

# CS612 – Algorithms in Bioinformatics

## Homework Assignment 2 – 03/03/2023

Aravind Haridas – 02071139

### 1(a): Multiple Sequence Alignment Using PSI-BLAST

No, we did not get the results. Below is the image.

BLAST® » blastp suite » results for RID-ZHYFK13A013

Job Title: QUERY1  
RID: ZHYFK13A013  
Program: Cite  
Database: pdb  
Query ID: IcdQuery\_79625  
Description: QUERY1  
Molecule type: amino acid  
Query Length: 457  
Other reports: [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

Filter Results  
Percent Identity:  to   
E value:  to   
Query Coverage:  to   
[Filter](#) [Reset](#)

No significant similarity found. For reasons why, [click here](#)

### b. 219 sequence found.

BLAST® » blastp suite » results for RID-ZX215NVM013

Job Title: QUERY1  
RID: ZX215NVM013  
Program: PSI-BLAST Iteration 1  
Database: nr  
Query ID: IcdQuery\_21986  
Description: QUERY1  
Molecule type: amino acid  
Query Length: 457  
Other reports: [Distance tree of results](#) [Multiple alignment](#) [MSA viewer](#)

Filter Results  
Organism:  ☐ exclude  
Type common name, binomial, taxid or group name  
[Add organism](#)  
Percent Identity:  to   
E value:  to   
Query Coverage:  to   
PSI-BLAST incl. threshold: 0.005  
[Filter](#) [Reset](#)

Run PSI-Blast iteration 2  
Number of sequences: 1000 [Run](#)

Descriptions | Graphic Summary | Alignments | Taxonomy

Sequences producing significant alignments

219 sequences selected

Sequences with E-value BETTER than threshold

select all 219 sequences selected

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession	Select for PSI blast	Used to build PSSM	Newly added
ORF457 [Sulfolobus islandicus]	Sulfolobus isl...	902	902	100%	0.0	100.00%	457	AAK06928.1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

number of sequences 1000 Run

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments Download Select columns Show 1000

