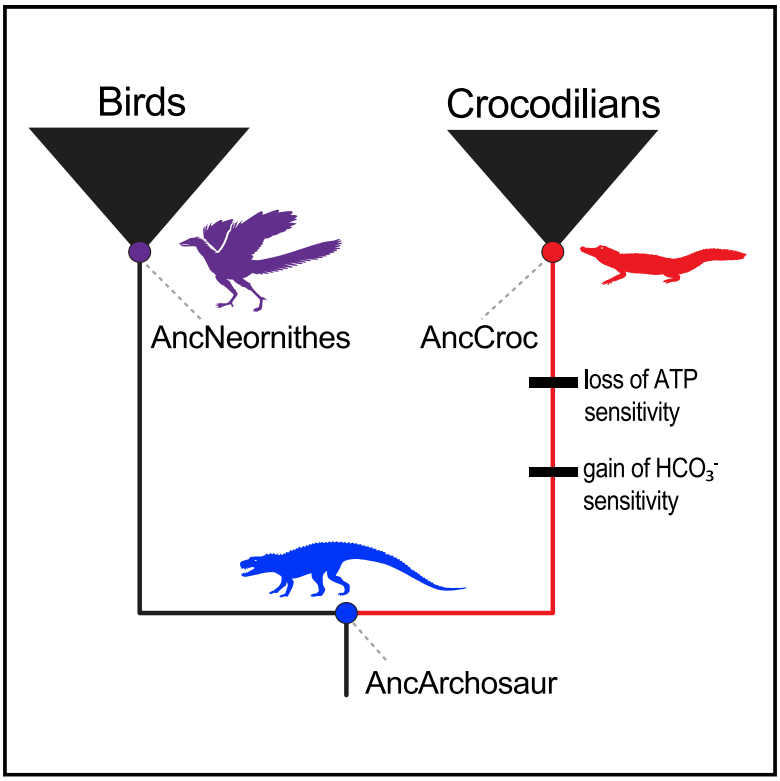
Article

Current Biology

Evolution and molecular basis of a novel allosteric property of crocodilian hemoglobin

Graphical abstract



Authors

Chandrasekhar Natarajan, Anthony V. Signore, Naim M. Bautista, Federico G. Hoffmann, Jeremy R.H. Tame, Angela Fago, Jay F. Storz

Correspondence

jstorz2@unl.edu

In brief

Natarajan et al. dissected the molecular basis of a unique mode of allosteric regulatory control in crocodilian hemoglobin. In addition to elucidating the evolution of a key biochemical adaptation, they demonstrate that gain of the novel protein function was attributable to numerous substitutions with indirect, context-dependent effects.

Highlights

- Crocodilians evolved a unique mode of allosterically regulating hemoglobin function
- Gain of new function and loss of ancestral function were not mechanistically coupled
- Gain of novel protein function evolved via numerous substitutions at disparate sites
- Many of the causative substitutions have highly indirect, context-dependent effects

Natarajan et al., 2023, Current Biology *33*, 98–108
January 9, 2023
https://doi.org/10.1016/j.cub.2022.11.049

CellPress

CellPress
OPEN ACCESS

Current Biology

Article

Evolution and molecular basis of a novel allosteric property of crocodilian hemoglobin

Chandrasekhar Natarajan,[1] Anthony V. Signore,[1,6] Naim M. Bautista,[1] Federico G. Hoffmann,[2,3] Jeremy R.H. Tame,[4] Angela Fago,[5] and Jay F. Storz[1,7,*]

[1]School of Biological Sciences, University of Nebraska, Lincoln, NE 68588, USA
[2]Department of Biochemistry, Molecular Biology, Entomology, and Plant Pathology, Mississippi State University, Mississippi State, MS 39762, USA
[3]Institute for Genomics, Biocomputing, and Biotechnology, Mississippi State University, Mississippi State, MS 39762, USA
[4]Drug Design Laboratory, Yokohama City University, Yokohama 230-0045, Japan
[5]Zoophysiology, Department of Bioscience, Aarhus University, DK-8000 Aarhus, Denmark
[6]Present address: Department of Biological Sciences, University of Manitoba, Winnipeg, MB R3T 2N2, Canada
[7]Lead contact
*Correspondence: jstorz2@unl.edu
https://doi.org/10.1016/j.cub.2022.11.049

SUMMARY

The extraordinary breath-hold diving capacity of crocodilians has been ascribed to a unique mode of allosterically regulating hemoglobin (Hb)-oxygenation in circulating red blood cells. We investigated the origin and mechanistic basis of this novel biochemical phenomenon by performing directed mutagenesis experiments on resurrected ancestral Hbs. Comparisons of Hb function between the common ancestor of archosaurs (the group that includes crocodilians and birds) and the last common ancestor of modern crocodilians revealed that regulation of Hb-O_2 affinity via allosteric binding of bicarbonate ions represents a croc-specific innovation that evolved in combination with the loss of allosteric regulation by ATP binding. Mutagenesis experiments revealed that evolution of the novel allosteric function in crocodilians and the concomitant loss of ancestral function were not mechanistically coupled and were caused by different sets of substitutions. The gain of bicarbonate sensitivity in crocodilian Hb involved the direct effect of few amino acid substitutions at key sites in combination with indirect effects of numerous other substitutions at structurally disparate sites. Such indirect interaction effects suggest that evolution of the novel protein function was conditional on neutral mutations that produced no adaptive benefit when they first arose but that contributed to a permissive background for subsequent function-altering mutations at other sites. Due to the context dependence of causative substitutions, the unique allosteric properties of crocodilian Hb cannot be easily transplanted into divergent homologs of other species.

INTRODUCTION

How do novel protein functions evolve in an incremental, step-by-step fashion and to what extent is the gain of new function mechanistically coupled with the loss of ancestral function? These are fundamental questions in molecular evolution, and they can be addressed with a protein-engineering approach that permits the identification and functional characterization of causative mutations. Here, we report discoveries regarding the molecular basis of a key physiological innovation during vertebrate evolution. We used ancestral protein resurrection in conjunction with site-directed mutagenesis experiments to dissect the molecular basis of a unique mechanism for allosterically regulating the O_2 affinity of crocodilian hemoglobin (Hb).

Crocodilians are able to remain submerged underwater for extraordinarily long periods of time—a physiological capacity that allows some species, such as the Nile crocodile (*Crocodylus niloticus*), to kill large, mammalian prey like wildebeest and zebras by dragging them underwater and drowning them.[1] This remarkable capacity for breath-hold diving has been attributed to the fact that the O_2-affinity of crocodilian Hb is regulated by the inhibitory effects of bicarbonate ions.[2–7] During submergence, there is a marked increase in the red cell concentration of bicarbonate (derived from the hydration of CO_2), which reduces Hb-O_2 affinity, thereby promoting O_2 unloading to the cells of metabolizing tissues.[6] This is a highly efficient physiological strategy since the circulatory O_2 store can be almost fully depleted and the CO_2 carrying capacity of the blood is maximized. As stated by Perutz,[4] "It is surprising that the crocodilian Hbs' simple and direct reciprocating action between oxygen and one of the end products of oxidative metabolism has not been adopted by other vertebrates." The bicarbonate sensitivity of crocodilian Hb may also play an important role in tissue O_2 delivery during the post-prandial "alkaline tide," as the secretion of hydrochloric acid into the stomach (which aids the digestion of skeletal material) involves an anion exchange across the gut lining that reduces the blood concentration of chloride while increasing the blood pH and bicarbonate concentration.[8–10]

The Hbs of jawed vertebrates are heterotetramers, consisting of two α-type subunits and two β-type subunits. Each of the four globin subunits contain a heme group with a ferrous (Fe^{2+}) iron atom that reversibly binds a single O_2 molecule. The Hb tetramer undergoes an oxygenation-linked transition in quaternary structure, whereby the two semi-rigid $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers rotate relative to one another during the transition between the deoxy (low-affinity [T]) conformation and the oxy (high-affinity [R]) conformation.[11] This oxygenation-linked shift in quaternary structure is central to the regulation of Hb-O_2 affinity by allosteric cofactors (non-heme ligands such as hydrogen ions, chloride ions, and organic phosphates) that are present in the red blood cell. Allosteric cofactors reduce Hb-O_2 affinity by preferentially binding to deoxyHb, thereby stabilizing the low-affinity T-state through the formation of additional hydrogen bonds and salt bridges within and between subunits. Among Hbs of jawed vertebrates, crocodilian Hb is unique in that its O_2 affinity is primarily modulated by bicarbonate ions rather than organic phosphates such as ATP.[2,3,5,7–9] The origin and mechanistic basis of this unique mode of allosteric regulatory control has remained a mystery.[4,12,13] Recent studies have made steps toward solving this mystery by demonstrating that bicarbonate and CO_2 are both potent allosteric effectors of crocodilian Hb and that they share common binding sites.[5,7] Thus, the effect of CO_2 can be conveniently studied as a proxy to investigate allosteric bicarbonate binding. Neither chloride ions nor ATP interfere with the strong CO_2 effect of adult crocodilian Hb,[7] suggesting that protein T-state binding sites for these allosteric effectors must be distinct.

On the basis of structural modeling, Perutz et al.[3] hypothesized that the binding site for organic phosphates in the positively charged central cavity between the β chains was co-opted for allosteric bicarbonate binding in crocodilian Hb, a shift that required a stereochemical reconfiguration of multiple residues. Perutz[4] suggested that the preferential binding of bicarbonate ions may have involved as few as three amino acid substitutions, two of which simultaneously eliminated residues directly involved in ATP binding. According to this hypothesis, two additional amino acid substitutions (1 in the α chain and 1 in the β chain) would be required to completely abolish the ancestral sensitivity to ATP. These hypotheses were tested by engineering crocodile-specific mutations into recombinant human Hb.[14] Contrary to Perutz's predictions, the study by Komiyama et al.[14] demonstrated that transplanting bicarbonate sensitivity into human Hb required a total of 12 amino acid substitutions (7 in the α chain, 5 in the β chain), most of which were concentrated in the symmetrical $\alpha_1\beta_2$ and $\alpha_2\beta_1$ intersubunit contact surfaces.

Although the study by Komiyama et al. provided important insights into possible structural mechanisms underlying bicarbonate sensitivity, one problem with the interpretation of such "horizontal" comparisons (swapping residues between homologous proteins of contemporary species) is that the focal mutations are introduced into a sequence context that may not be evolutionarily relevant. If mutations have context-dependent effects, then introducing crocodile-specific substitutions into human Hb may not recapitulate the functional effects of causative mutations on the genetic background in which they actually occurred during evolution (i.e., in the ancestor of modern crocodilians). Consequently, the particular substitutions required to transplant bicarbonate sensitivity into human Hb may be different from the substitutions that were responsible for the evolved bicarbonate sensitivity in the Hb of the crocodilian ancestor. An alternative "vertical" approach is to reconstruct and resurrect ancestral proteins to test the effects of historical mutations on the genetic background in which they actually occurred during evolution.[15–17] We used this approach to measure the functional effects of amino acid substitutions that occurred in the reconstructed ancestor of archosaurs, the group containing crocodilians and birds. Since bicarbonate sensitivity is a property specific to the Hbs of crocodilians, it must have evolved in the line of descent leading from the archosaur ancestor, which existed ∼240 million years ago in the Mid-Triassic, to the crocodilian ancestor, which existed ∼80 million years ago in the Late Cretaceous (Figure 1A).

