Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 8670–8675, July 1998
Evolution

Antarctic fish hemoglobins: Evidence for adaptive evolution at subzero temperature

LUCA BARGELLONI, STEFANIA MARCATO, AND TOMASO PATARNELLO*

Dipartimento di Biologia, Università di Padova, Via Ugo Bassi, 58/B I-35131 Padova, Italy

Communicated by Ernst Mayr, Harvard University, Cambridge, MA, May 20, 1998 (received for review November 15, 1997)

ABSTRACT **Notothenioids represent a large group of marine teleosts that are mostly endemic to the Antarctic Ocean. In this environment, the low metabolic demand and the high oxygen concentration reduce the need for hemoglobin(s) [Hb(s)]. The extreme condition is represented by the icefish (Channichthyidae, Notothenioidei), the only vertebrates that lack Hb. We obtained the nucleotide sequence coding for the β-globin chain of the single major Hb form in six red-blooded notothenioids. These included *Gymnodraco acuticeps*, one of the closest species to the Hb-less icefish, which is also the only known fish having a single Hb without Bohr effect. This species shows a higher rate of nonsynonymous substitutions (K_A), in contrast with the homogeneity of synonymous substitution (K_S) rates, and K_A/K_S ratios significantly greater than one in the majority of comparisons. These results are suggestive of positive selection, diversifying the single major Hb toward specialized functions. A single Hb that is free to diversify means that its role in routine oxygen transport can be reduced in the presence of a combination of physiological, ecological, and environmental factors. Although a reduced "routine" function for Hb, as is apparent in *G. acuticeps*, might, indeed, evoke the lack of Hb in icefish, evidence of diversifying selection reported here is at variance with the hypothesis of a simple trend from a single Hb toward the Hb-less condition.**

Over the past 55 million years, the temperature of the Southern Ocean has undergone a progressive reduction, from ≈20°C to the present-day extremes (−1.8°C; ref. 1). During the boundary between Oligocene and Miocene, about 22–25 million years ago, with the opening of the Drake passage and the formation of a circumpolar hydrographic barrier termed the Polar Front, the Antarctic marine environment became effectively isolated (2). This isolation greatly limited the opportunities for migration, forcing the shallow-water fish fauna to either adapt to the changing climate or become extinct. Although the fate of most species was, indeed, extinction, one group of teleost fish, the Notothenioidei, succeeded in adapting to the challenging environmental conditions, namely, low temperature, the presence of sea ice, habitat reduction, and extreme seasonality of primary production (3). Starting from a benthic ancestor, notothenioids underwent a remarkable radiation, which led them to fill numerous ecological niches, including the pelagic realm. Mostly endemic to the Southern Ocean, the notothenioid fish now dominate, in terms of biomass and number of species, the fish fauna inhabiting the waters surrounding Antarctica (4).

In this environment, the extremely low temperature greatly increases the viscosity of bodily fluids. To cope with the problem, notothenioids have evolved a diminished hematocrit and a reduced concentration of hemoglobin (Hb; ref. 5). This is made possible by a lower metabolic demand and a higher oxygen solubility, which are associated with the cold temperature. The tendency toward reduction of oxygen carriers reaches its extreme with the notothenioid family Channichthyidae (icefishes), which are the only vertebrates completely lacking Hb (6). Even red-blooded notothenioid fishes have a reduced amount of Hb, compared with temperate species, and generally show a single major form, accounting for 95–99% of total Hb (7). Multiplicity of Hb forms in fish is usually linked to the need for dealing with a mutable environment or different habitats (8). For this reason, the presence of a single Hb form has been regarded as the consequence of a narrowed role for oxygen carriers in notothenioids, possibly resulting from the benthic mode of life, as well as from the peculiarity of the environmental conditions (5, 9). The ecological importance of notothenioids in the Antarctic marine ecosystem and their remarkable adaptations to this extreme environment have meant that great efforts have been devoted to study notothenioid physiology, with special attention to the system of oxygen transport, and a great wealth of data at the anatomical, functional, and biochemical level has been produced. The next step was to address the issue of notothenioid Hb evolution by looking directly at globin genes.

The primary amino acid sequences of the α- and β-chains are known for the major Hb of several notothenioid species (7, 10), but no information is available at the DNA level, except for a single species, *Notothenia coriiceps neglecta* (11). By using a PCR-based approach (see *Methods*), we have obtained the complete nucleotide coding sequence of the major β-globin chain for additional six notothenioid species. Their phylogenetic relationships are known from both morphological (4) and molecular studies (ref. 12; Fig. 1*A*), and this knowledge offers the opportunity for a comparative approach to explore adaptive differences among lineages. Four of the species studied (*Pagothenia borchgrevinki*, *Trematomus newnesi*, *Trematomus bernacchii*, and *Trematomus hansoni*) are closely related, belonging to the subfamily Trematominae, but have largely diversified ecological behaviors. *T. hansoni* and *T. bernacchii* retain a benthic habit, whereas *T. newnesi* is semipelagic, living on the bottom but exploiting the water column, and *P. borchgrevinki* is cryopelagic, being associated with the immediate undersurface of ice, where it actively feeds. The other two species, *Cygnodraco mawsoni* and *Gymnodraco acuticeps*, are representative of the family Bathydraconidae, the closest sister group to the Hb-less icefish. Both are benthic, but *G. acuticeps* is a rather active ambush predator living in shallow water, in contact with anchor ice, whereas *C. mawsoni* is much less active and present at greater depth (110–300 m).

