THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 257, No. 2, Issue of January 25, pp. 643–650, 1982
Printed in U.S.A.

Amino Acid Sequences of the ϵ and α^E Globins of HbE, a Minor Early Embryonic Hemoglobin of the Chicken*

(Received for publication, April 24, 1981, and in revised form, September 9, 1981)

Barbara S. Chapman‡ and Leroy E. Hood§

From the Division of Biology, California Institute of Technology, Pasadena, California 91125

Allan J. Tobin¶

From the Department of Biology, University of California, Los Angeles, California 90024

We have determined amino acid sequences for the α-like and β-like globin components of HbE, one of the two minor hemoglobins in early chick embryos. The complete primary structure of the ϵ chain differs at 18 positions from the adult chicken β globin, but there are no changes in heme-binding residues, $\alpha_1\beta_2$ contact positions, or allosteric regulatory sites. By amino acid sequence analysis, we have identified a new α-like globin that we have called α^E. The α^E globin chain differs from the major adult α^A chain at 22 amino acid positions. This paper discusses the structural and implied functional characteristics of these globins and presents hypotheses regarding the possible role of minor embryonic hemoglobins.

Four hemoglobins, two major and two minor, are produced by yolk sac-derived erythroid cells of the early chick embryo (Bruns and Ingram, 1973; Brown and Ingram, 1974). Our previous analysis of the primary structures of globins composing the major early embryonic hemoglobins P and P′ suggests that members of the α globin family are more divergent from each other than are members of the β globin family in domestic chickens (Chapman *et al.*, 1980 and 1981). The α-like π and π′ globins (from Hbs P and P′) differ from each of the two adult chicken α-like globins in 43% of their amino acid residues, whereas the β-like ρ globin differs from the adult β globin by only 13%. We examined the possible functional significance of these amino acid substitutions by referring to the three-dimensional hemoglobin structure of Ladner *et al.* (1977). We inferred that unusual substitutions in π and π′ globins may account for the high oxygen affinity of the early embryonic hemoglobins as well as for their small Bohr effect. On the other hand, none of the substitutions in the ρ globin are likely to affect oxygen affinity or binding of ATP, which is the primary organophosphate in early embryonic red cells (Bartlett and Borgese, 1976).

In this paper, we examine the structure of HbE, one of the minor embryonic hemoglobins. We report here the complete amino acid sequence of the ϵ globin, the β-like chain of HbE, together with a partial sequence of the α^E globin covering 81 of its 141 residues. The partial α^E globin sequence, surprisingly, reveals a new chain, which differs from the previously reported adult α^A chain by multiple amino acid substitutions (Matsuda *et al.*, 1971; Paul *et al.*, 1974). The α^E globin amino acid residues not analyzed in these experiments have been determined from recombinant cDNA sequences (Salser *et al.*, 1979; Deacon *et al.*, 1980; Richards and Wells, 1980). As in the major early embryonic hemoglobins, we find that the minor β-like chains are not functionally different from the adult β chains, but there are significant structural and functional differences in the minor α-like chains.

Since this is the first report of complete primary structures for minor early embryonic hemoglobins, there are a number of important questions to be addressed. These questions regard the role of minor early embryonic hemoglobins in chicken development in particular and in vertebrate development in general. First, by analysis of these sequences in terms of published hemoglobin crystal structures, we will try to predict the functional properties of the native early embryonic hemoglobin molecules. Second, we will discuss the controversial suggestion that the α^E globin also serves as the major adult α-like globin in domestic chickens. Third, assuming that the α^E (embryonic) and α^A (adult) globins are distinct, we will attempt to identify the time, conditions, and mechanism through which α globin switching might occur. Fourth, we will consider the importance of α chain switching during embryonic development and present additional evidence for the hypothesis that embryos manage major shifts in hemoglobin oxygen affinity by α chain substitution, rather than by changes in allosteric effectors. Finally, we will discuss some rationales for the ubiquitous presence of minor hemoglobins in early vertebrate embryos.

EXPERIMENTAL PROCEDURES AND RESULTS[1]

DISCUSSION

Structures of the Globins of HbE Compared to Other Chicken Globin Chains—The complete sequence of the ϵ chain of HbE shown in Fig. 1 differs from the adult chicken β globin by 18 amino acid substitutions (Table I). Two-thirds

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of National Research Service Award HL05553. Present address, Department of Biochemistry, University of California, Berkeley, CA 94720. To whom inquiries may be addressed.

§ Recipient of National Institutes of Health Grant GM-06965.

¶ Recipient of National Science Foundation Grants PCM 76-02859 and 78-20767.

[1] Portions of this paper (including "Experimental Procedures," "Results," Figs. 1–9, and Table I) are presented in miniprint at the end of this paper. The abbreviations used are: CAM, carboxyamidomethylated; PTH, phenylthiohydantoin; HPLC, high pressure liquid chromatography; IHP, inositol hexaphosphate. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 81M-935, cite authors, and include a check or money order for $6.80 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

are conservative with respect to polarity and charge. Although there are four differences in predicted $\alpha_1\beta_1$ contact sites (Ladner *et al.*, 1977), three of which interact with changed residues in the α^E chain, there are no changes in Bohr effect or organophosphate binding site residues. For analysis of the ϵ globin organophosphate binding site, we have considered only positions 1, 2, 82, and 143, which are the mammalian DPG contacts. Additional positions 135 and 139, suggested by Arnone and Perutz (1974) as possible contacts to inositol pentaphosphate in adult chickens, are not considered here because inositol pentaphosphate is not present in early embryonic chick red cells (Bartlett and Borgese, 1976). Like the major embryonic ρ globin, the ϵ globin probably does not make a significant contribution to the elevated oxygen affinity or reduced Bohr effect characteristic of the early embryonic hemoglobins. Structural adaptation of the ϵ chains seems to be confined to the $\alpha_1\beta_1$ contact residues, which form a rigid interface with the α-like chains in hemoglobin dimers. In contrast, the α-like chains of HbE differ from the major adult α globin by many nonconservative changes, several of which are in functionally important positions.

We have determined an amino acid sequence for 81 residues of the α^E globin chain. These data show at least 12 differences between α^E chains and the α^A chains of adult chickens (Matsuda *et al.*, 1971). The partial α^E amino acid sequence is closely related to that determined from recombinant cDNA

