Cloning and structural characterization of a human non-erythroid band 3-like protein

Donald R. Demuth, Louise C. Showe, Marion Ballantine, Antonio Palumbo¹, Peter J. Fraser, Livia Cioe², Giovanni Rovera and Peter J. Curtis

The Wistar Institute of Anatomy and Biology, 36th Street at Spruce, Philadelphia, PA 19104, USA

¹Present address: Institute of Anatomy and Histology, School of Medicine, University of Torino, Torino, Italy

²Present address: Institute of Virology, University of Rome, 00185 Rome, Italy

Communicated by C. Weissmann

Polypeptides which are immunologically related to the erythrocyte anion transport protein have been identified in a variety of non-erythroid cells. We describe two cDNA clones encoding a human non-erythroid band 3 protein (HKB3) and the mouse erythrocyte band 3 (MEB3) and show that these proteins are structurally similar. Comparison of the predicted amino acid sequences from HKB3 and MEB3 reveals a high degree of sequence homology (71%) and conservation of the overall topography of the transmembrane domain. Similar levels of homology are also observed in comparisons with published amino acid sequence from the human erythrocyte band 3. In addition, specific residues which have been demonstrated to be involved in erythroid anion transport are conserved in HKB3, suggesting that this non-erythroid band 3 protein functions in this respect. Although protein sequence homology within the cytoplasmic domain is considerably lower (35%), three specific regions in HKB3 are conserved, one of which may represent an ankyrin binding site. Northern blot analysis reveals transcripts that cross-hybridize with the HKB3 cDNA in a variety of non-erythroid cell lines but not in cells of erythroid lineage.

Key words: anion transport/band 3/cDNA/non-erythroid

Introduction

The erythrocyte cytoskeleton confers and maintains the structural integrity of the mature erythrocyte and is composed of a variety of proteins including α - and β -spectrin, ankyrin and band 4.1. Proteins similar to those comprising the erythrocyte cytoskeleton exist in non-erythroid tissues (Aster *et al.*, 1984; Bennett *et al.*, 1982; Drenckhahn *et al.*, 1984; Kay *et al.*, 1984; Nelson and Lazarides, 1984), suggesting that many cell types may share common membrane skeletal features.

The erythrocyte band 3 is multifunctional, providing an anchor for the membrane cytoskeletal structure through its interaction with ankyrin (Bennett and Stenbuck, 1980), as well as catalyzing the $\text{Cl}^-/\text{HCO}_3^-$ anion transport system (Jennings, 1985; Knauf, 1979; Steck, 1974; Weith *et al.*, 1982). Characterization of this protein both by chemical modification and by proteolytic cleavage has shown these functions are carried out by distinct structural domains. The N-terminal 400 amino acid residues form the hydrophilic cytoplasmic domain which binds to ankyrin which in turn binds spectrin, thus anchoring

the cytoskeleton to the membrane (Bennett and Stenbuck, 1980). In addition, the major antigenic determinant of band 3 has also been mapped to this domain and appears to be in close proximity to the ankyrin binding site since antibodies raised against band 3 compete with labeled ankyrin for binding (Bennett and Stenbuck, 1980). Human erythrocyte band 3 also binds hemoglobin (Sayre and Fikiet, 1981), glycerol phosphate dehydrogenase (Yu and Steck, 1975) and aldolase (Strapazon and Steck, 1976). These interactions are mediated by the extreme N terminus of the protein (Kaul *et al.*, 1983).

The anion transport activity of the erythrocyte band 3 resides solely within the membrane domain which consists of ~500 amino acids at the C terminus of the protein. This activity functions independently since removal of the cytoplasmic domain by proteolytic cleavage with chymotrypsin results in little or no loss of transport activity (Grinstein *et al.*, 1979). Studies of the topography of the membrane domain of band 3 suggest that the polypeptide chain spans the plasma membrane several times (Jen-

Fig. 1. (A) Restriction maps of cDNAs encoding the murine erythrocyte band 3. Clones pLC5 and pLC14 were isolated by immunological screening of a cDNA expression library in the pUC system (Vieira and Messing, 1982). Clone MEB3-18 was isolated from a cDNA library in pBR322 derived from mouse anemic spleen probing with pLC5 and pLC14. Clone pMEB3-18 was digested with *Pst*I and subcloned into M13mp18 and mp19 (Norrander *et al.*, 1983) for sequencing. (B) Human cDNA clones isolated from a library in pBR322 derived from K562(S) cells using pMEB3-18 as probe. Fragments derived from clone pHKB3 were subsequently subcloned into M13mp18 or mp19 (Norrander *et al.*, 1983) for sequencing. The double-headed arrows denote fragments which were sequenced after progressive digestion with λ exonuclease III (Henikoff, 1984). Restriction endonuclease abbreviations are as follows: B, *Bam*HI; Hp, *Hpa*II; P, *Pst*I; Pv, *Pvu*II; S, *Sac*I; X, *Xba*I.

60
 CTG GAG ATC CCA GAT CGA GAC AGC GAA GAA GAA CTG GAG AAC ATA ATA GGA CAG ATA GCA TAT AGA GAC CTA ACC ATC CCT GTG ACC GAG ATG CAG GAC CCA GAA GCT CTT CCC ACA GAG
 leu Glu Ile Pro Asp Arg Asp Ser Glu Glu Glu Leu Glu Asn Ile Ile Gly Gln Ile Ala Tyr Arg Asp Arg Ser Glu Ile Pro Val Thr Glu MET Gln Asp Pro Glu Ala Leu Pro Thr Glu
 11 31 120

180
 CAA ACA GCC ACA GAC TAC GTC CCC AGC AGT ACC TCC ACA CCG CAC CCA AGC TCC GGT CAG GTC TAT GTG GAG CTT CAG GAA CTG ATG ATG GAC CAG AGG AAC CAG GAA CTA CAG TGG GTG
 Gln Thr Ala Thr Asp Tyr Val Pro Ser Ser Thr Ser Thr Pro His Pro Ser Ser Gly Gln Val Tyr Val Glu Leu Gln Glu Leu MET MET Asp Gln Arg Asn Gln Glu Leu Gln Trp Val
 51 71 240

300
 GAG GCA GCG CAC TGG ATA GGG CTG GAG GAA AAC CTT CGA GAG GAT GGT GTA TGG GGT CGC CCA CAT CTA TCT TAC CTG ACC TTC TGG AGC CTT CTA GAA CTG CAG AAG GTC TTC TCA AAA
 Glu Ala Ala His Trp Ile Gly Leu Glu Glu Asn Leu Arg Glu Asp Gly Val Trp Gly Arg Pro His Leu Ser Tyr Leu Thr Phe Trp Ser Leu Leu Glu Leu Gln Lys Val Phe Ser Lys
 91 111 360

420
 GGC ACC TTC CTT CTG GGT CTG GCA GAG ACA TCC CTG GCT GGG GTG GCC AAT CAC CTT CTA GAC TGC TTC ATC TAC GAG GAT CAG ATT CGG CCT CAG GAC CGA GAA GAA CTG CTC CGG GCT
 Gly Thr Phe Leu Leu Gly Leu Ala Glu Thr Ser Leu Ala Gly Val Ala Asn His Leu Leu Asp Cys Phe Ile Tyr Glu Asp Gln Ile Arg Pro Gln Asp Arg Glu Glu Leu Leu Arg Ala
 131 151 480

