Volume 312, number 2,3, 105–109 FEBS 11705 November 1992

Minireview

Unexpected intron location in non-vertebrate globin genes

Luc Moens[a], Jacques Vanfleteren[b], Ivo De Baere[a], Anna M. Jellie[c], Warren Tate[c] and Clive N.A. Trotman[c]

[a]*Department of Biochemistry, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium,* [b]*Laboratory of Animal Morphology and Systematics, University of Ghent, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium* and [c]*Department of Biochemistry, University of Otago, Dunedin, New Zealand*

Received 14 August 1992; revised version received 18 September 1992

The *Caenorhabditis elegans* and *Artemia* T4 globin sequences are highly homologous with other invertebrate globins.The intron/exon patterns of their genes display a single intron in the E and G helices respectively. Preceding introns in multirepeat globins are inserted in homologous positions. Comparison of the intron/exon patterns in the known globin gene sequences demonstrates that they are more diverse than first expected but nevertheless can be derived from an ancestral pattern having 3 introns and 4 exons.

Globin gene; Intron location; Evolution

1. INTRODUCTION

Vertebrate globin genes invariably have two introns resulting in three exons with boundaries between amino acid residues B12 and B13 and between G6 and G7 (helix notation). Plant globin genes have an extra intron between E14 and E15, making four exons, whereas most insect globin genes contain no intervening sequences [1,2]. It is currently thought that the 4 exons, 3 introns arrangement is ancestral and that evolution has sometimes led to the elimination of introns [1–3].

Recently the gene sequence of the internally duplicated globin from the nematode *Pseudoterranova decipiens* was described [4]. This sequence contains six introns and seven exons altogether. The first intron separates a secretory peptide leader sequence from the functional protein coding sequence (precoding intron). The next three introns are in the B, E and G helix coding sequences similar to plants. The fifth intron separates the two globin repeats (bridge intron) and the final one is found in the second repeat which has retained only this single intron in the B helix.

The globin of the clam *Barbatia reeveana* (Mollusca) also has a duplicated globin structure but has the vertebrate intron-exon pattern, together with a precoding intron before the first repeat and a bridge intron between the repeats [5].

This contrast suggests that the globin gene structure of invertebrates is more diverse than at first expected.

Here we examine the protein and gene structures of the monomeric haemoglobin of the nematode *Caenorhabditis elegans* and repeat T4 of the multi-repeat globin of the crustacean *Artemia* and we evaluate their globin gene structures from an evolutionary perspective.

2. ALIGNMENT OF GLOBIN AMINO ACID SEQUENCES

The proper alignment of the translated globin sequences is a prerequisite for the assessment of equivalent intron positions. The invertebrate globin sequences from which the genomic structures are known are aligned in Fig. 1.

Only the alignment of the *C. elegans* globin is discussed in detail below. The stereochemical interpretation of the other sequences has been published previously [11–17].

The proposed alignment recognizes the obligatory CD1 Phe and F8 His, as well as the conservation of the Gly at B6 and E8 associated with the crossing of the B and E helices [18].

Although E7 in *C. elegans* cannot be occupied by the usual His, the alignment is notably strong from E2 though E18. Moreover, the observed Gln at E7 is, in invertebrates, the most frequently seen alternative at this position [19,20].

The G helix is unusual in having two Trp residues at G5 and G9. However, alignment in the G helix is strengthened by the confidence with which we can lo-