Abbreviations: Hb, hemoglobin; RACE, rapid amplification of cDNA ends; MP, maximum parsimony; ML, maximum likelihood; TS, transition; TV, transversion; K_A, number of substitutions per nonsynonymous site; K_S, number of substitutions per synonymous site.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF067566–AF067571).

*To whom reprint requests should be addressed. e-mail: patarnel@civ.bio.unipd.it.

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

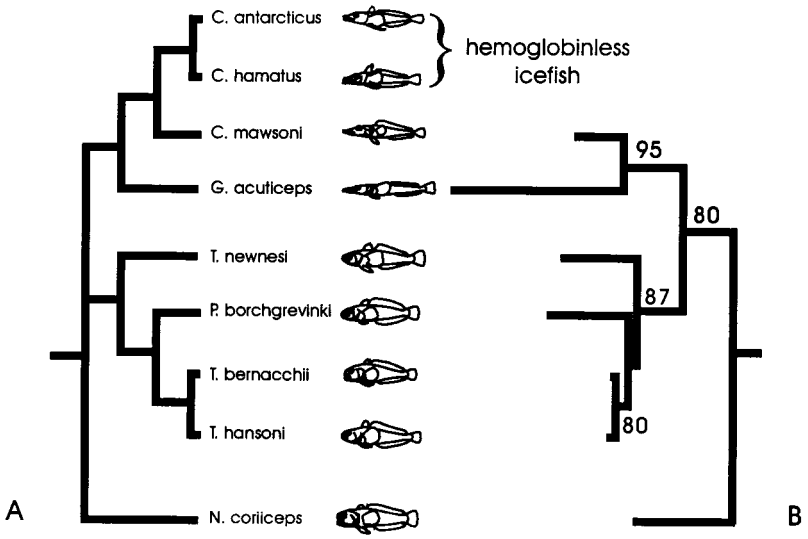

FIG. 1. (*A*) Phylogenetic relationships among the studied species, based on mitochondrial DNA sequences (12). Two species (*Chionodraco hamatus* and *Cryodraco antarcticus*) belonging to the Hb-less family Channichthyidae were also included. The systematic position of *N. coriiceps* was unresolved, as reflected by the basal polytomy. (*B*) Molecular phylogram of β-globin genes of the notothenioid species included in this study. The tree topology was identical under all the reconstruction methods used. The minimum evolution tree (based on pairwise nucleotide distances estimated with Jukes–Cantor correction for multiple hits) is presented. The tree was rooted with *S. salar* as outgroup. Bootstrap values higher than 50 are shown above branches.

These six species display interesting differences in their Hbs (7). Some of them, namely, *T. hansoni*, *T. bernacchii*, and *C. mawsoni*, have a single major Hb [Hb1, representing 95–97% of total Hb, formed by two identical α- and β-chains (α1 and β1, respectively)], with normal Bohr and Root effects. These effects are very weak for the Hb1 (70%, α1–β1) of *T. newnesi*; this species, however, has a second Hb component (HbC, 25% α1–β2) that shows a normal Bohr effect. *P. borchgrevinki* seems to have the highest multiplicity observed in notothenioids thus far, with five Hbs (Hb0, Hb1, Hb2, Hb3, and HbC), with the major component (Hb1, 70%) showing only weak Bohr and Root effects (13). The most striking, however, is *G. acuticeps*, which represents the only fish species with a single Hb (Hb1, 99%, α1–β1), completely lacking Bohr effect (7).

A comparative analysis of nucleotide sequences might shed light on the evolutionary processes experienced by notothenioid β-globins. Comparisons at the DNA level, in fact, enable the distinction between synonymous and nonsynonymous nucleotide changes. Synonymous ("silent") substitutions at coding genes make no change to the protein sequence encoded and appear to be largely neutral (particularly in vertebrate genes). In contrast, nonsynonymous substitutions result in amino acid replacements and are subjected to both negative and positive selection. Silent/replacement substitution patterns might, therefore, give helpful indications on the selective forces acting on the genes analyzed (14). Indeed, in the case of notothenioid β-globin, this kind of analysis provided evidence for positive selection operating to diversify the major Hb, possibly in relation to different ecological habits.

MATERIALS AND METHODS

Reverse Transcriptase–PCR with Degenerate Primers. Total RNA was extracted from spleen tissue with a commercial kit (RNA-Fast II; Molecular Systems). One to two micrograms of RNA was used to perform first strand cDNA synthesis with avian myeloblastosis virus (AMV) reverse transcriptase and random hexamers under standard conditions. Published amino acid sequences of notothenioid β-globin were aligned, and regions of high conservation were used to design PCR primers. To take into account genetic code degeneracy, inosine (at 4-fold degenerate sites) or multiple bases (at 2-fold degenerate sites) were included in the oligomers. PCR amplification with degenerate primers yielded a single product 350 bp long, representing the central portion of the β-globin coding sequence.