540
 CTG CTA CTC AAA CGC AGC CAT GCT GAG GAC CTA GGG AAT CTG GAG GGG GTG AAG CCC GCC GTC CTG ACC CGC TCT GGG GGC GCC TCT GAG CCT CTG CTT CCT CAC CAA CCA TCG CTG GAG
 Leu Leu Leu Lys Arg Ser His Ala Glu Asp Leu Gly Asn Leu Glu Gly Val Lys Pro Ala Val Leu Thr Arg Ser Gly Gly Ala Ser Glu Pro Leu Leu Pro His Gln Pro Ser Leu Glu
 171 191 600

660
 ACC CAG CTG TAC TGC GGA CAG GCA GAG GGG GGC TCA GAA GGA CCC TCA ACA TCT GGG ACT CTG AAG ATC CCC CCA GAT TCA GAA ACC ACA CTG GTG CTA GTG GGC CGG GCT AAT TTT CTG
 Thr Gln Leu Tyr Cys Gly Gln Ala Glu Gly Gly Ser Glu Gly Pro Ser Thr Ser Gly Thr Leu Lys Ile Pro Pro Asp Ser Glu Thr Thr Leu Val Leu Val Gly Arg Ala Asn Phe Leu
 211 231 720

780
 GAG AAG CCT GTA CTG GGC TTC GTG AGA CTG AAG GAG GCC GTG CCT CTG GAG GAC CTG GTG TTG CCA GAG CCT GTG GGC TTC CTT CTT GTT CTG CTG GGA CTT GAG GCT CCG CAT GTC GAC
 Glu Lys Pro Val Leu Gly Phe Val Arg Leu Lys Glu Ala Val Pro Leu Glu Asp Leu Val Leu Pro Glu Pro Val Gly Phe Leu Leu Val Leu Leu Gly Pro Glu Ala Pro His Val Asp
 251 271 840

900
 TAT ACC CAG CTG GGC AGG GCT GCA GCA ACT CTC ATG ACA GAA AGA GTG TTC CGC ATT ACC GCC TCC ATG GCG CAT AAC CGA GAG GAG CTG CTG CGC TCT CTG GAG AGC TTT CTG GAC TGT
 Tyr Thr Gln Leu Gly Arg Ala Ala Ala Thr Leu MET Thr Glu Arg Val Phe Arg Ile Thr Ala Ser MET Ala His Asn Arg Glu Glu Leu Leu Arg Ser Leu Glu Ser Phe Leu Asp Cys
 291 311 960

1020
 AGC CTG GTG CTG CCG CCC ACA GAC GCC CCT TCA GAA AAA GCT CTG CTT AAC CTG GTA CCG GTG CAG AAG GAG CTG CTT AGG AGG CGG TAC CTG CCC AGC CCC CGC AAG CCA GAC CCC AAC
 Ser Leu Val Leu Pro Pro Thr Asp Ala Pro Ser Glu Lys Ala Leu Leu Asn Leu Val Pro Val Gln Lys Glu Leu Leu Arg Arg Arg Tyr Leu Pro Ser Pro Ala Lys Pro Asp Pro Asn
 331 351 1080

1140
 TTG TAC AAC ACT CTA GAC TTG AAC GGG GGA AAA GGG GGC CCT GGT GAC GAA GAC GAC CCT CTA CGG CGC ACA GGC CGG ATC TTT GGA GGC CTG ATC CGT GAT ATC CGG CGC CGT TAT CCT
 Leu Tyr Asn Thr Leu Asp Leu Asn Gly Gly Lys Gly Gly Pro Gly Asp Glu Asp Asp Pro Ala Arg Arg Thr Gly Arg Ile Phe Gly Gly Leu Ile Arg Asp Ile Arg Arg Arg Tyr Pro
 371 391 1200

1260
 TAC TAC CTG AGT GAC ATC ACC GAT GCG CTC AGT CCC CAG GTG TTG GCT GCT GTC ATC TTC ATC TAC TTT GCT GCC TTG TCA CCT GCC GTC ACC TTC GGC GGC CTC CTG GGA GAA AAG ACC
 Tyr Tyr Leu Ser Asp Ile Thr Asp Ala Leu Ser Pro Gln Val Leu Ala Ala Val Ile Phe Ile Tyr Phe Ala Ala Leu Ser Pro Ala Val Thr Phe Gly Gly Leu Leu Gly Glu Lys Thr
 411 431 1320

1380
 CGG AAC TTG ATG GGG GTG TCG GAG CTG CTC ATC TCC ACA GCA GTG CAA AGC ATT CTC TTT GCT CTC CTG GGG GCA CAG CCC CTG CTT GTG CTA GGC TTA CCA GGA CCC CTG CTG GTG TTT
 Arg Asn Leu MET Gly Val Ser Glu Leu Leu Ile Ser Thr Ala Val Val Gln Ser Ile Leu Phe Ala Leu Leu Gly Ala Gln Pro Leu Leu Val Leu Gly Phe Ser Gly Pro Leu Leu Val Phe
 451 471 1440

1500
 GAA GAA GCC TTC TTC TCG TTC TGT GAA AGC AAT AAC CTG GAG TAT ATC GTG GGC CGC GCG TGG ATC GGC TTC TGG CTC ATC CTG CTC GTG ATG TTG GTG GTG GCC TTT GAA GGC AGC TTC
 Glu Glu Ala Phe Phe Ser Phe Cys Glu Ser Asn Asn Leu Glu Tyr Ile Val Gly Arg Ala Trp Ile Gly Phe Trp Leu Ile Leu Leu Val MET Leu Val Val Ala Phe Glu Gly Ser Phe
 491 511 1560

1620
 CTC GTC CAA TAC ATC TCC CGA TAC ACC CAG GAG ATC TTC TCC TTC CTC ATC TCC CTC ATC TTC ATC TAT GAG ACT TTC TCC AAG CTG ATC AAG ATT TTC CAG GAC TAC CCA CTA CAA CAG
 Leu Val Gln Tyr Ile Ser Arg Tyr Thr Gln Glu Ile Phe Ser Phe Leu Ile Ser Leu Ile Phe Ile Tyr Glu Thr Phe Ser Lys Leu Ile Lys Ile Phe Gln Asp Tyr Pro Leu Gln Gln
 531 551 1680

1740
 ACT TAT GCC CCT GTT GTG ATG AAG CCC AAA CCT CAG GGC CCT GTG CCC AAC ACA GCC CTC TTC TCT CTT GTG CTC ATG GCT GGC ACC TTC CTA CTT GCC ATG ACG CTA CGA AAG TTC AAG
 Thr Tyr Ala Pro Val Val MET Lys Pro Lys Pro Gln Gly Pro Val Pro Asn Thr Ala Leu Phe Ser Leu Val Leu MET Ala Gly Thr Phe Leu Leu Ala MET Thr Leu Arg Lys Phe Lys
 571 591 1800

1860
 AAC AGC ACC TAC TTC CCT GGC AAG CTG CGT AGA GTC ATT GGG GAC TTC GGG GTC CCC ATC TCC ATC CTG ATA ATG GTC CTG GTG GAT TCC TTC ATC AAG GGC ACC TAC ACA CAG AAA CTC
 Asn Ser Thr Tyr Phe Pro Gly Lys Leu Arg Arg Val Ile Gly Asp Phe Gly Val Pro Ile Ser Ile Leu Ile MET Val Leu Val Asp Ser Phe Ile Lys Gly Thr Tyr Thr Gln Lys Leu
 611 631 1920