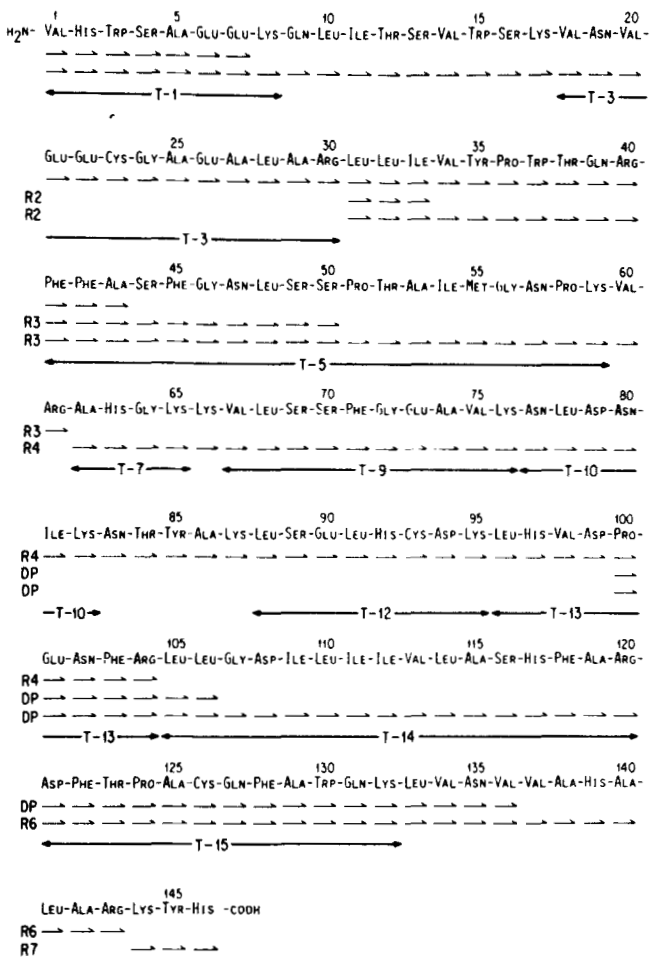


FIG. 1. **The complete amino acid sequence of the ϵ chain, the β-like globin of HbE.** →→, automated Edman degradation; ←*T*→, tryptic peptides; *R*, arginine cleavage fragment; *DP*, aspartic acid-proline cleavage fragment. Overlapping peptides were not obtained for residues 61 and 62 or for 143 and 144. Positions of the various fragments were determined as described in the text.

TABLE I

Comparison of amino acid substitutions among the four β-like polypeptides of the domestic chicken

The ρ globin sequence was obtained from Chapman *et al.* (1981), the β^H sequence from Dolan *et al.* (1981), the β globin sequence from Matsuda *et al.* (1973), and information on the three-dimensional structure from Ladner *et al.* (1977) and Arnone and Perutz (1974).

Position	Helix	ϵ	ρ	β^H	β	Function	Interacts with α
4	A1	Ser	Ser		Thr		
13	A10	Ser	Ser		Gly		
14	A11	Val	Val		Leu		
16	A13	Ser	Ser		Gly		
21	B3	Glu	Glu		Ala		
43	CD2	Ala	Asp	Ala	Ala	$\alpha_1\beta_2$ contact	92
44	CD3	Ser	Asn	Ser	Ser		
51	D2	Pro	Pro	Ala	Pro		
55	D6	Met	Ile	Ile	Leu	$\alpha_1\beta_1$ contact	119
59	E3	Lys	Lys	Met	Met		
69	E13	Ser	Ser	Ser	Thr		
73	E17	Glu	Glu	Glu	Asp		
83	EF6	Asn	Asn	Lys	Asn		
84	EF7	Thr	Thr	Ser	Thr		
85	F1	Tyr	Tyr	Phe	Phe		
86	F2	Ala	Ala	Ala	Ser		
87	F3	Lys	Lys	Gln	Gln		
90	F6	Glu	Glu	Lys	Glu		
94	FG1	Asp	Glu	Asp	Asp		
108	G10	Asp	Asn		Asp	$\alpha_1\beta_1$ contact	103
116	G18	Ser	Ala		Ala	$\alpha_1\beta_1$ contact	117, 114, 110
119	GH2	Ala	Thr		Ser	$\alpha_1\beta_1$ contact	111
120	GH3	Arg	Lys		Lys		
125	H3	Ala	Glu		Glu		
128	H6	Phe	Ala		Ala	$\alpha_1\beta_1$ contact	34, 35
135	H13	Asn	Ser		Arg	IHP[a] contact	
139	H17	His	Lys		His	IHP contact	

[a] IHP, inositol hexaphosphate.

clones pHb1003 (Salser *et al.*, 1979), pCGα-8 (Deacon *et al.*, 1980), and pCGα-3 (Richards and Wells, 1980). These clones were obtained from messenger RNA present in circulating blood cells of anemic chickens. Fig. 2 shows protein sequence data for α^E together with translations of nucleotide sequences for clones pHb1003, pCGα-8, and pCGα-3, and the sequence of the major adult α globin. The amino acid sequences predicted by the three cDNA clones are identical from the NH_2 terminus to residue 91 and from residue 122 to the COOH terminus. The recombinant DNA sequences differ from each other at amino acid residues 92–94, 110, and 120–121. The amino acid sequence predicted by clone pCGα-3 (Richards and Wells, 1980) is identical with our protein sequence.

The α^E globin differs by 22 amino acid substitutions from α^A chains of adult HbA (Table II). Fifty-nine per cent of these changes are not conservative, but there is no net difference in charge between α^A and α^E chains. These globins appear to be indistinguishable by standard electrophoretic or chromatographic methods. Although there are no differences between α^E and α^A among predicted Bohr-effect or heme-contact residues, there are six changes in $\alpha_1\beta_1$ contact residues. Three of these in α^E chains (34, 110, and 111) interact with substituted residues in ϵ chains. There is one change at $\alpha_1\beta_2$ contact position 38 involving a Pro for Thr substitution. At this interesting position, there also are amino acid replacements in the early embryonic α-like π and π' globins and in the α^D globin (Takei *et al.*, 1975). We have suggested that substitutions in this position may in part account for the elevated oxygen affinity of early embryonic hemoglobins (Chapman *et al.*, 1980). An Ala to Val substitution at position α 63 has been implicated in elevating the intrinsic oxygen affinity of adult hemoglobin in the barheaded goose (Oberthur et al., 1980). The α^E globin also has Val at position 63. Thus, there are at least two amino acid replacements in α^E chains relative to α^A

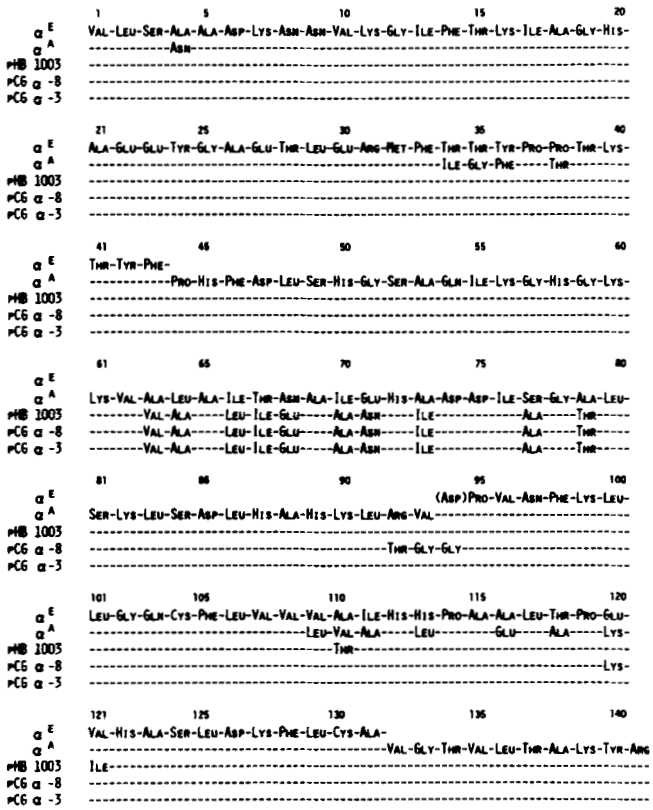


FIG. 2. **Amino acid sequence of the α^E globin chain.** This shows amino acid residues of α^E globin determined from sequenator analyses, together with amino acid sequences encoded by recombinant cDNA clones pHb1003 (Salser *et al.*, 1979), pCGα-8 (Deacon *et al.*, 1980), and pCGα-3 (Richards and Wells, 1980). The α^A chain sequence of Matsuda *et al.* (1971) is shown for comparison. Amino acid residues different from those in the *uppermost line* are indicated in the *lines below*.

chains that might increase the intrinsic oxygen affinity of chicken hemoglobins containing the α^E chain. We conclude that in the minor HbE, as in the major HbP and HbP', the α-like chains are primarily responsible for the high intrinsic oxygen affinity of the early embryonic hemoglobins.