219 sequences selected [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

Sequences with E-value BETTER than threshold

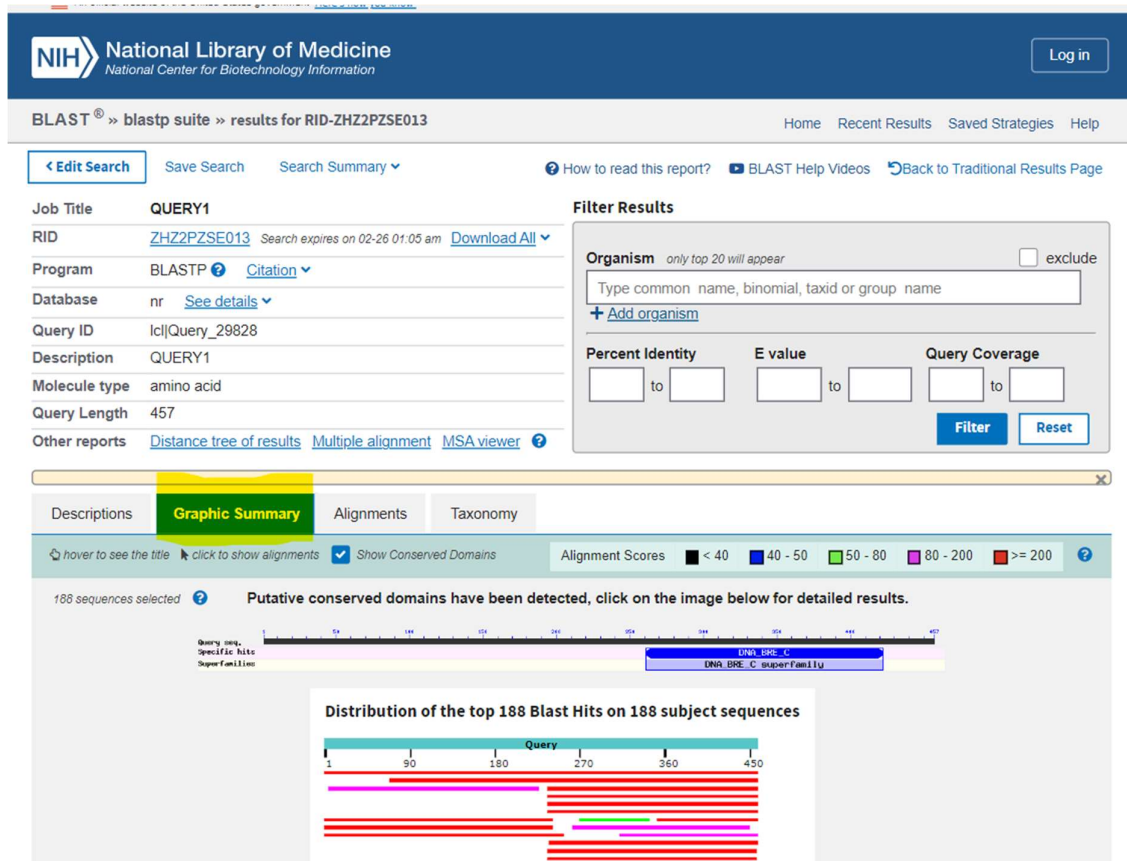
☒ select all 219 sequences selected

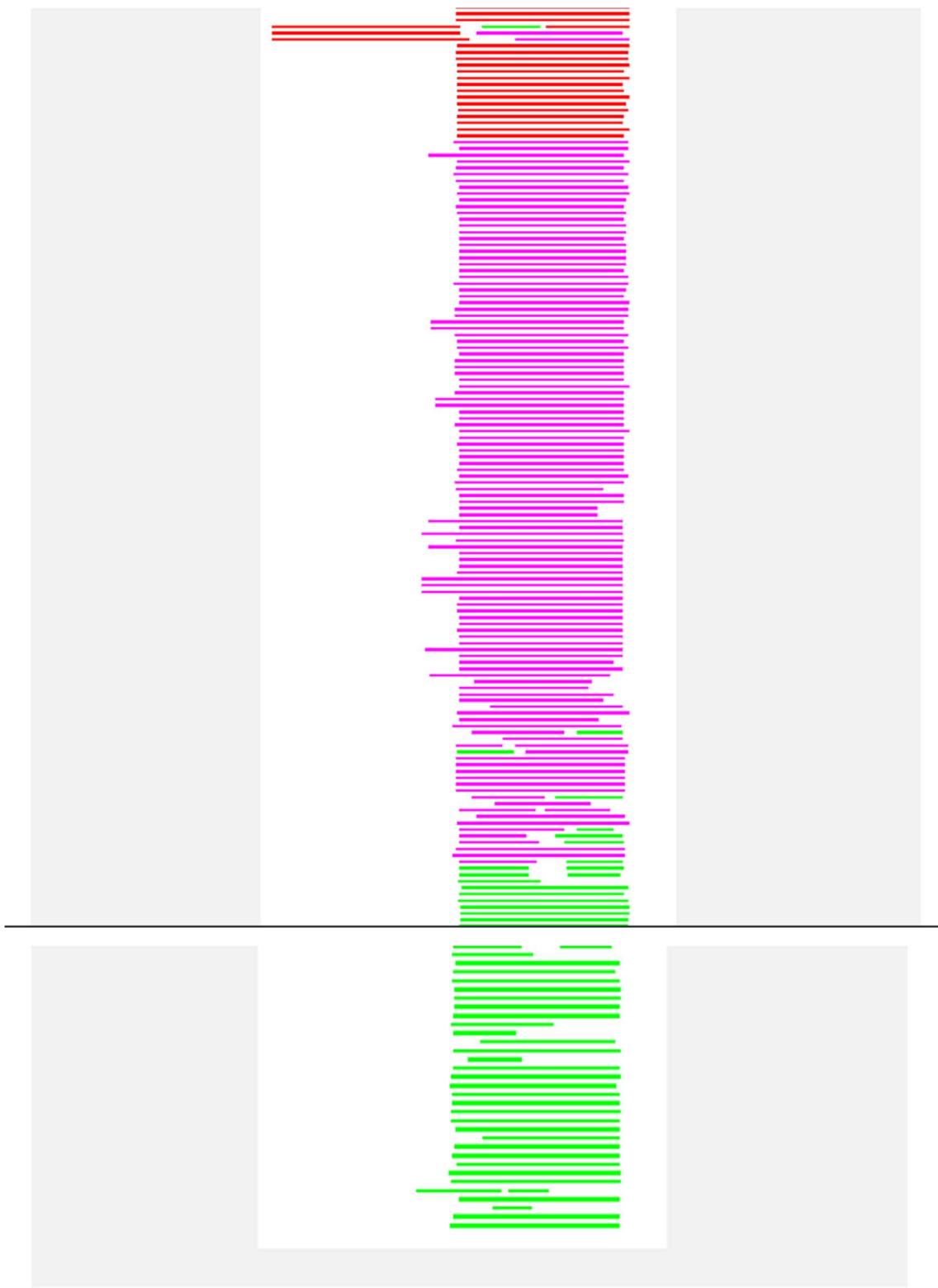
PSI-BLAST iteration 1

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession	Select for PSI blast	Used to build PSSM	Newly added
<input checked="" type="checkbox"/> ORF457 [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	902	902	100%	0.0	100.00%	457	<a href="#">AAK06928.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> site-specific integrase [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	773	773	84%	0.0	100.00%	388	<a href="#">WIP_252183452.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> tyrosine-type recombinase/integrase [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	412	412	48%	2e-137	88.74%	424	<a href="#">WIP_016730916.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> tyrosine-type recombinase/integrase [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	400	400	48%	1e-132	85.59%	424	<a href="#">WIP_012714036.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> site-specific integrase [Sulfuracidifex tepidarius]	<a href="#">Sulfuracidifex...</a>	399	399	48%	1e-132	84.23%	424	<a href="#">WIP_149564870.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> site-specific integrase [Saccharolobus shibatae]	<a href="#">Saccharolobu...</a>	393	393	48%	4e-130	84.23%	427	<a href="#">WIP_218260350.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> hypothetical protein [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	375	375	52%	2e-122	85.95%	467	<a href="#">WIP_012953077.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> hypothetical protein [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	370	370	52%	3e-120	84.30%	467	<a href="#">WIP_012714086.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> hypothetical protein [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	350	350	55%	2e-112	83.00%	475	<a href="#">WIP_012716467.