1980
 TCC GTG CCT GAT GGC CTC AAA GTG TCC AAC TCC TCC GCC CGG GGC TGG GTC ATC CAC CCT CTG GGG CTG TAT AGA CTT TTC CCC ACG TGG ATG ATG TTT GCC TCC GTT CTG CCC GCC CTG
 Ser Val Pro Asp Gly Leu Lys Val Ser Asn Ser Ser Ala Arg Gly Trp Val Ile His Pro Leu Gly Leu Tyr Arg Leu Phe Pro Thr Trp MET MET Phe Ala Ser Val Leu Pro Ala Leu
 651 671 2040

2100
 CTT GTC TTC ATT CTC ATT TTC CTT GAG TCT CAG ATC ACC ACG CTG ATT GTC AGT AAA CCA GAA CCG AAG ATG ATC AAG GGC TCT GGC TTC CAC CTG GAC CTG TTG GTC GTC GTT GGC ATG
 Leu Val Phe Ile Leu Ile Phe Leu Glu Ser Gln Ile Thr Thr Leu Ile Val Ser Lys Pro Glu Arg Lys MET Ile Lys Gly Ser Gly Phe His Leu Asp Leu Leu Leu Val Val Gly MET
 691 711 2160

2220
 GGT GGG GTG GCT GCC CTC TTT GGG ATG CCT TGG CTC AGT GCC ACC ACT GTA AGA TCT GTT ACC CAC GCC AAT GCC CTC ACA GTC ATG GGG AAG GCC AGT GGT CCA GGG GCT GCA GCC CAG
 Gly Gly Val Ala Ala Leu Phe Gly MET Pro Trp Leu Ser Ala Thr Thr Val Arg Ser Val Thr His Ala Asn Ala Leu Thr Val MET Gly Lys Ala Ser Gly Pro Gly Ala Ala Ala Gln
 731 751 2280

2340
 ATC CAA GAG GTC AAG GAA CAG CAG ATC AGT GGG CTC CTG GTC TCT GTG CTT GTG GGA CTG TCC ATC CTC ATG GAA CCT ATC CTG TCC CGC ATC CCC CTG GCT GTG CTC TTC GGC ATC TTC
 Ile Gln Glu Val Lys Glu Gln Arg Ile Ser Gly Leu Leu Val Ser Val Leu Val Gly Leu Ser Ile Leu MET Glu Pro Ile Leu Ser Arg Ile Pro Leu Ala Val Leu Phe Gly Ile Phe
 771 791 2400

2460
 CTG TAC ATG GGA GTC ACA TCC CTC AGT GGC ATC CAG CTC TTT GAC CGC ATC TTG CTC CTG TTC AAC CCA CCC AAG TAC CAC CCG GAT GTG CCC TTT GTC AAG AGG GTG AAG ACC TGG CGC
 Leu Tyr MET Gly Val Thr Ser Leu Ser Gly Ile Gln Leu Phe Asp Arg Ile Leu Leu Leu Phe Lys Pro Pro Lys Tyr His Pro Asp Val Pro Phe Val Lys Arg Val Lys Thr Trp Arg
 811 831 2520

2580
 ATG CAC CTC TTC ACG GGC ATC CAG ATC ATC TGC CTG GCT GTG CTG TGG GTG GTG AAG TCC ACT CCT GCC TCG CTG GCC TTG CCC TTC GTC CTC ATC CTC ACA GTG CCT CTC CGT CGT CTC
 MET His Leu Phe Thr Gly Ile Gln Ile Ile Cys Leu Ala Val Leu Trp Val Val Lys Ser Thr Pro Ala Ser Leu Ala Leu Pro Phe Val Leu Ile Leu Thr Val Pro Leu Arg Arg Leu
 851 871 2640

2700
 ATC CTC CCG CTC ATC TTC AGA GAG CTG GAA CTC CAG TGT CTG GAC GGT GAT GAT GCC AAA GTG ACC TTT GAC GAG GAG AAT GGC CTG GAT GAA TAT GAC GAA GTG CCC ATG CCT GTG TGA
 Ile Leu Pro Leu Ile Phe Arg Glu Leu Glu Leu Gln Cys Leu Asp Gly Asp Ala Lys Val Thr Phe Asp Glu Glu Asn Gly Leu Asp Glu Tyr Asp Glu Val Pro MET Pro Val
 891 911 2760

2820
 GCG GCA GGC CCA GGC CCC AAC CAC TCC ACT TTC CCA CCC TCC CAG GAA GAC CAC ACA TGG GCG TGT TAT GGC TGT GGG TGG GGA AGG AGG GGT GTT CTA GGA AAC CCT CTA TCT AAG GTA
 931 951 2880

2940
 CAC TGG ATT TTC TGG CTG GAG ACG ACC GGA AGC TAT ATC AGG GAA TTG GGG TCA CAG TCT CTT GCA ACT GTA TGG GCT GAG GCT GGA AAA
 971 991 2940

nings and Nicklish, 1984; Jennings, 1985), however, recent analysis of the complete nucleotide sequence of a cDNA encoding the murine erythrocyte band 3 reveals a structure which is more complex, traversing the membrane 12 or 13 times (Kopito and Lodish, 1985).

Proteins immunologically related to the erythrocyte band 3 have been detected in various non-erythroid tissues (Drenckhahn *et al.*, 1984; Kay *et al.*, 1984) and physiological studies have demonstrated that $\text{Cl}^-/\text{HCO}_3^-$ transport systems exist in non-erythroid cells (Frizzel *et al.*, 1976; Leslie *et al.*, 1973). Two potent inhibitors of anion transport in erythrocytes, 4,4'-di-isothiocyanodihydrostilbene-2,2'-disulfonate (H_2DIDS) and N-(4-azido-2-nitrophenyl)-2-aminoethyl sulfonate (NAP-taurine) (Jennings, 1984) have been shown to inhibit anion transport in rat liver (Cheng and Levy, 1980) and peptides immunologically related to the erythrocyte band 3 have been detected in this tissue (Kay *et al.*, 1984). Peptide mapping of a band 3-like protein from human peripheral blood lymphocytes indicates this protein is similar but not identical to the erythrocyte band 3 (Kay *et al.*, 1984). However, while involvement in anion transport has been suggested for these non-erythroid proteins, it remains to be demonstrated as to whether they also function in the anchoring of cytoskeletal structures to the membrane.

We report here the first definitive proof for a band 3 gene family in the characterization of a cloned cDNA (pHKB3) representing a human non-erythroid band 3 protein. Comparison of the predicted amino acid sequence from HKB3 with a cDNA encoding the mouse erythrocyte band 3 (MEB3) and with the known amino acid sequence from the human erythrocyte band 3 (HEB3) shows substantial homology within the membrane domains of these proteins and a similarity in the predicted topography of this domain in HKB3 and MEB3. Amino acid residues in the human erythrocyte band 3 which may play a role in anion transport are conserved in HKB3, suggesting that this non-erythroid band 3 could function in the transport process. While the overall amino acid homology between the cytoplasmic domains is comparatively low, specific regions are strictly conserved in all three proteins suggesting that some functions of the cytoplasmic domain of the erythroid band 3 are also conserved in the non-erythroid form. The location of one of these conserved regions is consistent with the tentative location assigned to the ankyrin binding site in HEB3 (Low *et al.*, 1984).