Correspondence address: L. Moens, Department of Biochemistry, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium. Fax: (32) (3) 8202248.

```
           H/LEADER SEQUENCE       HC/LINKER/PRE-A        A            AB              B             C          CD           D
                   1          2                 1             1                              1                           1
         12345678901234567890 1 1234 5678901234567   1234567890123456  1234567   123456789012 3456  1234567  1234567890   1234567

MbW                                             VL   SEGEWQLVLHVWAKVE  ------A   DVAGHGQDILIR LFKS  HPETLEK  FD--RFKHLK   TEAEMKA
Gm                                             VAF   TEKQDALVSSSFEAFK  ------A   NIPQYSVVFYTS▼ILEK  APAAKDL  FS---FLANG   -VDPT--
Ct4        #{KLLILALCFAAASA}                     L   TADQISTVQSSFAGVK  ------G   ----DAVGILYA VFKA  DPSIQAK  FT---F-AGK   DLDSIKG
Ct2                                            APL   SADEASLVRGSWAQVK  ------H   ----SEVDILYY IFKA  NPDIMAK  FP--QF-AGK   DLETLKG
Lt                                         DEHEHCC   SEEDHRIVQKQWDILW  RDTESSK   IKIGFGRLLLTK▼LAKD  IPDVNDL  FK--RV----   DIEHAE-
At                                      VADAVAKVCG   SEAIKGNLRRSWGVLM  -----SA   DIEATGLTYLAN▼LFTL  RPDTKTY  FT--RLGDVQ   KGKA---
At4   [DEFTVKSFNHVFEVIISFLNE]GLRQ▼ADI---VDPVTHL   TGRQKEMIKASWSKAR  ------T   DLRSLGQELFMR MFKA  HPEYQTL  FVNKGFADV-   PLVSLRE
At5   [DAASVEAWKKFFDVSIGVIAQ]GL-K▼VATSEEADPVTGL   YGKEIVALRQAFAAVT  ------P   RNVEIGKRVFAK LFAA  HPEYKNL  FK--KFEQY-   SVEELPS
Br1                             #▼MSVSAKLDEVT--   QPANKNLIRSTWNMMV  -------   GDRGNGVELMGL▼LFQR  APDSKID  FK--RLGDVS   -AENIPY
Br2   [-DDAIQAWWKLIDVICAVLKE]--SK▼KGVLETIEEVT--   KPANKGLIRETWNMIA  -------   GDRKNGVELMAL▼LFEM  APDSKKD  FR--RLGDVS   -PSNIPN
Pd1     #{MHSSIVLAIVLFVAIASA}  SK▼-------------   ---TRELCMKSLEHAK  --VGTSK   EAKQDGIDLYKH▼MFEH  YPAMKKY  FKHRENYT--   -PADVQK
Pd2   [DEPTKHAWQEIGKEFSHEISH]G-RH▼-------------   ---VRDHCMNSLEYIA  --IGDKE   HQKQNGIDLYKH▼MFEH  YPHMRKA  FKGRENFT--   -KEDVQK
Ce                         # SM N-------------   RQEISDLCVKSLEGRM  --VGTEA   QNIENGNAFFRY FFTN  FPDLRVY  FKGAEKYT--   -ADDVKK

                  E                     EF               F        FG          G                GH              H
                1          2                1              1                  1                                  1         2
      123 4567 89 01234 567890 1234567890123    1234567890  12345   123456 7890123456789   123456789   1234567890123456789012 1

MbW   SED LKKH GV TVLTA LGAILK  ---KKGHHEAE--   LKPLAQSHAT  KH-KI   PIKYLE▼FISEAIIHVLHSR   HPGDF----   GADAQGAMNKALELFRKDIAA
Gm    NPK LTGH AE KLFAL▼VRDSAG  ---ASGTVVAD--   -AALGSVHAQ  KA--V   TDPQFV▼VVKEALLKTIKAA   VQLK-GDKW   SDELSRAWEVAYDELAAAIKK
Ct4   SAD FSAH AN KIVGF FSKIIG  ----LPNIDGD--   VTTFVASHTP  RG--V   THDQLN NFRAGFVSYMKAH   -----DTD-   FAGAEAAWGATLDAFFGMVFA
Ct2   TGQ FATH AG▼RIVGF VSEIVA  MGNSANMPAMETL   IKDMAANHKA  RG--I   PKAQFN EFRASLVSYLQSK   ----L-VSW   NDSLGAAWTQGLDNVFNMMFS
Lt    GPK FSAH AL RILNG LDLAIN  LLDDPPALDAA--   LDHLAHQHEV  REG-V   QKAHFK▼KFGEILATGLPQV   ----L----   DDYDALAWKSCLKGILTKISS
At    NSK LRGH AI TLTYA LDWFVD  SLDDPSRLKCV--   VEKFAVNHIN  RK--I   SGDAFG▼SIIPEMKETLKAR   -MGSY----   SDDVGAAWVQAILGMQNAVLS
At4   DER FISH MA NVLGG FDTLLQ  -LDESSYFIYS--   LRNLGDAHIQ  RK--A   GTQHFR▼SFEAILIPILQES   -NQGL----   DAASVEAWKKFFDVSIGVIAQ
At5   TDA FHYH IS LVMNR FSSIGK  -IDDNVSFVYL--   LKKLGREHIK  RG--L   SRKQFD QFVELYIAEISSE   ---VL----   SDTGRNGLEKVLTFATGVIEQ
Br1   NRK LNGH GI TLWYA LMNFVD  DLDSKKDLEDV--   CRKFAVNHVI  RG--V   LDVKFG▼WIKEPMAELLRRK   ---CGNDC-   -DDAIQAWWKLIDVICAVLKE
Br2   NRK LNGH GI TLWYA LMNFVD  QLDNKIDLEDV--   CRKFAVNHVN  RG--V   LDVKFA▼WIKEPLAELLRRK   ---CGQTC-   TDQHIQAWWKLIDVVCAVLEE
Pd1   DPF FIKQ▼GQ NILLA CHVLCA  TYDDRETFDAY--   VGELMARHER  DHVKI   PNDVWN▼HFWEHFIEFLGSK   --TTL----   DEPTKHAWQEIGKEFSHEISH
Pd2   DAF FVKQ GH KILLA LRMLCS  SYDDEPTFDYF--   VDALMDRHIK  DDIHL   PQEQWH EFWKLFAEYLNEK   SHQHL----   TEAEKHAWSTIGEDFAHEADK
Ce    SER▼FDKQ GQ RILLA CHLLAN  VYTNEEVFKGY--   -RETINRHRI  --YKM   DPALWM AFFTVFTGYLESV   --CGL----   NDQQKAAWMALGKEFNAESQT

          HC
           1         2
      12345678901234567890

MbW   KYKELGYQG
Gm    A
Ct4   KM
Ct2   YL
Lt    RLNA
At    AL
At4   GLKV
At5   GL
Br1   SK
Br2   KK
Pd1   H
Pd2   HAKAEKDHHEGEHKEEHH
Ce    HLKNSNLPHV
```