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> putative transposase [Sulfurisphaera tokodaii str. 7]	<a href="#">Sulfurisphaer...</a>	221	221	48%	2e-65	51.13%	256	<a href="#">BAB65221.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> hypothetical protein [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	209	209	23%	7e-63	93.46%	108	<a href="#">WIP_238526210.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> tyrosine-type recombinase/integrase [Thaumarchaeota archaeon]	<a href="#">Thaumarchae...</a>	215	215	48%	1e-61	50.90%	347	<a href="#">MCL7389974.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> tyrosine-type recombinase/integrase [Thermoprotei archaeon]	<a href="#">Thermoprotei...</a>	215	215	48%	8e-61	48.88%	425	<a href="#">MCD6096157.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> tyrosine-type recombinase/integrase [Sulfolobus islandicus]	<a href="#">Sulfolobus isl...</a>	213	213	48%	6e-60	48.87%	427	<a href="#">WIP_012953053.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> tyrosine-type recombinase/integrase [Saccharolobus shibatae]	<a href="#">Saccharolobu...</a>	212	212	46%	1e-59	50.00%	426	<a href="#">WIP_218260876.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> tyrosine-type recombinase/integrase [Sulfolobus acidocaldarius]	<a href="#">Sulfolobus aci...</a>	209	209	48%	2e-58	47.96%	427	<a href="#">WIP_062088665.1</a>	<input checked="" type="checkbox"/>		
<input checked="" type="checkbox"/> integrase [Sulfolobus sp. SCGC AB-777.1.09]	<a href="#">Sulfolobus sp...</a>	207	207	46%	6e-58	49.53%	426	<a href="#">PVL059817.1</a>	<input checked="" type="checkbox"/>		

c) YES, In the summary we see the coverage from different significant distributed across the entire length of query sequence, we can clearly see from the above images we are getting them from the range of 180 to 450 and the fraction will be calculated.