RESULTS

Ancestral protein resurrection

We used a maximum likelihood (ML) approach to estimate the α- and β-globin sequences of three vertebrate ancestors: the most recent common ancestor of modern archosaurs ("AncArchosaur"), the ancestor of modern birds (the sister group of crocodilians; "AncNeornithes"), and the ancestor of modern crocodilians ("AncCroc") (Figures 1A, S1, and S2). The ML ancestral sequences were estimated with high levels of statistical confidence (Figures S3A–S3C).

We synthesized the ML α and β chain sequences of AncArchosaur, AncNeornithes, and AncCroc and cloned them into custom-designed expression plasmids.[18,19] We expressed and purified each of the three recombinant Hbs (rHbs), and we then measured oxygenation properties *in vitro*. The reconstructed AncArchosaur and AncCroc rHbs exhibited intrinsic O_2-affinities, cooperativities, and pH sensitivities (Bohr effect) that are typical of native Hbs from modern crocodilians and non-avian reptiles[5,7,20–23] (Table S1). To test whether functional inferences are robust to uncertainty in ancestral sequence estimates, we synthesized and tested alternative versions of AncArchosaur, AncNeornithes, and AncCroc that incorporated the alternative amino acids at each ambiguous site in the corresponding ML sequences. Specifically, we synthesized "AltAll" sequences by incorporating the "runner-up" amino acid at each site with a posterior probability < 0.80.[24] For AncArchosaur, AncNeornthes, and AncCroc, the AltAll α and β chain sequences differed from the corresponding ML ancestral sequences at 19, 4, and 11 sites, respectively (Figures S3A–S3C). In all cases, ML and AltAll versions of the same ancestral Hbs shared the same allosteric properties (Figures S3D–S3F), confirming that inferences about evolved changes in Hb function in the crocodilian lineage are robust to statistical uncertainty in the ancestral reconstructions. The key findings are that AncArchosaur and AncNeornithes rHbs exhibited sensitivity to ATP but not to bicarbonate, and conversely, the AncCroc rHb exhibited sensitivity to bicarbonate but not to ATP (Figures 1B and 1C). These results confirm that AncArchosaur and AncCroc bracket an evolutionary transition between two discrete, well-defined functional states.

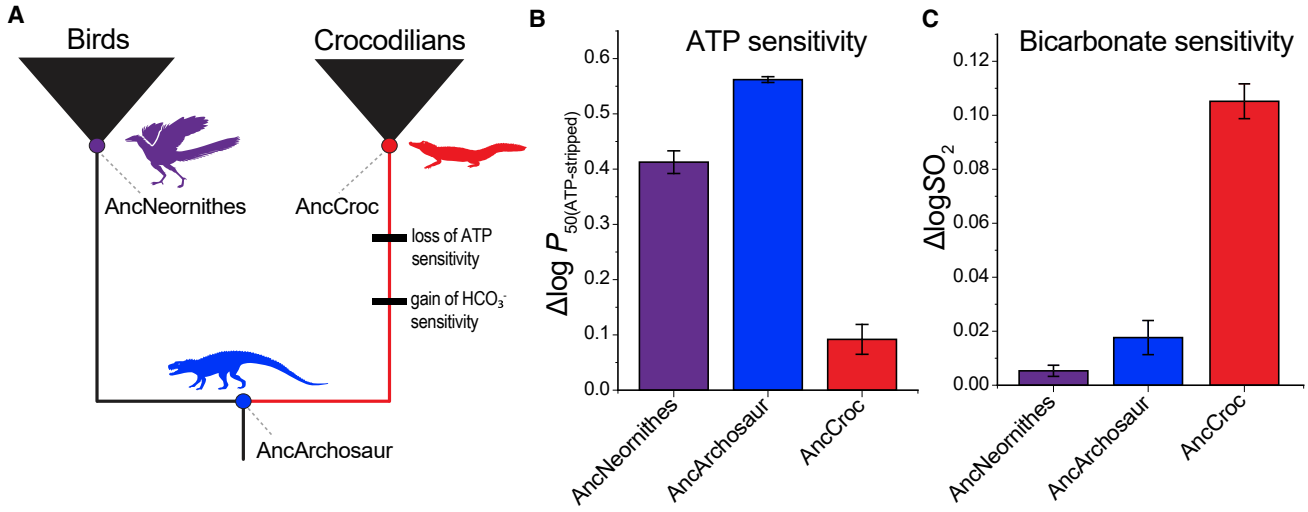


Figure 1. Resurrection and functional testing of ancestral hemoglobins

(A) Diagrammatic phylogeny of extant archosaurs, indicating nodes targeted for ancestral protein resurrection. The red branch connecting the common ancestor of modern archosaurs (AncArchosaur) with the ancestor of crocodilians (AncCroc) brackets the interval during which crocodilian Hb evolved the capacity to regulate O_2-affinity via allosteric binding of bicarbonate ions. In the ancestor of modern crocodilians, the gain of bicarbonate sensitivity was accompanied by a loss of ATP sensitivity (the ancestral mode of allosteric regulatory control in vertebrates), although the order in which these changes occurred is unknown.

(B and C) Relative to AncNeornithes and AncArchosaur, AncCroc evolved (B) a derived reduction in ATP sensitivity (measured as the difference in log-transformed values of P_{50} [the partial pressure of O_2 at which Hb is half-saturated] in the presence and absence of 0.2 mM ATP [25°C, 0.4 mM tetrameric Hb in 0.1 M HEPES, pH 7.4]) and (C) an increase in bicarbonate sensitivity (measured as the difference in log-transformed values of SO_2 [Hb-O_2 saturation] in the presence and absence of 1% CO_2 at a constant PO_2 equivalent to P_{50} in the absence of CO_2). Graphed values are means based on triplicate measurements (±SEM).

See also Figure S3 and Table S1.

Molecular changes underlying the loss of ATP sensitivity

The results of structural modeling suggest that ATP binding to the T-state of vertebrate Hb is mediated by direct interactions with a constellation of positively charged residues in the central cavity.[5,13,25] Structural comparisons between AncArchosaur and AncCroc Hbs suggested that the croc-specific loss of ATP sensitivity may have been caused by substitutions at a total of five β chain sites: 2, 82, 135, 139, and 143. We therefore introduced croc-specific amino acid states at each site on the AncArchosaur background, individually and in combination ("AncArchosaur+5"). Four of the five mutations produced modest reductions in ATP sensitivity when introduced individually on the AncArchosaur background, but the net effect of all five mutations did not reduce ATP sensitivity to the same level as AncCroc Hb (Figure 2A). We therefore introduced two additional croc-specific substitutions (Kβ87N and Kβ144E) that eliminated positive charges in the central cavity (Figure 2B). Although neither Kβ87 nor Kβ144 directly bind phosphate,[5,25] we hypothesized that the reduction in net positive charge caused by croc-specific substitutions at these two sites might further inhibit ATP binding. Consistent with this prediction, adding these two charge-changing mutations to the AncArchosaur+5 construct ("AncArchosaur+7") reduced ATP sensitivity to the same level as AncCroc (Figure 2A). Thus, a total of seven β chain substitutions in the central cavity (including two that are not directly involved in phosphate binding) are sufficient to explain the evolved loss of ATP sensitivity in the ancestor of crocodilians. The substitutions that were sufficient to eliminate ATP sensitivity did not simultaneously enhance bicarbonate sensitivity to the level observed in AncCroc (Figure S4A), indicating that the gain and loss of different allosteric interactions involve different substitutions.

Molecular changes underlying the gain of bicarbonate sensitivity

In considering the biophysical mechanism of bicarbonate binding to crocodilian Hb in light of the results of Komiyama et al.,[14] Perutz[13] suggested that two derived (croc-specific) amino acids, Kβ38 and Yβ41, bind a single bicarbonate ion in coordination with the highly conserved tyrosine at α42 (located on the α chain subunit of the opposing αβ dimer) (Figure 3). This mechanism predicts the binding of two bicarbonate ions per Hb tetramer, one at each of the two symmetrical interdimer interfaces ($\alpha_1\beta_2$ and $\alpha_2\beta_1$), and is consistent with experimental results for native crocodilian Hbs which documented that bicarbonate ions bind deoxyHb with a 2:1 stoichiometry.[2,7] To test the Perutz hypothesis, we synthesized four rHbs representing the requisite single- and double-mutant genotypes and measured the individual and combined effects of croc-specific mutations (Tβ38K and Fβ41Y) on the AncArchosaur background. Contrary to Perutz's prediction, introduction of croc-specific mutations at sites β38 and β41—individually and in pairwise combination—did not increase bicarbonate sensitivity relative to the ancestral reference value (Figure 4A). We then expanded the focus to the set of 12 amino acid states that Komiyama et al.[14] had identified as being necessary and sufficient to transplant bicarbonate sensitivity into human Hb, but we engineered these same changes into an evolutionary relevant background—namely, AncArchosaur. Since AncArchosaur and AncCroc share the same amino acid state at four of the 12 sites mutagenized by Komiyama et al.,

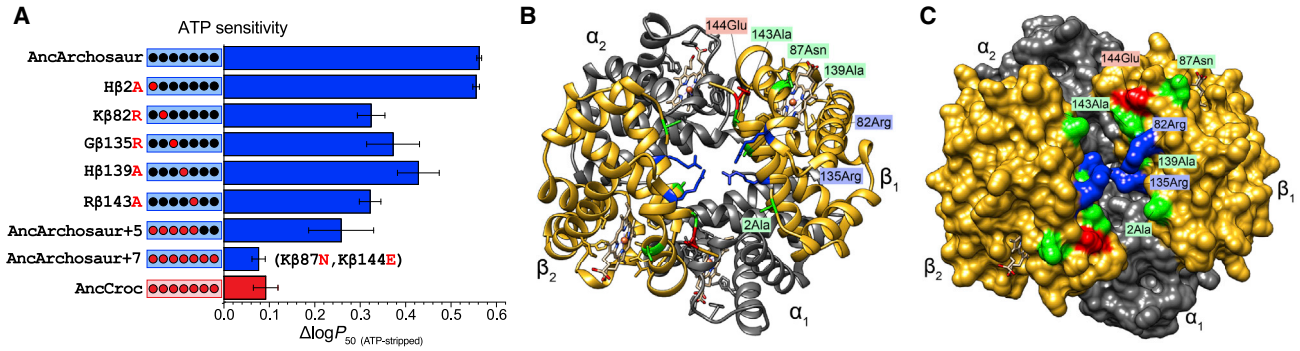


Figure 2. Experimental tests of hypotheses regarding the amino acid substitutions responsible for the loss of ATP sensitivity in the ancestor of crocodilians

(A) A total of seven croc-specific substitutions is sufficient to eliminate ATP sensitivity. Two of the substitutions, Kβ87N and Kβ144E, reduce the net positive charge in the central cavity but are not directly involved in phosphate binding. Graphed values are means based on triplicate measurements (±SEM).