Results

Isolation and sequencing of cDNA clones representing MEB3

A cDNA expression library was prepared by transferring inserts from an amplified cDNA library in pBR322 derived from mouse anemic spleen to the expression vectors pUC7, pUC8 and pUC9 (Cioe and Curtis, 1985). Plasmids were then introduced into *Escherichia coli* DH-1. Screening of these expression libraries with polyclonal antibodies raised against mouse erythrocyte membrane proteins yielded 17 positive cDNA clones, two of which also reacted with a polyclonal antibody against purified MEB3. These plasmids contained inserts of ~500 bp and ~1400 bp and were designated pLC5 and pLC14, respectively. Both inserts hybridized to a 4.4-kb mRNA in Northern blots of mouse anemic spleen RNA. Initially, pLC5 was sequenced and the predicted amino acid sequence from this fragment showed 70% homology

to the published N-terminal amino acid sequence of HEB3 (Kaul *et al.*, 1983). Inserts from pLC5 and pLC14 were subsequently used to screen a second mouse anemic spleen cDNA library in pBR322 (Curtis *et al.*, 1983), resulting in the identification and isolation of a third clone, pMEB3-18, containing an insert of 2969 bp (Figure 1A).

The compiled nucleotide and predicted amino acid sequences from the cDNA clones, pLC5, pLC14 and pMEB3-18 are shown in Figure 2. A single open reading frame consisting of 919 codons extends from nucleotide 1 to 2757. The predicted amino acid sequence shows substantial homology to several sequenced peptide fragments from HEB3. The presented nucleotide and predicted amino acid sequences are identical to the sequence recently determined by Kopito and Lodish (1985) for the murine erythrocyte band 3, although it lacks the sequence encoding the first 10 amino acids of the protein, and ~1300 bp of the 3'-untranslated region.

Isolation and sequencing of cDNA clones representing a human band 3-like protein

Southern blot analysis of human DNA with the pMEB3-18 probe indicated a sufficient degree of homology to allow screening of a human cDNA library. Clone pMEB3-18 was then used to screen a cDNA library prepared from the K562(S) cell line (Watt *et al.*, 1983) from which we have previously isolated a human α -spectrin cDNA (Curtis *et al.*, 1985). Screening of this library under non-stringent conditions ($6 \times \text{SSC}$, $5 \times \text{Denhardt's}$ solution) yielded six cDNA clones which cross-hybridized with pMEB3-18. These clones possessed common restriction fragments which cross-hybridized with the mouse band 3 cDNA indicating that they arose from a common mRNA (Figure 1B). The longest of these, pHKB3 containing an insert of 2760 bp, was subcloned into M13 for nucleotide sequencing. Northern blot analysis of K562(S) RNA, using labeled pHKB3 as probe indicates the corresponding mRNA is 4.4 kb in length, similar to the murine erythrocyte band 3 mRNA.

The entire nucleotide and predicted amino acid sequence of pHKB3 is shown in Figure 3. A single open reading frame extends for 2595 nucleotides from the 5' end, coding for 865 amino acids. Primer extension on poly(A)⁺ RNA from K562(S) cells indicates that ~500 nucleotides from the 5' end of the message are missing (data not shown) and based on the size of the mRNA, ~1100 nucleotides are also missing from the 3'-untranslated region. This long 3'-untranslated region would be consistent with the structure of the MEB3 message as described by Kopito and Lodish (1985). If the structure of the 5' end is also similar to MEB3, approximately the first 100 amino acids are missing from this sequence. The amino acid sequence predicted from the pHKB3 sequence shows that this clone encodes a protein containing two distinct structural domains. Comparison of the amino acid sequences from the HKB3 and MEB3 cDNAs, indicates substantial homology within the C-terminal half of the protein but much less in the N-terminal region. Further comparison (see below) of these sequences with available protein sequence from HEB3 shows that HKB3, while closely related to the erythroid band 3, clearly represents a different protein.

Comparison of the overall structure of HKB3 and MEB3

Hydrophobicity plots derived from the predicted amino acid sequences of HKB3 and MEB3 are shown in Figure 4. The plots

Fig. 2. Composite nucleotide and predicted amino acid sequences of cDNAs representing the MEB3. The nucleotide sequence is numbered above each line with position 1 being the 5' end of the cDNA. Amino acids are numbered below each line according to the complete sequence of Kopito and Lodish (1985). The cDNAs were sequenced by the chain termination method of Sanger *et al.* (1977) or according to the procedure of Maxam and Gilbert (1977). The entire sequence was obtained for both strands.

Fig. 3. Composite nucleotide and predicted amino acid sequences of pHKB3. The nucleotide sequence is numbered above each line with position 1 representing the 5' end of pHKB3. Amino acids are numbered below each line. Residue 1 corresponds to amino acid 123 of the MEB3 and to amino acid 110 from the HEB3 sequence (Kaul *et al.*, 1983). Sequencing was carried out using the chain termination method (Sanger *et al.*, 1977) or by chemical cleavage (Maxam and Gilbert, 1977). The entire sequence was obtained for both strands.

Fig. 4. Hydrophobicity plots for (A) MEB3 and (B) HKB3 amino acid sequences. Plots were made using an algorithm similar to that of Kyte and Doolittle (1982) with hydrophobicity values being averaged over seven amino acids. Hydrophobic regions in both sequences are designated a–j. Z represents the region in HKB3 containing a 29-amino acid insertion (see text).

are aligned from the C terminus and the amino acids numbered according to Kopito and Lodish (1985). The hydrophobicity profiles within the membrane-spanning domains of these proteins (residues 420–929) are remarkably similar. Both proteins contain 10 hydrophobic regions (a–j in MEB3; a'–j' in HKB3) representing membrane-spanning segments. Hydrophobic regions a', c', e', g' and h' of HKB3 predominantly contain hydrophobic amino acids and comply with existing criteria for producing membrane-spanning α -helices (Eisenberg, 1984) and the remain-

ing spanning segments contain both hydrophilic and hydrophobic residues and are predicted to exist as amphipathic helices. Regions b' and i' are each predicted to contain two amphipathic helices, consistent with the previously reported structure of MEB3 (Kopito and Lodish, 1985). It is not clear whether hydrophobic region j' spans the membrane once or twice and although j' shows a greater overall hydrophobic character than j (see Figure 4), both j and j' contain proline residues (Pro 878 and 886) which could prevent the formation of a second spanning helix. Therefore, the

Fig. 5. Comparison of the amino acid sequences within the membrane domains of HKB3, MEB3 and HEB3. Sequences were aligned using the PRTALN program of Wilbur and Lipman (1983). Numbering of the sequences is consistent with that used by Kopito and Lodish (1985) for the mouse erythrocyte band 3. For purposes of comparison the insertion present in the HKB3 sequence is not numbered. Potential glycosylation sites in both sequences are boxed and the arrow denotes the probable origin of the first transmembrane helix. A portion of the cytoplasmic domain of HKB3 (residues 376–422) is included to show that homology extends 25 amino acids into this domain before diverging considerably.

basic structure of the transmembrane domains of the HKB3 protein and MEB3 are highly conserved. The major difference observed is the presence of an extended hydrophilic loop in HKB3 (region Z in Figure 4).

A high degree of conservation is also seen at the level of specific residues contained within the membrane-spanning structure. Ten of the transmembrane helices of HKB3 are bounded

with arginine residues positioned at the internal or external surfaces of the membrane. These residues might function to stabilize the spanning helices by interaction with the negatively charged phospholipid head groups. Ten of these 13 arginine residues are conserved in analogous positions in MEB3. Also, five of the predicted membrane-spanning helices in both HKB3 and MEB3 contain a proline residue and may exist as bent helices similar to the spanning structures proposed for alamethicin (Fox and Richards, 1982).