Fig. 1. Alignment of non-vertebrate globin sequences showing the intron locations. Sequences were aligned using FASTA [9] and the template of Bashford et al. [10]. ▼; = intron position; # = preceeds coding sequence; {} = leader sequence; [ ] = H helix of preceeding repeat Mb = Sperm whale myoglobin [1]. Gm = *Glycine maxima* globin [14,33]. Ct4 = *Chironomus thummi* globin IV [11,32]. Ct2 = *Chironomus thummi* globin *IIb* [11,27]. Lt = *Lumbricus terrestris* globin [12]. At = *Anadara trapezia* minor globin [13,31]. AT4 = *Artemia* globin repeat T4 [17,23]. AT5 = *Artemia* globin repeat T5 [17,23]. Br1 = *Barbatia reeveana* globin repeat 1 [5]. Br2 = *Barbatia reeveana* globin repeat 2 (5). Pd1 = *Pseudoterranova decipiens* globin repeat 1 [4,16]. Pd2 = *Pseudoterranova decipiens* globin repeat 2 [4,16]. Ce = *Caenorhabditis elegans* globin [6,7]. (Deduced from genomic DNA sequence data obtained as a partial result of the *C. elegans* genome sequencing project (cosmid zk 637; PIR/NBRF Accesion number f39344) [6,7]).

cate F8 His and its flanking sequence to one side, and the H helix with its characteristic H8 Trp to the other side. The steric compensation of the bulky side chain of G9 Trp by the single hydrogen side chain of H12 Gly has been discussed previously [17,21]. The alignment in Fig. 1 is in accordance with Dixon et al. [4] but the unconventional nature of the 5-residue AB turn is further discussed below with the interpretation of the intron structure.

3. ORGANIZATION OF GLOBIN GENES

Cosmid zk 637 has an insert sequence that can be identified as globin like [6,7]. Analysis of this putative globin gene confirms the presence of a TATAA, a CAATA and a polyadenylation signal sequence as well as start and stop codons; therefore it contains the necessary elements of a functional gene. A single intron of 298 bases is present, having the characteristic nematode splicing consensus sequences [22] (Figs. 1 and 2).

The T4 repeat gene of *Artemia* globin contains a single intron of about 1,000 bases, in addition to interdomain introns located between domains T3 and T4 and between T4 and T5 (Figs. 1 and 2). All introns show the consensus splicing sequences [23].

Having correlated the gene sequence data with structural features of the protein, we can now evaluate the intron locations. Two classes of intron location are recognized, namely (A) intra-repeat and (B) inter-repeat.

