

Knitr Template

Zachary Foster

May 13, 2014

I decided to explore the structure of 354aa long protein of unknown function from *Pelagibacter* strain 7214 using a variety of structural prediction methods. I submitted the unknown sequence to the Phyre2 server using the normal mode. 62 residues (18% of the total sequence) hit a heptosyltransferase with 85.4% confidence (Figure 1). I also submitted the sequence to SWISS-MODEL using the Build Model function. The top hit was to a light-independent protochlorophyllide reductase subunit B, but was not estimated to be significant (Figure 2). Next, the query was submitted to iTASSER. The best hits were to PDB proteins with a hydroxycinnamoyltransferase function. All 10 top hits were some kind of transferase. The best hit had 98% coverage, but only 8% sequence identity (Figure 3). Function Prediction using COFACTOR and binding site prediction suggested similarity to acetyltransferases. I then searched ProSite for hits to the query using the ScanProsite tool, but there were no hits. PFAM was also searched and only one hit to a protein of unknown function was returned (Figure 4). I attempted to use SSCP (Secondary Structural Content Prediction), but I was unable to access the webpage; all links to it did not work. I searched PDB using the query sequence and no matches were returned.

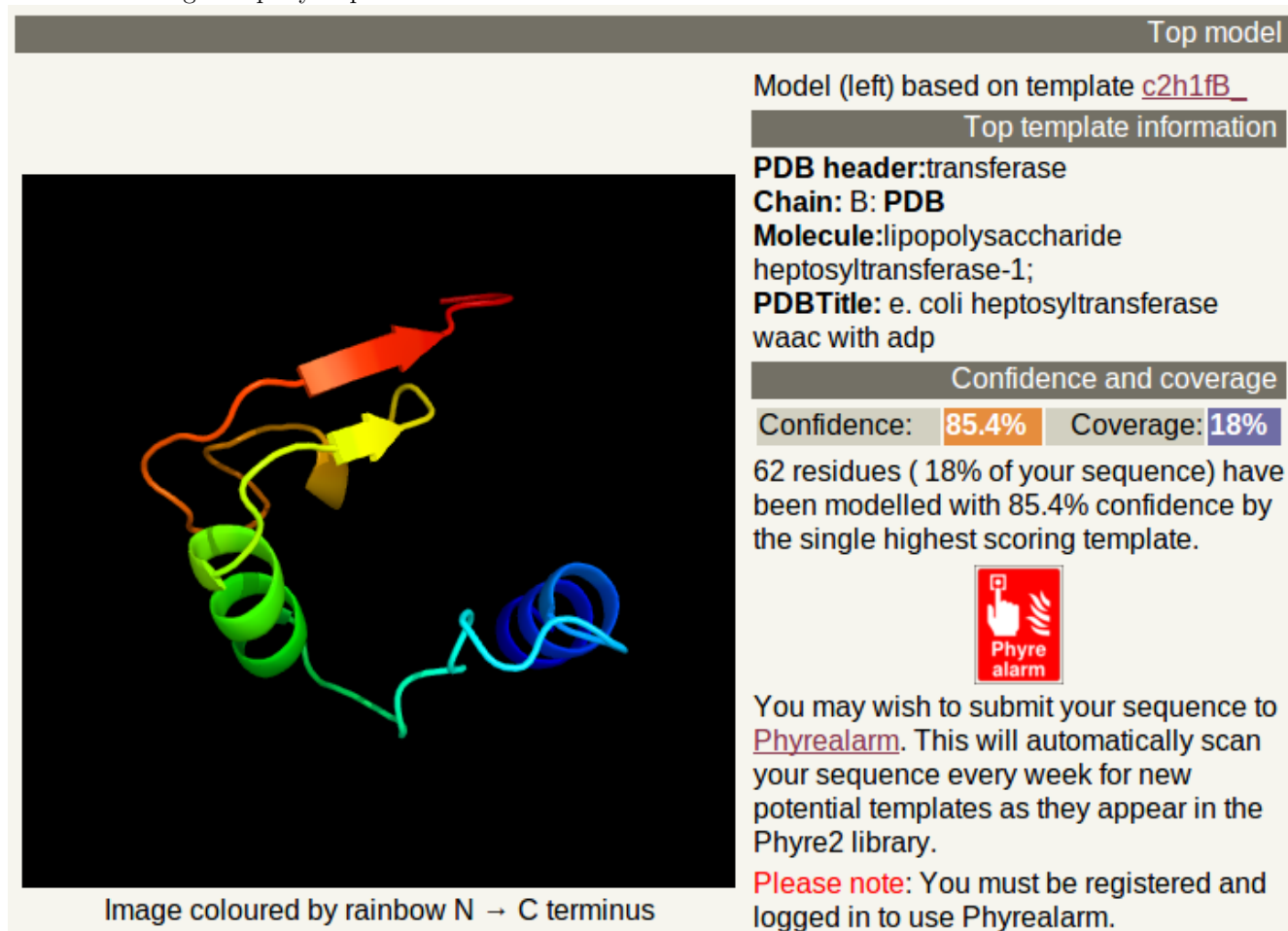


Figure 1: Phyre2 results

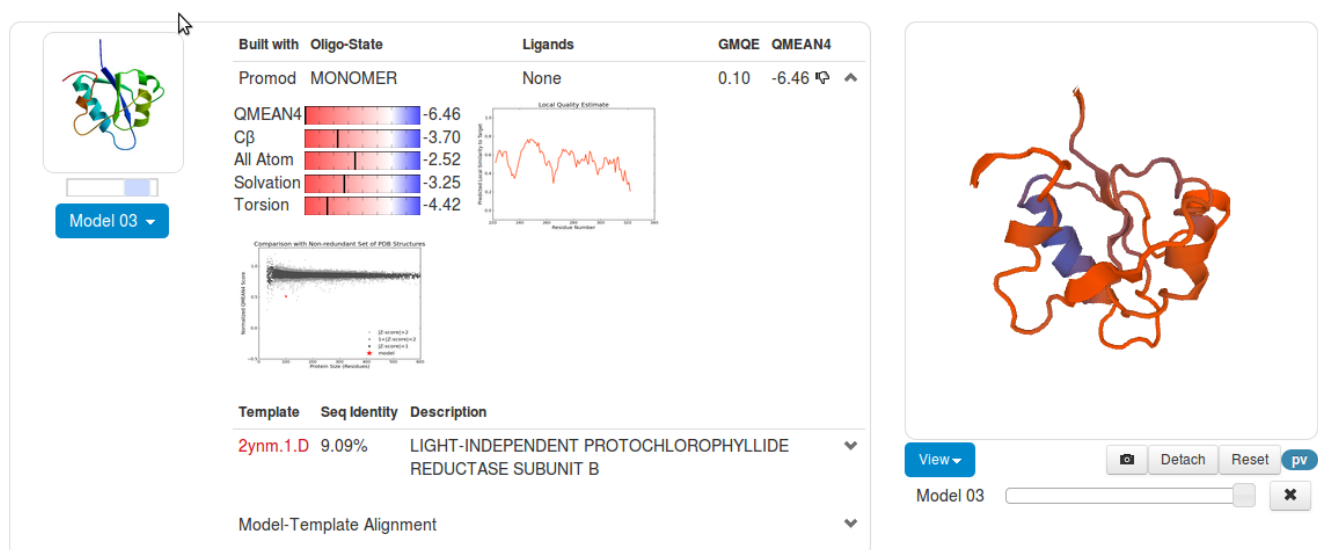


Figure 2: SWISS-MODEL results



Figure 3: Top iTASSER result 3D alignment

Sequence search results

[Show](#) the detailed description of this results page.

We found **1** Pfam-A match to your search sequence (**all** significant). You did not choose to search for Pfam-B matches.



[Show](#) the search options and sequence that you submitted.

[Return](#) to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

[Show](#) or [hide](#) all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
DUF563	Protein of unknown function (DUF563)	Family	n/a	113	329	114	323	2	201	207	117.0	9.7e-34	n/a	Show

Figure 4: PFAM result