(B) A ribbon model of AncArchosaur Hb showing the seven mutations that eliminate ATP sensitivity. The tetramer is viewed along the molecular dyad axis, with the α chains colored gray and the β chains colored yellow. For each of the seven mutated residues, charge of the derived (croc-specific) amino acid is denoted by color (red, negative; blue, positive; and green, neutral). Hemes are shown as stick models, with the iron atoms shown as brown spheres.

(C) A surface representation of the model shown in (B), using the same color scheme. Arg-β82 and Arg-β135 largely fill the central cavity between the two β chain subunits, where organic phosphates bind human deoxyHb.

See also Figure S4 and Table S1.

we only needed to introduce a total of eight new mutations on the AncArchosaur background. The results revealed that the eight mutations—individually and in combination—did not significantly increase bicarbonate sensitivity (Figure 4B), nor did they significantly alter ATP sensitivity (Figure S4B).

Why were the set of 12 croc-specific mutations sufficient to confer the bicarbonate effect in human Hb, but not in AncArchosaur Hb? One possible explanation is that one or more substitutions specific to human Hb may potentiate bicarbonate sensitivity in combination with croc-specific amino acid states at the 12 focal sites. To nominate candidates for such permissive substitutions, we identified 10 sites where human and AncCroc share the same amino acid state to the exclusion of AncArchosaur. We then synthesized a new version of AncArchosaur ("AncArchosaur+18") that contained the eight croc-specific substitutions tested by Komiyama et al. in combination with the 10 candidate permissive substitutions (Figure 4C). This set of 18 changes produced a significant increase in bicarbonate sensitivity, but it was still only 59% of the reference value for AncCroc Hb (Figure 4D). We then investigated an alternative biophysical mechanism by testing different multi-site combinations of 19 croc-specific substitutions that alter the net charge of the central cavity and/or alter proton dissociation constants of Lys or Arg side chains that could serve as binding sites for bicarbonate ions at a close-to-neutral pH. None of the tested combinations recapitulated the full bicarbonate sensitivity of AncCroc (Figure S5).

Insights from Hb isoform differentiation

The evolved transition between the bicarbonate-insensitive ancestor (AncArchosaur) and the bicarbonate-sensitive descendant (AncCroc) is the rationale for testing croc-specific forward mutations on the ancestral AncArchosaur background. Since none of our mutagenesis experiments on the AncArchosaur background successfully recapitulated the full bicarbonate sensitivity of AncCroc, we explored an alternative approach that enabled us to isolate the functional effect of β chain substitutions. Crocodilians possess multiple copies of β-type globin genes[26] and express structurally distinct Hb isoforms (isoHbs) at different stages of prenatal development, as is the case with all other amniotes.[13,25,27–29] We previously discovered that crocodilian embryos express two structurally distinct isoHbs: an embryonic isoHb that incorporates the β chain product of *HBB-T1* (HbI) and the normal adult isoHb that incorporates the β chain product of *HBB-T4* (HbII); both isoHbs share the same α chain subunit[23] (Figure 5A). *HBB-T1* and *HBB-T4* are products of a croc-specific duplication event[5,26]; hence, it is possible that embryonic HbI (with a β chain subunit distinct from that of the adult HbII) retains the ancestral bicarbonate-insensitive condition typical of all other amniote Hbs. If so, examination of isoform differences between embryonic HbI and adult HbII could provide a tractable means of identifying the molecular basis of bicarbonate sensitivity. To investigate isoHb differences in bicarbonate sensitivity, we isolated and purified native HbI and HbII from red blood cells of alligator embryos (*Alligator mississippiensis*) at day 40 of pre-hatching development. We confirmed that the adult isoform HbII is sensitive to bicarbonate but not to ATP (Figure 5B), as expected based on previous studies.[5,7,23] Importantly, we confirmed that the embryonic isoform HbI is not sensitive to bicarbonate but *is* sensitive to ATP (Figure 5B), just like AncArchosaur and the native Hbs of birds and all other amniotes. Since the embryonic and adult isoHbs share the same α chain, the observed isoHb differences in the mode of allosteric regulatory control indicate that bicarbonate sensitivity can evolve from an insensitive ancestral state via β chain substitutions alone.

To winnow down the list of candidate β-globin mutations for bicarbonate sensitivity, we reconstructed and synthesized the pre-duplication ancestor of the croc-specific *HBB-T1* and *HBB-T4* paralogs (β AncCrocT1/T4) (Figures 5C and S6), which enabled us to conduct a vertical comparison between the bicarbonate-insensitive (embryonic) Hb that incorporates the

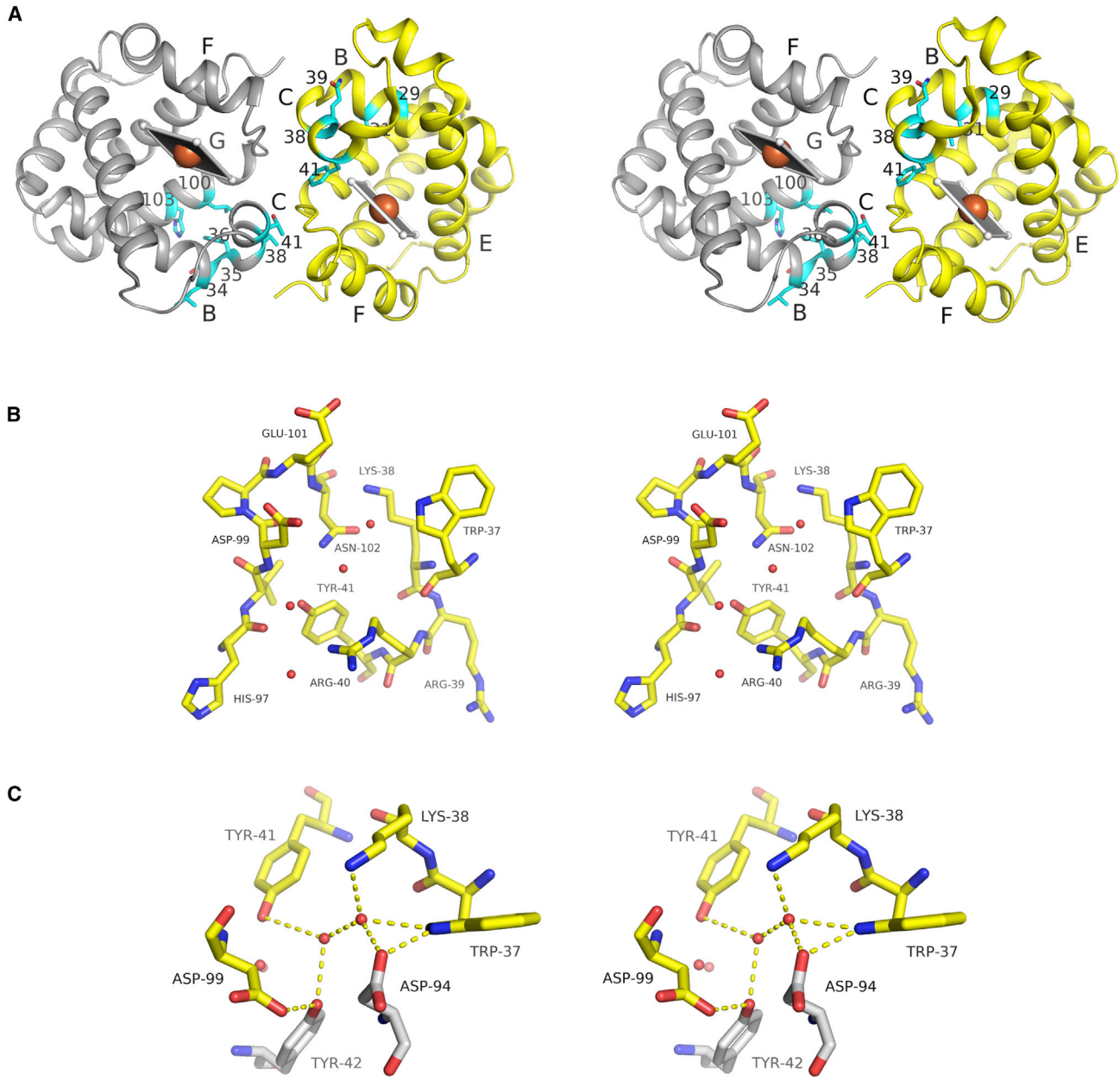


Figure 3. Stereo images showing Perutz's hypothesized binding site for bicarbonate ion at the symmetrical $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interdimeric interfaces of the Hb tetramer

Substitutions Tβ38K and Fβ41Y both occurred in the stem lineage of crocodilians, whereas Tyr-α42 is conserved among all modern archosaurs.

(A) The $\alpha_1\beta_2$ dimer of human deoxy Hb (PDB: 2dn2) is shown, using ribbons to depict the α- and β-type polypeptide chains (shown in gray and yellow, respectively). Select α-helices of both polypeptide chains are labeled, hemes are depicted as square planes with pseudo-atoms at the corners, and the iron atoms are depicted as brown spheres. The 12 residues mutated in the Komiyama et al.[14] study are shown in cyan, with the side chains shown as sticks.

(B) A closer view of key β chain residues in a model of Hb Scuba,[14] made by simple replacement of the Komiyama residues in deoxy HbA. Lys 38 is considerably larger than the threonine side chain it replaces, causing significant steric clashes with nearby residues such as Asn 102. Four water molecules are found in the experimental model in this region and are shown as red spheres.