PCR products were cloned into a vector with T-overhangs, and two to three clones of each insert were cycle-sequenced in both directions.

5′ and 3′ Rapid Amplification of cDNA Ends (RACE). On the basis of the obtained sequences, new primers were designed, with enhanced specificity. These oligomers were oriented to perform 3′ or 5′ RACE. Both reactions were performed with a commercial kit (3′ RACE kit or 5′RACE kit; BRL Life Sciences), following manufacturer's instructions, with slight modifications. Products of the 3′ RACE were directly sequenced (without prior cloning). For 5′ RACE products, cloning was needed to obtain reliable sequences.

PCR Amplification and Direct Sequencing of the Complete Coding Sequence. The sequence of the complete coding region of all species was obtained by first joining sequence information from reverse transcriptase–PCR, 3′ RACE, and 5′RACE steps. Then, by using novel primers, we amplified and directly sequenced the complete region, to confirm previous sequence information. Positions and sequences of all primers, as well as PCR conditions, are available from T.P. upon request.

Sequence Analysis. The obtained nucleotide sequences (GenBank accession nos. AF067566–AF067571) plus the β-globin cDNA sequence of *N. coriiceps* (U09187) and *Salmo salar* (X69958) were aligned with CLUSTALW (15). Only a single insertion–deletion event was present (corresponding to an additional amino acid residue in *S. salar*), which was excluded from subsequent analyses. Reconstructions of phylogenetic relationships among β-globin genes were performed with several tree-building methods: maximum parsimony (MP), maximum likelihood (ML), minimum evolution, and neighbor-joining. MP, ML, and minimum evolution were implemented in PAUP Version 4.0d64 (16), and neighbor-joining was implemented in MEGA Version 1.02 (17).

Different weighting schemes were used in MP accounting for differences between transitions (TSs) and transversions (TVs). Similarly, different nucleotide models in ML and different distances in minimum evolution and neighbor-joining were used, incorporating parameters such as base

frequencies and TV/TS ratio. The robustness of each node was assessed under all methods with a bootstrap approach (18), with 2,000 replicates. On the basis of the obtained topology, nucleotide sequences at internal nodes were reconstructed with a ML method, as implemented in PAML Version 4.3B (19).

Rate homogeneity was tested in pairwise comparisons at the nucleotide and amino acid level with different approaches. A relative rate test was conducted separately on synonymous and nonsynonymous nucleotide substitutions with two distinct methods. In the first, distances at synonymous and nonsynonymous sites, as well as their variances and covariances, were estimated between the two examined sequences and also in relation to a third reference sequence (20). Statistical significance of rate differences were assessed with a *t* test with infinite degrees of freedom. The second method is implemented in the program CODRATES (21) and is based on a likelihood ratio test between likelihood values obtained under the alternative hypotheses of constant or independently varying silent, as well as replacement rates in the frame of a ML model of codon evolution. In both cases, *N. coriiceps* β-globin was used as reference sequence.

Rate differences were also estimated with reconstructed sequences. In this case, there was no need to use a third sequence as reference or to calculate distance covariances along the common branch leading to the outgroup. Distances were estimated for each of the two compared sequences starting from the common ancestral sequence and statistically significant differences assessed with *t* tests.

A (two-taxa) relative rate test was also conducted directly on the deduced amino acid sequences, as implemented in the CLUSTER-TEST program of Takezaki *et al.* (22), with different amino acid distances.

With regard to the estimate of synonymous and nonsynonymous number of substitutions per site, several methods, including those of Nei and Gojobori (23), Li (24), Ina (25), and Muse (26), were used. Statistical deviation of the ratio of the number of substitutions per nonsynonymous site (K_A) to the number of substitutions per synonymous site (K_S) from the null hypothesis ($K_A/K_S = 1$) was assessed by calculating the deviation of $K_A - K_S$ from zero with a *t* test (27).

To determine K_A/K_S, we also used a different method. Our approach was based on a ML model of codon evolution (28–31). This likelihood approach should allow the incorporation of more realistic models of substitution, taking TS-TV bias, unequal base frequencies, and codon bias into account. A detailed discussion of the model is beyond the scope of this report; the reader should refer to Yang (30) and Yang and Nielsen (31) for a precise description. Briefly, as implemented in the program PAML (19), this ML method makes the estimation of K_A/K_S value(s) possible under different models, allowing for different levels of heterogeneity of K_A/K_S ratios among lineages. The simplest model assumes a single K_A/K_S ratio, which is the same for all branches of the phylogenetic tree. The most general model allows estimation of K_A/K_S values separately for each branch along the tree. Several intermediate models are possible between these two extremes, allowing only selected K_A/K_S values to be estimated independently. The fit of different models to the data might then be evaluated in a likelihood ratio test (30). In all elaborations of our data set, two different methods were used to estimate codon frequencies. Frequencies were either calculated from the average nucleotide frequencies at each of the three codon positions or left as free parameters to be estimated by the ML procedure (31).