For the fraction of the query sequence the significant hits match to 40% to 50%





d) In the blast sequence we will be getting the below one as

Query 241 YDSFKTPKIYQYAKVPVKLEEIKEVAKNIEHIPSKAYFVLLAESGLRPGELLNVSIENID  
300

<input type="checkbox"/>	site-specific integrase [Halonubrum ezzemoulense]	Halonubrum e...	189	189	48%	2e-53	23.87%	212	WP_094495686.1	✓	✓
<input type="checkbox"/>	site-specific integrase [Halonubrum litoreum]	Halonubrum lit...	190	190	47%	4e-53	26.03%	266	WP_152418664.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Candidatus Methanosuratincola sp.]	Candidatus M...	189	189	49%	4e-53	25.33%	255	MBC7126843.1	✓	✓
<input type="checkbox"/>	integrase/recombinase [Halonubrum litoreum JCM 13561]	Halonubrum lit...	189	189	47%	6e-53	26.03%	267	EMA59396.1	✓	✓
<input type="checkbox"/>	hypothetical protein D6D85_04625 [Candidatus Methanodesulfokores washburnensis]	Candidatus M...	190	190	52%	7e-53	24.37%	297	RSN76240.1	✓	✓
<input type="checkbox"/>	site-specific integrase [Halonubrum coriense]	Halonubrum c...	186	186	47%	8e-53	25.35%	187	WP_006114989.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Caldivirga sp.]	Caldivirga sp.	194	194	57%	1e-52	26.42%	426	NAZ27987.1	✓	✓
<input type="checkbox"/>	integrase [Sulfurisphaera ohwakuensis]	Sulfurisphaer...	193	193	46%	1e-52	34.88%	412	MBB5255172.1	✓	✓
<input type="checkbox"/>	hypothetical protein DRO26_02295 [Candidatus Bathyarchaeota archaeon]	Candidatus B...	190	190	56%	1e-52	25.87%	307	RLI05401.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Candidatus Methanodesulfokores washburnensis]	Candidatus M...	189	189	52%	1e-52	24.37%	280	WP_161969686.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Nitrososphaeria archaeon]	Nitrososphaer...	188	188	56%	2e-52	24.42%	252	NIN51843.1	✓	✓
<input type="checkbox"/>	integrase [Halonotus pteroides]	Halonotus pt...	189	189	57%	5e-52	22.64%	346	WP_120086309.1	✓	✓
<input type="checkbox"/>	hypothetical protein DRO26_00325 [Candidatus Bathyarchaeota archaeon]	Candidatus B...	187	187	57%	8e-52	22.85%	298	RLI06072.1	✓	✓
<input type="checkbox"/>	site-specific integrase [Halonubrum ezzemoulense]	Halonubrum e...	184	184	48%	1e-51	23.87%	212	WP_271963431.1	✓	✓
<input type="checkbox"/>	putative phage integrase [Saccharolobus shibatae B12]	Saccharolobu...	180	180	31%	5e-51	47.26%	147	QXJ27836.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [archaeon]	archaeon	185	185	49%	5e-51	28.32%	275	MCP8313130.1	✓	✓
<input type="checkbox"/>	site-specific integrase [Desulfurococcus archaeon]	Desulfurococ...	188	188	50%	5e-51	38.79%	386	MCD6428790.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Candidatus Bipolaricaulota bacterium]	Candidatus Bi...	186	186	57%	1e-50	23.37%	339	MBS3764919.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Candidatus Bathyarchaeota archaeon]	Candidatus B...	185	185	49%	2e-50	24.34%	340	MCJ7430628.1	✓	✓
<input type="checkbox"/>	site-specific integrase [Candidatus Bathyarchaeota archaeon]	Candidatus B...	183	183	55%	2e-50	24.41%	260	MCD6240993.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Candidatus Methanomethylicales archaeon]	Candidatus M...	182	182	52%	2e-50	26.25%	246	MBC7113232.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Candidatus Bathyarchaeota archaeon]	Candidatus B...	183	183	49%	3e-50	24.78%	286	MCJ7422901.1	✓	✓
<input type="checkbox"/>	hypothetical protein E3J74_03470 [Candidatus Bathyarchaeota archaeon]	Candidatus B...	182	182	47%	4e-50	27.40%	274	TET20319.1	✓	✓
<input type="checkbox"/>	hypothetical protein [Vulcanisaeta souinana]	Vulcanisaeta...	177	177	35%	9e-50	36.42%	164	WP_054843043.1	✓	✓
<input type="checkbox"/>	hypothetical protein AKJ66_04170 [candidate division MSBL1 archaeon SCGC-AAA259 ... candidate divi...	candidate divi...	183	183	57%	1e-49	22.64%	349	KXA92427.1	✓	✓
<input type="checkbox"/>	site-specific integrase [Candidatus Bathyarchaeota archaeon]	Candidatus B...	178	178	49%	1e-49	23.45%	207	MCJ7761658.1	✓	✓
<input type="checkbox"/>	tyrosine-type recombinase/integrase [Candidatus Bathyarchaeota archaeon]	Candidatus B...	181	181	55%	2e-49	23.41%	308	MCK5625814.1	✓	✓
<input type="checkbox"/>	hypothetical protein AKJ62_04995 [candidate division MSBL1 archaeon SCGC-AAA259 ... candidate divi...	candidate divi...	182	182	53%	3e-49	22.95%	334	KXA88517.1	✓	✓

