Globins are widely studied proteins because of their ability to transport and/or store diatomic oxygen. In invertebrate organisms, globin research is focused on the evolution from simple anaerobic organisms to complex aerobic organisms. Aerotaxis, the movement of an organism toward or away from oxygen, is an important aspect of this research, as aerotactic responses are triggered by a change in oxygen levels. Understanding the mechanisms behind aerotaxis will yield much insight as to how an organism adapts and evolves to changes in its environment. A subfamily of the globin-coupled sensors (GCSs), called heme-based aerotaxis transducers (HemATs), are responsible for aerotaxis. It is believed that the binding of oxygen to the transducer’s globin domain causes a conformational change in the protein’s structure, thus altering the activity of its signaling domain, allowing the organism to adapt to changes in oxygen levels in the surrounding environment. Two HemATs were recently identified in the thermophile *Anoxybacillus flavithermus*. In this study, these two proposed HemATs were expressed in BL21(DE3) *pLysS* E. coli and purified using Co²⁺ affinity chromatography.

INTRODUCTION

Globins are found in every kingdom of life and are especially important because of their ability to transport and store diatomic oxygen. Phylogenetic analyses and structure comparisons of invertebrate and vertebrate globin sequences reveal that globins have evolved from a common ancestor. Vertebrate globins, however, are much more widely studied than invertebrate globins and have revealed a wealth of information with respect to the relationship between structure and function, protein folding, interactions between subunits and ligands, and the effect of mutations on protein function. Invertebrate globins, though not as well understood as vertebrate globins, offer much in the study of the structural and functional evolution of globins.

In 2001, Hou et al identified a family of multi-domain proteins that they named globin-coupled sensors (GCSs). These proteins consist of a heme-binding sensor domain exhibiting a globin fold in the amino terminus and one or more signaling domain(s) in the carboxyl terminus. The GCS globin domain consists of approximately 195 amino acids and about 40 amino acids longer than contemporary myoglobins. Thus, it is believed that two protoglobins (Pgbs) found in the *Archaea Aeropyrum pernix* and *Methanosarcina acetivorans*, both approximately 195 amino acids in length, are the progenitors of the globin-coupled sensor.

Globin-coupled sensors are divided into two subfamilies: aerotaxis proteins and gene regulators. Two globin-coupled sensors belonging to the aerotactic subfamily have recently been identified in the thermophile *Anoxybacillus flavithermus* (formerly known as *Bacillus flavithermus*). These GCSs are heme-based aerotaxis transducers (HemAT) that contain a methyl-accepting chemotaxis protein signaling domain in the C-terminus. Aerotaxis is a response that either attracts or repels cells away from oxygen; thus, understanding the mechanisms that trigger an aerotactic response will reveal much about the adaptation of an organism to changes in oxygen levels and, on a broader level, the evolution of the organism. The objective of this project is to obtain purified samples of the proposed *A. flavithermus* aerotaxis transducers for further biochemical testing.

METHODS

Bioinformatics Analyses

Amino acid sequence of *hemAT-Bf1* and *hemAT-Bf2* were submitted to seven online biological databases to compare them to existing proteins. The databases used in this project are the National Center for Biotechnology Information’s Basic Local Alignment Search Tool (BLAST) for Proteins, Position-Specific Iterated BLAST (PSI-BLAST), Conserved Domain Database (CDD), and Clusters of Orthologous Groups (COGs), DAS Transmembrane Prediction Serve, the Simple Modular Architecture Research Tool (SMART), and Structure Assignment With Text Description (SAWTED).

Amplification and Cloning of *hemAT-Bf1* and *hemAT-Bf2*

The genomic DNA of *Anoxybacillus flavithermus* served as the template in the PCR amplification of *hemAT-Bf1* and *hemAT-Bf2*. Each PCR reaction...
subcloned 1µL DNA template, 5µL 10× Pfu buffer (Stratagene), 5µL 10mM dNTPs (Invitrogen), 3µL of dimethyl sulfoxide (DMSO, Invitrogen), 2µL of each primer (10 pmol/µL), 1µL Pfu DNA polymerase (Stratagene), and 31µL HPLC water. Primers contained NdeI and HindIII sites: HemAT-Bf1: Top 5’-AACATATGACTGGAGGG GAAATGATGTC-3’ Bottom 5’-AAAAGCTTCCGATGT TGCAACT CTGC-3’ HemAT-Bf2: Top 5’-AACATATGGCAC CCAAC ATGGGTG-3’ Bottom 5’-AAAAGCTT CTCGAG AAGGATGTC-3’

After a hot start, the mixture was denatured at 94°C for 30s, annealed for 30s at 55°C, and elongated at 72°C for 1 min for 25 cycles. The PCR products were cloned using the TOPO cloning vector kit (Invitrogen).

