Hemoglobins are ubiquitous in Eukarya and Bacteria but, until now, have not been found in Archaea. A phylogenetic analysis of the recently revealed microbial family of globin-coupled heme-based sensors suggests that these sensors descended from an ancient globin-only progenitor, or protoglobin (Pgb). Here, we report the discovery and characterization of two Pgb genes from the Archaea: ApPgb from the obligately aerobic hyperthermophile Aeropyrum pernix, and MaPgb from the strictly anaerobic methanogen Methanosarcina acetivorans. Both ApPgb and MaPgb bind molecular oxygen, nitric oxide, and carbon monoxide by means of a heme moiety that is coordinated to the protein through the F8 histidine (histidine 120). We postulate that these archaeal globins are the ancestors of contemporary hemoglobins.

Materials and Methods

Cloning and C-Terminal His-Tag Construction of the Pgb from A. pernix K1 (ApPgb). Genomic DNA from A. pernix served as the template for PCR amplification of the Pgb gene. NdeI and HindIII sites (underlined) were introduced in the primers for the PCR amplification (5′-CATATGACCAATGTATCTCC-3′; 5′-CTAAGCTTCCAGTCGTTTTCTTTTGTGTATGGA- TGCC-3′). The PCR product was cloned into the TOPO Blunt-PCR vector (Invitrogen) and then subcloned into the E. coli expression vector pET27 as a C-terminal His-tagged protein.

Expression and Purification of the Recombinant Globins in E. coli. Expression and one-step purification of the C-terminal His-tag Pgb were performed according to Pratt et al. (13). The MaPgb-MBP fusion protein was purified by affinity chromatography (amylose resin) according to the protocol from the manufacturer (New England Biolabs).

Preparation of Deoxy, Ferrous-Nitrosyl, and Ferric Globin. Both ApPgb and MaPgb purify in the ferric form; however, the last step in ApPgb purification involves elution with 250 mM imidazole, resulting in the imidazole-bound form. Removal of the imidazole was accomplished by mixing the proteins with concentrated dithionite (~200 mM), removing this reducing agent by gel filtration on a Sephadex G-25 (Pharmacia) column, and collecting the deoxy protein, all within an anaerobic glove box (Coy Laboratory Products, Ann Arbor, MI).Another form of preparation incorporates lower concentrations of dithionite (~10 mM) with methyl viologen as an electron shuttle. Ferrous-nitrosyl Pgb was produced by adding a 15:1 NO to protein molar ratio. Ferric Pgb was prepared either by allowing the reduced samples to autoxidize in air or by adding a 1:1 ratio of potassium ferricyanide to protein.

Absorption Spectra, Ligand Binding, and Autoxidation. Absorption spectra were measured with a Cary 4000 UV-Vis spectro-
photometer (Varian). Laser-flash photolysis and stopped-flow measurements were done with an LKS.60 laser kinetic spectrometer fitted with a PiStar stopped-flow drive unit (Applied Photophysics, Leatherhead, U.K.). For sample excitation, the LKS.60 spectrometer was coupled to a Quantel Brilliant B Nd:YAG laser with second-harmonic generation. Data acquisition was provided by an Agilent (Palo Alto, CA) 54830B digital oscilloscope for fast measurements or a 12-bit analog-to-digital converter (ADC) card within the workstation for slow measurements.

Ligand association and dissociation rates were measured for protein in 100 or 200 mM sodium phosphate (pH 7.0) at 25°C at the wavelength of maximum difference between the initial and final species. The carbon monoxide association rate was measured by laser-induced photodissociation of CO-saturated protein whereas the absorbance change was monitored at 421 nm. For the NO association rate, CO-saturated protein was photolyzed in the presence of dissolved NO, with absorbance changes monitored at 423 nm. At each ligand concentration, the observed association rate was measured at least three times.
times, and the true rate constants were calculated from linear plots of the observed rates, $k_{obs}$, vs. ligand concentration.