RESULTS AND DISCUSSION

In all the species analyzed, the β-globin coding region was the same size (438 nucleotides, 146 codons), whereas the 3′ untranslated region yielded a sequence that was variable in length [range, 103–120 bp, excluding the poly(A) sequence]. Only part of the 5′ untranslated region (15–40 bp, depending on the species) was obtained, probably because of premature termination in the reverse transcription during the 5′ RACE. When possible, protein sequences deduced from cDNA sequences were also compared with the amino acid sequences, which were available in the literature and obtained directly sequencing the protein. Generally a good agreement was found, with a single difference at position 134 of *C. mawsoni* β-globin sequence, where we observed a leucine instead of a methionine.

Because both untranslated regions were difficult to align with the homologous regions in *S. salar*, they were excluded from the data set, and only the complete coding region (438 bp) was used in the phylogenetic analyses.

To reconstruct the relationships among β-globin genes, we used several methods of phylogenetic inference (see *Methods*). The same gene tree was obtained (Fig. 1*B*) with all methods. This tree agreed with the tree resulting from the analysis of mitochondrial ribosomal RNA genes (Fig. 1*A*). The above evidence suggests that the globin genes examined are orthologous, i.e., divergence is not caused by gene duplication but traces back to a speciation event. On the basis of the obtained topology, ancestral sequences were also inferred by means of MP and ML methods and used in the subsequent analyses.

The constancy of evolutionary rate between lineages was then tested separately for synonymous and nonsynonymous substitutions at the nucleotide level. A rate test was performed on nucleotide substitutions with different methods, always with *N. coriiceps* as reference taxon. In all pairwise comparisons, silent mutation rates proved to be homogeneous. In contrast, when nonsynonymous substitution rates were compared between lineages, *G. acuticeps* showed a significantly higher rate of replacement changes with respect to *T. hansoni*, *T. bernacchii*, and *C. mawsoni* (Table 1). *P. borchgrevinki* and *T. newnesi* fell between these extremes, having nonsynonymous rates that were higher (though not significantly) than those of *T. hansoni*, *T. bernacchii*, and *C. mawsoni* but lower than that of *G. acuticeps*.

When sequences were analyzed at the protein level, *P. borchgrevinki* also displayed a significantly higher rate of substitution with respect to *T. hansoni* and *T. bernacchii* (Table 1).

Given the overall constancy of synonymous rates, it can be reliably assumed that all lineages have a similar "neutral" substitution rate. Under this hypothesis, differences in the rate of amino acid substitutions are only explained as a consequence of selective forces acting on the same gene in a different way, depending on the species examined.

Table 1. Pairwise comparisons of substitution rates

Species compared	Replacement substitution rate	Amino acid substitution rate	K_A/K_S ratio
Ga vs. Cm	Ga > Cm*†‡§	Ga > Cm**¶‖	5.0**
Ga vs. Tb	Ga > Tb*†‡§	Ga > Tb**¶‖	2.3*
Ga vs. Th	Ga > Th*†‡§	Ga > Th**¶‖	2.3*
Pb vs. Tb	Pb > Tb*§	Pb > Tb*¶‖	1.3
Pb vs. Th	Pb > Th*§	Pb > Th*¶‖	1.0

Only statistically significant results are shown (*, $P < 0.05$; **, $P < 0.01$).
Ga, *G. acuticeps*; Cm, *C. mawsoni*; Tb, *T. bernacchii*; Th, *T. hansoni*; Pb, *P. borchgrevinki*.
†Relative rate test according to Wu and Li (20).
‡ML test using CODRATES according to Muse and Gaut (21).
§Rate test using reconstructed ancestral sequences.
¶Relative rate test with amino acid distance (Poisson correction).
‖Relative rate test with amino acid gamma distance (parameter $a = 0.3$). K_A/K_S ratios relative to pairwise comparisons are also reported.