Is the α^E Chain Also the Major Adult α-like Chain?—Recently, there has been some speculation that α^E globin may be the major adult chicken α-like globin (Dodgson *et al.*, 1981). This idea is supported by a number of findings. First, all fully analyzed adult chicken cDNA clones have been of the α^E globin-type (Salser *et al.*, 1979; Deacon *et al.*, 1980; Richards and Wells, 1980; Dodgson *et al.*, 1981). Second, amino acid sequences of major adult α-like globins from the bar-headed goose, the ostrich, and one analysis from the greylag goose are of the α^E globin type (Table III). Third, only one major α-like gene has been characterized in chicken chromosomal DNA, and it appears to encode the α^E globin (Dodgson *et al.*, 1981). On the other hand, it seems to us unlikely that the amino acid sequence prepared by Matsuda *et al.* (1971) and confirmed in part by NH_2-terminal sequenator analysis (Paul *et al.*, 1974), could contain so many serious errors. Table III shows that there is a discrepancy of similar magnitude between two amino acid sequences for the major α-like globin of the greylag goose. The α^E and α^A globin sequences cannot represent allelic variants because there are too many differences and because genetic data from mitochondrial DNA analysis show that domestic chickens form a recently derived, homogeneous group (Glaus *et al.*, 1980). In addition, we note a preliminary report describing major adult α globin cDNA

TABLE II

Amino acid substitutions in the early embryonic α^E chain compared with the adult α^A chain residues

The α^A sequence was obtained from Matsuda *et al.* (1971) and the three-dimensional locations from Ladner *et al.* (1977).

Position	Helix	α^E	α^A	Function	Interacts with β
4	A2	Asn	Ala		
34	B15	Thr	Ile	$\alpha_1\beta_1$ contact	128
35	B16	Thr	Gly	$\alpha_1\beta_1$ contact	128
36	C1	Tyr	Phe	$\alpha_1\beta_1$ contact	131, 104
38	C3	Pro	Thr	$\alpha_1\beta_2$ contact	97
63	E12	Val	Ala		
64	E13	Ala	Leu		
66	E15	Leu	Ile		
67	E16	Ile	Thr		
68	E17	Glu	Asn		
70	E19	Ala	Ile		
71	E20	Asn	Glu		
73	EF1	Ile	Ala		
77	EF5	Ala	Ser		
79	EF7	Thr	Ala		
109	G16	Val	Leu		
110	G17	Ala	Val	$\alpha_1\beta_1$ contact	112, 116
111	G18	Ile	Ala	$\alpha_1\beta_1$ contact	119, 115
113	GH1	His	Leu		
116	GH4	Ala	Glu		
118	H1	Thr	Ala	$\alpha_1\beta_1$ contact	30
120	H3	Glu	Lys		

TABLE III

Comparison of six avian α globin amino acid sequences

Data were obtained from the following: 1) Matsuda *et al.* (1971); 2) Richards and Wells (1980); 3) Debouverie (1975); 4) Braunitzer and Oberthur (1979); 5) Oberthur *et al.* (1980); and 6) Oberthur *et al.* (1980).

	Amino acid differences[a]					
	1)	2)	3)	4)	5)	6)
1) Chicken α^A		22	17	30	31	28
2) Chicken α^E	15.6		28	16	17	18
3) Greylag goose	12.1	19.9		28	30	30
4) Greylag goose	21.3	11.3	19.9		3	15
5) Barheaded goose	22.0	12.1	21.3	2.1		16
6) Ostrich	19.9	12.8	21.3	10.6	11.3	

[a] Above diagonal line, number of differences; below line, percentage of differences.

clones that appear distinct from previously reported α^E cDNA clones (Reynaud *et al.*, 1980) and a report of a peculiar anemic duck α-like globin (Paddock and Gaubatz, 1981). It thus seems premature to conclude that our α^E globin sequence, rather than that previously published, is the major adult α-like globin in domestic chickens.

Is There α Globin Switching in Adult Birds?—We have considered the possibility that adult birds may switch α^A and α^E globins under certain physiological conditions. If α^E chains confer higher intrinsic oxygen affinity on hemoglobin than do α^A chains, then the net effect of switching would be elevated peripheral blood oxygen affinity. Based on experiments with mammals, Eaton *et al.* (1973) have suggested that "increased, rather than decreased, oxygen affinity is an effective mode of short-term adaptation to markedly reduced environmental oxygen pressures." Birds may use such a strategy to adapt to lowered oxygen tensions and high tissue demand encountered during flight.

Chicken α Globin Switching Might Have a Cellular Mechanism—In a preliminary analysis of the globin produced by chicken bone marrow cells following phenylhydrazine injection,[2] we observed that normal adult α^A globin was decreased

[2] B. Chapman and A. Tobin, unpublished.

or undetectable in bone marrow hemolysates from anemic chickens. These analyses of chicken globins were carried out by polyacrylamide gel electrophoresis of hemolysates under denaturing conditions (Rovera *et al.*, 1978; Tobin *et al.*, 1979). As expected, peripheral blood hemolysates from anemic adult chickens contained greatly reduced or undetectable levels of α^A globin whenever α^A globin was missing from bone marrow hemolysates. The amount of α^D globin (from the minor adult chicken HbD) was not depressed by phenylhydrazine treatment but instead appeared to increase, as previously observed (Stino and Washburn, 1970). However, we found normal levels of an α^A-like globin in peripheral blood from two- to three-week-old chicks, even when no α^A globin was detectable in their bone marrow samples. This anomalous result appeared to be correlated with the presence of large numbers of circulating erythroblasts in young chicks (Godet et al., 1970).[2] Although we recognize that our preliminary results and those of earlier workers may be artifactual, we hypothesize that α^A globin is mainly synthesized by bone marrow erythroid cells and that electrophoretically indistinguishable α^E globin may be preferentially produced in circulating erythroblasts. It is possible that α^E globin may be produced during postembryonic life under conditions of anemia or hypoxia. Our unpublished results suggest that if there is α globin switching in adult birds, it may be accomplished by substitution of circulating cell types.