The CO off rate was determined by oxidation of CO-saturated protein with potassium ferricyanide (100 $\mu$M). The entire process was monitored at 422 nm over a time course of $\approx$6 min in the Cary 4000 spectrophotometer.

Oxy-protopglobin spectra were obtained in the Cary spectrophotometer and verified by performing flash-flow experiments on CO-bound protein injected with O$_2$-saturated buffer. In addition, spectra were obtained with a fiber optic spectrometer from Ocean Optics (Orlando, FL) by using a detector with a 2048-element linear silicon charge-coupled-device (CCD) array that gives almost instantaneous spectra, although lower quality (data not shown). Autoxidation rates were measured after the addition of air-saturated buffer to deoxy ApPgb in air.

Sequence Alignment and Phylogenetic Analysis. Pgbs and representative globin protein sequences were aligned in multiple stages. Groups of similar globins were first aligned separately and then combined together in a profile-based alignment by using CLUSTALX (14). The globin sequences were then manually adjusted in DNASTAR’s MEGALIGN (DNASTAR, Madison, WI) to preserve the integrity of both the solved and PRED predicted (15) secondary structures. Consensus sequences were calculated with CONSENSE (www.bork.embl-heidelberg.de/Alignment/consenseq.html). The HemAT globin domain sequence limits are based on the crystal structure of HemAT–Bs (Protein Data Bank ID code 1OR6). Bootstrapped (10,000 replicates) trees were created using CLUSTALX, and exported from TREEVIEW (16) and modified in Adobe ILLUSTRATOR (Adobe Systems, Mountain View, CA) for presentation. Sequences outside the traditional globin fold (A-H helices) were not included in the construction of the trees.

3D Homology Model. GENEMINE was used to create the ApPgb 3D homology model (truncated to 161 residues) based on the structure of the HemAT–Bs sensor domain (PDB ID code 1OR6) as guided by the alignment in Fig. 1A (10; www.bioinformatics.ucla.edu/~genemine/). Lack of a suitable structural template precluded us from including the N-terminal 44 residues that are not part of the traditional globin fold. HYPERCHEM was used to correct manually the bonds in the heme group, including energy minimization by using the Polak-Ribiere algorithm, terminating with a 0.1 kcal/(Å mol) gradient (www.hyper.com). Illustrations were created by using MOLSCRIPT and RASTER3D.

Molecular Dynamics Simulations (MDS) (Force Field and ab InItio MDS). Force field molecular dynamics used the CHARMM force field in the VMD/NAMD package (www.cs.uiuc.edu/Research/vmd/ and www.cs.uiuc.edu/Research/namd/) from the Theoretical and Computational Biophysics Group at the University of Illinois. Ab Initio molecular dynamics using the atom-centered density matrix propagation (ADMP) Method with 5,000-fs steps was used for characterization NO, CO, and H$_2$S vibrational bending (available in GAUSSIAN 03, www.Gaussian.com). Density Functional Theory (Becke 3 Lee Yang Parr or B3LYP) with 6-31G basis set was applied to the Fe and ligand, and Universal Force Field was applied to the rest of the heme group and a few of the surrounding protein residues, including the proximal Histidine. The protein model with His-120 covalently attached to the heme Fe was solvated in a water bath cube of $\approx$9,700 water molecules with cyclic boundary conditions throughout the simulation with cysteines 45 and 102 joined by a disulfide link. NO, CO, and H$_2$S all formed stable bonds with Fe at B3LYP/6-31G (see Figs. 6 and 7, which are published as supporting information on the PNAS web site).