In most cases, replacement substitutions are likely to be deleterious and subjected to purifying selection. In a few cases, however, some amino acid replacements are favorable and positively selected. A higher rate of nonsynonymous substitutions can be interpreted as the consequence of either positive selection or reduced negative selection due to relaxed constraints. A possible way to discriminate between these two alternative hypotheses is to estimate K_A and K_S. In most protein-coding genes, negative selection on replacement substitutions is usually prevalent; hence, K_A is expected to be lower than K_S, and their ratio (K_A/K_S) is expected to be less than one. However, if advantageous amino acid mutations occur, they might reach fixation faster than neutral substitutions, inflating K_A. Hence, a K_A/K_S ratio greater than one is indicative of strong adaptive selection (10). This is the case for *G. acuticeps*, in which replacement substitutions always outnumber the silent ones (K_A/K_S falls in the range of 1.2–5.0), and K_A/K_S is significantly greater than one in the direct comparison with its sister species ($K_A/K_S = 5.0$; $P < 0.01$). This evidence suggests that the increased rate of amino acid substitutions of *G. acuticeps* β-globin was driven by positive selection toward diversification. Though the most impressive, the case of *G. acuticeps*, however, is not unique. Also, *P. borchgrevinki* shows an amino acid substitution rate that is significantly higher ($P < 0.05$) than those of *T. hansoni* and *T. bernacchii*, with $K_A/K_S \geq 1$ (range, 1–1.3). In general, most pairwise comparisons yielded K_A/K_S ratios higher than one. The lowest values were observed when *T. hansoni* and *T. bernacchii* were compared ($K_A/K_S = 0.29$). Values less than one were obtained also for the K_A/K_S ratios between *N. coriiceps* and all other species (average $K_A/K_S = 0.8$) when *G. acuticeps* was excluded from the analysis to avoid the overwhelming effect of the large variation displayed by this species. This pattern is further confirmed by the analysis on ancestral sequences. K_A and K_S were, in fact, estimated for each branch with the reconstructed sequences at ancestral nodes, to partition molecular variation along the globin gene tree (Fig. 2). Again, K_A was equal or greater than K_S for all branches with two exceptions: the branch leading to *T. hansoni* and the part of the tree connecting *N. coriiceps* β-globin to the ancestral node common to all the other notothenioids. In the latter case, the inferred K_A/K_S ratio was one of the lowest found here ($K_A/K_S = 0.45$).

All the results presented are based on the method of Nei and Gojobori (23) to calculate K_A and K_S because it best fits our data (low divergence, no TS-TV bias) and has the lowest variances. However, it should be noted that the results did not depend on the method used to estimate K_A and K_S. Different methods are known to produce different results, when distantly related sequences are compared, but when the distance is lower than 0.1–0.2, the result is insensitive to the effect of potential sources of error such as multiple hits, TS-TV, or codon bias (32). This is the case of notothenioid β-globin sequences, in which pairwise distances are almost always lower than 0.1 (mean $K_S = 0.044$, mean $K_A = 0.062$).

ML analysis of synonymous and nonsynonymous rate ratios provides further support to the above evidence. The ML approach evaluates the entire codon as unit of evolution, instead of the single nucleotide or amino acid site (28, 30, 31). This method allows the direct incorporation and, therefore, the correction for effects such as TS-TV and codon bias that are usually ignored with the approximate methods described above.

On the basis of the topology obtained for the β-globin genes (Fig. 1*B*), likelihood ratio tests were applied to test for the constancy of nonsynonymous to synonymous rate ratios among evolutionary lineages.

Different models can be tested with the ML approach. The two most extreme models assume either a unique K_A/K_S rate ratio for all branches of the tree (one-ratio model) or an independent K_A/K_S value for each branch (free-ratio model). Between these extremes, several intermediate models can be tested, based on different assumptions, depending on which branches are allowed to have a "free" rate ratio. When β-globin genes were analyzed, a three-ratio model was found to fit data significantly better than one-ratio model. With the one-ratio model, the log likelihood (l) value was $l_{one} = -974.9$, whereas under a model with three free ratios, we obtained $l_{three} = -971.6$. Because the latter model involves the estimation of three values, whereas the one-ratio model assumes a single K_A/K_S ratio, twice the log likelihood difference, $2\Delta l = 2(l_{three} - l_{one}) = 6.6$, can be compared with a χ^2 distribution with 2 (3 − 1) degrees of freedom. This allows a test of which model better fits the data. The difference between the two models is significant ($P < 0.05$) in favor of the three-ratio model, thus confirming that a significant heterogeneity in K_A/K_S ratios is present in notothenioid β-globins. The three-ratio model used allows the separate estimation three K_A/K_S values along the β-globin tree, as shown in Fig. 2: one value for the branch