(C) A similar view to that of (B) but rotated by 90° about the horizontal axis, bringing Tyr 42α and Asp 94α into view. Potential hydrogen bonds are shown as yellow dotted lines. Tyr 41β lies close to the backbone of His 97 and Asp 99 of the same subunit and may form hydrogen bonds with them.

pre-duplication ancestor of *HBB-T1/T4* and the bicarbonate-sensitive (adult) Hb that incorporates the product of *HBB-T4* (=β AncCroc) (Figures 5C and 5D). As in the case with the AncArchosaur versus AncCroc comparison, AncCrocT1/T4 versus AncCroc also constitutes a vertical comparison between a bicarbonate-insensitive ancestor and a bicarbonate-sensitive descendant, but the ancestral and descendant genotypes differ at fewer sites (49 versus 74), and all substitutions are restricted to

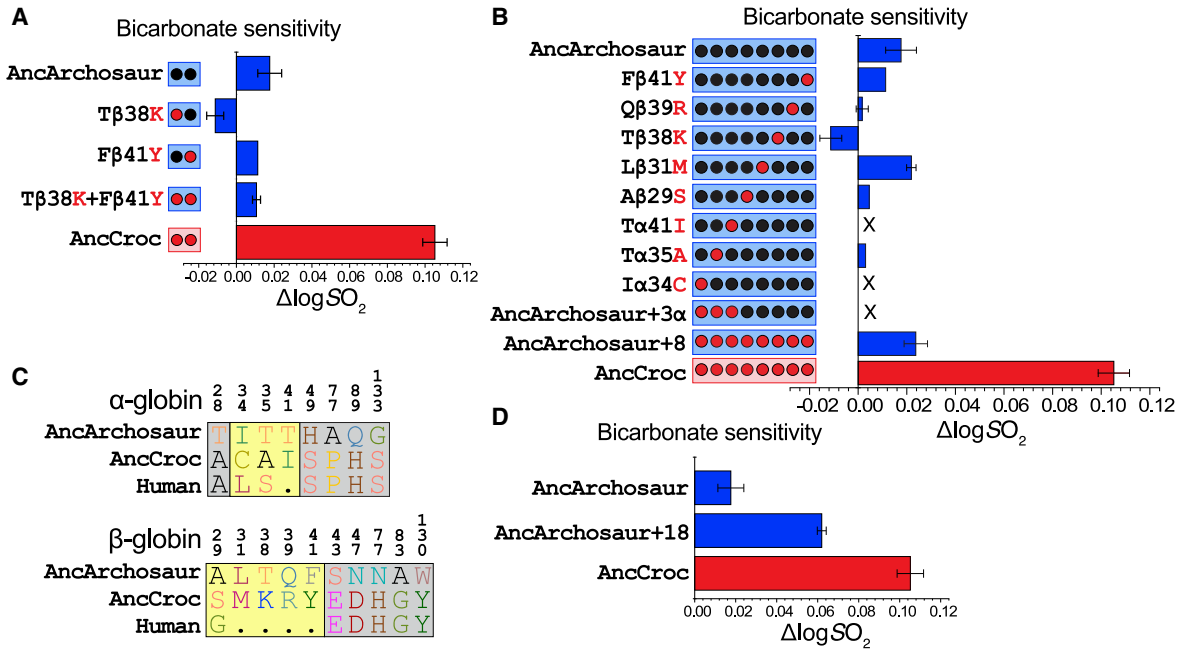


Figure 4. Experimental tests of hypotheses regarding the amino acid substitutions responsible for the gain of bicarbonate sensitivity in the ancestor of crocodilians

(A) Contrary to the hypothesized mechanism of bicarbonate binding shown in Figure 3, introducing croc-specific mutations Kβ87N and Kβ144E on the AncArchosaur background—individually and in combination—did not augment bicarbonate sensitivity (measured as the difference in log-transformed SO_2 [Hb-O_2 saturation] in the presence and absence of 1% CO_2).

(B) Effects of individual mutations at 8 candidate sites on bicarbonate sensitivity. When introduced individually on the AncArchosaur background, two α chain mutations, I34C and T41I (located in $\alpha_1\beta_1/\alpha_2\beta_2$ and $\alpha_1\beta_2/\alpha_2\beta_1$ interfaces, respectively) resulted in nonfunctional protein (denoted by an "X"). The same was true when both mutations were combined with T35A as part of a trio of croc-specific α chain mutations ("AncArchosaur+3α"). None of the 8 candidate mutations significantly increased bicarbonate sensitivity above the AncArchosaur reference value, individually or in combination ("AncArchosaur+8").

(C) The set of candidate mutations for bicarbonate sensitivity include eight that were previously tested by Komiyama et al.[14] (yellow background) in combination with 10 candidate permissive substitutions where AncCroc and human share amino acid states to the exclusion of AncArchosaur (gray background).

(D) The combination of 18 mutations on the AncArchosaur background produced a level of bicarbonate sensitivity equal to 59% of the AncCroc reference value. Graphed values are means based on triplicate measurements (±SEM).

See also Figures S4 and S5 and Table S1.

the β chain. Using the same AltAll experimental approach described previously,[24] we confirmed that the inferred transition in the mode of allosteric regulation is robust to statistical uncertainty in the ancestral sequence estimate (Table S1).

As a first step, we introduced a total of 13 AncCroc-specific mutations into the β AncCrocT1/T4 background: the eight β chain mutations tested by Komiyama et al. in combination with five charge- or polarity-changing mutations concentrated in the $\alpha_1\beta_2/\alpha_2\beta_1$ interface (Figure S7), including the Tβ38K and Fβ41Y mutations that are central to the Perutz hypothesis (Figures 3B and 3C). This combination of mutations ("AncCrocT1/T4+13") produced a 61% increase in bicarbonate sensitivity, relative to that of AncCroc (Figure 6A). Reverting these mutations to the ancestral amino acid states on the AncCroc background eliminated bicarbonate sensitivity (Figure 6A). Follow-up experiments revealed that the increased bicarbonate sensitivity could be explained by the five croc-specific $\alpha_1\beta_2/\alpha_2\beta_1$ interface mutations alone ("AncCrocT1/T4+5," which included Tβ38K and Fβ41Y) and, again, reverting those same five mutations on the AncCroc background was sufficient to eliminate sensitivity (Figure 6A). Although the "+5" set of mutations in the $\alpha_1\beta_2/\alpha_2\beta_1$ interface may be necessary to confer bicarbonate sensitivity, those five changes alone are clearly not sufficient to recapitulate the full sensitivity of AncCroc. To test whether some combination of other croc-specific β chain mutations need to be combined with the +5 mutations to confer full sensitivity, we divided the AncCrocT1/T4 β chain sequence into equal thirds, and we synthesized rHb constructs containing all croc-specific substitutions in the first, second, and third parts (residues 1–49, 50–98, and 99–146, respectively). The "Part I" construct contained the +5 mutations in combination with 16 additional croc-specific mutations concentrated in the N terminus and $\alpha_1\beta_2/\alpha_2\beta_1$ and $\alpha_1\beta_1/\alpha_2\beta_2$ interfaces. Experimental results demonstrate that the set of 21 croc-specific Part I mutations successfully recapitulates the full bicarbonate sensitivity of AncCroc (Figure 6B).

Having identified a set of 21 croc-specific substitutions that are sufficient to confer full bicarbonate sensitivity on the AncCrocT1/T4 background, we then performed additional mutagenesis experiments in an attempt to further winnow the number of necessary substitutions. We synthesized new β AncCrocT1/T4 constructs that divided Part I into two equal halves, Part IA and Part IB (containing all croc-specific mutations between β chain residues 1-18 [positions NA1 to AB1] and 19-49 [positions C1 to CD8], respectively), and Part IB1

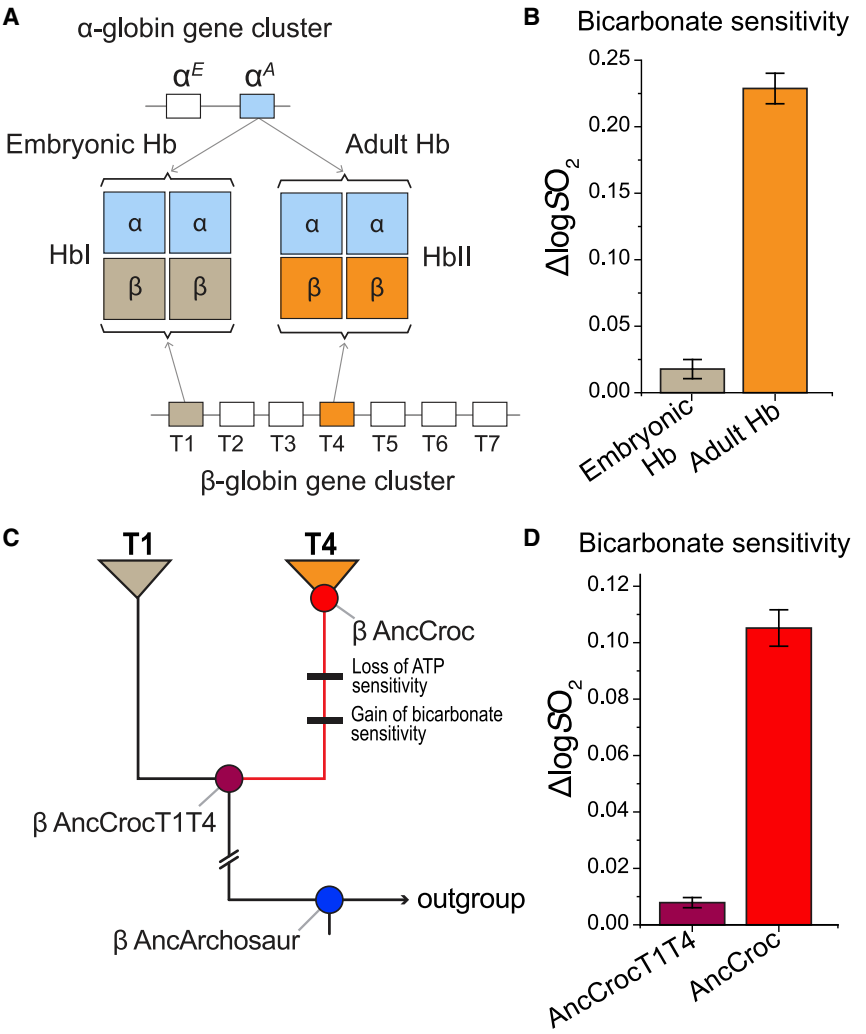


Figure 5. Evolved differences in allosteric properties of embryonic and adult crocodilian isoHbs provide an alternative means of investigating the molecular basis of bicarbonate sensitivity

(A) Embryonic and adult isoHbs of crocodilians: HbI and HbII, respectively. The two isoHbs share identical α chains but incorporate structurally distinct β chain products of the *HBB-T1* and *HBB-T4* paralogs.

(B) In contrast to the native adult HbII isoform, native embryonic HbI is insensitive to bicarbonate (the ancestral condition, shared by Hbs of all other amniotes).

(C) Phylogeny of croc-specific β-globin paralogs showing internal nodes targeted for ancestral protein resurrection.

(D) Experiments on recombinant, resurrected ancestral proteins confirmed that the β chain substitutions that conferred bicarbonate sensitivity occurred in the post-duplication branch (red) that connects "β AncCrocT1/T4" (the single-copy, pre-duplication ancestor of the croc-specific *HBB-T1* and *HBB-T4* paralogs) to "β AncCroc." β AncCrocT1/T4 and AncCroc therefore bracket an evolutionary transition between ancestral and derived functional states. Graphed values are means based on triplicate measurements (±SEM).