3.1. *Intra-repeat introns*

The intron positions in invertebrate globin B and G helices, determined so far, are precisely conserved at the B12/B13 and the G6/G7 location (Fig. 1). In contrast

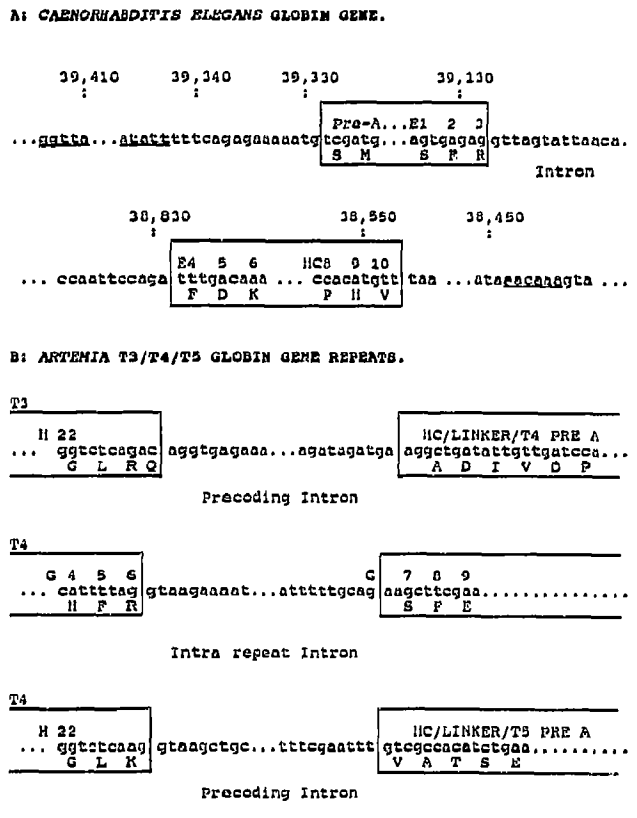

Fig. 2. Gene organization of the *C. elegans* globin and of repeat T4 of *Artemia*. (A) Partial nucleotide sequence of the *C. elegans* globin gene showing regulatory sites (underlined) and splice junctions. Exons are boxed. Numbers refer to the sequence of the template strand insert in cosmid zk 637 [6,7]. Deduced coding strand sequence is shown for convenience. Helix notification as in Fig. 1. Amino acids are given by the single letter code underneath the second base of each codon. (B) Partial nucleotide sequence of *Artemia* T3/T4/T5 globin repeats showing splice junctions. Exons are boxed. Helix notification as in Fig. 1.

the location of the intron within the E helix seems to be variable. The central intron in plants is inserted at E14–15, which is 4 residues after the position predicted by Gō [24,25]. This intron separates the haem binding region into two structural units, F2 and F3 that make contact with haem from opposite sides [24]. In *P. decipiens* the equivalent central intron is inserted at E7-E8 (CAA/Q G[intron]GT/G CAA/Q) [24]. The equivalent *C. elegans* intron is inserted at a still more anterior position E3-E4. It is tempting to speculate that it has migrated secondarily from its ancestral more distal position, which from E7 now reads: (CAA/Q GGC/G CAA/Q) and is coincident with the splice site in the *P. decipiens* gene. Similar intron sliding is observed in other genes such as triose phosphate isomerase [26]. In these two globins the nucleotide sequences differ by only one base in this region.

Intron insertion may differ by as much as six codons [2], however the conservation of the insertion positions in the B and G helices contrasts with the variation in the E helix.

In contrast to all other *Chironomus* globin genes sequenced so far, a single intron is present within the E helix (E9–E10) of the gene coding for globin IIB [1,2,27]

Single introns are also present in the B helix of *P. decipiens* repeat 2 and in the G-helix of *Artemia* repeat T4 [23].

3.2. *Inter-repeat introns*

In addition to the intra-repeat introns described above, the multi-repeat globin genes of *Artemia*, *Barbatia* and *Pseudoterranova* have inter-repeat introns separating the individual globin units (Fig. 1). From an evolutionary point of view, intra- and inter-repeat introns are not equivalent. Intra-repeat introns represent functional regions of the protein in the sense of the mini-gene hypothesis of Gilbert [28] and thus refer to the primordial gene organization. Inter-repeat introns have accompanied gene duplication events and can thus be considered as secondarily acquired and much more recent.

