978) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A34*, 289–295.
Ziegler, D. M., & Poulsen, L. L. (1977) *Trends Biochem. Soc. (Pers. Ed.) 2*, 79–81.

Amino Acid Sequence of Sialic Acid Binding Lectin from Frog (*Rana catesbeiana*) Eggs†

Koiti Titani,*,‡ Koji Takio,‡ Manabu Kuwada,‡,§ Kazuo Nitta,‖ Fusao Sakakibara,‖ Hiroaki Kawauchi,‖ Giichi Takayanagi,‖ and Sen-itiroh Hakomori*,⊥

Department of Biochemistry, University of Washington, Seattle, Washington 98195, Program of Biochemical Oncology/Membrane Research, Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, and Department of Biochemistry, Cancer Research Institute, Tohoku College of Pharmaceutical Sciences, Sendai 983, Japan

Received June 2, 1986; Revised Manuscript Received December 17, 1986

ABSTRACT: The complete amino acid sequence of sialic acid binding lectin from frog (*Rana catesbeiana*) egg is presented. The 111-residue sequence was determined by the analysis of peptides generated by digestion of the S-carboxymethylated protein with *Achromobacter* protease I, chymotrypsin, or cyanogen bromide. The sequence is unique and not homologous to any known protein sequence. The protein may represent a new type of lectin.

Lectins are multivalent carbohydrate binding proteins with the ability to agglutinate erythrocytes, other normal and cancer cells, and microorganisms. They were first discovered in plants, but recently they have also been found in animals and other organisms [cf. review by Barondes (1981)].

Kawauchi et al. (1975) found that eggs of frog (*Rana catesbeiana*) contain two lectins that show different agglutinating activities, one with human blood group A erythrocytes and the other with mouse Ehrlich ascites carcinoma cells or rat ascites hepatoma cells (AH-109A). These two lectins were purified to homogeneity, and their binding specificity, agglutinatability, and physical and chemical properties have been studied. One of the lectins showed a preferential agglutination of cancer cells and displayed a specific binding to sialyl glycoprotein (Sakakibara et al., 1977), and consequently, it was designated sialic acid binding protein.

In the present paper, we present the complete amino acid sequence of 111 residues of the frog egg sialic acid binding lectin and discuss the structural features in relation to the biological function.

Materials and Methods

Sialic acid binding lectin was prepared as described by Sakakibara et al. (1977). After extraction from acetone powder of eggs of *Rana catesbeiana* with saline, it was purified by chromatography on Sephadex G-75, DEAE-cellulose, hy-

† This work was supported by Research Grants HL-29595 (to K. Titani), GM-15731 (to K. Walsh), and GM-23100 (to S.H.) from the National Institutes of Health. K. Titani is also a visiting professor of the Fujita-Gakuen Health University, Toyoake-shi, Japan. K. Takio is a senior associate of the Howard Hughes Medical Institute Laboratory at the University of Washington, Seattle.

* Address correspondence to K. Titani at the School of Medicine, Fujita-Gakuen Health University, Toyoake-shi, Japan, or to S. Hakomori at the Fred Hutchinson Cancer Research Center.
‡ University of Washington.
§ Present address: Eisai Co. Tsukuba Research Laboratories, 1-3, Tokodai 5-chome, Toyosato-machi, Tsukuba-gun, Ibaragi, Japan.
‖ Tohoku College of Pharmaceutical Sciences.
⊥ Fred Hutchinson Cancer Research Center.