Comparison of the amino acid sequences of pHKB3 and MEB3

Membrane domain. The amino acid sequences derived from the membrane domains of pHKB3 and MEB3 are shown in Figure 5. For ease of comparison, the sequences are numbered as described by Kopito and Lodish (1985) for MEB3. Therefore the single insertion in HKB3 is not numbered. Within this domain (amino acids 420–929) these sequences are 71% homologous. Similar levels of homology are also observed when HKB3 is compared with published sequences from peptide fragments generated from the membrane domain of the HEB3 (Brock *et al.*, 1983; Mawby and Finlay, 1982). A single insertion of 29 amino acids at position 582 in the HKB3 sequence was introduced to obtain optimal alignment with the mouse band 3. Based on the topological model of HKB3 this inserted region (Z in Figure 4) exists as an external loop between two membrane-spanning regions of the protein. Three potential glycosylation sites, shown boxed in Figure 5, are present in this inserted region, while the two glycosylation sites identified in the murine erythrocyte band 3 (Kopito and Lodish, 1985) are not conserved in the HKB3 sequence.

Chemical modification of amino acid residues has aided in determining the topography of the membrane domain of HEB3 and in identifying residues which may be important in the transport process. Three external lysine residues present in a 17-kD chymotrypsin fragment from human band 3 have been shown to be sensitive to chemical modification (Jennings, 1985), and one of these appears to be the main residue modified by H_2DIDS , a potent inhibitor of anion transport (Jennings, 1984). Sites sensitive to H_2DIDS have been suggested to be high affinity chloride binding sites in the human erythrocyte band 3 (Falke *et al.*, 1983). These lysines have been identified (Lys 449, 558 and 561) in MEB3 (Kopito and Lodish, 1985) and are conserved in the non-erythroid HKB3 sequence. Also, an extracellular tyrosine residue (Tyr 646) which is susceptible to radioiodination (Jenkins and Tanner, 1977) is conserved in the human and mouse sequences shown in Figure 5.

Brock *et al.* (1983) have shown that three lysine residues (Lys 608, 610, 618) which are present at the inner membrane surface are located in a 72-amino acid peptide fragment from HEB3. One of these residues appears to be important in the anion transport process (Brock *et al.*, 1983; Kempf *et al.*, 1981). Two of the lysines (Lys 608 and 610) identified by Brock *et al.* (1983) are conserved in HKB3 (see Figure 5) and are predicted to reside in the cytoplasmic loop located between hydrophobic regions e' and f'. Therefore, one of these two lysines must represent the residue important in anion transport.

Cytoplasmic domain. The amino acid sequence derived from the cytoplasmic domain of HKB3 is compared in Figure 6 with the corresponding region of MEB3 and with a published sequence from the N-terminal domain of HEB3 (Kaul *et al.*, 1983). This comparison clearly shows that the cytoplasmic domain of the non-erythroid HKB3 sequence has diverged considerably from its erythroid counterparts. However, three discrete regions within

Fig. 6. Comparison of the amino acid sequence of the cytoplasmic domain of HKB3 with the mouse (MEB3) and human (HEB3) erythrocyte band 3. Sequences were aligned using the PRTALN program of Wilbur and Lipman (1983) and are numbered according to the mouse band 3 sequence of Kopito and Lodish (1985). Insertions within the HKB3 sequence are not numbered. The proposed ankyrin binding region is boxed and the proline-rich hinge region is underlined.

Fig. 7. Analysis of the expression of HKB3. RNA isolated from various cell lines was hybridized with a probe specific for the 5' end of pHKB3. 20 μ g of total RNA was electrophoresed on 1% agarose gels containing 2.2 M formaldehyde (Maniatis *et al.*, 1983), transferred to nitrocellulose and hybridized at 50°C in 50% formamide, 0.9 M NaCl, 0.05 M Hepes, pH 8.0, 0.005 M EDTA, 1 $\times$ Denhardt's solution, 0.2 mg/ml salmon sperm DNA. HEB3 was detected in anemic spleen poly(A)⁺ RNA by hybridization under non-stringent conditions (6 $\times$ SSC, 5 $\times$ Denhardt's at 68°C) and probing with a 1200-bp *Pst*I fragment from pHKB3 encoding a portion of the transmembrane domain. (A) Lane 1, uninduced K562(S) cells; lane 2, induced K562(S) cells; lane 3, human anemic spleen poly(A)⁺ RNA; lane 4, human anemic spleen poly(A)⁺ RNA, hybridized under non-stringent conditions. (B) Expression of HKB3 in leucocyte-derived cell lines. Lane 1, uninduced K562(S); lane 2, Jurkat (T cell lymphoma); lane 3, U937 (monocytic lymphoma); lane 4, HL60 (myelocytic); lane 5, BL-2 (Burkitt lymphoma); lane 6, GM1056 (lymphoblastoid). (C) Expression of HKB3 in tissue-derived cell lines. Lane 1, HepG2 (hepatoma); lane 2, P152 (neuroblastoma); lane 3, SkBr5 (breast carcinoma); lane 4, Kato III (pancreatic carcinoma); lane 5, NB (SV40-transformed renal culture).

the HKB3 sequence show significant homology to erythroid band 3. The 20-amino acid stretch from positions 157 to 177 in HKB3 is 75% homologous to both HEB3 and MEB3. Within this region, the 10 amino acids from 167 to 177 show 90% homology with MEB3 and are identical to the corresponding stretch in the HEB3 sequence. The levels of homology observed drop considerably

on either side of this site. This conserved sequence, lying between a proline-rich region centered around residue 189 in HKB3 and an area (residues 125–138) suggested to be the IgG binding site in MEB3 (Kopito and Lodish, 1985), could represent an ankyrin binding site. The location of this sequence is consistent with the observations of Low *et al.* (1984) which have ten-

tatively positioned the ankyrin binding site in close proximity to these two regions and suggests that the HKB3 protein may function to bind ankyrin in non-erythroid cells.

Two additional regions in HKB3 show significant homology to the MEB3 sequence. The region defined by residues 232–336 is 50% homologous to MEB3 and the area between residues 398 and 423 is also highly homologous (80%), however, no specific functions have been assigned to these regions of the erythroid band 3.

Expression of HKB3

To determine whether HKB3 is expressed in a tissue-specific manner, RNA from cell lines originating from a variety of tissues was analyzed by Northern blotting (Figure 7). A 698-bp *PstI/SacI* fragment derived from the 5' end of the cDNA which has minimal homology with the erythroid band 3 was used as a probe and hybridizations were carried out in 50% formamide at 50°C. This HKB3-specific probe detected a 4.4-kb transcript in the uninduced K562(S) cell line from which the clone was isolated and in cell lines derived from B and T cell lymphomas, kidney, liver, pancreas, breast and nerve cells. In marked contrast, no cross-hybridizing transcripts were observed in K562(S) which had been induced to differentiate along the erythroid lineage with sodium butyrate or in anemic spleen from a patient with β^0 -thalassemia. HEB3 mRNA was detected only in anemic spleen by hybridizing under non-stringent conditions (see Materials and methods) with a 601-bp *BamHI* fragment representing a segment of the transmembrane domain of HKB3. Significant levels of β -actin mRNA were detected in all RNA samples (data not shown) indicating that our inability to detect cross-hybridizing transcripts in these samples was not a result of insufficient loading of RNA. It appears then that HKB3 is not expressed in anemic spleen, which is primarily erythroid in nature, or in induced K562(S) cells, but is expressed in a variety of non-erythroid cells. A second cross-hybridizing mRNA (~8 kb) was also detected in RNA from many of the cell lines. This message was specifically detected with the probe derived from the cytoplasmic domain of HKB3, and is not observed when these RNAs are probed with a 601-bp *BamHI* fragment described above. If this message represents another member of the band 3 gene family, the membrane-spanning region has diverged considerably from that of the HKB3 gene or is absent altogether.