Results and Discussion

Using the BLAST algorithm, all available complete and incomplete microbial genomes were queried with a consensus sequence corresponding to the heme-binding domain of the globin–coupled sensors (11). Initially, an $\approx$23-kDa (194-residue) protein was identified in the thermophilic cyanobacterium Thermosynechococcus elongatus. This sequence was further used to query the National Center for Biotechnology Information’s (nr) protein database with PSI-BLAST. After two iterations, the search results identified the globin–coupled sensors, the flavohemoglobins, and additional proteins with molecular masses ranging from $\approx$23 kDa (194 residues) to $\approx$27 kDa (227 residues). The additional proteins belonged to the archaea A. pernix and M. acetivorans, an Actinobacterium Thermobifida fusca, and the Green non-sulfur bacterium Chloroflexus aurantiacus (Fig. 1A). The archaeal Pgbs, ApPgb and MaPgb, were exactly 195 aa long: the minimum length found to be required for heme and reversible O$_2$ binding by the globin–coupled aerotaxis transducer HemAT–Hs (9).

To determine experimentally whether ApPgb and MaPgb bind heme, both genes were expressed in E. coli from recom
binant vectors, and the corresponding proteins were purified by His-tag and MBP affinity chromatography, respectively, according to Piatibratov et al. (13) (Fig. 2). Electronic absorption spectra of the liganded and unliganded ferrous (Fe\textsuperscript{2+}/H\textsubscript{11001}) (Fig. 3A and E) and ferric (Fe\textsuperscript{3+}/H\textsubscript{11001}) (Fig. 3B–D and F) forms of ApPgb absorb in the near ultra-violet and visible regions, a characteristic of heme proteins. Difference spectra (reduced minus oxidized) resulted in a peak at 564 nm, indicative of a \(b\)-type heme (data not shown). Similar results were found for MaPgb (data not shown).

Ferrous deoxy ApPgb bound molecular oxygen (O\textsubscript{2}), carbon monoxide (CO), and nitric oxide (NO) to form the corresponding oxy-, carbonmonoxy- (Fig. 3A) and nitrosyl-protoglobin (Fig. 3E) species. Addition of O\textsubscript{2} to the deoxy species caused the Soret band to shift from 433 nm to 410 nm (Fig. 3A), autoxidizing rapidly with half-lives of 4.3 (ApPgb) and 3.6 (MaPgb) minutes (Fig. 4C). In contrast, typical globins are known to autoxidize with half-lives on the order of hours to days (17, 18), and both the full-length and truncated versions of the HemAT-Hs protein have been shown to form stable O\textsubscript{2}-bound complexes (9). The relatively rapid oxidation of ApPgb in air suggests that this protein may have a role in transferring electrons to O\textsubscript{2} under oxic conditions, possibly as a strategy for O\textsubscript{2} detoxification.

Fig. 2. Electrophoresis of purified ApPgb and MaPgb. Lane 1, protein molecular weight markers; lane 2, ApPgb\textsubscript{pernix} from A. pernix; lane 3, MaPgb-MBP fusion from M. acetivorans; lane 4, MaPgb-MBP digested with Factor Xa; lane 5, purified MaPgb. The apparently larger molecular weight of ApPgb\textsubscript{pernix} is attributed to the His tag construction.

Fig. 3. Absorption spectra of ApPgb in different redox states and bound to various ligands. (A) Ferrous (Fe\textsuperscript{2+}) ApPgb in the CO-bound (dotted line) and the O\textsubscript{2}-bound (dashed and dotted line) forms. (B) Ferric (Fe\textsuperscript{3+}) met (solid line) and azido (N\textsubscript{3}, dashed line) ApPgb. (C) Cyano (CN\textsuperscript{-}, dotted line) ApPgb. (D) Imidazole-bound (dashed and dotted line) ApPgb. (E) Ferrous deoxy (dotted line) ApPgb saturated with NO to give the Fe\textsuperscript{2+}–NO (solid line) species. (F) Ferric met (dotted line) and the Fe\textsuperscript{3+}–NO (solid line) species. When NO is added to ferrous deoxy ApPgb in approximately a 20:1 ratio, a spectrum identical to the Fe\textsuperscript{3+}–NO spectrum results. All ligands were added to saturation. MaPgb absorption spectra are virtually identical except for a 2-nm blue shift at each \(\lambda_{\text{max}}\). (G) Comparison of wild-type ApPgb bound with imidazole (solid line) to the H120A mutant in the presence of imidazole (dashed and dotted line). The H121A mutant possesses a wild-type spectrum (data not shown). All measurements were done at 25°C and pH 7.0 in either 100 or 200 mM sodium phosphate buffer.