[Download](#) [GenPept](#) [Graphics](#) [newly added](#) [Next](#) [Previous](#) [Descriptions](#)

**site-specific integrase [Halonubrum coriense]**  
Sequence ID: [WP\\_006114989.1](#) Length: 187 Number of Matches: 1  
[See 1 more title\(s\)](#) [See all Identical Proteins \(IPG\)](#)

Range 1: 1 to 187 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
186 bits(473)	8e-53	Composition-based stats.	55/217(25%)	98/217(45%)	30/217(13%)

Query 241 YDSFKTPKIYQYAKVPVKLEEIKEVAKNIEHIPSKAYFVLLAESGLRPGELLNVSIENID 300

D+F+ P EE++ +E+ KA F+ + SGLR EL ++++ID

Sbjct 1 MDAFEMPSKSAKPTDIPSKKEELQTFYSTLENPKYKAALFASTSGLRSSELCQLTMDDD 60

Query 301 LKARIIWINKETQTKRAYFSFFSRKTAEFLEKVVLPAREEFIRANEKNIAKLAAANQE 360

+ R++ KE+ TK+ + +F++ + AE E + +

Sbjct 61 QEKRLVPEKESSTQKTWLTFFYNDEAAEAFAKPEYKPD----- 100

Query 361 IDLEKWKALFPYKDDVLRRIYEAMDRALGKRFFELALRRHFATYMQLKVPPLAINIL 420

+ +F +L RK + + G R + LRR FAT M V I+

Sbjct 101 -----DERVFQTSKQLLNRR-FRHVSEDGVRVTVQKLRRWFATMSRCGVDKAYIDAF 153

Query 421 QGRVGPNEFRILKENYTVFTIEDLRKLYDEAGLVLE 457

GR + +L+++Y ++ E L+++YDEAG+ VLE

Sbjct 154 CGRTPSS---VLEQHYLDYSPERLKEIYDEAGITVLE 187

**Related Information**  
[Identical Proteins](#) - Identical proteins to WP\_006114989.1

e) 4A8E\_A This is the known protein sequence found in the PDB code.