Discussion

In this report we have presented the nucleotide and predicted amino acid sequences of two cDNAs, one representing murine erythrocyte band 3 and the other encoding a protein that is similar but not identical to the human erythroid band 3, which is expressed in K562(S) cells and in a variety of non-erythroid cell lines.

Although the K562(S) cell line has been described as erythroid-like, these cells express granulocytic antigens (Drew *et al.*, 1977; Marie *et al.*, 1981) as well as erythroid markers (Gahmberg *et al.*, 1979; Marie *et al.*, 1981). This cell line appears to undergo only limited erythroid differentiation upon butyrate induction resulting in the production of small amounts of fetal and embryonic hemoglobins (Cioe *et al.*, 1981). Fukuda *et al.* (1980) have suggested that the erythrocyte band 3 is not expressed at significant levels until the later stages of erythroid maturation which may explain our inability to detect HEB3 mRNA in the induced K562(S) cells. The fact that we can detect HKB3 mRNA in uninduced K562 cells but not in induced cultures or in anemic spleen, which is rich in erythroid precursors, clearly shows that HKB3 is not an erythrocyte form of band 3.

The structures of HKB3 and erythrocyte band 3 are remarkably similar. Hydrophobicity plots of HKB3 and MEB3 indicate that the polypeptide chain traverses the membrane 12 or possibly 13 times and that the topography of the transmembrane domain of HKB3 is essentially the same as that previously described for MEB3 (Kopito and Lodish, 1985). Virtually all of the hydrophilic residues contained within amphipathic membrane-spanning helices are conserved between the MEB3 and HKB3 sequences. The amino acid sequence derived from the membrane domain of HEB3, which corresponds to hydrophobic regions b', e' and f' in HKB3, also shows the conservation of hydrophilic residues within spanning helices, implying that these residues are functionally important. It has been suggested that these amphipathic helices associate to form an aqueous pore and thereby facilitate anion transport (Kopito and Lodish, 1985). However, all of the charged residues contained within the amphipathic helices are acidic, and negatively charged, which precludes a transport mechanism involving the passage of Cl^- through the membrane by simple ion pairing. At least one of these acidic residues is important since chloride transport is inhibited by protonation of a group with a pK_a of ~5 (Milanick and Gunn, 1982) and by water-soluble carbodiimides (Weith *et al.*, 1982). This carboxylate(s) appears to reside within the 72-amino acid papain fragment sequenced by Brock *et al.*, (1983). Three acidic residues reside in this fragment from HEB3 and all are conserved in HKB3. Two of these residues lie in an amphipathic helix, with the third being located in the external loop between hydrophobic regions f' and g'. Lysine residues which may be important in anion transport are also associated with amphipathic helices, but it is not clear what role these helices might play in the anion transport process. However, the extensive conservation of structure and of specific residues associated with anion transport strongly suggests that HKB3 functions in this respect.

The 59 amino acid external loop (region Z in Figure 4) in HKB3 represents the major structural difference observed within the transmembrane domains of this protein and the erythrocyte band 3 sequences. This region shows little homology to the analogous loop in MEB3 with the exception of 10 residues (558–568) which contain a lysine (Lys 558 or 561) important in the anion transport process. Three glycosylation sites unique to HKB3 are present in this loop, but the extent of glycosylation has not been determined. However, the presence of these three new sites coupled with the absence of the erythrocyte glycosylation sequence suggest that glycosylation of HKB3 would be involved in cellular functions which are not shared with the erythroid protein.

Comparison of the cytoplasmic domain of HKB3 with both the MEB3 and HEB3 sequences shows the region from Asp 157 to His 177 to be highly conserved. Within this span, the sequence L R A L L K H S H is virtually identical, however, a single change from His to Arg at residue 175 is present in the mouse. This region is in close proximity to both the putative IgG binding site (residues 125–138) and the proline-rich hinge region (residues 188–204) in MEB3 and is consistent with the tentative location assigned to the ankyrin binding site in HEB3 (Low *et al.*, 1984). Also, the regions flanking this conserved sequence are <35% homologous to the erythrocyte band 3 sequences, suggesting that this region defines a functional site which is conserved in both erythroid and non-erythroid forms of band 3. Chou and Fasman (1973) analysis predicts this region to be α -helical and the amino acid sequence shows the putative helix would be highly charged. Therefore it is possible that binding occurs through charge interactions with a similar region contained in

ankyrin. Ankyrin can be eluted from band 3 in erythrocyte membranes by extraction in high salt, i.e. 1 M KCl (Weaver *et al.*, 1984), consistent with an association based on charge interactions. However, we cannot rule out the possibility that other regions within the cytoplasmic domain may also play a role in ankyrin binding.

The presence of non-erythroid forms of spectrin (Bennett *et al.*, 1982; Goodman *et al.*, 1981), ankyrin (Davis and Bennett, 1984; Nelson and Lazarides, 1984) and band 4.1 (Aster *et al.*, 1984; Goodman *et al.*, 1984) has led to speculation as to whether these proteins are anchored to the membrane by a protein homologous to the erythrocyte band 3. Reports by Cox *et al.* (1985) have identified a band 3-related RNA and polypeptide in avian kidney cells using an avian erythroid band 3 cDNA probe or antibodies prepared against the erythrocyte band 3. Due to the inability to demonstrate cross-hybridizing species in other avian cell types, Cox *et al.* (1985) have speculated that anchoring of the membrane cytoskeleton in other cell types is not mediated by a band 3-like protein. HKB3 is clearly expressed in non-erythroid cells and, furthermore, sequence homology within the putative ankyrin binding region suggests that the ankyrin binding function may be intact. The inability of Cox *et al.* (1985) to identify band 3 transcripts in other tissues might have resulted if the avian cDNA probe was primarily derived from the 5' end of the erythroid band 3, since the sequence analysis we present here demonstrates the low levels of homology between erythroid and non-erythroid forms of band 3 in this region. Similarly, since antibodies produced against band 3 are primarily directed against the major antigenic determinant of band 3 which is located in the N-terminal region of the protein (Low *et al.*, 1984) it would also be unlikely that they would recognize non-erythroid forms of band 3.

The transmembrane orientation of band 3 differs from many other membrane-spanning proteins in that its N terminus lies within the cytoplasm. The insertion of band 3 into the membrane and the subsequent modifications made in the rough endoplasmic reticulum (ER) and Golgi apparatus follows a pathway similar to that of certain other membrane proteins (Braell and Lodish, 1981) and band 3 appears to be inserted into the rough ER membrane in the same orientation as it exists in the mature state (Braell and Lodish, 1981). However, the signal required for insertion into the membrane must be internal and as such cannot be cleaved off after insertion. Kopito and Lodish (1985) have suggested that a hairpin loop formed by the first two transmembrane helices of the MEB3 might function as the signal for insertion. The degree of conservation observed within the membrane domains of the mouse and HKB3 sequences suggests that such a hairpin, if it forms, would do so in both cases. However the 23 amino acids (residues 398–420) immediately preceding the first transmembrane helix are also highly conserved, exhibiting 80% homology to MEB3. This region, containing five arginine residues, might bind to the inner surface of the membrane by interaction with phospholipid head groups and in this manner anchor the cytoplasmic domain to the membrane to allow the co-translational insertion of the remainder of the protein. This type of interaction could also stabilize the membrane domain in the mature protein. Arginine residues capable of participating in similar interactions are positioned around 10 of the 13 transmembrane helices in HKB3 and most of these residues (10 of 13) are conserved in the MEB3. The existence of this type of interaction is consistent with the observations of Chicken and Sharon (1984) that band 3 appears to immobilize a large amount of

phospholipid when inserted into unilamellar phosphatidylcholine vesicles.

