BLAST Project

Tyler Stack
tms2659
Oct. 10, 2013

Seq. #0: DVSEYTKADL FQPGKVTPLA
Catalase, from *Corynebacterium glutamicum*

>gi|19551506|ref|NP_599508.1| catalase [Corynebacterium glutamicum ATCC 13032]
MSEKSAADQIVDRGMRPKLSGNTTRHNGAPVPSENISATAGPQGPNVLNDIHLIEKLAHFNRENVPERIPHAKGHGAFGELHITEDVSEYTKADLFQPGKVTPLAVRFSTVAGEQGSPD
TWRDVHGFALRFYTEEGNYDIVGNNTPTFFLRDGMKFPDFIHSQKRLNKNGLRDADMQWDFWTRAPESAHQVTYLMGDRGTPKTSRHQDGFGSHTFQWINAEGKPVWVKYHFKTRQGWD
CFTDAEAAKVAGENADYQREDLYNAIENGDFPIWDVKVQIMPFEDAENYRWNPFDLTKTWSQKDYPLIPVGYFILNRNPRNFFAQIEQLALDPGNIVPGVGLSPDRMLQARIFAYADQQ
RYRIGANYRDLPVNRPINEVNTYSREGSMQYIFDAEGEPSYSPNRYDKGAGYLDNGTDSSSNHTSYGQADDIYVNPDPHGTDLVRAAYVKHQDDDDFIQPGILYREVLDEGEKERLADN
ISNAMQGISEATEPRVYDYWNNVDENLGARVKELYLQKKA

>gi|58036263:274324-275874 Corynebacterium glutamicum ATCC 13032, complete genome
ATGTCTGAGAAGTCAGCAGCAGACCAGATCGTAGATCGCGGAATGCGTCCAAAGCTTTCTGGAAACACTACCCGCCACAACGGAGCACCAGTTCCATCTGAGAACATCTCCGCAACCGC
AGGCCCACAGGGTCCAAACGTTCTCAATGACATTCACCTCATTGAAAAGCTCGCACACTTTAACCGCGAGAACGTTCCAGAGCGTATCCCTCACGCAAAGGGCCACGGCGCTTTCGGTG
AGCTGCACATCACCGAGGACGTATCCGAATACACCAAGGCAGACCTGTTCCAGCCTGGTAAGGTCACCCCGCTGGCTGTTCGCTTCTCTACTGTTGCTGGTGAGCAGGGCTCCCCAGAT
ACCTGGCGCGACGTCCACGGCTTCGCTCTTCGCTTCTACACCGAAGAGGGCAACTACGACATCGTGGGTAACAACACCCCAACCTTCTTCCTTCGTGACGGCATGAAGTTCCCGGACTT
CATCCACTCACAGAAGCGTCTCAACAAGAACGGTCTGCGCGATGCAGACATGCAGTGGGATTTCTGGACCCGCGCACCTGAATCTGCACACCAGGTGACCTACCTGATGGGTGACCGCG
GTACCCCTAAGACCTCCCGCCACCAGGACGGCTTCGGCTCCCACACCTTCCAGTGGATTAACGCTGAAGGTAAGCCAGTTTGGGTTAAGTACCACTTCAAGACCCGCCAGGGCTGGGAT
TGCTTCACCGATGCAGAAGCAGCAAAGGTTGCAGGCGAGAACGCTGACTACCAGCGCGAAGACCTCTACAACGCTATTGAAAACGGCGACTTCCCAATCTGGGACGTCAAGGTTCAGAT
CATGCCTTTCGAGGATGCAGAGAACTACCGCTGGAACCCATTCGACCTGACCAAGACCTGGTCCCAGAAGGATTACCCACTGATCCCAGTCGGTTACTTCATCCTGAACCGCAACCCAC
GCAACTTCTTCGCTCAGATCGAGCAGCTTGCACTGGATCCAGGCAACATCGTTCCTGGCGTCGGCCTGTCCCCAGACCGCATGCTCCAGGCACGTATCTTCGCATACGCTGACCAGCAG
CGTTACCGCATCGGCGCTAACTACCGCGACCTGCCAGTGAACCGTCCAATCAACGAGGTCAACACCTACAGCCGCGAAGGTTCCATGCAGTACATCTTCGACGCTGAGGGCGAGCCTTC
CTACAGCCCTAACCGCTACGACAAGGGCGCAGGCTACCTGGATAACGGTACGGATTCCTCCTCCAACCACACCTCCTACGGCCAGGCTGATGACATCTACGTCAACCCAGACCCACACG
GCACCGACCTGGTTCGTGCTGCTTACGTCAAGCACCAGGATGATGACGACTTCATCCAGCCAGGCATCCTATACCGCGAGGTCCTGGATGAGGGCGAGAAGGAGCGATTGGCAGACAAC
ATCTCCAACGCAATGCAGGGCATCTCTGAGGCAACCGAGCCACGCGTCTACGACTACTGGAACAACGTTGATGAGAACCTCGGCGCTCGCGTCAAGGAGCTTTACCTCCAGA
AGAAGGCTTAA