Regulation of α^E Chain Synthesis during Ontogeny—Since tryptic peptide maps and amino acid compositions suggest that the α components of HbA in 18-day embryos are α^A chains (Bruns and Ingram, 1973; Moss and Hamilton, 1974), α^E chains may be replaced by α^A chains between the 5th and 18th days of chick embryonic development. If this α globin switch occurs during ontogeny and if the cellular mechanism discussed above is also used by the embryo, then we have some clues as to when the α^E to α^A globin switch might occur.

Three lines of evidence indicate that a shift from α^E to α^A chain synthesis could occur on the 14th day of development. Large numbers of erythroblasts and reticulocytes circulate until day 14 of ontogeny, then give way to mature cells during the next 2 days (Bruns and Ingram, 1973; Chapman and Tobin, 1979). If embryonic α^E chain synthesis occurs in immature peripheral blood cells, then the amount of α^E globin may decline along with numbers of these cells. Secondly, erythropoiesis is initiated on day 14 in bone marrow (Sandreuter, 1951; Godet, 1974). If embryonic α^A globin is produced by erythroid cells residing in bone marrow, then α^A chains should increase when bone marrow becomes a major erythropoietic site. Thirdly, the ratio of HbD to HbA changes from nearly 1.0 at day 7 to about 0.4 after day 14 (Bruns and Ingram, 1973). Since HbD and HbA contain the same β chain (Vandecasserie *et al.*, 1975), any change in the ratio of these hemoglobins represents a change in relative amounts of α^D and α^A chains. The α^E globin, however, cannot be distinguished from α^A globin, so the reported HbD to HbA ratio thus reflects the α^D/α^E globin chain ratio until α^A globin replaces α^E globin. Stabilization after day 14 of the α chain ratio may indicate a switch from α^E to α^A chain synthesis at day 14.

Possible Role for ϵ and α^E Globins in Chick Development—We have proposed that hemoglobin switching provides a way for vertebrate embryos to adjust their blood oxygen affinity to meet changing physiological conditions during development. In the chick embryo, a graded decrease in oxygen-carrying capacity seems to be achieved by sequential substitution of α-like globin chains having decreasing intrinsic oxygen binding affinities. The highly conserved structure of the χ and ζ chains of mammalian embryos suggests that these major early embryonic α-like chains play a similar physiological role at the beginning of development (Melderis et al., 1974; Jelkmann and Bauer, 1978). It has been suggested that elevated oxygen affinity is useful to the rapidly metabolizing early embryo in its relatively anoxic environment and that the major early hemoglobins may be specially adapted for functioning in the nucleated primary erythroblast (Jelkmann and Bauer, 1978).

The early embryonic chicken β-like chains, on the other hand, appear to function essentially as substitutes for late embryonic β chains until they appear, differing from the β chains with respect to amounts produced and duration of synthesis. In the chicken and also in rabbits, mice, pigs, and humans, there are two kinds of early embryonic β-like chains: a minor one that forms minor hemoglobins with adult-type α chains and a major one that combines with the divergent early α-like chains (Steinheider *et al.*, 1972; Fantoni *et al.*, 1967; Steer and Braunitzer, 1980; Gale *et al.*, 1979). These mammals differ from the chicken during the middle and late embryonic periods, however, in that their strategy for maintaining elevated oxygen affinity depends on the properties of the β-like γ chain, rather than on switching α-like globins.

Although both major and minor early hemoglobins are replaced at the end of the first developmental period by a new set of hemoglobins in the chick embryo, α^D and, presumably, α^E globins persist as the α-like chains in the major hemoglobins of the next developmental period. Similarly, α and γ chains of mice, rabbits, and humans are initially present in minor early embryonic hemoglobins, then later are found in major hemoglobins (Fantoni *et al.*, 1967; Steinheider *et al.*, 1972; Gale *et al.*, 1979). Some evolutionary advantages accrue to systems organized in this way. At the time of globin gene switching, no more than one new major globin gene need be turned on. Also, the evolutionary process by which α- to-β chain ratios are balanced must act only on one locus at a time.

The complexity of early embryonic hemoglobin switching seems to reflect the interplay of at least two factors. One is evolutionary selection for specific hemoglobin function during embryogenesis. Another is the genetic mechanism by which expression of multigene family members is regulated (Hood *et al.*, 1975).

Acknowledgments—We sincerely thank Dr. James Schilling for teaching one of us (B. S. C.) amino acid sequencing techniques, Margaret Kowalczyk for illustration, and Lori Erdley for patient assistance in the preparation of this manuscript. The critical comments of Dr. Allan C. Wilson and Sandra Martin were invaluable. We especially appreciate the many stimulating discussions with Dr. J. D. Engel and thank him for sharing unpublished manuscripts with us.

REFERENCES

Arnone, A., and Perutz, M. F. (1974) *Nature* **249,** 34–36
Bartlett, G. R., and Borgese, T. A. (1976) *Comp. Biochem. Physiol. A. Comp. Physiol.* **55,** 207–210.
Braunitzer, G., and Oberthür, W. (1979) *Hoppe-Seyler's Z. Physiol. Chem.* **360,** 679–683
Brown, J. L., and Ingram, V. M. (1974) *J. Biol. Chem.* **249,** 3960–3972
Bruns, G., and Ingram, V. M. (1973) *Philos Trans. R. Soc. Lond. B Biol. Sci.* **266,** 225–305
Chapman, B., and Tobin, A. (1979) *Dev. Biol.* **69,** 375–387
Chapman, B. S., Tobin, A. J., and Hood, L. (1980) *J. Biol. Chem.* **255,** 9051–9059
Chapman, B. S., Tobin, A. J., and Hood, L. (1981) *J. Biol. Chem.* **256,** 5524–5531
Dayhoff, M. O. (1976) *Atlas of Protein Sequence and Structure,* National Biomedical Research Foundation, Silver Spring, MD
Deacon, N., Shine, J., and Naora, H. (1980) *Nucleic Acids Res.* **8,** 1187–1199
Debouverie, D. (1975) *Biochimie* **57,** 569–578
Dodgson, J. B., McCune, K. C., Rusling, D. J., Krust, A., and Engel, J. D. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 5998–6002