The presence of inter-repeat introns may help us to assign structure to the N-terminal region of the nematode sequences, which align well with each other but are difficult to reconcile with other globins in the A helix and AB turn. A 5-residue AB turn is not known structurally in globins and the resulting alignment places an intron illogically in the A-helix [4]. The emerging precedent is for introns to be located between functional repeats, e.g. in the *Artemia* globin gene at both ends of domain T4 where the clear alignment of the H and A helices and the length of the linkers leave no doubt that this is the case. An attractive alignment is obtained by modelling the nematode sequences on *Chironomus* globin III in which the start of the B helix is shortened by 4 residues (Table I). This places favourable AsX residues at the start of the B helix and logically positions the inter-repeat intron before an inter-repeat linker of about 6 residues

4. GLOBIN GENE EVOLUTION

The globin gene organization found in invertebrates is compatible with the idea of an ancestral gene containing 4 exons and 3 introns, with a tendency for introns to have been lost during evolution (Fig. 3) [1,2]. This arrangement appears to have been retained in known plant globins and in the first repeat of the *P. decipiens* gene whereas all introns except one have been secondarily deleted from the second repeat. Only a single intron has been retained in the E helix of *C. elegans* and in the *Chironomus IIb* gene. This intron is equivalent to the ancestral and plant central introns.

It might be argued that these introns are a more recent acquisition and that the ancestral globin gene contains two introns only (in the B and G helices). However novel intron insertion would be opposed to the general evolutionary tendency of intron reduction.