The initial BLAST result from the 20-mer provided four hits with 100% query cover and identity, all part of the sequence for catalase from the organism *Corynebacterium glutamicum*. The protein and the coding DNA sequences are provided above. Catalase is an important enzyme, found in many aerobic organisms, catalase protects the cell from reactive oxygen species by combining two molecules of hydrogen peroxide to form water and molecular oxygen, normally using a heme cofactor.[1] This protein is diffusion controlled, and can perform this dismutation reaction at one of the highest efficiencies known for an enzyme.[2] Catalase from certain organisms is capable of binding $NADP^+$ if these enzymes contain a Rossman fold, and the current reasoning for this binding is currently under debate.[3] It is proposed that $NADP^+$ binding protects the enzyme from inactivation.[4]

In order to clone this gene and express the protein recombinately in *E. coli*, the gene will be cloned into pET-14b, which will provide a His_6-Tag on the N-terminus. In order to clone this gene, suitable primers will be designed to amplify the gene of interest with proper restriction sites (NdeI and XhoI) so the gene may be inserted into pET-14b. An online tool (http://tools.neb.com/NEBcutter2) confirmed that the gene does not contain the recognition sites for either NdeI or XhoI. The forward primer will have the sequence 5' TGGTGT**CATATG**TCTGAGAAGTCAGC 3' and the reverse primer will have the sequence 5' TAATAT**CTCGAG**TTAAGCCTTCTTCTTGG 3', where the restriction sites have been bolded. An online primer calculator tool (http://www6.appliedbiosystems.com/support/techtools/calc/) predicts the primers to have a melting temperature of 62 °C.

An image of this protein is shown in Figure 1. The protein is a tetramer and binds heme and $NADP^+$ (shown in Figure 1A). Catalase is "remarkably stable" and part of the reason is because of the unique interface where different subunits form a knot-like structure (Figure 1B).[1] The active site heme group uses Tyr353 as a lower ligand, and His71 and Asn143 sit above the planar heme group and are important for catalysis (Figure 1C).[1] A multiple sequence alignment was prepared to compare the protein from *C. glutamicum* and other organisms, ranging from bacteria to humans. The multiple sequence alignment is shown below in Figure 2, and there are many residues that are conserved across the different organisms. Some residues are found around the heme group and form the active site and likely orient and activate the dismutation reaction as mentioned above or bind with the heme (Arg107, Arg122, Arg349, Arg360). There is a central ß-barrel that makes the core of the protein, and a number of aromatic residues (Phe108, Phe131, Tyr132, Trp272, Trp298) are conserved and make the core of this barrel, and Thr102 makes hydrogen bond interactions in place of an amide bond to continue the ß-strand. Other conserved residues include those involved in the knot, making tight turns in loops (Gly41, Gly44, Pro45, Pro336, Gly337) and making intersubunit contacts (His59, Phe60, Glu63) and those involved in binding $NADP^+$ by providing positive charges to interact with the phosphates (His230, Lys232).