See also Figures S6 and S7 and Table S1.

(which contains a subset of 7 croc-specific substitutions in Part IB that span the $\alpha_1\beta_2/\alpha_2\beta_1$ interface) (Figure 6C). Experimental results revealed that the set of croc-specific substitutions concentrated at the β chain N terminus (Part IA) produced no significant effect, whereas the 13 croc-specific mutations in Part IB conferred a level of bicarbonate sensitivity equal to 83.3% of the reference value for AncCroc Hb (Figure 6D). When that set of 13 croc-specific mutations was pared down to a subset of 7 mutations in $\alpha_1\beta_2/\alpha_2\beta_1$ interface (Part IB1), bicarbonate sensitivity decreased to 70.2 % of the reference value (Figure 6D). It is clear that full bicarbonate sensitivity requires numerous substitutions that are distributed across multiple subdomains of the protein.

In summary, mutagenesis experiments revealed a set of 21 β chain substitutions that are sufficient to confer full bicarbonate sensitivity on the ancestral AncCrocT1/T4 background, and a small subset of substitutions in the $\alpha_1\beta_2/\alpha_2\beta_1$ interface make a major contribution. In contrast to the results of previous efforts to transplant bicarbonate sensitivity into human Hb,[14] introducing our set of 21 croc-specific mutations onto the AncCrocT1/T4 background successfully conferred bicarbonate sensitivity without severely compromising intrinsic O_2-affinity or cooperativity (Table S1). The two croc-specific substitutions highlighted by Perutz, Tβ38K, and Fβ41Y, appear to be necessary to confer the bicarbonate effect on the T1/T4 background, but those two substitutions alone are not sufficient. The full sensitivity observed in AncCroc is only achieved when the interface mutations are combined with numerous other croc-specific amino acid states within the N-terminal 1/3 of the β chain. It may be that the triad of residues Kβ38, Yβ41, and Tα42 constitute the only site for oxygenation-linked binding of bicarbonate (Figures 3B and 3C), but the requisite stereochemical conformation of the three residues may depend on a higher-level (quaternary) conformation that is determined by the amino acid states of structurally remote sites.

Tests of positive selection

Given that the gain of bicarbonate sensitivity in crocodilian Hb is attributable to β chain substitutions that occurred in the adult β-globin gene (*HBB-T4*) following duplication of the single-copy ancestor of *HBB-T1* and *HBB-T4* (β AncCroc T1/T4) (Figure 5C), we performed a molecular evolution analysis to test for evidence of positive selection in the post-duplication branch of the gene tree that connects β AncT1/T4 to β AncCroc (Figure S6). This test revealed evidence for a strikingly accelerated rate of nonsynonymous substitution in the branch that connects β AncT1/T4 to the single-copy, pre-duplication ancestor of the croc-specific paralogs *HBB-T4* (= β AncCroc) and *HBB-T6* (which may encode β chain subunits of an isoHb expressed in

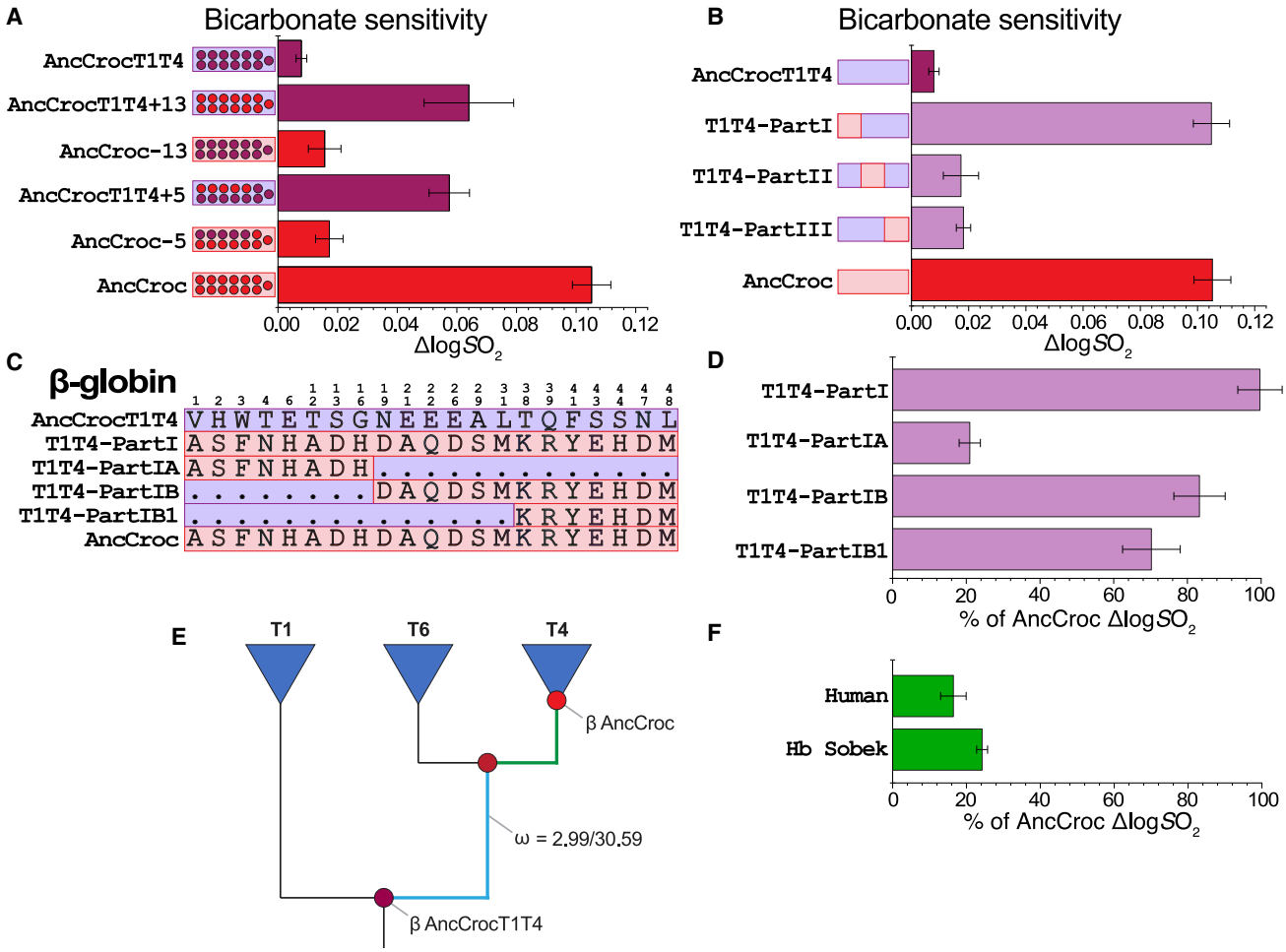


Figure 6. Testing candidate mutations for bicarbonate sensitivity on the AncCrocT1/T4 background

(A) At 13 candidate sites, forward mutations on the AncCrocT1/T4 background (blue) produce a significant enhancement of bicarbonate sensitivity that falls short of the AncCroc reference value. The observed enhancement can be explained by the effects of just five croc-specific mutations in the $\alpha_1\beta_2/\alpha_2\beta_1$ interface. Reverting such mutations on the AncCroc background (red) eliminates bicarbonate sensitivity.

(B) To test whether some combination of other croc-specific β chain mutations need to be combined with the +5 mutations to confer full sensitivity, we divided the AncCrocT1/T4 β chain sequence into equal thirds, and we tested constructs containing all croc-specific substitutions in each of the three parts. "Part I" contained the +5 mutations in combination with 16 additional croc-specific mutations concentrated in the N terminus and $\alpha_1\beta_2/\alpha_2\beta_1$ and $\alpha_1\beta_1/\alpha_2\beta_2$ interfaces. Results demonstrate that the set of 21 croc-specific Part I mutations successfully recapitulates the full bicarbonate sensitivity of AncCroc.

(C) We synthesized additional constructs representing further subdivisions of Part I.

(D) The smaller subsets of Part I mutations do not fully recapitulate bicarbonate sensitivity, although the set of mutations in the $\alpha_1\beta_2/\alpha_2\beta_1$ interface (including Tβ38K and Fβ41Y) make the largest contribution.

(E) Diagrammatic phylogeny of crocodilian β-globin paralogs showing a significantly accelerated rate of nonsynonymous substitution (ω [= d_N/d_S] in branch model/clade model) in the post-duplication branch connecting β AncT1/T4 to the single-copy, pre-duplication ancestor of the croc-specific *HBB-T4* and *HBB-T6* paralogs.

(F) No significant difference in bicarbonate sensitivity between human Hb and the mutant Hb Sobek that contains the same 21 croc-specific substitutions that were sufficient to confer full bicarbonate sensitivity on the AncCrocT1/T4 background (T1T4-Part I).

Graphed values in (A), (B), (D), and (F) are means based on triplicate measurements (±SEM).

See also Tables S1 and S2.

early embryogenesis) (Figure 6E; Table 1), and a clade test revealed significantly accelerated substitution rates among *HBB-T4* orthologs of different crocodilian lineages (Table S2). This molecular evolution analysis provides an independent line of evidence that the bicarbonate sensitivity of crocodilian Hb represents an adaptive trait that evolved via strong positive selection.

Can bicarbonate sensitivity be transplanted into human Hb?

Having discovered a set of 21 croc-specific mutations that are sufficient to confer bicarbonate sensitivity when introduced into ancestral crocodilian Hb, we tested whether the same set of mutations produce the same phenotypic effect when introduced into a divergent genetic background. Specifically, we

Table 1. Branch model statistics for crocodilian β-globin paralogs

Model	Np	−lnL	ω estimates	LRT
Model 0 (null model)	61	2,887.22	$\omega = 0.2245$	–
AncCrocT1/T4 to AncCrocT4/T6 branch as foreground	62	2,884.14	$\omega_0 = 0.21$ $\omega_1 = 2.99$	$p < 0.01$
AncCrocT1/T4 to β AncCroc as foreground	62	2,887.21	$\omega_0 = 0.22$ $\omega_1 = 0.23$	NS
Stem of *HBB-T4* as foreground	62	2,887.22	$\omega_0 = 0.24$ $\omega_1 = 0.14$	NS
HBB-T4 clade as foreground (stem + crown)	62	2,872.1	$\omega_0 = 0.14$ $\omega_1 = 0.50$	$p < 0.001$

$\omega = d_N/d_S$ (the rate of nonsynonymous substitutions per nonsynonymous site relative to the rate of synonymous substitutions per synonymous site). Models considered the following branches as foreground: (1) the branch connecting β AncCrocT1/T4 to the pre-duplication ancestor of *HBB-T4* and *HBB-T6*, (2) the branch connecting β AncCrocT1/T4 to *HBB-T4* (= β AncCroc), (3) the stem of the *HBB-T4* clade, and (4) the entire *HBB-T4* clade, stem plus crown, relative to the rest of the tree (Figure S6). NS, not significant. See also Table S2.

used site-directed mutagenesis to introduce the 21 Part I croc-specific mutations into recombinant human Hb (named Hb "Sobek" after the half-crocodile/half-human Egyptian deity). The experiment revealed that the 21 mutations that were sufficient to confer bicarbonate sensitivity on the crocodilian genetic background produced a negligible increase in bicarbonate sensitivity when introduced into the structural context of human Hb, well below the full allosteric regulatory capacity of AncCroc (Figure 6F).