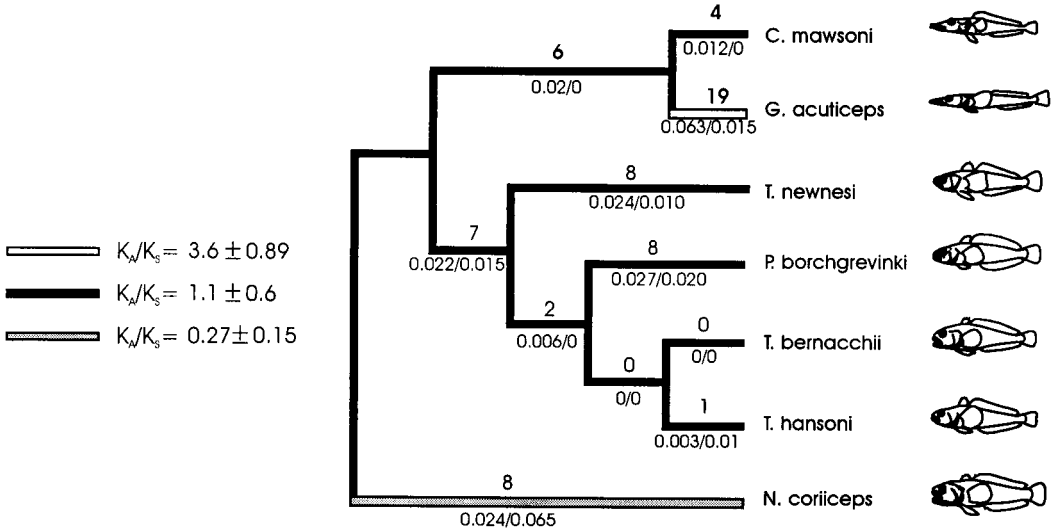

FIG. 2. K_A and K_S values were estimated for each branch of the globin tree with the reconstructed sequences at ancestral nodes. Numbers above lineages indicate the minimum number of amino acid substitutions to explain differences between reconstructed sequences. K_A/K_S ratios are shown below branches. Results from codon analysis with PAML (see text) are reported on the left. Different shades on branches (white, black, and gray) refer to the different K_A/K_S values estimated along the tree.

leading to *G. acuticeps*, one for the branch connecting *N. coriiceps* to the rest of globin sequences, and another for all the remaining branches. The three estimated values (and relative SEs) of K_A/K_S are 3.6 ± 0.8 (*G. acuticeps*), 0.27 ± 0.15 (*N. coriiceps*), and 1.1 ± 0.6 (the rest of the lineages). A model with all values free to vary was not significantly better than the three-ratio one ($P = 0.5$), and the simpler three-ratio model was preferred.

In summary, all the analyses we performed showed that replacement rates are not homogeneous among globin genes of different notothenioid species and that K_A/K_S values are larger than one in several branches, being significantly higher in *G. acuticeps*.

Hb plays a pivotal role in all vertebrates, providing oxygen delivery to peripheral tissues. It has often been maintained that, especially in fish, selection acts differently on Hbs of different species to match the different environmental and ecological requirements. For instance, the rainbow trout shows three major Hb forms (HbIV, HbI, and HbII); HbIV displays a strong Bohr effect, whereas HbI and Hb II are insensitive to pH (i.e., the Bohr effect is completely absent). It is thought that, during fast movement, the pH at the gills might drop too low for efficient oxygen uptake by those Hbs that exhibit the Bohr effect (8). In this case, the presence of an additional, pH-independent Hb component is, therefore, regarded as a specialized "adaptive" function for oxygen delivery at low pH conditions. Moreover, a recent study (33) examined the effect of temperature on the oxygen-binding properties of the Hbs of three notothenioid fish species. A reduced thermal sensitivity of oxygen binding has been proposed (7) as a general adaptive strategy of the oxygen transport system to subzero temperatures and interpreted as an "energy-saving" mechanism because only small amounts of energy are liberated during oxygenation–deoxygenation cycles. Fago *et al.* (33) found that Hbs lacking Bohr and Root effects show a heat of oxygenation that is largely independent on temperature. A reduced or absent Bohr effect, therefore, should not be regarded as a "loss of function." However, the adaptive value of such peculiar physiological conditions often remains elusive. Although, in some cases, it has been possible to assign functional "shifts" to single amino acid replacements (34), the structural basis of Hb properties such as Bohr and Root effects is still unclear, especially in fish Hbs. The same effect might be obtained in different species under different conditions, and a large number of amino acid residues are likely to be involved in either long-range or local interactions (35). Hence, the need arises for the study of adaptation at the molecular level, correlating results of DNA analyses with specific functional and ecological properties.

In the case of notothenioids β-globins, our study revealed significant variation in nonsynonymous rates of substitution. Evidence for different rates in β-globin sequence evolution also correlates quite well with the anatomy, physiology, and ecology of notothenioid species. In fact, higher rates result for those species having diversified their habitat (*P. borchgrevinki* and *T. newnesi*) and their physiology (*G. acuticeps*). In addition, replacement substitutions occur more frequently than silent ones in most lineages, reaching an extreme with *G. acuticeps.* In the latter case, the K_A/K_S ratio is significantly greater than one ($K_A/K_S = 3.6 \pm 0.8$), providing evidence for positive selection. Moreover, all methods used to estimate K_A/K_S assume that this ratio is constant over all codon sites, i.e., all amino acid positions of the β-globin sequence experience the same selective pressure. This assumption is extremely unrealistic because several amino acid residues are conserved to maintain the structure and function of the protein. Therefore, the criterion defining positive selection ($K_A/K_S > 1$) is often exceedingly restrictive. Functional constraints likely set an upper limit to amino acid mutations, whereas synonymous substitutions occur nearly equally at different phylogenetic distances. When distant phylogenetic groups are compared, negative selection usually obscures any evidence of adaptive events. This is the case for the value observed on the branch leading to *N. coriiceps* ($K_A/K_S = 0.27$ or 0.45, depending on the method used, see Fig. 2), indicative of the importance of performing K_A-K_S analyses in the appropriate evolutionary time frame. Furthermore, the assumption of constant rate of amino acid change across all protein sites is rejected when two alternative models are compared: a constant model of rates among sites *versus* a γ model with four categories of rates. The latter is significantly better at explaining variation at the amino acid level in the notothenioid β-globins ($P < 0.001$), with an estimated γ parameter $a = 0.24$. This suggests a remarkable heterogeneity in rates among sites. A reconstruction of mutation rate at each site (Fig. 3) revealed invariant, moderately, and fast-evolving positions. Invariant sites exactly corresponded to amino acid residues known to be essential for a proper Hb function, whereas sites showing a high rate of evolution should be regarded as possible target of positive selection.