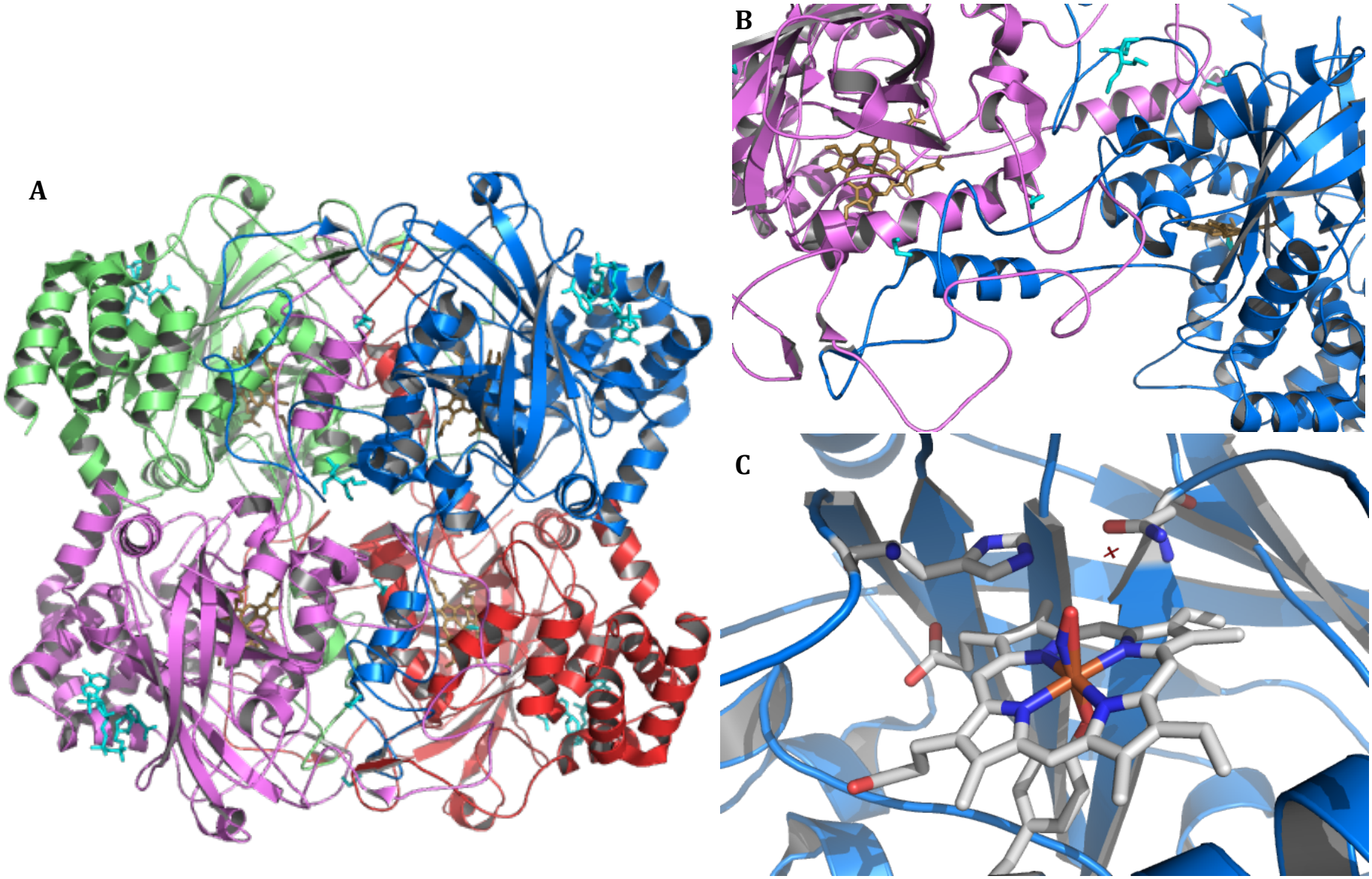

Figure 1: Structure of catalase from *C. glutamicum* (PDB 4B7F). **A.** Full tetramer of catalase. Each subunit is a different color. Heme is colored brown, while $NADP^+$ is colored cyan. **B.** An example of the knot that forms between two subunits. Coloring is similar to previous image. **C.** The active site of catalase. The heme group is liganded by Tyr353 and NO, while the active site residues His71 and Asn143 sit above the plane making hydrogen bonds with a water molecule. The water molecule is shown as a red cross, and the heme group, the NO molecule and the protein side chains are colored by element: carbon, nitrogen, oxygen, and iron are colored gray, blue, red, and brown, respectively. All images were made with PyMOL.