0006-2960/87/0426-2189$01.50/0

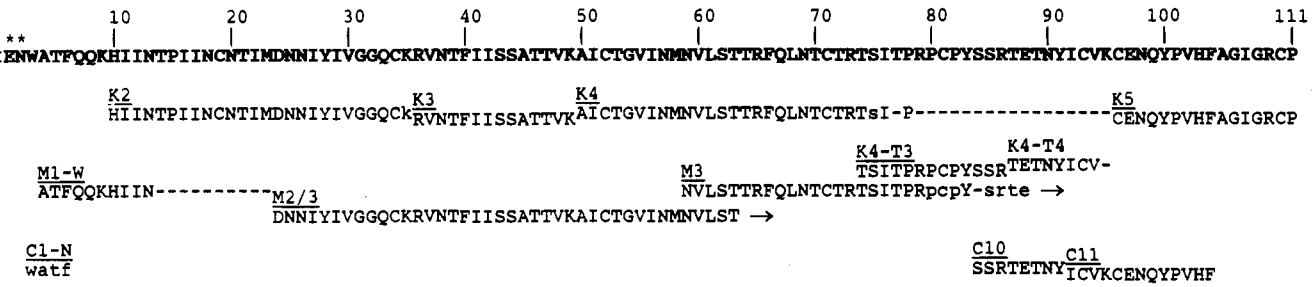

FIGURE 1: Detailed summary of the proof of sequence of sialolectin from *R. catesbeiana* eggs. The proven sequences of specific peptides are given in one-letter code (Table I) below the summary sequence (bold type). The symbol <E indicates 5-pyrrolidone-2-carboxylic acid. Prefixes K, M, and C denote peptides generated by cleavage of the CM-protein at lysyl and methionyl bonds and with chymotrypsin, respectively. Subpeptides are identified by hyphenated suffixes, with the following code indicating the subdigesting agent: D, W, or T, cleavage with dilute acid, BNPS-skatole, or trypsin. Sequences written in upper case letters were proven by Edman degradation; those in lower case letters indicate tentative identification or deduced from amino acid compositions. Those not identified are shown by dashes or by arrows. Currently tentative residues are indicated by an asterisk above the summary sequence.

droxylapatite, and finally on (carboxymethyl)cellulose (CM-cellulose)[1] columns. As described by Sakakibara et al. (1977), the sialic acid binding agglutinin was completely separated in terms of activity by DEAE-cellulose chromatography from the other lectin, which is also present in the eggs and shows a specific reactivity with human blood group A erythrocytes.

TPCK-treated trypsin and α-chymotrypsin were obtained from Worthington. *Achromobacter* protease I, which specifically cleaves lysyl bonds (Masaki et al., 1981), was a generous gift of Dr. T. Masaki of the Department of Agricultural Chemistry, Ibaraki University, Ibaraki, Japan.

The protein was reduced and S-carboxymethylated with iodo[^{3}H]acetic acid (New England Nuclear) by a slight modification of Crestfield et al. (1963). The CM-protein was cleaved at room temperature with cyanogen bromide in 70% formic acid as described by Gross (1967). A digest of CM-protein with *Achromobacter* protease I was prepared at 37 °C in 50 mM Tris-HCl, pH 9.0, in the presence of 2 M urea; the protein was dissolved in a small volume of 8 M urea and then diluted with 3 volumes of Tris buffer, pH 9.0. Other enzymatic digestions were performed at 37 °C in 0.1 M ammonium bicarbonate, pH 8.0. The chymotryptic digest included 10 mM *p*-aminobenzamidine to prevent slight cleavages at lysyl or arginyl bonds with contaminating trypsin. Cleavage with BNPS-skatole or with dilute acid followed the procedure of either Omenn et al. (1970) or Inglis (1983).

Peptides were separated primarily by size-exclusion HPLC with a Varian 5000 liquid chromatograph, using three TSK gel columns connected in series (G3000SW–G2000SW–G2000SW, each 7.5 × 600 mm, generously provided by Toyo Soda Co.), and purified by reversed-phase HPLC on a column of SynChropak RP8 (4.1 × 250 mm, SynChrom, Inc.) (Mahoney & Hermodson, 1980) or separated directly by reversed-phase HPLC.

An Applied Biosystems Model 430A peptide synthesizer was used to prepare two nonapeptides corresponding to residues 1–9 with either Glu or Gln as the amino-terminal residue. Synthetic peptides and their derivatives were analyzed by reversed-phase HPLC with a Hewlett Packard Model 1090M on an RP-300 (2.1 × 30 mm, Brownlee) column.

Amino acid analysis was performed with a Dionex D500 amino acid analyzer or by reversed-phase HPLC of phenylthiocarbamyl derivatives (Bidlingmeyer et al., 1984) using a Waters pico-tag system. Automated sequence analysis was performed with an Applied Biosystems 470A protein sequencer using a program adapted from Hewick et al. (1981). PTH derivatives of amino acids were identified in a semiquantitative manner by two complementary HPLC systems (Ericsson et al., 1977; Glajch et al., 1985).

Computerized protein data base search used the SEARCH program by Dayhoff et al. (1983).

RESULTS

General Strategy of Sequence Analysis. Sequence analysis of the intact CM-protein (2 nmol) yielded no PTH's in five cycles of Edman degradation, clearly indicating that the amino terminus of the protein is blocked. Most of the sequence was obtained by analysis of five nonoverlapping peptides generated by *Achromobacter* protease I (K1–K5) and their subpeptides. The remainder of the sequence and overlaps of the peptides were obtained by analysis of cyanogen bromide and chymotryptic peptides. The proof of the 111-residue sequence of the protein is summarized in Figure 1.