In light of the above arguments, evidence of rate heterogeneity and high values of K_A/K_S should, therefore, be interpreted as an indication that adaptive selection had an important role in shaping the evolution of the major β-globin in

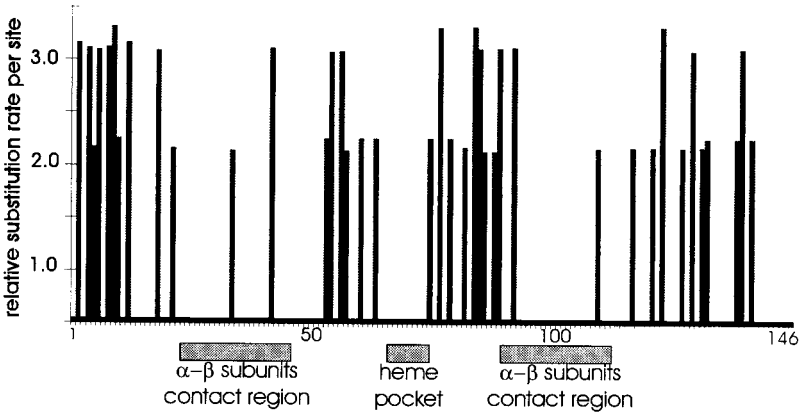

FIG. 3. Differences among sites in the rate of amino acid substitutions were examined generating a variability profile along the β-globin sequence. First, we used an MP approach to estimation of the parameter *a* of the γ distribution (36) with four categories of rate, based on the obtained topology (19). Second, substitution rates per site were estimated with a discrete γ model, as implemented in PAML. Evolutionary rates were scaled such that the mean of rates across sites was one. On the horizontal axis are reported the amino acid residues of the major notothenioid Hb (1–146). On the vertical axis are indicated the "relative" values of substitution rate at each site. Gray boxes, protein regions that are essential for proper Hb function.

several lineages besides *G. acuticeps*, possibly in relation to different ecological and/or physiological characters.

The hypothesis of diversifying selection acting on Hb, in the Antarctic environment, where the need of oxygen carriers is reduced might seem counterintuitive and relaxed selection could appear a more likely evolutionary scenario. However, evidence from the present study suggests caution against simplistic interpretations. As mentioned earlier, Hbs showing weak or absent Bohr and Root effects might have an adaptive role in peculiar conditions. In temperate fish, this occurs in the presence of several different Hb forms. On the contrary, a single Hb is generally present in notothenioids. Under the extreme Antarctic conditions, there might be a reduced role in routine oxygen transport for the single major Hb, which is, therefore, free to adapt to more specialized functions. In *T. newnesi* and *P. borchgrevinki*, this seems to be partially counterbalanced through the expression of other, minor, Hb forms. In *G. acuticeps,* despite the existence of a second conserved gene, only a single Hb (99%) is present. Nevertheless, its β-globin experienced a surprisingly rapid evolution.

To understand this striking evidence, it should be recalled that this species is one of the closest relatives to the Hb-less icefish. It has a scaleless skin, allowing cutaneous respiration, like channichthyids, has pseudochoanae, and has a particular ecological behavior (ambush predator), being usually sluggish. These peculiarities likely contributed to greatly reducing the need of routine oxygen transport. On the contrary, in agreement with the molecular evidence of diversifying selection on the β-globin gene, specialized Hb functions might still be required to face the consequences of the sudden bursts of activity due to its preying behavior and to provide the necessary thermal adaptation (*G. acuticeps* lives in contact with anchor ice, experiencing extremely low temperatures, as suggested by the high level of circulating antifreeze glycopeptides).

In conclusion, our comparative analysis reveals the unexpected complexity of notothenioid β-globin evolution, demonstrating that selection acts differently on different species. Although a reduced "routine" function for Hb, as apparent in *G. acuticeps*, might, indeed, evoke the lack of Hb in icefish, evidence of diversifying selection reported here is at variance with the hypothesis of a simple trend from a single Hb toward the loss of Hb.

Recently, Cocca and colleagues (11) found genomic remnants of α-globin genes in icefish but no traces of β-globin genes. This led to the hypothesis that a sudden deletion of the β-globin locus might have been the event leading to the Hb-less condition (11).