Dolan, M., Sugarman, B. J., and Engel, J. D. (1981) *Cell* **24,** 669–677
Eaton, J. W., Skelton, T. D., and Berger, E. (1973) *Science* **183,** 743–744
Fantoni, A., Bank, A., and Marks, P. (1967) *Science* **157,** 1327–1329
Gale, R., Clegg, J., and Huehns, E. (1979) *Nature* **280,** 162–164
Glaus, K. R., Zassenhaus, H. P., Fechheimer, N. S., and Perlman, P. S. (1980) in *The Organization and Expression of the Mitochondrial Genome* (Kroon, A. M., and Saccone, C., eds) pp 131–135, Elsevier/North Holland Biomedical Press, Amsterdam
Godet, J. (1974) *Dev. Biol.* **40,** 199–207
Godet, J., Schurch, D., Blanchet, J. P., and Nigon, V. (1970) *Exp. Cell Res.* **60,** 157–165
Hood, L., Campbell, J., and Elgin, S. (1975) *Annu. Rev. Genet.* **9,** 305–353
Hunkapiller, M. W., and Hood, L. (1978) *Biochemistry* **17,** 2124–2133
Jelkmann, W., and Bauer, C. (1978) *Pfluegers Arch. Eur. J. Physiol.* **377,** 75–80
Johnson, N., Hunkapiller, M., and Hood, L. (1979) *Anal. Biochem.* **100,** 335–339
Ladner, C., Heidner, E., and Perutz, M. (1977) *J. Mol. Biol.* **114,** 385–414
Matsuda, G., Takei, H., Wu, K., and Shiozawa, T. (1971) *Int. J. Protein Res.* **3,** 173–174
Matsuda, G., Maita, T., Mizumo, K., and Ota, H. (1973) *Nature New Biol.* **244,** 244
McKean, D., Potter, M., and Hood, L. (1973) *Biochemistry* **12,** 749–759
Melderis, H., Steinheider, G., and Ostertag, W. (1974) *Nature* **250,** 774–776
Moss, B. A., and Hamilton, E. A. (1974) *Biochim. Biophys. Acta* **371,** 379–391
Oberthür, W., Voelter, W., and Braunitzer, G. (1980) *Hoppe-Seyler's Z. Physiol. Chem.* **361,** 967–975
Z. Physiol. Chem. **361,** 967–975
Paddock, G. V., and Gaubatz, J. (1981) *Eur. J. Biochem.* **117,** 269–273
Paul, C., Vandecasserie, C., Schnek, A. G., and Leonis, J. (1974) *Biochim. Biophys. Acta* **371,** 155–158
Piszkiewicz, D., Landon, M., and Smith, E. (1970) *Biochem. Biophys. Res. Commun.* **40,** 1173–1178
Reynaud, C. A., Tahar, S. B., Krust, A., Franco, M. P., Goldenberg, S., Gannon, F., and Scherrer, K. (1980) *Gene* **11,** 259–269
Richards, R. I., and Wells, J. R. E. (1980) *J. Biol. Chem.* **255,** 9306–9311
Rossi-Fanelli, A., Antonini, E., and Caputo, A. (1958) *Biochim. Biophys. Acta* **30,** 608–615
Rovera, G., Magarian, C., and Borun, T. (1978) *Anal. Biochem.* **85,** 506–518
Salser, W., Cummings, I., Liu, A., Strommer, J., Padayatty, J., and Clarke, P. (1979) in *Cellular and Molecular Regulation of Hemoglobin Switching* (Stamatoyannopoulos, G., and Nienhuis, A., eds) pp. 621–645, Grune and Stratton, New York
Sandreuter, A. (1951) *Acta Anat.* **11,** (Suppl. 14) 1–72
Steer, W., and Braunitzer, G. (1980) *Hoppe-Seyler's Z. Physiol. Chem.* **361,** 1165–1169
Steinheider, G., Melderis, H., and Ostertag, W. (1972) *Haematol. Bluttransfus.* **10,** 225–235
Stino, F. K., and Washburn, K. W. (1970) *Poultry Sci.* **49,** 101–114
Takei, H., Ota, Y., Wu, K. C., Kiyohara, T., and Matsuda, G. (1975) *J. Biochem. (Tokyo)* **77,** 1345–1347
Tobin, A., Chapman, B., Hansen, D., Lasky, L., and Selvig, S. (1979) in *Cellular and Molecular Regulation of Hemoglobin Switching* (Stamatoyannopoulos, G., and Nienhuis, A., eds) pp. 205–212, Grune and Stratton, New York
Vandecasserie, C., Paul, C., Schnek, A. G., and Leonis, J. (1975) *Biochimie* **57,** 843–844

Supplementary material to "Amino Acid Sequences of the ε and α^E Globins of HbE" by B.S. Chapman, L.E. Hood, and A.J. Tobin.

EXPERIMENTAL PROCEDURES

Isolation of ε^E and α^E globins

Hemolysates prepared from 5 day old White Leghorn embryos (Chapman and Tobin, 1979) were separated into their component hemoglobins by ion exchange chromatography on CM-Sephadex (Pharmacia) as described by Brown and Ingram (1974). HbE was obtained by pooling fractions from the last peak of hemoglobin to be eluted by the phosphate buffer gradient (data not shown). These fractions were concentrated against solid sucrose, dialyzed against distilled water and converted to globin by acid acetone precipitation (Rossi-Fanelli et al., 1958). Reduced HbE globin was alkylated with Iodoacetamide (McManan et al., 1973), desalted by gel filtration on Biogel P-10 (Biorad), and lyophilized.

To separate α-like from β-like globins, approximately 10 mg of alkylated globin were dissolved in 0.2 M formic acid and applied to CM-cellulose (1.5 x 10 cm column of CM-52, Whatman). After a 50 ml wash with starting buffer, the globins were eluted with a linear gradient composed of 150 ml of 0.2 M formic acid and 150 ml of 1.8 M formic acid adjusted to pH 1.9 with redistilled pyridine (Hoss and Hamilton, 1974). Fractions containing α-like globin, which eluted first, and β-like globin, which eluted later, were diluted with distilled water and lyophilized.

Automated Edman degradation and PTH-amino acid identification

Globin was sequenced with a modified Beckman 890B sequenator (Hunkapiller and Hood, 1978). All reagents, solvents and procedures were as previously described (Hunkapiller and Hood, 1978; Chapman et al., 1980). In one experiment fluorescamine was used to block NH_2-terminal sequences contaminating aspartic acid-proline cleavage fragments. Fluoram (Roche) was dissolved to 1 mg/ml in heptane. Approximately 25 nmol of globin were dried in the spinning cup, and 200 µl of fluorescamine were added and dried under vacuum. Quadrol coupling buffer was introduced. After 2 min, the sample was dried and washed with benzene and ethyl acetate. At this point the normal Edman degradation cycle was initiated. For all sequenator runs, phenylthiohydantoin (PTH) amino acid derivatives were identified by reverse phase high pressure liquid chromatography (HPLC) (Johnson et al., 1979).

Arginine and aspartic acid-proline cleavages

Carboxyamidomethylated (CAM) globin was specifically cleaved at the aspartic acid-proline peptide bond using mild acid hydrolysis (Piszkiewicz et al., 1970; Chapman et al., 1980). The NH_2-terminal and C-terminal portions of the cleaved globin were separated on the basis of their sizes by gel filtration on a column of Biogel P-10 (1.5x90 cm in 0.5% formic acid).

Arginine cleavage fragments were prepared by trypsin digestion of succinylated globin (Chapman et al., 1980), and separated on the basis of size under non-denaturing conditions using a 1.5 x 75 cm column of P-10 equilibrated with 0.1 M ammonium bicarbonate. Separated fragments were pooled and lyophilized.

Tryptic peptide compositions

Purified globin was digested with TPCK-treated trypsin and separated in two dimensions on Kodak thin layer polyamide sheets (Brown and Ingram, 1974). Peptides were identified by ninhydrin staining, eluted, and hydrolyzed (Chapman et al., 1980). Peptides migrating rapidly in the chromatography dimension were split into two portions after elution and hydrolyzed for 24 and 72 hrs. This procedure allowed improved accuracy in analyses of peptides containing several consecutive hydrophobic residues. Quantitative amino acid analyses were performed with a Durrum D-500 amino acid analyzer.