Generation of Peptides. Six hundred micrograms of the [^{3}H]CM-protein (ca. 50 nmol) was dissolved in 50 μL of 8 M urea and, after dilution with 150 μL of 50 mM Tris-HCl, pH 9.0, digested at 37 °C with *Achromobacter* protease I for 6 h. The digest was separated into eight fractions by size-exclusion HPLC as shown in Figure 2A. Four major fractions were further purified on a SynChropak RP-8 column by reversed-phase HPLC (Figure 2B–E). Five major peptides designated as K1–K5, accounting for the entire protein, were thus isolated (Table I) and subjected to sequence analyses.

The digest of the [^{3}H]CM-protein (40 nmol) with cyanogen bromide was separated in a similar manner to those shown in Figure 2 by a combination of size-exclusion and reversed-phase HPLC. Three peptides isolated, M1, M2/3, and M3, were analyzed to provide overlaps of four K peptides, K1, K2, K3, and K4. One of the chymotryptic peptides, C11, which was isolated from a digest (20 nmol) directly by reversed-phase HPLC (not shown), was analyzed to provide the last overlap of the K peptides, K4 and K5.

Sequence Analysis. The analysis of five major K peptides (K1–K5) and their subpeptides provided most of the sequence (92%) as shown in Figure 1. Of the five peptides, only K1 was N^{α} blocked, as was the intact protein, indicating that this peptide was derived from the amino-terminal portion of the protein. Digestion of K1 with *Staphylococcal* V8 protease generated no subpeptide, indicating that all the three glutamyl (glutaminyl) residues are likely to be the amide form. Dilute acid treatment of the blocked peptide C1 also generated no

[1] Abbreviations: HPLC, high-performance liquid chromatography; CM, carboxymethyl; TPCK, L-1-(tosylamido)-2-phenylethyl chloromethyl ketone; BNPS-skatole, bromine adduct of 2-[(2-nitrophenyl)-sulfenyl]-3-methylindole; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; PTH, phenylthiohydantoin; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; AUFS, absorbance units full scale.

Table I

Amino Acid Composition of R. catesbeiana Sialo-Lectin and the Peptides Used for the Sequence Analysis[a]

Peptide		K1	K2	K3	K4	K5	K4-T3[b]	K4-T4	M1	M2/3	M3	C1	C10	C11	Whole CM-Protein		
															Previous[c] Analysis	Present Analysis	
Residues		1-9	10-35	36-49	50-95	96-111	74-86	87-95	1-23	24-111	59-111	1-6	84-91	92-104	1-111		
CM-Cys	(C)		1.0(2)		1.3(4)	2.1(2)	0.5(1)	0.3(1)	(1)	5.4(7)	1.4(5)			1.6(2)	8.1	8.3	(8)
Asp/Asn	(D/N)	0.9(1)	6.1(6)	1.0(1)	4.3(4)	1.0(1)	1.4(0)	1.1(1)	4.8(4)	8.7(9)	5.3(4)	1.3(1)	0.9(1)	1.2(1)	12.9	13.7	(13)
Thr	(T)	1.0(1)	2.0(2)	2.5(3)	6.9(9)		2.0(2)	2.1(2)	2.5(3)	10.5(12)	7.1(8)	0.9(1)	1.9(2)		17.7	15.3	(15)
Ser	(S)			1.8(2)	3.7(4)		1.8(3)			5.4(6)	3.6(4)		1.9(2)		4.8	4.9	(6)
Glu/Gln	(E/Q)	3.0(3)	0.5(1)		1.8(2)	2.0(2)	0.7(0)	0.9(1)	3.0(3)	4.9(5)	4.1(4)	1.2(1)	1.0(1)	2.2(2)	8.1	8.8	(8)
Pro	(P)		1.3(1)		3.1(3)	1.9(2)	2.1(3)		1.0(1)	5.8(5)	4.1(5)			0.9(1)	5.6	7.0	(6)
Gly	(G)		2.1(2)		1.5(1)	2.2(2)				5.5(5)	3.5(2)				4.8	5.2	(5)
Ala	(A)	1.0(1)		1.0(1)	1.2(1)	1.0(1)			1.0(1)	3.0(3)	1.8(1)	1.0(1)			3.2	4.0	(4)
Val	(V)		1.0(1)	1.9(2)	2.8(3)	1.0(1)	0.5(0)	1.0(1)		5.7(7)	3.2(3)			2.0(2)	8.1	6.7	(7)
Met	(M)		1.1(1)		1.0(1)				0.3(1)[d]	0.8(2)					1.6	2.0	(2)
Ile	(I)		6.6(7)	1.7(2)	4.0(4)	1.2(1)	1.5(1)	1.0(1)	4.3(5)	7.6(9)	4.3(3)			1.0(1)	12.9	12.4	(14)
Leu	(L)				2.2(2)					2.0(2)	2.0(2)				1.6	2.2	(2)
Tyr	(Y)		0.9(1)		1.9(2)	0.6(1)	1.0(1)	0.9(1)		3.4(4)	2.2(3)		0.8(1)	1.0(1)	3.2	3.6	(4)
Phe	(F)	1.0(1)		1.0(1)	1.1(1)	1.0(1)			1.0(1)	2.9(3)	2.2(2)	0.9(1)		1.0(1)	4.0	4.0	(4)
His	(H)		1.0(1)			0.8(1)			0.8(1)	0.8(1)	0.9(1)			1.0(1)	1.6	2.0	(2)
Lys	(K)	0.9(1)	1.0(1)	1.0(1)	1.1(1)		0.4(0)	0.8(1)	0.5(1)	1.6(3)	1.4(1)			1.1(1)	4.8	4.1	(4)
Arg	(R)			1.0(1)	4.0(4)	0.8(1)				5.5(6)	4.4(5)		1.0(1)		6.5	6.5	(6)
Trp[e]	(W)	N.D.(1)	N.D.(0)	N.D.(0)	N.D.(0)	N.D.(0)	N.D.(0)	N.D.(0)	N.D.(1)	N.D.(0)	N.D.(0)	N.D.(1)	N.D.(0)	N.D.(0)	0.8	N.D.	(1)
Total Residues		9	26	14	46	16	13	9	23	88	53	6	8	13			111
Yield (%)		19	12	15	10	27	--	--	23	24	37	--	--	--			