✓	tyrosine-type recombinase/integrase [Chitinophagaceae bacterium]	Chitinophaga...	85.4	85.4	54%	8e-14	17.93%	374	MCB0777386.1	✓	🔍
✓	tyrosine-type recombinase/integrase [Leptospira yanagawae]	Leptospira ya...	85.4	85.4	44%	8e-14	17.79%	368	WP_135633918.1	✓	🔍
✓	hypothetical protein CV691_09410 [Candidatus Bathyarchaeota archaeon]	Candidatus B...	85.4	85.4	52%	8e-14	22.92%	383	PVX23891.1	✓	🔍
✓	integrase [Candidatus Poribacteria bacterium]	Candidatus P...	82.7	82.7	46%	8e-14	19.16%	233	MBM3212642.1	✓	🔍
✓	tyrosine-type recombinase/integrase [Scorosaibacterium faourense]	Scorosaibact...	85.0	85.0	57%	9e-14	16.85%	353	WP_243414876.1	✓	🔍
✓	tyrosine recombinase XerC [Candidatus Margulisbacteria bacterium]	Candidatus M...	84.2	84.2	46%	9e-14	18.72%	298	MBU0687569.1	✓	🔍
✓	TPA recombinase XerC [Thermofilum sp.]	Thermofilum sp.	83.8	83.8	45%	9e-14	21.84%	285	HDI32008.1	✓	🔍
✓	TPA integrase [Planctomycetaceae bacterium]	Planctomycet...	83.8	83.8	55%	9e-14	16.47%	286	HBO44539.1	✓	🔍
✓	integrase [Firmicutes bacterium]	Firmicutes ba...	83.8	83.8	57%	9e-14	16.48%	273	MTI47672.1	✓	🔍
✓	tyrosine-type recombinase/integrase [bacterium]	bacterium	83.8	83.8	47%	9e-14	18.89%	284	MBL7994304.1	✓	🔍
✓	The structure of a dimeric Xer recombinase from archaea [Pyrococcus abyssi]	Pyrococcus a...	83.8	83.8	46%	9e-14	20.66%	292	4A8E_A	✓	🔍
✓	tyrosine-type recombinase/integrase [Metallosphaera cuprina]	Metallosphae...	83.8	83.8	44%	9e-14	22.06%	284	WP_013737213.1	✓	🔍
✓	tyrosine-type recombinase/integrase [Desulfurobacterium atlanticum]	Desulfurobact...	85.0	85.0	57%	9e-14	20.30%	366	WP_089323339.1	✓	🔍
✓	site-specific integrase [Flavobacterium sp. 83]	Flavobacteriu...	84.6	84.6	62%	9e-14	21.23%	347	WP_035670805.1	✓	🔍
✓	tyrosine recombinase [Spirochaeta sp.]	Spirochaeta sp.	84.6	84.6	46%	9e-14	19.35%	344	MQY76625.1	✓	🔍
✓	tyrosine-type recombinase/integrase [Polaribacter atindae]	Polaribacter a...	85.0	85.0	45%	9e-14	21.26%	374	WP_068449836.1	✓	🔍
✓	tyrosine recombinase XerC [Acidobacteriota bacterium]	Acidobacterio...	84.2	84.2	45%	9e-14	21.86%	319	PYS42645.1	✓	🔍
✓	tyrosine-type recombinase/integrase [Cytobacillus oceanisediminis]	Cytobacillus o...	83.8	83.8	51%	1e-13	19.58%	278	QOK29793.1	✓	🔍
✓	tyrosine-type recombinase/integrase [Candidatus Woesearchaeota archaeon]	Candidatus W...	83.5	83.5	44%	1e-13	21.84%	269	MBL7050751.1	✓	🔍
✓	tyrosine-type recombinase/integrase [Formosa sp. HeII_31_208]	Formosa sp....	85.0	85.0	56%	1e-13	21.35%	375	WP_091424817.1	✓	🔍
✓	site-specific tyrosine recombinase XerD [Thermosyntropha lipolytica]	Thermosyntro...	83.8	83.8	54%	1e-13	19.53%	296	WP_073093510.1	✓	🔍
✓	putative tyrosine recombinase XerC-like protein [Candidatus Methanofastidiosum methyl...]	Candidatus M...	83.8	83.8	45%	1e-13	23.33%	284	KYC53593.1	✓	🔍
✓	site-specific recombinase XerD [Natranaerovirga hydrolytica]	Natranaerovir...	85.0	85.0	50%	1e-13	19.48%	356	TCK92654.1	✓	🔍
✓	tyrosine-type recombinase/integrase [Methanotrichaceae archaeon]	Methanotrich...	83.8	83.8	55%	1e-13	18.25%	296	MCJ7444010.1	✓	🔍
✓	site-specific integrase [Cytophagales bacterium]	Cytophagales ...	83.5	83.5	55%	1e-13	18.36%	268	MCG8322767.1	✓	🔍