DISCUSSION

Our experimental results provide answers to several fundamental questions about protein evolution and biochemical adaptation. First, are adaptive changes in protein function attributable to few substitutions at key sites or many substitutions that have individually minor effects? Perutz[4] suggested that adaptive modifications of Hb function will typically be attributable to "…a few replacements in key positions," a scenario consistent with the evolution of substrate-specificity in metabolic enzymes[30] and ligand-specificity of allosterically regulated steroid receptors.[31,32] With regard to the loss of ATP sensitivity, we identified several croc-specific mutations that directly eliminated phosphate-binding sites, but the full loss of sensitivity required the indirect effect of two additional mutations that altered the net positive charge in the central cavity. With regard to the gain of bicarbonate sensitivity, we identified two croc-specific mutations in the $\alpha_1\beta_2/\alpha_2\beta_1$ interface (Tβ38K and Fβ41Y) that appear to play a direct role in the allosteric binding of bicarbonate ion, but they confer their allosteric effect only in conjunction with mutations at numerous other sites that are structurally remote from the intersubunit interface. Thus, evolved changes in allosteric regulation involved the direct effect of few substitutions at key positions in combination with indirect effects of numerous other substitutions at structurally disparate sites. Such indirect interaction effects suggest that the evolution of novel functions may be highly dependent on neutral mutations that produce no adaptive benefit when they first arise but which contribute to a permissive background for subsequent function-altering mutations at other sites.

A second key question concerns the extent to which ancestral protein functions can be maintained during the evolutionary transition to a novel function. With regard to the evolution of crocodilian Hb, the question is whether it is possible to maintain the capacity for allosteric regulatory control by ATP binding (the ancestral condition for the Hb of jawed vertebrates) while simultaneously evolving bicarbonate sensitivity. The results of our experiments indicate that it is possible to maintain sensitivity to both ATP and bicarbonate and that there is no ineluctable trade-off between the two modes of allosteric regulation. The minimal set of amino acid mutations that are sufficient to eliminate ATP sensitivity on the AncArchosaur background do not simultaneously affect bicarbonate sensitivity. Likewise, none of the mutations implicated in the gain of bicarbonate sensitivity affect allosteric ATP binding.

A third key question concerns the context-dependence of mutations (epistasis) and its implications for adaptive protein evolution. Our experimental attempt to transplant bicarbonate sensitivity into human Hb provides a striking example of intramolecular epistasis: the same 21 mutations that conferred full bicarbonate sensitivity when introduced into ancestral crocodilian Hb produced no effect in human Hb. This context-dependence highlights the importance of using a vertical approach in mutagenesis experiments (introducing forward mutations on ancestral backgrounds) to identify evolved mechanisms of functional change.[15] The result also suggests an important role for historical contingency in the evolution of novel protein functions, as it indicates that some adaptive solutions are not equally accessible from all possible ancestral starting points. The fact that causative substitutions in crocodilian Hb have no effect on the divergent genetic background of human Hb suggest that bicarbonate sensitivity may not be a mutationally accessible design solution for Hbs of other vertebrates or that the same allosteric property would have to evolve via a completely different molecular mechanism.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY

- Lead contact
- Materials availability
- Data and code availability

- METHOD DETAILS
  - Estimation of ancestral globin sequences
  - Vector construction and site-directed mutagenesis
  - Protein expression and purification
  - Collection of blood samples and preparation of hemolysates
  - Purification of native Hbs
  - Measurement of oxygen equilibration curves
  - Measurement of bicarbonate sensitivity using CO_2
  - Tests of positive selection

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.11.049.

ACKNOWLEDGMENTS

We thank A. Kumar and E. Petersen for assistance in the lab. This work was funded by National Institutes of Health (NIH) grant R01HL087216 (J.F.S.) and National Science Foundation (NSF) grant OIA-1736249 (J.F.S.).

AUTHOR CONTRIBUTIONS

J.F.S., A.F., and C.N. designed the research; C.N., A.V.S., N.M.B., and F.G.H. performed the research; C.N., A.V.S., N.M.B., F.G.H., J.R.H.T., A.F., and J.F.S. analyzed data; and J.F.S. and C.N. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 7, 2022
Revised: October 25, 2022
Accepted: November 22, 2022
Published: December 21, 2022

REFERENCES

1. Grigg, G., and Kirshner, D. (2015). Biology and Evolution of Crocodylians (Cornell University Press).
2. Bauer, C., Forster, M., Gros, G., Mosca, A., Perrella, M., Rollema, H.S., and Vogel, D. (1981). Analysis of bicarbonate binding to crocodilian hemoglobin. J. Biol. Chem. *256*, 8429–8435.
3. Perutz, M.F., Bauer, C., Gros, G., Leclercq, F., Vandecasserie, C., Schnek, A.G., Braunitzer, G., Friday, A.E., and Joysey, K.A. (1981). Allosteric regulation of crocodilian haemoglobin. Nature *291*, 682–684.
4. Perutz, M.F. (1983). Species adaptation in a protein molecule. Mol. Biol. Evol. *1*, 1–28.
5. Fago, A., Natarajan, C., Pettinati, M., Hoffmann, F.G., Wang, T., Weber, R.E., Drusin, S.I., Issoglio, F., Martí, M.A., Estrin, D.A., and Storz, J.F. (2020). Structure and function of crocodilian hemoglobins and allosteric regulation by chloride, ATP, and CO_2. Am. J. Physiol. Regul. Integr. Comp. Physiol. *318*, R657–R667.
6. Bautista, N.M., Damsgaard, C., Fago, A., and Wang, T. (2021). Carbon dioxide and bicarbonate accumulation in caiman erythrocytes during diving. J. Exp. Biol. *224*, jeb242435.
7. Bautista, N.M., Malte, H., Natarajan, C., Wang, T., Storz, J.F., and Fago, A. (2021). New insights into the allosteric effects of CO2 and bicarbonate on crocodilian hemoglobin. J. Exp. Biol. *224*, jeb242615.
8. Weber, R.E., and White, F.N. (1986). Oxygen binding in alligator blood related to temperature, diving, and alkaline tide. Am. J. Physiol. *251*, R901–R908.
9. Weber, R.E., and White, F.N. (1994). Chloride-dependent organic phosphate sensitivity of the oxygenation reaction in crocodilian hemoglobins. J. Exp. Biol. *192*, 1–11.
10. Busk, M., Overgaard, J., Hicks, J.W., Bennett, A.F., and Wang, T. (2000). Effects of feeding on arterial blood gases in the American alligator *Alligator mississippiensis*. J. Exp. Biol. *203*, 3117–3124.
11. Perutz, M.F. (1972). Nature of haem-haem interaction. Nature *237*, 495–499.
12. Perutz, M.F. (1997). Science Is Not a Quiet Life: Unravelling the Atomic Mechanism of Haemoglobin (Imperial College Press).
13. Storz, J.F. (2019). Hemoglobin: Insights into Protein Structure, Function, and Evolution (Oxford University Press).
14. Komiyama, N.H., Miyazaki, G., Tame, J., and Nagai, K. (1995). Transplanting a unique allosteric effect from crocodile into human haemoglobin. Nature *373*, 244–246.
15. Harms, M.J., and Thornton, J.W. (2010). Analyzing protein structure and function using ancestral gene reconstruction. Curr. Opin. Struct. Biol. *20*, 360–366.
16. Harms, M.J., and Thornton, J.W. (2013). Evolutionary biochemistry: revealing the historical and physical causes of protein properties. Nat. Rev. Genet. *14*, 559–571.
17. Hochberg, G.K.A., and Thornton, J.W. (2017). Reconstructing ancient proteins to understand the causes of structure and function. Annu. Rev. Biophys. *46*, 247–269.
18. Natarajan, C., Jiang, X., Fago, A., Weber, R.E., Moriyama, H., and Storz, J.F. (2011). Expression and purification of recombinant hemoglobin in *Escherichia coli*. PLoS One *6*, e20176.
19. Natarajan, C., Signore, A.V., Kumar, V., and Storz, J.F. (2020). Synthesis of recombinant human hemoglobin with NH_2-terminal aetylation in *Escherichia coli*. Curr. Protoc. Protein Sci. *101*, e112. https://doi.org/10.1002/cpps.112.
20. Damsgaard, C., Storz, J.F., Hoffmann, F.G., and Fago, A. (2013). Hemoglobin isoform differentiation and allosteric regulation of oxygen binding in the turtle, *Trachemys scripta*. Am. J. Physiol. Regul. Integr. Comp. Physiol. *305*, R961–R967.
21. Weber, R.E., Fago, A., Malte, H., Storz, J.F., and Gorr, T.A. (2013). Lack of conventional oxygen-linked proton and anion binding sites does not impair allosteric regulation of oxygen binding in dwarf caiman hemoglobin. Am. J. Physiol. Regul. Integr. Comp. Physiol. *305*, R300–R312.
22. Storz, J.F., Natarajan, C., Moriyama, H., Hoffmann, F.G., Wang, T., Fago, A., Malte, H., Overgaard, J., and Weber, R.E. (2015). Oxygenation properties and isoform diversity of snake hemoglobins. Am. J. Physiol. Regul. Integr. Comp. Physiol. *309*, R1178–R1191.
23. Bautista, N.M., Petersen, E.E., Jensen, R.J., Natarajan, C., Storz, J.F., Crossley, D.A., and Fago, A. (2021). Changes in hemoglobin function and isoform expression during embryonic development in the American alligator, *Alligator mississippiensis*. Am. J. Physiol. Regul. Integr. Comp. Physiol. *321*, R869–R878.
24. Eick, G.N., Bridgham, J.T., Anderson, D.P., Harms, M.J., and Thornton, J.W. (2017). Robustness of reconstructed ancestral protein functions to statistical uncertainty. Mol. Biol. Evol. *34*, 247–261.
25. Grispo, M.T., Natarajan, C., Projecto-Garcia, J., Moriyama, H., Weber, R.E., and Storz, J.F. (2012). Gene duplication and the evolution of hemoglobin isoform differentiation in birds. J. Biol. Chem. *287*, 37647–37658.
26. Hoffmann, F.G., Vandewege, M.W., Storz, J.F., and Opazo, J.C. (2018). Gene turnover and diversification of the α- and β-globin gene families in sauropsid vertebrates. Genome Biol. Evol. *10*, 344–358.
27. Hoffmann, F.G., Storz, J.F., Gorr, T.A., and Opazo, J.C. (2010). Lineage-specific patterns of functional diversification in the α- and β-globin gene families of tetrapod vertebrates. Mol. Biol. Evol. *27*, 1126–1138.