[a]In the present analysis, the whole CM-protein was analyzed by the conventional method with a Dionex amino acid analyzer and the peptides by the Pico-Tag system with a Waters HPLC apparatus. The results are expressed as residues per molecule by analysis or, in parentheses, from the sequence.

[b]The peptide was contaminated with K4-T4 (see the text).

[c]Taken from K. Nitta, H. Kawauchi, G. Takayanagi, & S. Hakomori (unpub. data). Their results were recalculated on the basis of a molecular weight of 12,500.

[d]Determined as homoserine.

[e]Trp was not determined (N.D.) in the present experiment.

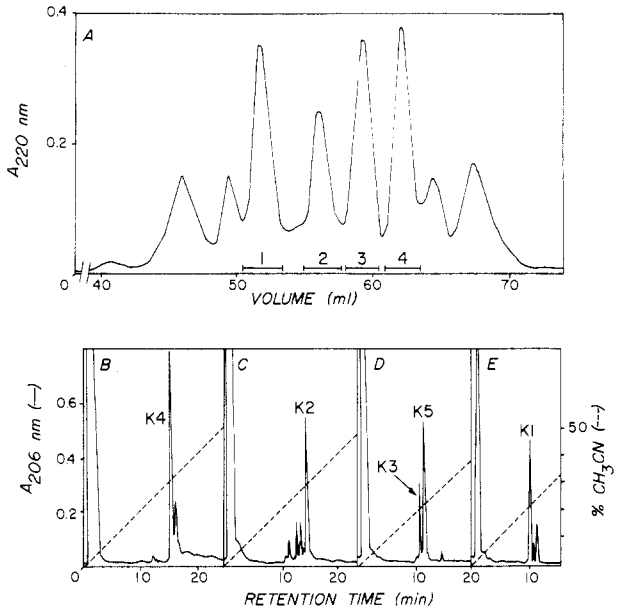

FIGURE 2: Separation of peptides generated by digestion of the CM-protein with *Achromobacter* protease I. (A) Primary separation of the digest (50 nmol) on three TSK columns connected in series (G3000SW–G2000SW–G2000SW, each 7.5 × 600 mm) and equilibrated in 6 M guanidine hydrochloride–50 mM potassium phosphate, pH 6.5, at a flow rate of 0.5 mL/min. Peptides were monitored at 220 nm and 2.0 AUFS and collected in 0.5-mL fractions. (B–E) Purification of pools 1–4 on SynChropak RP-8 column (4.1 × 250 mm) using a TFA–acetonitrile system (Mahoney & Hermodson, 1981). Purified peptides are identified by the prefix K as in Figure 1.