Fig. 4. Parameters for binding of ligands to ferrous ApPgb at 25°C and pH 7.0. (A) Association rates for binding of CO (triangles) and NO (circles), as determined by laser-flash photolysis. (B) Dissociation rate of CO, as determined by oxidation of CO-saturated protein with potassium ferricyanide. (C) Autoxidation of ApPgb on exposure of the deoxy form to air. Details are in Materials and Methods.
Carbon monoxide binding to ApPgb produced a typically intense Soret absorption band at 422 nm (Fig. 3A). The rate constants for CO association and dissociation at 25°C and pH 7 were 11 μM⁻¹s⁻¹ (Fig. 4A) and 47 ms⁻¹ (Fig. 4B and Table 1), respectively. Similar values for those parameters have been measured for Hbs from plants and other species. For example, the CO association rate constant for ApPgb falls within the range reported for rice Hb (7.2 μM⁻¹s⁻¹) and Paraspongilla Hb (14 μM⁻¹s⁻¹) (19, 20). The CO dissociation rate constant for ApPgb is similar to values reported for the heme-based sensor kinase BjFixL from Bradyrhizobium japonicum (45 ms⁻¹) and the phosphodiesterase ArPEDE1 from Acetobacter xylidum (60 ms⁻¹) (21, 22). The equilibrium association constant for binding of CO to ApPgb (230 μM⁻¹) falls between the values measured for elephant Hb (140 μM⁻¹) and barley Hb (520 μM⁻¹) (17, 23).

The ferrous forms of the Pgbs were also found to bind NO, which is now known to be a physiological ligand of bacterial as well as eukaryotic Hbs (24–26). Binding of NO to deoxy ApPgb caused a blue-shift of the Soret absorption band from 433 nm to 419 nm (Fig. 3E). The association rate constant for NO binding to ferrous ApPgb was 59 μM⁻¹s⁻¹ (Fig. 4A and Table 1), approximately three times higher than for sperm-whale Mb (SWMb) but three times lower than for leghemoglobin (27, 28). Ferric ApPgb also found to bind NO, yielding an absorption spectrum with a Soret band at 425 nm and clear α and β bands at 574 and 542 nm, respectively (Fig. 3F). Similar results in ligand binding were obtained with MaPgb (data not shown).

The Pgbs feature three residues that are thought to have an ancient origin: a proximal histidine (His F8), a cysteine at the E19 position (Cys E19), and a distal tyrosine (Tyr B10). In an alignment of SWMb with the truncated Hbs, the proximal histidine (His F8) was the only residue found to be absolutely conserved (29). In the ApPgb and MaPgb proteins, there are two histidine residues near the site of heme binding, His 120 and His 121. To determine which of those histidines coordinates to the heme iron, each of those two histidine residues in ApPgb was independently changed to alanine by site-directed mutagenesis. The H121A ApPgb retained a wild-type absorption spectrum whereas the H120A ApPgb almost entirely lost its visible absorption (Fig. 3G). A 3D homology model, created for ApPgb based on the ferric unliganded HemAT-Bs sensor domain (PDB ID code 1OR6) (Fig. 1B) and refined by force-field molecular-dynamics simulations, confirmed the identity of the proximal histidine as His 120.