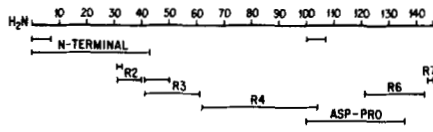


Figure 1. Scheme for automated sequence analysis of the ε globin chain. The extent of sequenator analysis for each fragment is shown by a bar. Analyses indicated on the same line were carried out simultaneously, e.g. R3 and R6. Parts of R2, R3, and the aspartic acid-proline cleavage fragment were analyzed three times. Details of fragment preparation and residue assignment are discussed in the text.

RESULTS

The strategy described below (and illustrated in Fig. 1) was designed to determine the complete amino acid sequence of the β-like chain and determine if α^A was the α-like chain of HbE. This was done using about 10 mg of pure globin. Hemolysates from five day embryos were fractionated by ion exchange chromatography on CM-Sephadex. The late-eluting HbE peak was concentrated, converted to globin by acid acetone precipitation, and alkylated. An NH_2-terminal analysis using the automatic sequenator was obtained from 0.5 mg of pure HbE globin with both α and β chains present (Fig. 2). An additional 0.5 mg of HbE was cleaved by mild acid hydrolysis at the single aspartic acid-proline bond present in α-like and β-like globins. The large NH_2-terminal and small C-terminal fragments were separated by gel filtration on Biogel P-10 (data not shown). The small fragments were analyzed with the sequenator in the presence and absence of fluorescamine to block contaminating NH_2-terminal peptides. This analysis of the small C-terminal fragments provided sequence data for α and β chains simultaneously (Fig. 3). Approximately 8 mg of globin were separated into α^E and ε chains by ion exchange chromatography on CM-cellulose (Fig. 4). Internal fragments of ε globin were prepared by cleavage of pure ε globin at arginine residues, and separation of the peptides by gel filtration under non-denaturing conditions (Fig. 5). The resulting fragments were characterized by sequenator analyses (Figs. 6-9). The remaining globin was digested with trypsin before separation on thin layer peptide maps. Tryptic peptide compositions were used to confirm residue assignments based on sequenator analysis (Table I).

NH_2-terminal analysis

Sequenator analysis of whole HbE globin showed that α-like and β-like chains were not blocked (Fig. 2). Forty-three residues of each chain were assigned. Amino acids 1-8 and 18-30 of the ε chain form tryptic peptides whose composition matches peptides T-1 and T-3 obtained from peptide mapping (Table I). Subtraction of the sequence for α^A (Matsuda et al., 1971) gives a single sequence for ε chain residues 9-17. Tryptic peptide T-2 from the ε chain was mixed with another peptide on the peptide map and was not useful in confirming the sequenator data. Assignments of ε chain residues 31-43 were confirmed by independent analysis of arginine cleavage fragment R-2, which was prepared from purified ε globin. Although the α-like chain of HbE has been reported to be α^A (Brown and Ingram, 1974), NH_2-terminal sequenator analysis of HbE globin reveals five amino acid differences between α^E and the published α^A sequence (Matsuda et al., 1971). These differences occur at positions 4, 74-76, and 38, and have been predicted by the DNA sequences of three recombinant cDNA clones (Salser et al., 1979; Deacon et al., 1980; Richards and Wells, 1980).

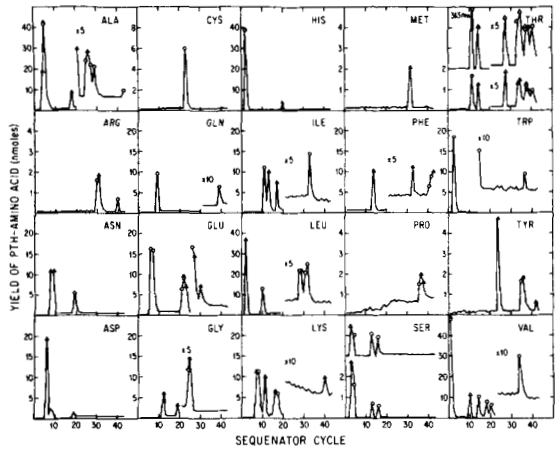


Figure 2. NH_2-terminal sequences of the α-like and β-like chains of HbE. Fifty nmol of CAM-HbE globin were degraded through 43 cycles at a repetitive yield of 93%. PTH-amino acids in each cycle were analyzed by HPLC. At each cycle, peak heights of each PTH amino acid were measured, nmol yields calculated from standard chromatograms, and the values plotted. The following symbols are used: (Δ) α^E chain and (o) ε chain. Residues were assigned to each chain as described in the text. Signals for PTH-serine and PTH-threonine derivatives at 313 nm absorbance are shown in upper plots. In order to show the analysis on a linear scale, nmol yields for later cycles are multiplied by the factors indicated.

Analysis of the aspartic acid-proline cleavage fragments of HbE.

Whole HbE globin was cleaved at the aspartic acid-proline bonds of ε and α^E chains by mild acid hydrolysis. The large NH_2-terminal and small C-terminal fragments were separated by gel filtration under non-denaturing conditions (data not shown). Sequenator analysis of the small fragments (Fig. 3) showed a 25% residual contamination with unblocked NH_2-terminal fragments. Thirty-seven residues of the α^E and ε chains were assigned. Amino acids 100-104, 105-120 and 121-132 of the ε chain were consistent with the compositions of tryptic peptides T-13, T-14 and T-15 respectively (Table I). Amino acids 121-136 for the ε chain also were confirmed by the sequence of arginine cleavage fragment R-6, which was obtained from purified ε globin (Figs. 6 and 7). Slow cleavage of the initial proline residue in these Asp-Pro cleavage fragments, combined with slow cleavage of α^E globin prolines at cycles 20 and 25 and an ε globin proline at cycle 25, caused sequenator "lag" to be greater than signal during later cycles. Thus the sequence may appear to be one residue out of phase.

Analysis of the small Asp-Pro cleavage fragment reveals seven amino acid differences in the α-like chain of HbE from the published α^A globin sequence. These occur at positions 109-111, 113, 116, 118 and 120. These differences cannot be assigned to ε globin, nor can they be explained as contamination by NH_2-terminal fragments. In the latter case, the NH_2-terminal sequence could not produce the observed amino acid residues where they occur. The presence of these changes, and the five differences found in the NH_2-terminal analysis, establishes that the α-like chain of HbE is different from α^A.

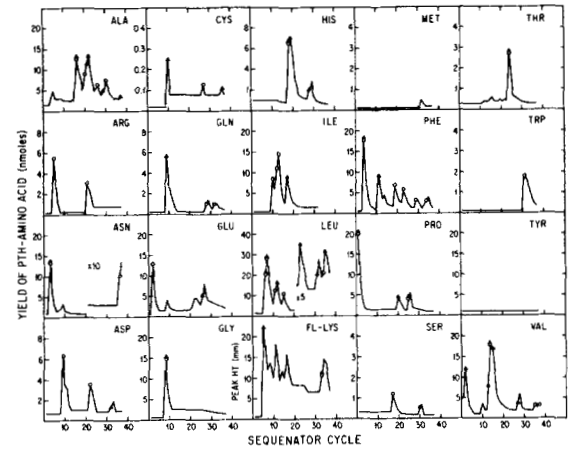


Figure 3. Sequenator analysis of the two C-terminal cleavage fragments of the α-like and β-like globins of HbE. CAM-globin prepared from HbE was hydrolyzed in 10% acetic acid, adjusted to pH 2.5 with pyridine, and 7M guanidine-HCl at 46° C for 4 days. Small (C-terminal) and large (NH_2-terminal) fragments were separated on Biogel P-10 in 0.5% formic acid. Thirty nmol of small fragments were treated with fluorescamine in the sequenator cup as described in Materials and Methods, then degraded through 37 cycles at 95% repetitive yield. The following symbols are used: (Δ) α^E chain and (o) ε chain. As discussed in Results, the slow cleavage of several proline residues led to a "lag" in the detection of these and subsequent residues, which caused serious problems in interpretation of data after the 37th cycle. Residues seen as small, unlabeled peaks are NH_2-terminal amino acids contributed by contaminating NH_2-terminal fragments.