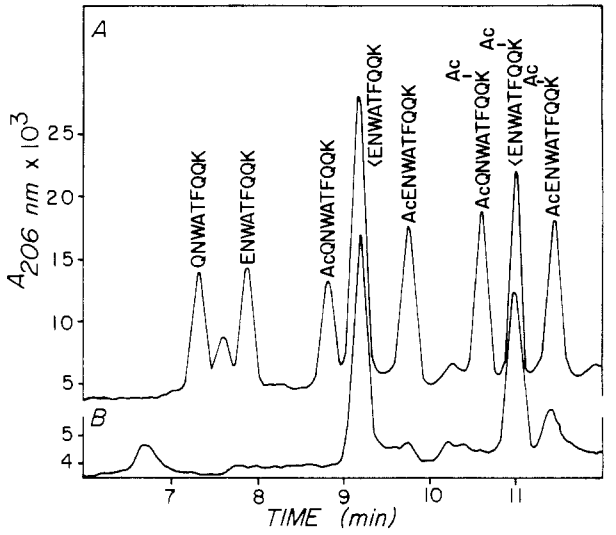

FIGURE 3: Elution profile of the blocked amino-terminal peptide K1 and synthetic peptides on a reversed-phase HPLC column (solvent A, 0.1% TFA; solvent B, 80% acetonitrile; gradient, 12–28% B in 12 min, at a flow rate of 0.3 mL/min). (A) Synthetic peptides and their derivatives. (B) Mixture of peptide K1 and acetylated peptide K1.

subpeptides at milder conditions (at 114 °C for 2 h) but yielded a weak sequence of X-Ala-Thr-Phe when it was treated at more drastic conditions (114 °C for 14 h), indicating that the second residue is not aspartic acid but is likely to be asparagine. Treatment of N^{α}-blocked cyanogen bromide peptide M1 with BNPS-skatole generated a peptide with an amino-terminal sequence of Ala-Thr-Phe---, indicating the presence of a tryptophanyl residue at the third residue. This peptide provided the overlap of peptides K1 and K2. From all of these data, it appears that the sequence of K1 is N^{α}-blocked Gln-Asn-Trp-Ala-Thr-Phe-Gln-Gln-Lys. The amino acid composition of N-blocked chymotryptic peptide C1 (residues 1–6) is in accord with this sequence (B 1.3, Z 1.2, T 0.9, A 1.0, and F 0.9).

Peptides K2 and K3 were analyzed to the carboxyl terminus. Those results provided the sequence of residues 10–35 and 36–49, respectively. The analysis of cyanogen bromide overlapping peptide M2/3 confirmed the carboxy-terminal sequence of K2 (residues 24–35), the entire sequence of K3 (residues 36–49), and the amino-terminal sequence of K4 (residues 50–63) and provided overlaps of K2, K3, and K4. The analysis of K4 extended the sequence up to residue 78, though Ser-75 was tentatively identified and residue 77 was not clearly identified in this analysis. Two subpeptides of K4, K4-T3 and K4-T4, generated with trypsin and isolated by reversed-phase HPLC, and cyanogen bromide peptide M3 were analyzed to provide the rest of the sequence of K4, although the overlap of residues 86 and 87 was rather weak. Peptide K5, which was assumed to be derived from the carboxyl-terminal portion of the whole protein because it lacks lysine in the composition (Table I), was analyzed to the end and provided the sequence of residues 96–111.

The remaining two overlaps of K4-T3 to K4-T4 and of K4 to K5 were finally provided by the analysis of two chymotryptic peptides, C10 and C11.

Finally, the nature of the N-blocking group was examined by using two synthetic nonapeptides corresponding to the sequence of K1 with either Glu or Gln at the amino terminus. Peptides were either acetylated with acetic anhydride and pyridine or treated at 110 °C for 10 h at neutral pH under nitrogen. The latter treatment resulted in formation of pyroglutamyl peptides regardless of the starting peptide. As shown in Figure 3, all eight synthetic peptides (free amino-terminal, N^{α}-acetylated, N^{α},N^{ϵ}-bisacetylated, pyroglutamyl, and N^{ϵ}-acetylated pyroglutamyl peptides) were completely resolved by reversed-phase HPLC under the analytical conditions employed. The amino terminus of the protein is likely to be a pyroglutamic acid since the blocked amino-terminal peptide K1 behaves in exactly the same manner as the synthetic pyroglutamyl peptide before and after acetylation on the reversed-phase column.

Discussion

A few types of lectin with specific binding affinity to sialic acids have been described: wheat germ agglutinin (Greenaway & LeVine, 1973) and hemolymph of horseshoe crab (Roche et al., 1975; Mohar et al., 1982), slug (Miller et al., 1982), and lobster (Noguchi, 1903; Hall et al., 1972, 1974; Ravine & Dranath, 1985). One of the egg lectins from *Rana catesbeiana* has been shown to display preferential agglutination of a large variety of animal and tumor cells including normal lymphocytes, erythrocytes, and fibroblasts of animal and human origin. Since the agglutination was inhibited by sialic acid containing glycoproteins, particularly a high molecular weight sialomucin, and the inhibitory activity of those sialic acid containing compounds can be abolished by desialylation, the agglutination was considered to be based on the specific affinity of lectin toward sialic acid residues of cell surface glycoproteins (Nitta et al., 1986). A lectin with a similar affinity to sialic acid containing compounds was previously isolated and partially purified from *Rana japonica* eggs (Sakakibara et al., 1978).