Molecular modeling of ApPgb also identified Cys-102 as being analogous to the E19 cysteine near the E-helix terminus of both the Ascaris suum Hb (AscarisHB) and the H2S-binding annelid Hb from Riftia pachyptila. This modeling further indicated that Cys-102 and Cys-45 at the A-B helical junction are precisely positioned to form a disulfide bridge (Fig. 1B). Such a disulfide linkage should contribute to the stability of those thermophilic Pgbs. Interestingly, a Neighbor-Joining tree (Fig. 5) shows that the Pgbs and the Hbs from Ascaris suum and R. pachyptila group within the same branch. In the As. suum Hb, the positioning of the Cys E19 near the distal pocket has been shown to play a role in its NADPH-dependent NO-activated deoxygenase function (24). In the annelid Hb from R. pachyptila, however, Cys E19 is thought to be important for H2S binding (30). The possibility of cysteine thiol binding atypical ligands and bringing about diseased states has been suggested as a possible driving force for the evolutionary loss of H2S-binding in Hbs from organisms living in sulfide-free habitats (31). This hypothesis is consistent with ancient globins working to detoxify sulfide and nitric oxide, and the Cys E19 being absent from modern globins that have adapted to oxic environments.

In the Mycobacterium tuberculosis Hb, HbN, a distal tyrosine (Tyr B10) is thought to be an ancient adaptation for the scavenging and detoxification of NO, and its gradual loss during evolution has been proposed as a possible requirement for efficient O2 transport by Hbs (32). In the As. suum Hb, the globin-coupled sensors, and several other known globins, Tyr B10 is thought to play a central role in O2 binding due to stabilization of the bound O2 by hydrogen bonding (6, 7, 11, 18, 33–37). Replacement of the Tyr B10 with phenylalanine in the As. suum Hb resulted in rapid autoxidation of this protein (33). For both the ApPgb and MaPgb Pgbs, molecular modeling indicates that, although a Tyr B10 is present, this residue is unlikely to stabilize O2 binding because the lowest-energy orientation of the side chain is parallel to the heme plane rather than directed at the site of ligand coordination. This result is consistent with the high autoxidation rate observed for the Pgbs when exposed to O2 (Fig. 4C).

In summary, we have identified globins from Archaea that bind O2, CO, and NO. Mutagenesis experiments coupled with our 3D homology model of ApPgb indicate that heme is covalently bound to the apo-globin by means of His-120. In addition, the model predicts an intramolecular disulfide bridge aiding in ApPgb thermostability. Moreover, the B10 distal tyrosine is oriented parallel to the heme plane rather than directed toward the bound ligand, decreasing the Pgb’s capacity to bind O2. In summary, we conclude that ApPgb and

Table 1. Ligand binding properties of ApPgb compared with sperm whale myoglobin and the Ascaris hemoglobin

<table>
<thead>
<tr>
<th></th>
<th>CO</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_{on}$, $\mu M^{-1} s^{-1}$</td>
<td>$k_{off}$, s⁻¹</td>
</tr>
<tr>
<td>ApPgb</td>
<td>11*</td>
<td>0.047*</td>
</tr>
<tr>
<td>SWMb</td>
<td>0.55*</td>
<td>0.019*</td>
</tr>
<tr>
<td>Ascaris Hb</td>
<td>0.17*</td>
<td>0.018*</td>
</tr>
</tbody>
</table>

*These studies: pH 7, 25 °C. Error is ±15%.
†Within a species.
‡Ref. 38; pH 7, 20 °C.
§Ref. 17; pH 7, 37 °C (in air).
¶Ref. 39; pH 7, 20 °C.
‖Ref. 40; pH 6, 12 °C (in air).

Freitas et al.
MaPgbs are prototypes for contemporary Hbs. They may serve several biological functions, including protection from nitrosative and oxidative stress.

This work is dedicated to Dr. Oskar Zaborsky. We thank Dr. Yutaka Kawarabayashi (Center for Glycoscience, Advanced Industrial Science and Technology, Japan) for providing the *A. pernix* genomic DNA, and Dr. William W. Metcalf (Department of Microbiology, University of Illinois at Urbana–Champaign) for providing the *M. acetivorans* genomic DNA. This investigation was supported by National Science Foundation Grant MCB0080125 and by a University of Hawaii intramural grant (to M.A.) and by U.S. Public Health Service Grant HL-64038 (to M.-A.G.-G.).