The comparison of cDNA clones encoding two closely related members of the band 3 gene family from mouse and human has indicated that an erythroid and non-erythroid form of the band 3 protein are structurally and functionally similar. The two distinct structural domains characteristic of the erythrocyte band 3 are also present in HKB3, a human non-erythroid form of band 3. The basic structure of the transmembrane domain and the specific amino acid residues associated with anion transport in the erythrocyte are conserved in HKB3, strongly suggesting that the HKB3 protein functions in anion transport. Also, the conservation of specific regions within the divergent cytoplasmic domain of HKB3 further suggests that the ability to interact with a membrane cytoskeletal complex via a conserved ankyrin binding site has been retained. The availability of these two band 3-related sequences will allow us to study in greater detail the role of the band 3 protein in erythroid and non-erythroid cells.

Materials and methods

Bacterial strains and vectors

Plasmids pUC7, pUC8 and pUC9, M13mp18, M13mp19 and *E. coli* JM103 were kindly supplied by Dr J. Messing (University of Minnesota). *E. coli* strains DH1 and JM103 were transformed according to the procedure of Hanahan (1983) or by treatment with CaCl_2 (Mandel and Higa, 1970).

Cell lines

The following cell lines were used for Northern blot analysis: K562(S) (Cioe *et al.*, 1981); JuE, T cell lymphoma (Gillis and Watson, 1980); U937, histiocytic lymphoma (Sundström and Nilsson, 1976); HL-60, myelogenous leukemia (Gallagher *et al.*, 1979); BL-2, Burkitt's lymphoma (Bertrand *et al.*, 1981); HepG2, hepatoma (Aden *et al.*, 1979); Kato III (ATCC HTB 103); NB, SV40-transformed renal culture (Shein and Enders, 1962); GM1056, Epstein–Barr virus immortalized lymphoblastic (obtained from the Human Genetic Mutant Cell Depository, Camden, NJ); SKBr5, breast carcinoma and P152, neuroblastoma (kindly supplied by Dr Zenon Stepelwski, Wistar Institute, Philadelphia, PA).

K562(S) was grown and induced with butyric acid as previously described (Cioe *et al.*, 1981).

Construction and screening of the cDNA libraries

Construction of the cDNA expression library in the pUC system (Vieira and Messing, 1982) and the cDNA libraries derived from mouse anemic spleen (Curtis *et al.*, 1983) and K562(S) cells (Watt *et al.*, 1983) in pBR322 were previously described. The cDNA expression library was immunologically screened as described by Helfman *et al.* (1983) using antibodies against mouse erythrocyte membrane proteins (Cioe and Curtis, 1985) as well as with anti-mouse band 3 (kindly supplied by Dr M.M.B. Kay, Texas A & M University, Temple, TX). The mouse anemic spleen and K562(S) cell cDNA libraries were screened according to the procedure of Grunstein and Hogness (1975). Filters were hybridized in $6 \times \text{SSC}$, $5 \times \text{Denhardt's}$ at 68°C for 24 h and washed three times with $2 \times \text{SSC}$, 0.1% SDS at 60°C for 30 min. Approximately 60 000 colonies from each of the cDNA libraries were plated and screened using probes derived from positive colonies obtained from the immunological screening of the expression libraries. Inserts were labeled using calf thymus primers and *E. coli* DNA polymerase I (Boehringer Mannheim) as described by Summers (1975).

DNA sequencing

Restriction fragments obtained from positive clones isolated from the various cDNA libraries were subcloned into M13mp18 or mp19 (Norlander *et al.*, 1983) and sequenced by the dideoxy chain termination method of Sanger *et al.* (1977). Larger restriction fragments were sequenced entirely after progressive digestion with λ exonuclease III (N.E. Biolabs) as described by Henikoff (1984). Essentially, $20 \mu\text{g}$ of the M13 clone to be sequenced were digested to completion with the appropriate restriction enzymes followed by digestion with 1000 U of λ exonuclease III for 20 min at 25°C . Aliquots were removed at 1–2 min intervals, quenched with 0.5 M NaCl, 0.2% SDS, 0.015 M EDTA at 65°C for 5 min and subsequently treated with S1 nuclease (Boehringer Mannheim) and *E. coli* DNA polymerase, Klenow fragment (Boehringer Mannheim) prior to religation. Under these conditions, exonuclease III digested at a rate of ~ 80 nucleotides/min. Additional sequences were determined by the procedure of Maxam and Gilbert (1977) using 3' end-labeled restriction fragments or by primer extension from synthetic oligonucleotide primers synthesized on a DNA synthesizer (Applied Biosystems, Inc., Model 380A).

Isolation and electrophoresis of RNA

RNA was prepared from various cell lines by the guanidinium isothiocyanate method (Chirgwin et al., 1979). Total RNA (~20 µg) was fractionated on 1% agarose gels in the presence of 2.2 M formaldehyde (Maniatis et al., 1983). Transfer to nitrocellulose, pre-hybridization and hybridization were essentially as previously described (Cioe and Curtis, 1985). Non-stringent hybridization and washing conditions were as described for the cDNA library screening. 5'- and 3'-specific probes were generated by restriction enzyme digestion and gel electrophoresis purification. Probes were labeled utilizing calf thymus primers (Summers, 1975).

Acknowledgements

We wish to thank Drs Leila Diamond, Carlo Croce and Zenon Stepiewski for supplying cell lines, Dr Saul Surrey for providing samples of human anemic spleen and Dr M.M.B. Kay for supplying antisera to mouse band 3. We also thank Dr Michael Showe for his assistance with the plotting of the hydrophobicity profiles. We are grateful to the Editorial staff of The Wistar Institute for preparation of this manuscript and to Marina Hoffman for editorial assistance. We would also like to thank Dr A. Caton for critical reading of the manuscript. This work was supported by Public Health Service Grants GM-32583 (to L.S.) and CA-10815 from the National Institutes of Health. L.C. was a recipient of a Special Fellowship from the Leukemia Society of America and P.F. is a U.S. Public Health Service Pre-doctoral Trainee (CA-09171).