In general, the proof of the structure of sialic acid binding lectin from *Rana catesbeiana* eggs involved two primary sets of overlapping peptides derived by cleavage at either lysyl or methionyl residues. The third set of peptides generated with chymotrypsin finally confirmed two weak overlaps with the

two primary sets of peptides. A special problem was presented by the blocked amino terminus. The blocking group is likely to be a pyroglutamyl residue (cyclized glutaminyl residue). Since this conclusion is based on indirect evidence, the nature of the blocking group and the order of the first two residues should be regarded as tentative.

Although most residues were identified by two complementary systems, residues 3, 4, and 71–73 were analyzed by only one system. Of the total 111 residues, 93 were analyzed at least twice in different peptides. The portions with single analysis were residues 14–23 and 105–111. The proof of the sequence is rigorous except for the amino-terminal two residues and those indicated above.

From the present study, it is clear that the protein is composed of a single subunit of 111 residues. It appears most likely that it is not a glycoprotein, though Sakakibara et al. (1977) reported that the purified protein contains about 0.6% carbohydrate. The subunit molecular weight is calculated to be 12 450 from the amino acid sequence presented herein, assuming that (1) the amino terminus is a pyroglutamyl residue, (2) all of the half-cystinyl residues are oxidized disulfide bridges, and (3) the protein contains no covalently bound carbohydrate. Although this molecular weight value is slightly lower than the 13 900 estimated by amino acid analysis (K. Nitta, H. Kawauchi, G. Takayanagi, and S. Hakomori, unpublished data), it is in good agreement with the value estimated by gel filtration on TSK columns in the presence of 6 M guanidine after reduction and S-carboxymethylation and by SDS–PAGE (data not shown). Under the above assumptions, the isoelectric point of the protein is calculated to be 12.20, the high content of Asx and Glx being present as the amide forms in the sequence. This agrees with the results that the sum of asparagine and glutamine residues of sialic acid binding agglutinin is 15–18 among 21 residues recovered as aspartic acid and glutamic acid in the acid hydrolysate, that the isoelectric point of the agglutinin is 11.80, and that the protein runs by disc polyacrylamide gel electrophoresis in a position similar to those of basic proteins such as protamine and histone (K. Nitta, H. Kawauchi, G. Takayanagi, and S. Hakomori, unpublished data).

The sequence is unique as well as is the activity. A sequence homology search (segment length 41) did not find a significant relation to any of 3450 protein sequences known, including those of plant or animal lectins. Most of the top 50 sequences showed scattered identity, and some clustered identity (up to 4 residues) appeared to be selected by rather unreasonably high scores given to Cys/Cys matches. We are, therefore, unable to speculate about the site of cell binding of the protein by homology.

Our unpublished results indicate that the modification of histidine, methionine, or tyrosine in the protein with either iodoacetic acid or iodoacetamide has little effect on the agglutinating activity for mouse Ehrlich ascites carcinoma or rat ascites hepatoma cells. Among several chemical modifications that we tested, the blocking of ϵ-amino groups of lysyl residues by acetylation, maleylation, or carbamylation appears to destroy the agglutinating activity, but that of arginyl with butanedione had little effect, implying that the ϵ-amino group of the lysyl residues may play an important role for the biological activity. At present, it is difficult to speculate which specific lysyl residue is most likely involved in the cancer cell binding site of the protein.

Sialic acid binding lectins in the hemolymph of invertebrates have been regarded as functionally similar to vertebrate antibodies in the defense mechanism. In vitro studies have shown that invertebrate lectins do facilitate phagocytosis (McKay & Jenkins, 1970; Pauley et al., 1971). The sialic acid binding lectin in frog egg may well be involved in fertilization and development in the frog embryo, although no studies in this direction have been made. Interestingly, however, the sialic acid binding lectin described in this paper inhibits tumor growth in vivo through a possible activation of natural killer cells and a macrophage system directed toward tumor cells (K. Nitta, F. Sakakibara, and H. Kawauchi, unpublished observation). In fact, leach lectin induces macrophage activation in vitro. Consequently, tumor cell growth was inhibited in vitro (Nakajima et al., 1986). A special organization of amino groups provided by the primary structure of this lectin may provide key interaction between highly sialylated tumor cell receptors and those present in macrophages. Since the primary structure of the lectin is now elucidated, an approach using a synthetic peptide based on the essential sequence of the protein involved in such macrophage/immunocyte activation may provide an important basis for understanding tumoricidal mechanisms triggered by the lectin.