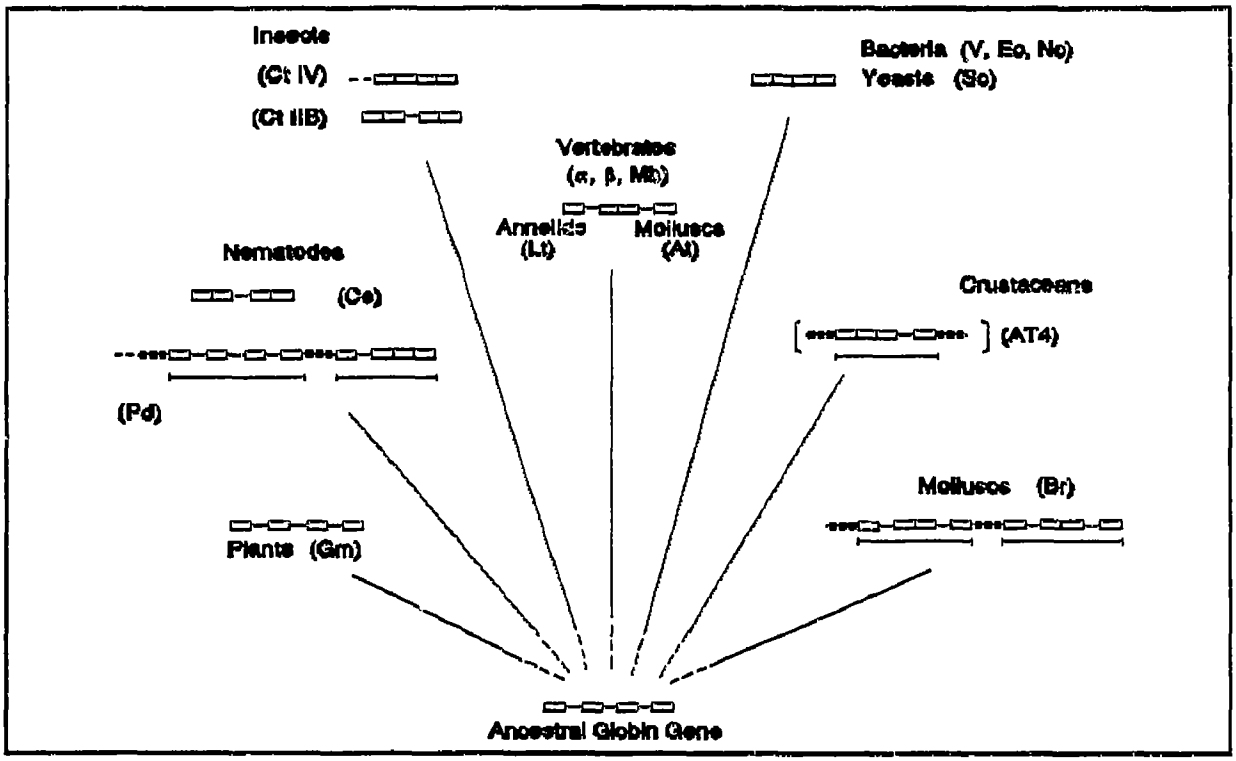

Fig. 3. Evolution of globin gene arrangement. This figure illustrates how the array of globin gene structures found in extant species can all be derived from a single ancestral 4 exon/3 intron con-figuration. A general tendency of rejecting introns is obvious. Note that the diagram depicted is not a phylogenetic representation. Signal sequences = ---. Exon = closed box. Intron: intra-repeat intron = ---, Inter-repeat intron = -.-.-. In multi-repeat structures the repeats are underlined. Vertebrates (α,β,Mb) [2]. Bacteria (V = *Vitreoscilla*) [34]. (Ec = *Echerichia coli*) [30]. (Nc = *Nostoc commune*) [35]. Yeast (Sc = *Saccharomyces cerevesiae*) [29]. Other abbreviations as in Fig. 1.

Moreover, the occurrence of the central intron conceivably hinders the concerted evolution of the haem binding region. Therefore, it is more likely that it would be eliminated in the course of evolution instead of being acquired.

The conservation of intra-repeat introns in the B and G helices of *Pseudoterranova* (repeat 2) and *Artemia* (T4) genes at the conserved positions supports their derivation from the ancestral pattern (Fig. 3). Intron deletion in the globins of bacteria, yeasts and *Chironomus* (except *IIb*) has reached the stage where there are none left [2,29,30,35].

The ancestral Hb molecule was presumably an intracellular molecule, since it is present in yeasts, bacteria and protozoa. Derivation of extracellular globins would have required particular adaptations: e.g., the recruit-

Table I

Alternative alignment of nematode A and B helices based on the location of inter-repeat introns

	HC/Linker/Pre-A	A	AB	B
	1	1		1
	123456789012	1234567890123456		1234567890123456
Mbw	V-L	SEGEWQLVLHVWAKVE	A	DVAGHGQDILIRLFKS
Ct3	L	SADQISTVQASFDKVK	G	----DPVGILYAVFKA
As1	NKT-RELCM	KSLEHAKVDTSNEARQ	-	----DGIDLYKHMFEN
As2	NKHGRHQCM	RSLQHIDIGHSETAKQ	-	----NGIDLYKHMFEN
Pd1	SKT-RELCM	KSLEHAKVGTSKEAKQ	-	----DGIDLYKHMFEH
Pd2	RHSVRDHCM	NSLEYIAIGDKEHQKQ	-	----NGIDLYKHMFEH
Ce	SMNRQEISDLCV	KSLEGRMVGTEAQNIE	-	----NGNAFFRYFFTN
Tc	AKSDEEIRKDAL	SALDVVPLGSTPEKLE	-	----NGREFYKYFFTN

Ct3, *Chironomus III*; As1 and 2, *Ascaris suum* repeat 1 and 2; Tc, *Trichostrongylus colubriformis* globin (36); rest of abbreviations as in Fig. 1.

ment of a leader sequence and an increase in M_r to minimise excretion.

The equivalence of the inter-repeat introns suggests that they are derived from an original intron separating the leader sequence from the globin coding sequence. A potential mechanism has been described for the *Barbatia* globin gene [5].

Acknowledgements: We thank Dr. Waterston and Dr. Sulston for making the *C. elegans* globin gene sequence available to us before publication. The Belgian National Science Foundation (NFWO) is greatly acknowledged for grants to JV and LM. CNAT thanks the NZ lottery grants board and AJM thanks the NZ health research council for grants.