References

- Aden, D.P., Fogel, A., Plotkin, S.A., Damjanov, I. and Knowles, B.B. (1979) *Nature*, **282**, 615–616.
- Aster, J.C., Welsh, M.J., Brewer, G.J. and Maises, H. (1984) *Biochem. Biophys. Res. Commun.*, **119**, 726–734.
- Bennett, V. and Stenbuck, P.J. (1980) *J. Biol. Chem.*, **255**, 6424–6432.
- Bennett, V., Davis, J. and Fowler, W.G. (1982) *Nature*, **299**, 126–131.
- Bertrand, S., Berger, R., Philip, T., Bernheim, A., Bryon, P.A., Bectoglia, J., Dore, J.F., Brunat-Mentigny, M. and Lenoir, G.M. (1981) *Eur. J. Cancer*, **17**, 577–583.
- Braell, W.A. and Lodish, H.F. (1981) *J. Biol. Chem.*, **256**, 11337–11344.
- Brock, C.J., Tanner, M.J.A. and Kempf, C. (1983) *Biochem. J.*, **213**, 577–586.
- Cheng, S. and Levy, D. (1980) *J. Biol. Chem.*, **255**, 2637–2640.
- Chicken, C.A. and Sharon, F.J. (1984) *Biochim. Biophys. Acta*, **774**, 110–118.
- Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochemistry*, **18**, 5294–5299.
- Chou, P.Y. and Fasman, G. (1973) *J. Mol. Biol.*, **74**, 263–281.
- Cioe, L. and Curtis, P. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1367–1371.
- Cioe, L., McNab, A., Hubbell, H.R., Meo, P., Curtis, P. and Rovera, G. (1981) *Cancer Res.*, **41**, 237–243.
- Cox, J.V., Moon, R.J. and Lazarides, E. (1985) *J. Cell Biol.*, **100**, 1548–1557.
- Curtis, P.J., Withers, E., Demuth, D., Watt, R., Venta, P.J. and Tashian, R.E. (1983) *Gene*, **25**, 325–332.
- Curtis, P.J., Palumbo, A., Ming, J., Fraser, P., Cioe, L., Meo, P., Shane, S. and Rovera, G. (1985) *Gene*, **36**, 357–362.
- Davis, J.Q. and Bennett, V. (1984) *J. Biol. Chem.*, **259**, 13550–13559.
- Drenckhahn, D., Zinke, K., Schauer, U., Appell, K.C. and Low, P.C. (1984) *Eur. J. Cell Biol.*, **34**, 144–150.
- Drew, S.I., Terasaki, P.I., Billing, R.J., Bergh, O.J., Minowada, J. and Kelen, E. (1977) *Blood*, **49**, 715–718.
- Eisenberg, D. (1984) *Annu. Rev. Biochem.*, **53**, 595–623.
- Falke, J.J., Pace, R.J. and Chan, S.I. (1984) *J. Biol. Chem.*, **259**, 6472–6480.
- Fox, R.O. and Richards, F.M. (1982) *Nature*, **300**, 325–330.
- Frizzel, R.A., Koch, M.J. and Schultz, S.G. (1976) *J. Membr. Biol.*, **27**, 297–316.
- Fukuda, M., Fukuda, M.N., Papayannopoulou, T. and Hakomori, S.-I. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3474–3478.
- Gahmberg, C.G., Jokinen, M. and Andersson, L.C. (1979) *J. Biol. Chem.*, **254**, 7443–7448.
- Gallagher, R.S., Collins, S. and Trujillo, J. (1979) *Blood*, **54**, 713–733.
- Gillis, S. and Watson, J. (1980) *J. Exp. Med.*, **152**, 1709–1719.
- Goodman, S.R., Zagon, I.S. and Kulikowski, R.R. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 7570–7574.
- Goodman, S.R., Casoria, L.A., Coleman, D.B. and Zagon, I.S. (1984) *Science*, **224**, 1433–1436.
- Grinstein, S., Ship, S. and Rothstein, A. (1979) *Biochim. Biophys. Acta*, **507**, 294–304.
- Grinstein, M. and Hogness, D. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 3961–3964.
- Hanahan, D. (1983) *J. Mol. Biol.*, **166**, 557–580.
- Helfman, D.M., Feramisco, J.R., Fiddles, J.C., Thomas, G.P. and Hughes, S.H. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 31–35.
- Henikoff, S. (1984) *Gene*, **281**, 351–359.
- Jenkins, R.E. and Tanner, M.J.A. (1977) *Biochem. J.*, **161**, 139–147.
- Jennings, M.L. (1984) *J. Membr. Biol.*, **80**, 105–117.
- Jennings, M.L. (1985) *Annu. Rev. Physiol.*, **47**, 519–533.
- Jennings, M.L. and Nicklish, J.S. (1984) *Biochemistry*, **23**, 6432–6436.
- Kaul, R.K., Prasanna-Murthy, S.N., Giridhar-Reddy, A., Steck, T.L. and Kohler, H. (1983) *J. Biol. Chem.*, **258**, 7981–7990.
- Kay, M.M.B., Tracey, C.M., Goodman, J.R., Cone, J.C. and Bassell, P.S. (1984) *Proc. Natl. Acad. Sci. USA*, **80**, 6882–6886.
- Kempf, C., Brock, C., Sigrist, H., Tanner, M.J.A. and Zahler, P. (1981) *Biochim. Biophys. Acta*, **641**, 88–98.
- Knauf, P.A. (1979) *Curr. Top. Membr. Transp.*, **12**, 249–363.
- Kopito, R.R. and Lodish, H.F. (1985) *Nature*, **316**, 234–238.
- Kyte, J. and Doolittle, R.L. (1982) *J. Mol. Biol.*, **157**, 105–132.
- Leslie, B.R., Schwartz, J.H. and Steinmütz, P.R. (1973) *Am. J. Physiol.*, **225**, 610–617.
- Low, P.S., Westfall, M.A., Allen, D.P. and Appell, K.C. (1984) *J. Biol. Chem.*, **259**, 13070–13076.
- Mandel, M. and Higa, A. (1970) *J. Mol. Biol.*, **53**, 154–169.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1983) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
- Marie, J.P., Izaguirre, C.A., Civin, C.I., Mirro, J. and McCullock, E.A. (1981) *Blood*, **58**, 708–711.
- Mawby, W.J. and Finlay, J.B.C. (1982) *Biochem. J.*, **205**, 465–475.
- Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
- Milanick, M.A. and Gunn, P.B. (1982) *J. Gen. Physiol.*, **79**, 87–113.
- Nelson, W.J. and Lazarides, E. (1984) *Cell*, **39**, 309–320.
- Norrander, J., Kempe, T. and Messing, J. (1983) *Gene*, **26**, 101–106.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Sayare, M. and Fikiet, M. (1981) *J. Biol. Chem.*, **256**, 13152–13158.
- Shein, H.M. and Enders, J.F. (1962) *Proc. Natl. Acad. Sci. USA*, **48**, 1164–1172.
- Steck, T.L. (1974) *J. Cell Biol.*, **62**, 1–19.
- Strapazon, E. and Steck, T.L. (1976) *Biochemistry*, **15**, 1421–1424.
- Summers, J. (1975) *J. Virol.*, **15**, 946–953.
- Sundström, C. and Nilsson, K. (1976) *Int. J. Cancer*, **17**, 565–577.
- Vieira, J. and Messing, J. (1982) *Gene*, **19**, 259–268.
- Watt, R., Stanton, L.W., Marcu, K.B., Gallo, R.C., Croce, C.M. and Rovera, G. (1983) *Nature*, **303**, 725–728.
- Weaver, D.C., Pasternack, G.R. and Marchesi, V.T. (1984) *J. Biol. Chem.*, **259**, 6170–6175.
- Weith, J.O., Anderson, O.S., Brahm, J., Pjerrum, P.J. and Borders, C.L., Jr. (1982) *Phil. Trans. R. Soc. Lond. Ser. B*, **299**, 383–399.
- Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 726–730.
- Yu, J. and Steck, J.L. (1975) *J. Biol. Chem.*, **250**, 9176–9184.

Received on 12 March 1986