Table I. Tryptic peptide compositions of the ε globin from HbE [a]

	T-1 1-8	T-3 18-30	T-5 41-59	T-7 62-65	T-9 67-76	T-10 77-82	T-12 88-95	T-13 96-104	T-14 105-120	T-15 121-132
Asx		1.1 (1)	2.2 (2)			2.9 (3)	1.1 (1)	2.3 (2)	1.1 (1)	0.9 (1)
Thr			0.8 (1)							0.8 (1)
Ser	1.1 (1)		2.8 (3)		2.0 (2)		1.1 (1)		1.0 (1)	
Glx	1.8 (2)	2.9 (3)			1.3 (1)		1.1 (1)	1.3 (1)		2.4 (2)
Pro			2.2 (2)					0.7 (1)		0.8 (1)
Gly		1.2 (1)	2.0 (2)	1.0 (1)	1.0 (1)				0.9 (1)	
Ala	1.3 (1)	2.9 (3)	1.7 (2)	1.0 (1)	0.9 (1)				2.0 (2)	2.3 (2)
Cys		b (1)					b (1)			b (1)
Val	0.7 (1)	1.7 (2)			2.0 (2)			1.3 (1)	1.0 (1)	
Met			0.5 (1)							
Ile			0.6 (1)			0.8 (1)			2.9 (3)	
Leu		0.9 (1)	1.2 (1)		0.9 (1)	1.3 (1)	1.5 (2)	0.7 (1)	4.3 (4)	
Phe			3.1 (3)		1.0 (1)			0.8 (1)	0.8 (1)	1.9 (2)
His	1.0 (1)			1.1 (1)			1.1 (1)	0.9 (1)	0.9 (1)	
Lys	1.0 (1)		0.8 (1)	0.9 (1)	0.7 (1)	1.2 (1)	1.1 (1)			0.8 (1)
Arg		1.0 (1)						0.8 (1)	1.1 (1)	
Trp	ND (1)									ND (1)
yield nmol	0.3	0.2	0.3	0.2	0.3	0.1	0.2	0.1	0.3	0.2
charge	0	-2	+1	+2	0	0	0	0	+1	0
NH_2-term. residues	Val 8	Val 13	Phe 19	Ala 4	Val 10	Asn 6	Leu 8	Leu 9	Leu 16	Asp 12

[a] Peptides are numbered from the NH_2 terminus. Values shown are calculated molar ratios of amino acids recovered from thin layer peptide maps. Numbers in parentheses are the expected values based on sequenator analysis of large fragments. Residues present at less than 0.3 mole per mole peptide are not shown.

[b] Cysteine derivatives were present on analyzer chromatograms but the amount of material was not determined.

Sequence of internal peptides of the ε chain

The α-like and β-like chains of HbE globin were separated by ion exchange chromatography on CM-cellulose with a formic acid gradient (Fig. 4, Moss and Hamilton, 1974). Purified ε globin was lyophilized, succinylated and digested with trypsin. The resulting arginine cleavage fragments were separated by gel filtration (Fig. 5). Fragments R3 and R6 were not resolved by this procedure, so they were analyzed as a mixture (Figs. 6 and 7). The first 16 residues of R6 (pos. 121-136) overlapped and confirmed the sequence of the Asp-Pro cleavage fragment discussed above. The first 12 residues of R6 agree with the composition of tryptic peptide T-15 (Table I). The last 7 residues of R6 were assigned by taking advantage of the increasing difference in PTH-amino acid yield from the two fragments (Fig. 7). Excellent quantitative data from HPLC allowed assignments to be made with a high degree of certainty.

Residues of the R3 fragment were identified as follows (see Fig. 1). The first three amino acids overlap the NH_2-terminal sequenator run. Ten residues (41-50) also were found as a minor sequence during the analysis shown in Fig. 9. Assignments for residues 41-59 correctly predict the composition of tryptic peptide T-5 (Table I). The Val-Arg dipeptide at position 60-61, indicated by sequenator data (Fig. 6), was not otherwise confirmed. A valine residue at position 60 is invariant in vertebrate β globins (Dayhoff, 1976). Arginine is found at position 61 in the adult chicken β globin (Matsuda et al., 1973), and is expected as the terminal residue of the arginine fragment.

Figure 8 shows a sequenator analysis of one of the two small arginine cleavage peaks. The major sequence is that of R7, the C-terminal tripeptide. Lys-Tyr-His is the typical C-terminal tripeptide of vertebrate β globins. In the arginine cleavage preparation, only this peptide lacks a terminal arginine, and accordingly, it must be the C-terminal fragment. After completion of the R7 peptide, sequenator analysis of the contaminating R2 and R3 fragments continued. The R2 sequence (positions 31-40) was found to be the same in the ε chain as in the adult β globin (Matsuda et al., 1973) and the major early embryonic p chain (Chapman et al., 1981). The R2 sequence assignments also agreed with the NH_2-terminal sequence analysis of residues 31-40. The R3 sequence in this mixture matched the R3 sequence constructed earlier by subtraction of the R6 sequence from the mixture of R3 and R6.

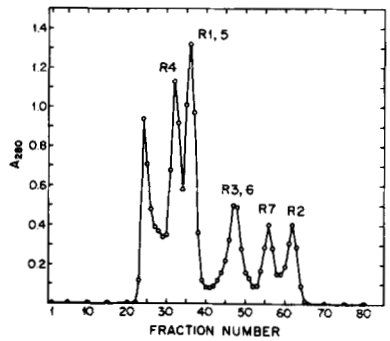


Figure 5. Separation of ε chain arginine cleavage fragments. Two mg of ε globin were succinylated and digested with trypsin. The digest was loaded directly onto Biogel P-10 (1.6 x 72 cm) equilibrated with 0.1 M ammonium bicarbonate. Fractions of 2.2 ml were collected at 8.8 ml/hr. Four fractions were pooled from each peak. Fragments were identified by amino acid composition and sequenator analysis.