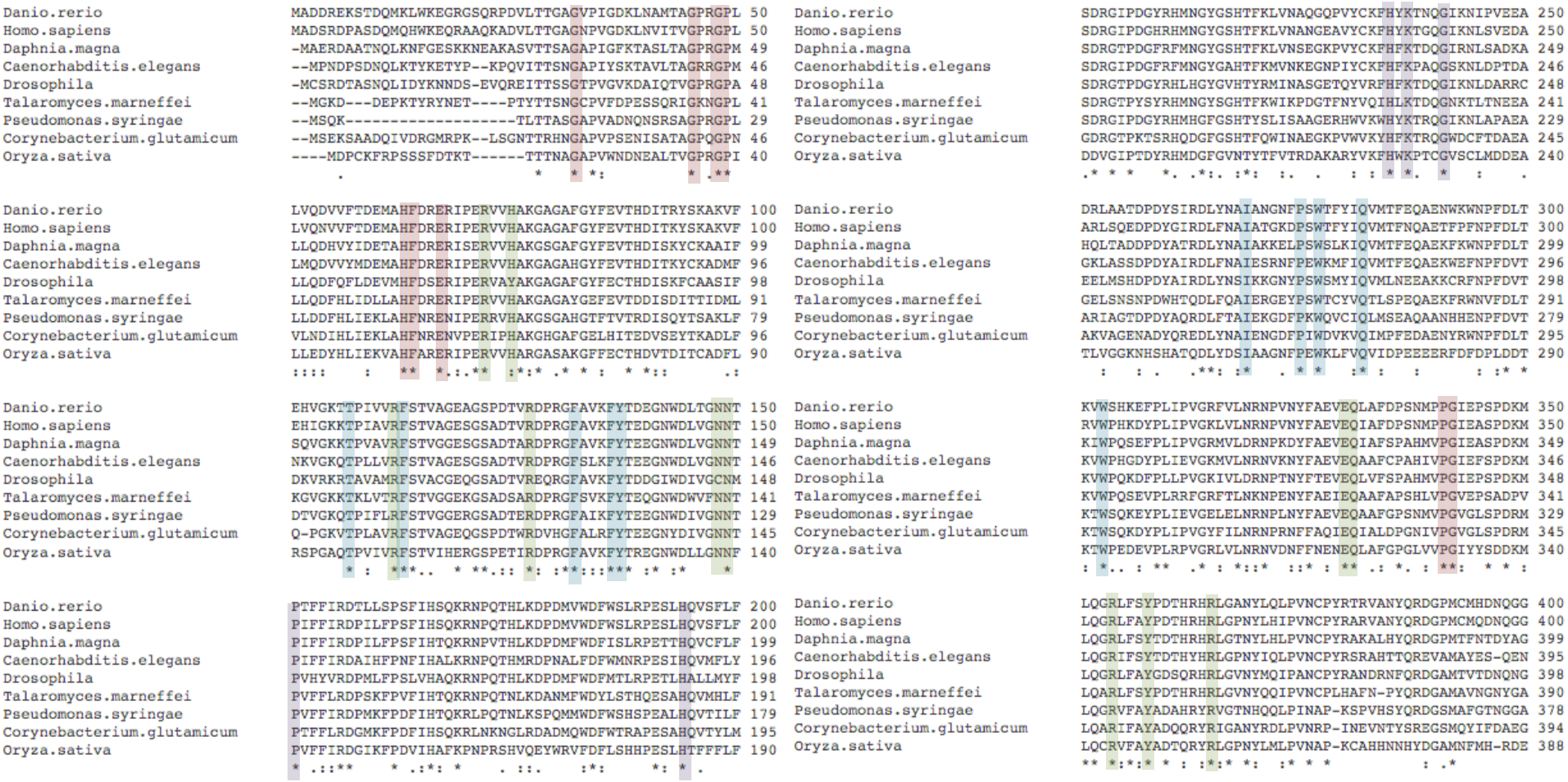

Figure 2: Multiple sequence alignment of catalase from *Corynebacterium glutamicum* and other species. Across the sequences, an identical residue is designated with an asterisk (*) while conservative mutations are shown by a colon (:), while moderately conservative mutations are shown with a period (.), and the amino acid residue number is shown to the right. Some of the conserved residues are categorized, where green residues are involved in the active site, blue residues are important for forming the hydrophobic interior, purple residues make contact with $NADP^+$, and red residues are involved in the 'knot' formed between subunits.

References.

1. Bravo J, Mate MJ, Schneider T, Switala J, Wilson K, Loewen PC, Fita I. Structure of catalase HPII from Escherichia coli at 1.9 A resolution. *Proteins*. **1999**, 155-166.

2. Vainshtein BK, Melik-Adamyan WR, Barynin VV, Vagin AA, Grebenko AI. Three-dimensional structure of the enzyme catalase. *Nature* **1981**, 411 - 412

3. Fita I, Rossmann MG. The NADPH binding site on beef liver catalase. *Proc. Natl. Acad. Sci. USA* **1985**, 1604-1608.

4. Zamocky M, Furtmüller PG, Obinger C. Evolution of catalases from bacteria to humans. *Antioxid Redox Signal.* **2008,** 1527-1548.