Subcloning of hemAT-Bf1 and hemAT-Bf2

Plasmids containing the hemAT-Bf1 and hemAT-Bf2 inserts were isolated using the Promega Plasmid Isolation Kit. An EcoRI mini-digestion was run to confirm the presence of the inserts. Each reaction contained 3µL plasmid DNA, 0.5µL of EcoRI restriction enzyme, 1.5µL 10× Buffer H, and 10µL HPLC water. The reactions were incubated in a water bath for one hour at 37°C and run on an ethidium bromide 1% agarose gel. A preparative digestion was then performed to prepare the target genes and pET21a(+) (Novagen) vectors containing a universal His-tag for ligation. The plasmid DNA and the expression vector were digested separately with the reaction mixture as follows: 6µL DNA, 3µL 10× Buffer D, 1µL NdeI restriction enzyme, 1µL BamHI restriction enzyme, and 19µL HPLC water. The reactions were incubated for two hours in a water bath at 37°C and run on a 1% agarose gel. After one hour, 3µL 10× alkaline phosphatase buffer and 1µL alkaline phosphatase enzyme (Promega) were added to the reactions containing the vector. The digests were excised from the gel and purified using the GeneClean Kit (QBiogene). The purified inserts and vectors were then ligated at 14°C for 8 hours using 8µL DNA mix (insert and vector), 1µL T4 ligase, and 1µL 10× T4 ligase buffer. The ligation reaction was then transformed into TOP10 competent cells (Invitrogen). A mini-digestion as described above confirmed the presence of the ligated genes.

Expression and Purification of hemAT-Bf1 and hemAT-Bf2

hemAT-Bf1 and hemAT-Bf2 ligations transformed into BL21(DE3) Escherichia coli pLysS cells (Novagen), a host for high expression of recombinant proteins. Small-scale induction of the GCs was then performed. The cells were incubated into 1mL LB media containing ampicillin and chloramphenicol and incubated at 37°C in a shaker until turbid. 200µL of the culture was transferred to 5mL LB medium with ampicillin and chloramphenicol and incubated until the turbidity reached approximately 0.5 OD600. Gene expression was induced by the addition of 3µL 1M isopropyl-beta-D-thiogalactopyranoside (IPTG), and the cultures were incubated for another two hours. 300µL of the cultures were removed and 30 µL of trichloroacetic acid (TCA) were added to each sample. The mixtures were vortexed and then centrifuged for 10 min at 4°C at 14000 rpm, to harvest the cells. The supernatant was aspirated and the pellets were resuspended in 400µL of 50% acetone. Again, the suspensions were centrifuged for 10 min at 4°C at 14000 rpm and the supernatant was aspirated. The pellets were dried for 20 min, and 30µL of 1× sample buffer was added to each pellet. The mixtures were heated at 97°C in an Eppendorf Thermomixer R at 1000 rpm for 5 min and a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was run with the heated mixtures to confirm the presence of the expressed GCs. Having confirmed that hemAT-Bf1 and hemAT-Bf2 were expressed, large-scale induction was then performed. Cells were grown in 1L LB medium containing ampicillin and chloramphenicol and induced with 600µL 1M IPTG and a heme precursor mixture containing 17mg amnolevulinic acid and 100mg ferrous sulfate. After two hours of incubation, the bacterial cells were harvested by centrifugation. The pellet was resuspended in 20mL 200mM NaCl 50mM Na2HPO4 buffer and sonicated on ice for six cycles of 3.5 pulses per second at 20s intervals for two minutes to lyse the cells. The suspension was centrifuged for 20 min at 14000 rpm and the supernatant was filtered through a 0.20µm pore size Millipore filter. Affinity chromatography was performed using the supernatant according to Piatibratov et al.