However, our evidence suggests that, despite being set free from routine functions in some species, Hb in red-blooded notothenioids may be finely tuned to match different physiological and ecological needs, and therefore, its adaptive importance should not be overlooked. Hence, rather than being the starting point, β-globin gene deletion is likely to be the end of the evolutionary process that led the icefish to completely abolish even the need of specialized Hb functions and to attain their unique Hb-less condition.

We thank Drs. B. Battaglia, A. Clarke, A. DeVries, G. di Prisco, A. Meyer, and L. Zane for useful discussions and comments about the manuscript. We also thank Drs. Y. Ina, S. Muse, M. Nei, D. Swofford, and Z. Yang for providing the computer programs used for data analysis. This work was supported by the Italian Program for Antarctic Research.

1. Clarke, A. (1983) *Oceanogr. Mar. Biol. Annu. Rev.* **21,** 341–453.
2. Kennett, J. P. (1982) *Marine Geology* (Prentice-Hall, Englewood Cliffs, NJ).
3. Clarke, A. & Johnston, I. A. (1996) *TREE* **11,** 212–218.
4. Eastman, J. T. (1993) *Antarctic Fish Biology* (Academic, San Diego).
5. Macdonald, J. A. & Wells, R. M. G (1991) in *Biology of Antarctic Fish*, eds. di Prisco, G., Maresca, B. & Tota, B. (Springer, Berlin), pp. 163–178.
6. Ruud, J. T. (1954) *Nature (London)* **173,** 848–850.
7. di Prisco, G., D'Avino, R., Caruso, C., Tamburrini, Camardella, L., Rutigliano, B., Carratore, V. & Romano, M. (1991) in *Biology of Antarctic Fish*, eds. di Prisco, G., Maresca, B. & Tota, B. (Springer, Berlin), pp. 269–274.
8. Perutz, M. F. (1983) *Mol. Biol. Evol.* **1,** 1–28.
9. MacDonald, J. A., Montgomery, J. C. & Wells, R. M. G. (1987) in *Advances in Marine Biology*, eds. Blaxter, J. H. S. & Southward, A. J. (Academic, London), Vol. 21, pp. 321–388.
10. Tamburrini, M., D'Avino, R., Fago, A., Carratore, V., Kunzmann, A. & di Prisco, G. (1996) *J. Biol. Chem.* **271,** 23780–23785.
11. Cocca, E., Ratnayake-Lecamwasam, M., Parker, S., Camardella, L., Ciaramella, M., di Prisco, G. & Detrich, H. W. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1817–1821.
12. Bargelloni, L., Ritchie, P. A., Patarnello, T., Battaglia, B., Lambert, D. L. & Meyer, A. (1994) *Mol. Biol. Evol.* **11,** 854–863.
13. di Prisco, G. (1997) in *Antarctic Communities: Species, Structure and Survival*, eds. Battaglia, B., Valencia, J. & Walton, D. W. H. (Cambridge Univ. Press, Cambridge, U.K.), pp. 251–260.
14. Sharp P. M. (1997) *Nature (London)* **385,** 111–112.
15. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
16. Swofford, D. PAUP*: Phylogenetic Analysis Using Parsimony Methods, Version 4.0d64, in press.
17. Kumar, S., Tamura, K. & Nei, M. (1993) MEGA: Molecular Evolutionary Genetics Analysis (Pennsylvania State Univ., University Park, PA), Version 1.02.
18. Felsenstein, J. (1985) *Evolution* **39,** 783–791.
19. Yang, Z. (1997) *Comput. Appl. Biosci.* **13,** 555–556.
20. Wu, C. I. & Li, W. H. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1741–1745.
21. Muse, S. V. & Gaut, B. S. (1994) *Mol. Biol. Evol.* **11,** 715–724.
22. Takezaki, N., Rzhetsky, A. & Nei, M. (1995) *Mol. Biol. Evol.* **12,** 823–833.
23. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
24. Li, W.-H. (1993) *J. Mol. Evol.* **36,** 96–99.
25. Ina, Y. (1995) *J. Mol. Evol.* **40,** 190–226.
26. Muse, S. V. (1996) *Mol. Biol. Evol.* **13,** 105–114.
27. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335,** 167–170.
28. Goldman, N. & Yang, Z. (1994) *Mol. Biol. Evol.* **11,** 725–736.
29. Nielsen, R. & Yang, Z. (1998) *Genetics* **148,** 929–936.
30. Yang, Z. (1998) *Mol. Biol. Evol.* **15,** 568–573.
31. Yang, Z. & Nielsen, R. (1998) *J. Mol. Evol.* **46,** 409–418.
32. Ina, Y. (1996) *J. Genet.* **75,** 91–115.
33. Fago, A., Wells, R. M. G. & Weber, R. E. (1997) *Comp. Biochem. Physiol.* **118B,** 319–326.
34. Golding, G. B. & Dean, A. M. (1998) *Mol. Biol. Evol.* **15,** 355–369.
35. Ito, N., Komiyama, N. H. & Fermi, G. (1995) *J. Mol. Biol.* **250,** 648–658.
36. Yang, Z. (1996) *TREE* **11,** 367–372.