ACKNOWLEDGMENTS

We thank Roger D. Wade, Santosh Kumar, and Maria Harrylock for excellent technical assistance and Mary K. Woods for her patience and skill in manuscript preparation.

REFERENCES

Barondes, S. H. (1981) *Annu. Rev. Biochem. 50*, 207.
Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. L. (1984) *J. Chromatogr. 336*, 93.
Crestfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem. 245*, 1895.
Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983) *Methods Enzymol. 91*, 514.
Ericsson, L. H., Wade, R. D., Gagnon, J., McDonald, R. M., Granberg, R. R., & Walsh, K. A. (1977) *Solid Phase Methods in Protein Sequence Analysis* (Previero, A., & Coletti-Previero, M. A., Eds.) p 137, Elsevier/North-Holland, Amsterdam.
Glajch, J. L., Gluckman, J. C., Charikofsky, J. G., Minor, J. M., & Kirkland, J. J. (1985) *J. Chromatogr. 318*, 2.
Greenaway, P. J., & LeVine, D. (1973) *Nature (London), New Biol. 241*, 191.
Gross, E. (1967) *Methods Enzymol. 11*, 238.
Hall, J. L., & Rowlands, D. T., Jr. (1974) *Biochemistry 13*, 821.
Hall, J. L., Rowlands, D. T., Jr., & Nilsson, U. R. (1972) *J. Immunol. 109*, 816.
Hewick, R. M., Hunkapiller, M. W., Hood, L. E., & Dreyer, W. J. (1981) *J. Biol. Chem. 256*, 7990.
Inglis, A. S. (1983) *Methods Enzymol. 91*, 324.
Kawauchi, H., Sakakibara, F., & Watanabe, K. (1975) *Experientia 13*, 364.
Mahoney, W. C., & Hermodson, M. A. (1980) *J. Biol. Chem. 255*, 11199.
Masaki, T., Tanabe, M., Nakamura, K., & Soejima, M. (1981) *Biochim. Biophys. Acta 660*, 44.
McKay, D., & Jenkin, C. R. (1970) *Aust. J. Exp. Biol. Med. Sci. 48*, 139.
Miller, R. L., Collawn, J. F., Jr., & Fish, W. W. (1982) *J. Biol. Chem. 257*, 7574.
Mohan, S., Dorai, D. T., Srimal, S., & Bachhawat, B. K. (1982) *Biochem. J. 203*, 253.
Nakajima, Y., Suzuki, H., Sakakibara, F., Kawauchi, H., Mizuno, D., & Yamazaki, M. (1986) *Jpn. J. Exp. Med. 56*, 19.

Nitta, K., Takayanagi, G., & Kawauchi, H. (1983) *Chem. Pharm. Bull. 31*, 315.
Noguchi, H. (1903) *Zentralbl. Bakteriol., Parasitenkd., Infektionskrankh. Hyg., Abt. 1: Orig. 133*, 362.
Omenn, G. S., Fontana, A., & Anfinsen, C. B. (1970) *J. Biol. Chem. 245*, 1895.
Pauley, G. B., Granger, G. A., & Krassner, S. M. (1971) *J. Invertebr. Pathol. 18*, 207.
Ravindranath, M. H., Higa, H. H., Cooper, E. L., & Paulson, J. C. (1985) *J. Biol. Chem. 50*, 8850.
Roche, A.-C., Schauer, R., & Mosigny, M. (1975) *FEBS Lett. 57*, 245.
Sakakibara, F., Takayanagi, G., Ise, H., & Kawauchi, H. (1977) *Yakugaku Zasshi 97*, 855.
Sakakibara, F., Kawauchi, H., Takayanagi, G., & Ise, H. (1979) *Cancer Res. 39*, 1347.

Proton Nuclear Magnetic Resonance and Spectrophotometric Studies of Nickel(II)–Iron(II) Hybrid Hemoglobins†

Naoya Shibayama,*,‡ Toshiro Inubushi,‡ Hideki Morimoto,§ and Takashi Yonetani‡

Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6089, and Department of Biophysical Engineering, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan

Received September 19, 1986; Revised Manuscript Received December 12, 1986

ABSTRACT: Ni(II)–Fe(II) hybrid hemoglobins, $\alpha(Fe)_2\beta(Ni)_2$ and $\alpha(Ni)_2\beta(Fe)_2$, have been characterized by proton nuclear magnetic resonance with Ni(II) protoporphyrin IX (Ni-PP) incorporated in apoprotein, which serves as a permanent deoxyheme. $\alpha(Fe)_2\beta(Ni)_2$, $\alpha(Ni)_2\beta(Fe)_2$, and NiHb commonly show exchangeable proton resonances at 11 and 14 ppm, due to hydrogen-bonded protons in a deoxy-like structure. Upon binding of carbon monoxide (CO) to $\alpha(Fe)_2\beta(Ni)_2$, these resonances disappear at pH 6.5 to pH 8.5. On the other hand, the complementary hybrid $\alpha(Ni)_2\beta(Fe\text{–}CO)_2$ showed the 11 and 14 ppm resonances at low pH. Upon raising pH, the intensities of both resonances are reduced, although these changes are not synchronized. Electronic absorption spectra and hyperfine-shifted proton resonances indicate that the ligation of CO in the β(Fe) subunits induced changes in the coordination and spin states of Ni-PP in the α subunits. In a deoxy-like structure, the coordination of Ni-PP in the α subunits is predominantly in a low-spin ($S = 0$) four-coordination state, whereas in an oxy-like structure the contribution of a high-spin ($S = 1$) five-coordination state markedly increased. Ni-PP in the β subunits always takes a high-spin five-coordination state regardless of solution conditions and the state of ligation in the partner α(Fe) subunits. In the β(Ni) subunits, a significant downfield shift of the proximal histidyl $N_\delta H$ resonance and a change in the absorption spectrum of Ni-PP were detected, upon changing the quaternary structure of the hybrid. Furthermore, the proximal histidyl resonance of $\alpha(Fe\text{–}CO)_2\beta(Ni)_2$ undergoes a slight upfield shift upon raising the pH from 6.5 to 8.5, while the 14 ppm resonance, the marker signal of a deoxy quaternary structure, hardly changes in this pH region. In addition, ring current shifted resonances in both hybrid Hbs were examined. The chemical shifts were analyzed in terms of the E11-Val methyls vs. the porphyrin rings in hybrid Hbs.

Even though Hb[1] is a well-investigated allosteric protein, the mechanism of cooperative oxygenation is not fully understood. This is caused by the difficulty in characterizing the intermediate species directly in the course of oxygenation, since the intermediate species of ligation are present in low concentrations in a cooperative system. Thus, many kinds of hybrid Hbs, which were regarded as models for the intermediate species, have been artificially prepared and examined (Banerjee & Cassoly, 1969; Ikeda-Saito et al., 1977; Blough & Hoffman, 1984; Simolo et al., 1985).

We have surveyed metal-substituted hybrid Hbs, $\alpha(Fe)_2\beta(M)_2$ and $\alpha(M)_2\beta(Fe)_2$, in which the hemes in either α or β subunits are substituted with porphyrins containing iron-series transition metal ions (M). The oxygen equilibrium properties of these hybrid Hbs indicated that Ni(II) protoporphyrin IX, Cu(II) protoporphyrin IX, and Zn(II) protoporphyrin IX behave like a permanent deoxyheme when the first oxygen molecule binds to these hybrid Hbs (Shibayama et al., 1986a; N. Shibayama et al., unpublished results; Miyazaki et al., unpublished results). We further carried out several chemical modifications such as des-Arg(α141), *N*-ethylsuccinimide–β93-Cys [NES(β93-Cys)], and NES(β93-Cys)-des-Arg(α141) on Ni–Fe hybrid Hbs (Shibayama et al., 1986b). By comparison with the oxygen equilibrium data between modified

†This work was supported by a grant from the NIH (HL 14508 to T.Y.). All NMR spectra were taken in the medical school NMR facility at the University of Pennsylvania, which is partially supported by Grants SO7-RR-05415 and SO7-RR-07083 from the NIH.

‡University of Pennsylvania.
§Osaka University.

[1] Abbreviations: Hb, hemoglobin; M–Fe hybrid Hb, hybrid hemoglobin in which hemes in either the α or β subunit are substituted with an iron-series transition metal ion (M); $\alpha(Fe)_2\beta(Ni)_2$, hybrid hemoglobin containing ferrous protoporphyrin IX in the α subunits and nickel(II) protoporphyrin IX in the β subunits; $\alpha(Ni)_2\beta(Fe)_2$, hybrid hemoglobin complementary with $\alpha(Fe)_2\beta(Ni)_2$; NiHb, hemoglobin containing nickel(II) protoporphyrin IX in both the α and β subunits; Ni-PP, nickel(II) protoporphyrin IX; NMR, nuclear magnetic resonance; CO, carbon monoxide; Bistris, 2-[bis(2-hydroxymethyl)amino]-2-(hydroxymethyl)-1,3-propanediol; Tris, tris(hydroxymethyl)aminomethane; FeHb, hemoglobin containing ferrous protoporphyrin IX in both the α and β subunits; IHP, inositol hexaphosphate; ZnHb, hemoglobin containing zinc(II) protoporphyrin IX in both the α and β subunits; SH, sulfhydryl.

0006-2960/87/0426-2194$01.50/0