28. Opazo, J.C., Hoffmann, F.G., Witt, C.C., Berenbrink, M., and Storz, J.F. (2015). Gene turnover in the avian globin gene family and evolutionary changes in hemoglobin isoform expression. Mol. Biol. Evol. *32*, 871–887.

29. Storz, J.F., Hoffmann, F.G., Opazo, J.C., Sanger, T.J., and Moriyama, H. (2011). Developmental regulation of hemoglobin synthesis in the green anole lizard *Anolis carolinensis*. J. Exp. Biol. *214*, 575–581. https://doi.org/10.1242/jeb.050443.

30. Boucher, J.I., Jacobowitz, J.R., Beckett, B.C., Classen, S., and Theobald, D.L. (2014). An atomic-resolution view of neofunctionalization in the evolution of apicomplexan lactate dehydrogenases. eLife *3*, e02304.

31. Harms, M.J., Eick, G.N., Goswami, D., Colucci, J.K., Griffin, P.R., Ortlund, E.A., and Thornton, J.W. (2013). Biophysical mechanisms for large-effect mutations in the evolution of steroid hormone receptors. Proc. Natl. Acad. Sci. USA *110*, 11475–11480.

32. Siddiq, M.A., Hochberg, G.K.A., and Thornton, J.W. (2017). Evolution of protein specificity: insights from ancestral protein reconstruction. Curr. Opin. Struct. Biol. *47*, 113–122.

33. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*, 1586–1591.

34. Hanson-Smith, V., Kolaczkowski, B., and Thornton, J.W. (2010). Robustness of ancestral sequence reconstruction to phylogenetic uncertainty. Mol. Biol. Evol. *27*, 1988–1999.

35. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534.

36. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods *14*, 587–589.

37. Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y., Faircloth, B.C., Nabholz, B., Howard, J.T., et al. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. Science *346*, 1320–1331.

38. Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K., and Mooers, A.O. (2012). The global diversity of birds in space and time. Nature *491*, 444–448.

39. Hackett, S.J., Kimball, R.T., Reddy, S., Bowie, R.C., Braun, E.L., Braun, M.J., Chojnowski, J.L., Cox, W.A., Han, K.L., Harshman, J., et al. (2008). A phylogenomic study of birds reveals their evolutionary history. Science *320*, 1763–1768.

40. Green, R.E., Braun, E.L., Armstrong, J., Earl, D., Nguyen, N., Hickey, G., Vandewege, M.W., St John, J.A., Capella-Gutiérrez, S., Castoe, T.A., et al. (2014). Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. Science *346*, 1254449.

41. Natarajan, C., Inoguchi, N., Weber, R.E., Fago, A., Moriyama, H., and Storz, J.F. (2013). Epistasis among adaptive mutations in deer mouse hemoglobin. Science *340*, 1324–1327.

42. Tufts, D.M., Natarajan, C., Revsbech, I.G., Projecto-Garcia, J., Hoffmann, F.G., Weber, R.E., Fago, A., Moriyama, H., and Storz, J.F. (2015). Epistasis constrains mutational pathways of hemoglobin adaptation in high-altitude pikas. Mol. Biol. Evol. *32*, 287–298.

43. Natarajan, C., Hoffmann, F.G., Weber, R.E., Fago, A., Witt, C.C., and Storz, J.F. (2016). Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. Science *354*, 336–339.

44. Kumar, A., Natarajan, C., Moriyama, H., Witt, C.C., Weber, R.E., Fago, A., and Storz, J.F. (2017). Stability-mediated epistasis restricts accessible mutational pathways in the functional evolution of avian hemoglobin. Mol. Biol. Evol. *34*, 1240–1251.

45. Natarajan, C., Jendroszek, A., Kumar, A., Weber, R.E., Tame, J.R.H., Fago, A., and Storz, J.F. (2018). Molecular basis of hemoglobin adaptation in the high-flying bar-headed goose. PLoS Genet. *14*, e1007331.

46. Natarajan, C., Signore, A.V., Kumar, V., Weber, R.E., Fago, A., and Storz, J.F. (2020). Effect of NH_2-terminal acetylation on the oxygenation properties of vertebrate haemoglobin. Biochem. J. *477*, 3839–3850.

47. Fago, A., Malte, H., and Dohn, N. (1999). Bicarbonate binding to hemoglobin links oxygen and carbon dioxide transport in hagfish. Respir. Physiol. *115*, 309–315.

48. Yang, Z.H., Kumar, S., and Nei, M. (1995). A new method of inference of ancestral nucleotide and amino acid sequences. Genetics *141*, 1641–1650.

49. Zhang, J., Nielsen, R., and Yang, Z.H. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. *22*, 2472–2479.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli DH5α competent cells	Thermo Fisher Scientific	Cat#18265017
Escherichia coli XL 10-Gold competent cells	Agilent Technologies La Jolla, CA, USA	Cat#200314
Escherichia coli JM109(DE3)	Fisher Scientific	Cat# PR-P9801
Chemicals, peptides, and recombinant proteins		
Hemin, bovine	Sigma-Aldrich	Cat# H9039-100G
IPTG (Isopropyl β-D-1-thiogalactopyranoside)	Invitrogen	Cat#15529019
Ampicillin	Sigma-Aldrich	Cat#A0166-25G
Kanamycin	Sigma-Aldrich	Cat#60615-25G
Terrific broth	RPI	Cat#T15100
Sodium chloride	Sigma-Aldrich	Cat#S9888-1KG
Trisbase	Sigma-Aldrich	Cat#93362-1KG
Polyethyleneimine Solution	Sigma-Aldrich	Cat#181978
HiTrap Q HP-sepharose column	GE Healthcare	Cat#17-5159-01
HiTrap SP HP-sepharose column	GE Healthcare	Cat#17-5161-01
Amicon Ultra-15 centrifugal filter (30 kDa cutoff)	Fisher Scientific	Cat#UFC903024
PD-10 column	GE Healthcare	Cat#GE17-0851-01
Mono Q 5/50 GL	GE Healthcare	Cat#GE17-5166-01
Glucose	Thermo Fisher Scientific	Cat# 15023021
Sodium hydrosulfite (sodium dithionite)	Sigma-Aldrich	Cat#157953
HEPES	Sigma-Aldrich	Cat#H3375-1KG
CAPS	Sigma-Aldrich	Cat#C2632-1KG
Critical commercial assays		
QuikChange II XL Site-Directed Mutagenesis Kit	Agilent Technologies, La Jolla, CA, USA	Cat#200521
GeneJET Plasmid Miniprep Kit	Thermo Fisher Scientific	Cat# K0503
Deposited data		
Synthetic gene cassettes	GeneArt gene synthesis, Thermo Fisher Scientific	GenBank: ON959522-ON959565
Oligonucleotides		
Site-directed mutagenesis primers	This study (Table S3)	N/A
Recombinant DNA		
pGM plasmid	Natarajan et al.[18]	https://doi.org/10.1371/journal.pone.0020176
pCOMAP plasmid	Natarajan et al.[18]	https://doi.org/10.1371/journal.pone.0020176
Software and algorithms		
Sigma plot 11.0	Systat Software	https://systatsoftware.com/
Origin 2018 (9.5)	Originlab Corporation	https://www.originlab.com/
PAML version 4.8	Yang[33]	http://abacus.gene.ucl.ac.uk/software/paml.html
Lazarus set of Python scripts	Hanson-Smith et al.[34]	http://markov.uoregon.edu/software/lazarus/
IQTree version 2.1.3	Minh et al.[35]	http://www.iqtree.org/
ModelFinder	Kalyaanamoorthy et al.[36]	http://www.iqtree.org/ModelFinder/
Geneious Pro 5.4.7	Biomatters	https://www.geneious.com/
PyMOL v2.4.0	Schrödinger, LLC	https://pymol.org/2/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact: Jay F. Storz (jstorz2@unl.edu).

Materials availability

All plasmid sequences used in this study have been deposited to GenBank (see key resources table).

Data and code availability

- All reconstructed sequences that we used for ancestral protein resurrection were deposited in GenBank (ON959522-ON959565) and are publicly available as of the date of publication.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Estimation of ancestral globin sequences

We estimated α- and β-type globin sequences of ancestral hemoglobins (Hbs) representing three internal nodes in the archosaur phylogeny (AncNeornithes, AncArchosaur, and AncCroc) (Figures S1 and S2). We estimated all ancestral sequences using the maximum likelihood (ML) approach implemented in PAML version 4.8,[33] with the aid of the Lazarus set of Python scripts for parsing results.[34] For each analysis, we collected a set of non-redundant α- and β-globin sequences that provided balanced phylogenetic coverage of extant archosaur diversity (Figures S1 and S2). We constructed supertrees for avian sequences by starting with a backbone provided by the total-evidence phylogeny from Jarvis et al.[37] Sequences from all species could be unambiguously assigned to their appropriate branches in this backbone tree. Subtrees for avian sequences were taken from Jetz et al.,[38] which was constructed using the Hackett et al.[39] backbone. Relationships among the major groups of sauropsids were based on the phylogeny in Green et al.[40] and assignments of orthology were based on Hoffmann et al.[26] In all cases, we used fully annotated globin genes from high-coverage genome assemblies along with newly generated sequences from a phylogenetically diverse set of crocodilian species.[5]

To estimate the ancestral α-globin sequences of AncArchosaur and AncNeornithes Hbs, we included orthologous sequences from representatives of each of the major crocodilian and avian lineages (Figure S1A) along with two testudine sequences as outgroups. To estimate the ancestral β-globin sequences of AncArchosaur and AncNeornithes Hbs, we expanded the phylogenetic coverage of paralogous β-type globins due to the history of lineage-specific duplication in both birds and crocodilians.[26,27] The avian β-globin paralogs include σ-, ε-, β^H-, and β^A-globin,[27] and those of crocodilians include *HBB-T1*, *HBB-T4*, *HBB-T6*, and *HBB-T7*.[5,26] We included β-globin paralogs from *Anolis* and Chinese softshell turtle as outgroups (Figure S1B). To estimate the ancestral α-globin sequence of AncCroc Hb, we expanded phylogenetic coverage of the adult-expressed crocodilian α-globin gene that is orthologous to avian α^A-globin.[25,26,28] To estimate the ancestral β-globin sequence of AncCroc Hb, we used an expanded supertree that included the full complement of croc-specific β-globin paralogs from representatives of each of the three extant families of crocodilians (Alligatoridae [alligators and caimans], Crocodylidae [true crocodiles], and Gavialidae [gharial and false gharial]). As outgroups, we included paralogous β-globin sequences from human, *Coelacanth*, *Xenopus*, and spotted gar (Figure S2B). We used this same set of β-globin sequences to estimate the pre-duplication, single-copy ancestor of the crocodilian *HBB-T1* and *HBB-T4* paralogs ('β AncCrocT1/T4') (Figure S6). All globin sequences used for ancestral reconstructions are listed in Table S4, and all reconstructed sequences that we used for ancestral protein resurrection were deposited in GenBank (ON959522-ON959565).