REFERENCES

[1] Dickerson, R.E. and Geis, I. (1983) in Hemoglobin: Structure, Function, Evolution and Pathology (Dickerson, R.E. and Geis, I., Eds.) Benjamin Cummings, Menlo Park, California.
[2] Hardison, R.C. (1991) in Evolution at the Molecular Level (Selander, R.K., Clark, A.G. and Whittam, T.S., Eds.) pp. 272–290, Sinauer Associates, Sunderland.
[3] Lewin, R. (1985) Science, 226, 328.
[4] Dixon, B., Walker, B., Kimmings, W. and Pohajdak, B. (1992) J. Mol. Evol. 35, 131–136.
[5] Naito, Y., Riggs, C., Vandergon, T.L. and Riggs, A. (1991) Proc. Natl. Acad. Sci. USA 88, 6672–6676.
[6] Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R. and Waterston, R. (1992) Nature 356, 37–41
[7] Waterston, R. (1991) (personal communication).
[8] Bray, J.A. (1991) MSc Thesis, University of Otago
[9] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
[10] Bashford, D., Chothia, C. and Lesk, A.M. (1987) J. Mol. Biol. 196, 199–216
[11] Goodman, M., Braunitzer, G., Kleinschmidt, T. and Aschauer, M. (1983) Hoppe-Seyler's Z. Physiol. Chem. 364, 205–217.
[12] Jhiang, S.M. and Riggs, A.F. (1989) J. Biol. Chem. 264, 19003–19008.
[13] Fisher, W.K., Gilbert, A.T. and Thompson, E.O.P. (1984) Aust. J. Biol. Sci. 37, 191–203.
[14] Vainstein, B., Harutyunyan, E., Kuranova, I.P., Borisov, V.V. Sosfenovn N.I., Pavlovsky, A.G., Grebenko, A.I. and Nebrasov, Y.B. (1978) *Kristallografiya* 23, 517–526.
[15] Riggs, C.K. and Riggs, A.F. (1990) in: Invertebrate Dioxygen Carriers (Préaux, G. Ed.) Leuven University Press, Leuven, Belgium, pp. 57–60.
[16] Dixon, B., Walker, B., Kimmings, W. and Pohajdak, B. (1991) Proc. Natl. Acad. Sci. USA 88, 5655–5659.
[17] Trotman, C.N.A., Manning, A.M., Moens, L. and Tate, W.P. (1991) J. Biol. Chem., 266, 13789–13795.
[18] Richmond, T.J. and Richards, F.M. (1978) J. Mol. Biol. 119, 537–555.
[19] Goodman, M., Pedwaydon, J., Czelusniak, J., Suzuki, T., Gotoh, T., Moens, L., Shishikura, F., Walz, D. and Vinogradov, S. (1988) J. Mol. Evol. 27, 236–249.
[20] Peeters, K., De Baere, I. and Moens, L. (1992) in: International Congress on Invertebrate Dioxygen Carriers, Book of Abstracts, II-P07.
[21] De Baere, I., Liu, L., Moens, L., Van Beeumen, J., Gielens, C., Richelle, J., Trotman, C., Fine, J., Gerstein, M. and Perutz, M. (1992) Proc. Natl. Acad. Sci. USA 89, 4638–4642.
[22] Emmons, S. (1988) in: The Nematode *Caenorhabditis elegans*, Cold Spring Harbor Laboratory, pp. 47–79.
[23] Jellie, A. (1992) M.Sc. Thesis, University of Otago.
[24] Gō, M. (1981) Nature 291,90–92.
[25] Gō, M. (1983) Proc. Natl. Acad. Sci. USA 80, 1964–1968.
[26] Lewin, B. (1990) Genes IV, Oxford University Press, New York and Cell Press, Cambridge, MA.
[27] Kao, W. and Bergstrom, G. (1992) in: International Congress on Invertebrate Dioxygen Carriers, Book of Abstracts, II-L4.
[28] Gilbert, W. (1987) Cold Spring Harbor Symp. Quant. Biol. 52, 901–905.
[29] Zhu, J. and Riggs, A.F. (1992) Proc. Natl. Acad. Sci. USA 89, 5015–5019.
[30] Vasudevan, S.G., Armarego, W.L.F., Shaw, D.C., Lilley, P.E., Dixon, N.E. and Poole, R.K. (1991) Mol. Gen. Genet. 226, 49–58.
[31] Titchen, D.A., Glenn, W.K., Nassif, N., Thompson, A.R., and Thompson, E.O.P. (1991) Biochim. Biophys. Acta 1089, 61–67.
[32] Antoine, M. and Niessing, J. (1984) Nature, 310, 795–798.
[33] Hyldig-Nielsen, J., Jenssen, E.O., Paludan, K., Wiborg, O., Garret, R., Jorgensen, P. and Marker, K.A. (1982) Nucleic Acids Res. 10, 689–701.
[34] Wakabayashi, S., Matsubara, H. and Webster, D.A. (1986) Nature 322, 481–483
[35] Potts, M., Angeloni, S.V., Ebel, R.E. and Bassam, D. (1992) Science 256, 1690–1692.
[36] Frenkel, J.M., Dopheide, T.A., Wagland, B.M. and Ward, C.W. (1992) Mol. Biochem. Parasitol. 50, 27–36.