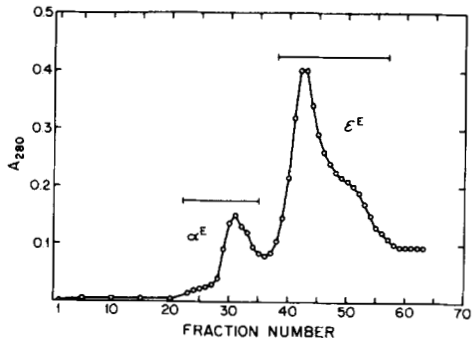


Figure 4. Separation of the α^E and ε chains of HbE by ion exchange chromatography on CM-cellulose. Eight mg of CAM-globin were dissolved in 0.2 M formic acid and applied to a 1.8x10 cm column of Whatman CM-52 equilibrated with the same buffer. Chains were eluted at 45 ml/hr in 5 ml fractions using a gradient of 0.2 M formic acid and 1.8 M formic acid brought to pH 1.9 with pyridine. Fractions were pooled as indicated by the bars and lyophilized.

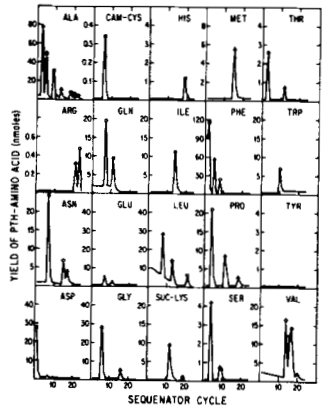


Figure 6. Sequenator analysis of ε chain arginine cleavage fragments R3 and R6. One hundred nmol of a mixture of R3 and R6 fragments were degraded through 25 cycles at a repetitive yield of 84 and 89% respectively (see Fig. 7). Residues were assigned to the appropriate chain as described in the text. The following symbols are used: (○) R6 and (△) R3.

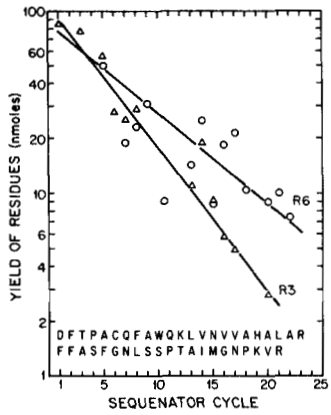
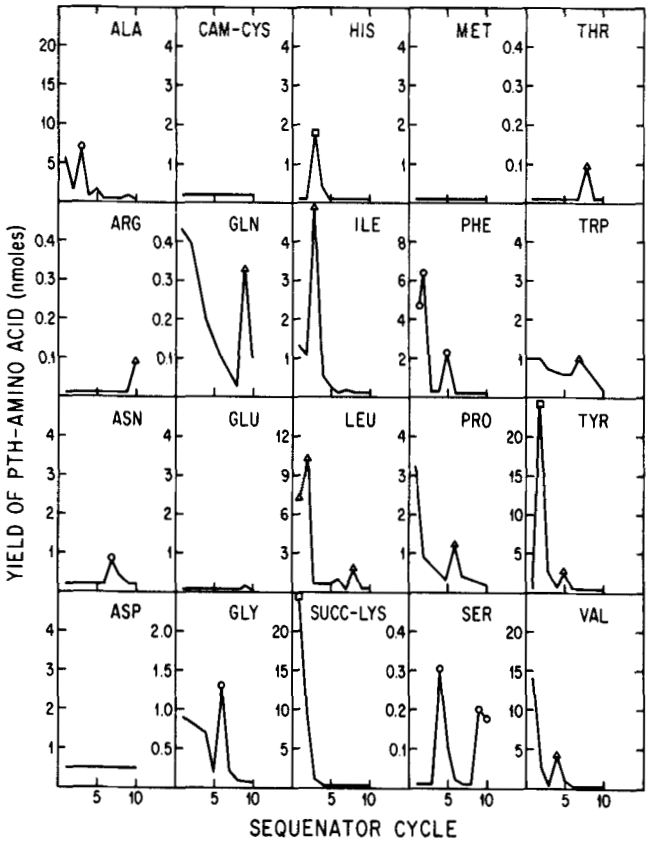


Figure 7. Repetitive yield analysis of the R3 and R6 arginine fragments from the ϕ chain. Quantitative data from HPLC for each identified residue of R3 and R6 are plotted on a semilog scale. Repetitive yields were calculated by linear regression analysis. The following symbols are used: (○) R6, 89% and (△) R3, 84%. The difference in repetitive yields is probably due to the difference in length of the two peptides. Sequences are shown below.

Sequenator analysis of the R4 fragment is shown in Fig. 9. This 43 residue peptide was uncontaminated with other fragments and allowed unambiguous assignment of residues 62-104. The sequence is consistent with the compositions of tryptic peptides T-7, T-9, T-10, T-12 and T-13 (Table I). It overlaps by five residues the Asp-Pro cleavage fragment. Although not overlapped to R3, which ends at arg 61, the ϕ chain R4 fragment is 88% homologous to its counterpart in the adult β sequence (Matsuda et al., 1973). A five residue tryptic peptide covering positions 83-87 is identical with a peptide in the p globin (Chapman et al., 1981). We did not obtain overlapping peptides for positions 61 and 62 or for positions 143 and 144. However, the complete structure is consistent with data on ϕ globin molecular weight, amino acid composition, and peptide map (data not shown).



R2 LEU-LEU-ILE-VAL-TYR-PRO-TRP-THR-GLN-ARG
R3 PHE-PHE-ALA-SER-PHE-GLY-GLN-LEU-SER-SER-
R7 LYS-TYR-HIS

Figure 8. Sequenator analysis of the R7 fragment from the ϕ chain. Fifty nmol of arginine cleavage fragments (Fig. 5) were degraded through 11 cycles. At the first cycle, approximately 30 nmol of R7, 10 nmol of R2, and 5 nmol of R3 were recovered. Residues were assigned to the various fragments as described in the text. Repetitive yields for R2 (△) and R3 (○) were 65%. Repetitive yield for R7 (□) was very poor because of its short length and composition. Although the R7 tripeptide was the major peptide in this preparation and was of primary interest, this sample was analyzed through 10 cycles to obtain useful information about the contaminating R2 and R3 fragments. Other contaminants, seen most notably in the Gln, Trp, Pro, Gly, and Val panels, did not obscure residues of R7, R2, and R3. Sequences are shown below.

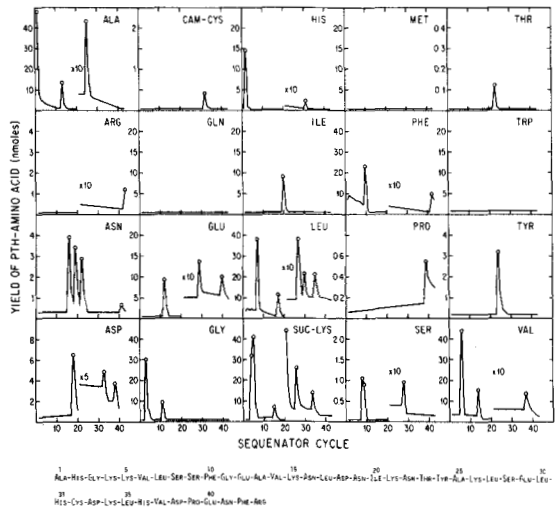


Figure 9. Sequenator analysis of the R4 fragment from the ϕ chain. Sixty nmol of this fragment were degraded through 44 cycles at 89% repetitive yield. There were no contaminating peptides, nor was there appreciable background as indicated by the plots. The sequence is shown below.