Vector construction and site-directed mutagenesis

The reconstructed amino acid sequences of AncArchosaur, AncNeornithes, AncCroc, and AncCrocT1T4 were reverse-translated to DNA sequence and were optimized for *E. coli* codon preferences. The α- and β-globin gene cassettes were synthesized by GeneArt Gene Synthesis (Thermo Fisher Scientific). The gene cassettes were subcloned into the pGM custom expression system as described previously.[18,19,41–46] Codon changes were engineered using the QuikChange II XL Site-Directed Mutagenesis kit (Agilent Technologies)(primers listed in Table S3), with all mutated sites verified by Sanger sequencing.

Protein expression and purification

Expression of recombinant Hbs (rHbs) was carried out in the JM109(DE3) *E. coli* strain (New England Biolabs). To ensure that N-terminal methionines were post-translationally cleaved from the nascent globin chains, host cells were co-transformed with a plasmid (pCOMAP) that contained an additional copy of the methionine aminopeptidase gene (MAP) along with a kanamycin resistance gene.[18,19] Both pGM and pCOMAP plasmids were co-transformed and subject to dual antibiotic selection in an LB agar plate

containing ampicillin and kanamycin. The expression of each rHb mutant was carried out in 1.5-2 L of TB medium. The cells were grown in an orbital shaker (New Brunswick) at 37°C, 200 rpm, until their absorbance values reached 0.6-0.8 at 600 nm. rHb expression was induced by the addition of 0.2 mM IPTG and the cultures were then supplemented with hemin (50 μg/mL) and glucose (20 g/L). Culture conditions and the protocol for preparing cell lysates were described previously.[18,19,41–46] The overnight culture was saturated with CO for 15 min, and the cells were centrifuged for 45 min at 15000*g* and 4°C. Pelleted cells were stored at -80°C until further use. The pellets were resuspended with Tris lysis buffer (50 mM Tris, 0.5 mM DTT, 1 mM EDTA) with lysozyme (1 mg/g of cells) for proper lysis before sonication. We used sonication cycles of 10 s pulse-on/20 s pulse-off for 15 min at 3.0 output (QSonicaQ500).[18,19] We then added polyethyleneimine solution to the crude lysate to a final concentration of 0.25% to 1% to precipitate the bacterial nucleic acids. The crude lysate was then centrifuged for 45 min at 15000*g* and 4°C, and the clarified supernatant was dialyzed overnight against the anion-exchange buffer for chromatography.

Recombinantly expressed proteins were purified using a two-step ion-exchange chromatography using the ÄKTA Start system (GE Healthcare) and Q-Sepharose columns (HiTrap QHP, 5 mL, 17-5159-01; GE Healthcare) pre-equilibrated with Tris/CAPS buffer (Table S5). For each rHb, the buffer pH and other purification conditions were optimized on the basis of isoelectric point (pI). Each rHb was eluted using a linear gradient of 0-1.0 M NaCl. The eluted sample was desalted and dialyzed overnight against the second column buffer. We used prepackaged SP-Sepharose columns (HiTrap SPHP, 5 mL, 17-516101; GE Healthcare) equilibrated with HEPES/phosphate buffer (Table S5). Each rHb sample was eluted with a linear gradient of 0-1.0 M NaCl with the corresponding buffer and pH. The purified rHb samples were analyzed using 4-20% SDS-polyacrylamide gel electrophoresis to assess the purity of the fractions prior to measurements of O_2-binding properties.

Collection of blood samples and preparation of hemolysates

We sampled blood of alligator embryos at the 40% stage of pre-hatching development by catheterizing the major artery of the chorioallantoic membrane, as described by Bautista et al.[23] Upon completion of catheter placement, 100 to 500 μL of blood was collected into heparinized microhematocrit tubes and was then centrifuged at 20,854g for 5 min to separate red cell and plasma fractions. Individual hemolysates were prepared from thawed blood or erythrocyte samples (20-50 μL) by adding an approximately fourfold volume of ice-cold 10 mM HEPES, 0.5 mM EDTA, pH 7.4, followed by incubation on ice for 30 min. Cell debris was separated from the hemolysate via centrifugation at 4°C (12,000 g, 15 min). Clarified hemolysates were stored at -80°C prior to the isolation and purification of Hb fractions used for the functional experiments.

Purification of native Hbs

Thawed hemolysates were desalted on a PD-10 column (GE Healthcare) equilibrated with 20 nM HEPES, 0.5 nM EDTA, pH 7.65. To separate embryonic and adult isoHbs of American alligator (HbI and HbII, respectively),[23] we subjected the hemolysates to anion-exchange chromatography on a Mono Q 5/50 GL column (1 mL) connected to an ÄKTA Pure Chromatography System (GE Healthcare). The column was equilibrated with 20 mM HEPES, 0.5 mM EDTA, pH 7.65, and isoHbs were eluted with a 0-0.1 M NaCl linear gradient at a flow rate of 1 mL/min. Absorbance was monitored at 415 and 280 nm to identify heme-containing proteins. Eluted peaks containing HbI and HbII were dialyzed against 10 mM HEPES, 0.5 mM EDTA, pH 7.4, and was then concentrated by ultrafiltration (Millipore) at 4°C to a final heme concentration of >0.05 mM.

Measurement of oxygen equilibration curves

Solutions of purified rHbs were desalted by passing samples through PD-10 desalting columns (GE Healthcare) that were equilibrated with 0.01 M HEPES, 0.5 mM EDTA, pH 7.4, and concentrated using Amicon Ultra-4 Centrifugal Filter Units (MilliporeSigma). O_2-equilibrium curves for Hb solutions (0.4 mM Hb tetramer in 100 mM HEPES, 0.5 mM EDTA buffer) were measured at 25°C using a Blood Oxygen Binding System (BOBS; Loligo Systems). BOBS is a temperature-controlled gas diffusion chamber connected to a programmable high precision Gas Mixing System pump (GMS – Loligo Systems) with a built-in spectrophotometer that allows the recording of changes in absorbance of a sample. Each Hb solution was sequentially equilibrated with 3-5 different partial pressures of O_2 (PO_2's) at saturation levels between 30 to 70%, while absorbance was continually monitored at 430 nm (deoxy peak) and 421 nm (oxy/deoxy isosbestic point). Hb-O_2 saturation was estimated at each equilibrium step by comparing the absorbance at 430 nm to fully oxygenated and deoxygenated baselines. Values of the PO_2 at half-saturation (P_{50}) and the cooperativity coefficient (n_{50}) were then estimated from linear Hill plots (log[fractional saturation/[1-fractional saturation]] vs. logPO_2). O_2-equilibrium curves were measured in the absence (stripped) and presence of Cl^- ions (0.1 M KCl) and organic phosphates (0.2 mM adenosine triphosphate [ATP]). O_2-equilibrium curves were also measured at three different pH levels, where the pH of working solutions was adjusted with NaOH to 7.2, 7.4, or 7.6 and were measured with an Orion Star A211 pH meter and an Orion PerpHecT ROSS Combination pH Micro Electrode. Linear regressions were fit to plots of logP_{50} vs. pH, and the resulting equation was used to estimate P_{50} values at pH 7.40 (± SE of the regression coefficient). The n_{50} values are presented as mean ± SEM of triplicate measurements.

Measurement of bicarbonate sensitivity using CO_2

Bicarbonate sensitivity of the ancestral and mutant rHbs was measured by using CO_2 gas as previously described.[7,47] We measured time-course changes in O_2 saturation of the samples using 3 μL samples (0.14 mM heme, 0.1 M HEPES, pH 7.2, 0.5 mM EDTA at 25°C) using BOBS. The system allows for the control of discrete O_2 and CO_2 tensions – balanced with ultrapure N_2 – delivered to the surface of the sample in the chamber. Each rHb sample was equilibrated to the partial pressure of O_2 needed to reach its

half-saturation (P_{50}), estimated from previously measured O_2 equilibrium curves. After equilibration to a PO_2 that matched the P_{50} of the sample, 1% CO_2 was added to the gas mixture while maintaining PO_2 constant, and the reduction in Hb-O_2 saturation (SO_2) was recorded. In this procedure, addition of 1% CO_2 did not change the pH of the buffer significantly,[7] whereby CO_2-induced changes in Hb-O_2 saturation were not due to the Bohr effect. Bicarbonate sensitivity was quantified as the log-transformed difference between SO_2 before and after the addition of 1% CO_2. These experiments were performed in triplicate and estimated quantities are reported as the mean ± 1 SEM.

Tests of positive selection

Using a maximum likelihood (ML) framework, we tested for evidence of positive selection in the crocodilian β-globin genes with codon-based models, as implemented in the *codeml* program in PAML v4.9.[33] We estimated phylogenetic relationships among crocodilian β-globin genes using IQTree version 2.1.3[35] under the best-fitting model selected by the ModelFinder subroutine[36] and we used chicken β^A-globin as an outgroup (Figure S6). We used the branch-site and clade models to examine variation in ω, the ratio of the rate of nonsynonymous substitution per nonsynonymous site, d_N, to the rate of synonymous substitution per synonymous site, d_S. Using both the branch and branch-site models,[48,49] we tested for changes in ω in the post-duplication line of descent leading from the single-copy progenitor of the *HBB-T1* and *HBB-T4* genes (β AncCrocT1/T4, which is also the common ancestor of all crocodilian β-type globin paralogs) to the most recent common ancestor of *HBB-T4* orthologs from all extant crocodilians (= the β-chain of AncCroc Hb)(Figure S6). We examined (*i*) the branch connecting β AncCrocT1/T4 to the pre-duplication ancestor of *HBB-T4* and *HBB-T6*, (*ii*) the stem of the *HBB-T4* clade, (*iii*) the branch connecting β AncCrocT1/T4 to β AncCroc, and (*iv*) the entire *HBB-T4* clade, stem plus crown, relative to the rest of the tree. These 2-ω branch models were compared to a model where all branches share the same ω (M0). We followed a similar procedure for the branch-site models, setting the same 4 sets of branches as foreground. In background branches, sites evolve under purifying selection ($\omega_0 < 1$) or under a neutral regime ($\omega_1 = 1$), whereas in foreground branches, some sites are allowed to switch from purifying selection to positive selection ($\omega_{2a} > 1$) or from a neutral regime to positive selection ($\omega_{2b} > 1$). The null models for these analyses set $\omega_2 = 1$.