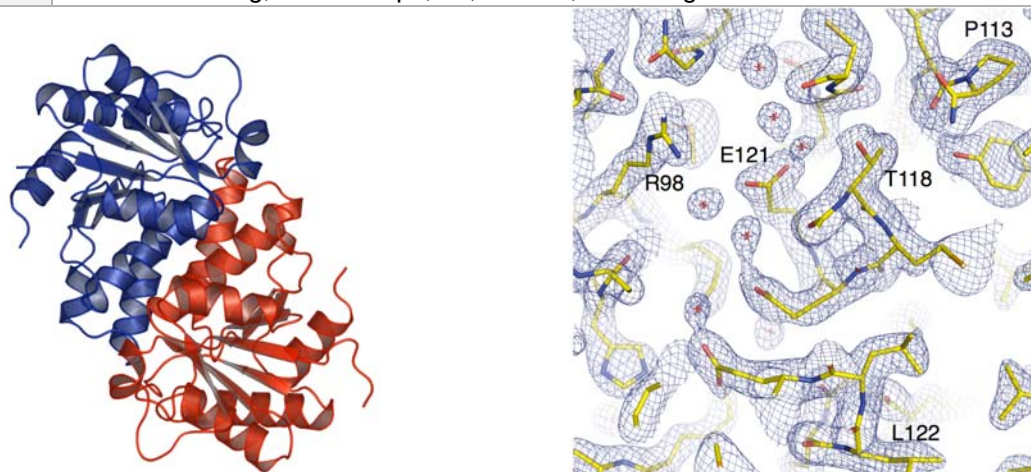


Center for Eukaryotic Structural Genomics

Protein Structure Data Summary

Target ID	GO.19912	
Source Organism	<i>Arabidopsis thaliana</i>	
Target Name	At5g11950.1	
PDB Entry	1YDH	Deposition: 23-Dec-2004
Function	Lysine decarboxylase-like protein (FF/Refine: 2Q4D)	
Produced From	<i>E. coli</i> B834 (DE3) p(LACI+RARE) pVP-13	
Structure by X-ray	Resolution: 2.15 Å	R-value (R-free): 16.2% (21.3%)
	No. of Residues/ASU: 365	Monomers/ASU: 2
Data Collected At	Advanced Photon Source 19-BM 15-Dec-2004	
Authors	G.E. Wesenberg, G.N. Phillips, Jr., E. Bitto, C.A. Bingman	



Structural Features

The gene At5g11950.1 from *Arabidopsis thaliana* encodes a 24 kDa protein of unknown function. Based on a sequence similarity, At5g11950.1 was annotated as lysine decarboxylase (LDC)-like protein. The crystal structure of At5g11950.1 revealed an α/β protein fold with the central seven-stranded β -sheet surrounded by eight α -helices (1). Two At5g11950.1 subunits associate to form a tight dimer in the crystalline asymmetric unit. The interface between the two monomers is mostly hydrophobic and stabilized by contact of helices $\alpha 5$ and $\alpha 6$. The closest structural homologs of At5g11950.1 identified by the VAST search are putative LDC-like proteins from *Bacillus subtilis*, *Thermotoga maritima* and *Thermus thermophilus* (PDB ID codes 1T35, 1RCU, 1WEK, and 1WEH). Despite the low sequence homology (20%), all these homologs share a consensus motif PGGxGTxxE that forms a part of a putative active site cleft. The residues Arg98, Thr118, and Glu121 located at the bottom of the cleft are highly conserved among LDC-like proteins and are likely to play important role in catalysis.

The tertiary structure of At5g11950.1 is almost identical to that of At2g37210.1, another CESC protein target deposited in PDB under the code 2A33. The two structures superimpose with 0.9Å root mean square deviation over 167 aligned residues. The only difference between the two structures is an absence of the short $\alpha 3$ helix in the At2g37210.1 structure that may control the access of the substrate to the active site. The two homologs share 71% sequence identity and are likely to have a similar function in *A. thaliana*.

References: (1) Jeon, W.B., Allard, S.T.M., Bingman, C.A., Bitto, E., Han, B.W., Wesenberg, G.E. (2006) X-ray crystal structures of the conserved hypothetical proteins from *Arabidopsis thaliana* gene loci At5g11950 and At2g37210. *Proteins* 65(4):1051-4.

Percent Identity with Nearest PDB Structure at Time Solved	40% (1T35)
Pfam Cluster	Lysine_decarboxylase
Sequence Cluster Size	484

Center for Eukaryotic Structural Genomics (CESG), University of Wisconsin-Madison Biochemistry Department, 433 Babcock Drive, Madison, WI 53706-1549; phone: 608.263.2183; fax: 608.890.1942; email: cesginfo@biochem.wisc.edu; website: <http://www.uwstructuralgenomics.org>. This research funded by NIH / NIGMS Protein Structure Initiative grants U54 GM074901 and P50 GM064598.