Absorption Spectra of hemAT-Bf1 and hemAT-Bf2

Purified samples of hemAT-Bf1 and hemAT-Bf2 were concentrated by centrifugation in a 4mL Millipore centrifugal filter and tube for proteins with a 10000 MWCO. A Varian Cary 1E UV-Visible Spectrophotometer was used to measure the absorption of the two HemATs. Spectra were also taken after the two HemATs were deoxygenated with sodium dithionite and after they were reoxygenated by exposure to oxygen.

Results and Discussion

Data collected determined that hemAT-Bf1 and hemAT-Bf2 contain an amino terminus globin domain and a carboxyl terminus methyl-accepting chemotaxis protein (MCP) signaling domain. The amino acid sequence alignment of hemAT-Bf1, hemAT-Bf2, and hemAT-Bs, from Bacillus subtilis in
MegAlign (DNASTar), shows the proximal histidine and the phenylalanine between the C and D helices conserved in all globins in the N-terminus and high homology in the C-terminus MCP signaling domain.

Homology models of hemAT-Bf1 and hemAT-Bf2, generated by NCBI’s Cn3D 4.1 modeling program through the BLASTP and CDD searches, were based on the heme-binding domain of hemAT-Bs from Bacillus subtilis and the methyl-accepting chemotaxis protein domain from the Escherichia coli serine chemoreceptor, Tsr (Figure 2). These preliminary results validate the hypothesis that hemAT-Bf1 and hemAT-Bf2 are heme-binding proteins.

hemAT-Bf1 and hemAT-Bf2 were expressed in BL21(DE3) Escherichia coli pLysS and purified through 6-histidine tag affinity chromatography with imidazole as the eluent. The purified samples collected were run on an SDS-PAGE to confirm the presence of the two hemATs. Absorption spectra showed that hemAT-Bf1 and hemAT-Bf2 are capable of reversibly binding oxygen. Absorption maxima in the oxy state of hemAT-Bf1 and hemAT-Bf2 occur at approximately 410nm (Soret), 580nm (α-band), and 540nm (β-band). When deoxygenated with the introduction of sodium dithionite to the protein solutions, the Soret band shifts to approximately 430nm and the α- and β-bands merge at about 560nm. After reoxygenation of the two HemATs, the bands shift back to the original positions.

CONCLUSION

Preliminary bioinformatics studies identified hemAT-Bf1 and hemAT-Bf2 as globin-coupled sensors containing a globin domain in the amino terminus and a methyl-accepting chemotaxis protein signaling domain in the carboxyl terminus. hemAT-Bf1 and hemAT-Bf2 were then expressed in BL21(DE3) E. coli pLysS competent cells and purified using the Co3+ affinity chromatography method devised by Piatibratov et al. Spectral analyses of the purified protein samples have confirmed that the two HemATs are heme-based proteins and can reversibly bind oxygen, as shown in the heme-based aerotaxis transducers found in Archaea Halobacterium salinarum, the Bacterium Bacillus subtilis, and myoglobin. Currently, chimeric constructions of the two HemATs are currently being synthesized using the amino termini of the hemAT-Bf1 and hemAT-Bf2 and the signaling domain of the E. coli aerotaxis transducer Tsr. Motion analysis, a method of tracking the swimming behavior of cells, is being performed to determine whether the globin domain triggers the signaling domain. These tests will help to reveal the mechanisms behind oxygen sensing and the adaptation of the organism to changes in oxygen levels.

ACKNOWLEDGMENTS

The author thanks Dr. Maqsudul Alam for his mentorship; Dr. Shaobin Hou, Xuehua Wan, Jennifer Saito, for their help in conducting the research; Tracey Freitas, Jimmy Saw, Jing Liu, and Valerie Kawabata for their assistance and support; and Dr. George Hui for his guidance.

Thank you to the National Institutes of Health, the National Institute of Diabetes and Digestive and Kidney Diseases, and the Charles R. Drew University of Medicine and Science for the opportunity to continue conducting research in a professional institution.

This research was supported by the National Science Foundation Grant MCB0080125.

REFERENCES