GenPept ▾

## Chain A, Probable Tyrosine Recombinase Xerc-like

PDB: 4A8E\_A

[Identical Proteins](#) [FASTA](#) [Graphics](#)

Go to: 📄

LOCUS **4A8E\_A** 292 aa linear BCT 01-DEC-2020

DEFINITION Chain A, Probable Tyrosine Recombinase Xerc-like.

ACCESSION 4A8E\_A

VERSION 4A8E\_A

DBSOURCE pdb: molecule 4A8E, chain A, release Jan 15, 2014;  
deposition: Nov 21, 2011;  
class: Cell Cycle;  
source: Mmdb\_id: [105434](#), Pdb\_id 1: 4A8E;  
Exp. method: X-Ray Diffraction.

KEYWORDS .

SOURCE Pyrococcus abyssi

ORGANISM [Pyrococcus abyssi](#)  
Archaea; Euryarchaeota; Thermococci; Thermococcales;  
Thermococcaceae; Pyrococcus.

REFERENCE 1 (residues 1 to 292)

AUTHORS Serre,M.C., El Arnaout,T., Brooks,M.A., Durand,D., Lisboa,J., Lazar,N., Raynal,B., van Tilbeurgh,H. and Quevillon-Cheruel,S.

TITLE The carboxy-terminal alphaN helix of the archaeal XerA tyrosine recombinase is a molecular switch to control site-specific recombination

JOURNAL PLoS ONE 8 (5), E63010 (2013)

PUBMED [23667562](#)

REMARK Publication Status: Online-Only

REFERENCE 2 (residues 1 to 292)

AUTHORS Brooks,M.A., Elarnaout,T., Duranda,D., Lisboa,J., Lazar,N., Raynal,B., Vantilbeurgh,H., Serre,M. and Quevillon-Cheruel,S.

TITLE Direct Submission

JOURNAL Submitted (21-NOV-2011)

COMMENT The structure of a dimeric Xer recombinase from archaea.

FEATURES Location/Qualifiers



```

        /note="helix 7"
Region      join(145..188,242..292)
            /region_name="Domain 3"
            /note="NCBI Domains"
SecStr      146..150
            /sec_str_type="sheet"
            /note="strand 1"
SecStr      151..157
            /sec_str_type="sheet"
            /note="strand 2"
SecStr      164..169
            /sec_str_type="sheet"
            /note="strand 3"
SecStr      171..183
            /sec_str_type="helix"
            /note="helix 8"
SecStr      193..196
            /sec_str_type="sheet"
            /note="strand 4"
SecStr      199..202
            /sec_str_type="sheet"
            /note="strand 5"
SecStr      205..219
            /sec_str_type="helix"
            /note="helix 9"
SecStr      225..239
            /sec_str_type="helix"
            /note="helix 10"
SecStr      243..250
            /sec_str_type="helix"
            /note="helix 11"
SecStr      262..272
            /sec_str_type="helix"
            /note="helix 12"

ORIGIN
    1 meereervrd dtieefatyl elegksrntv rmytyyiskf feeghsptar dalrflaklk
    61 rkgystsrsln lviqalkayf kfegldseae klktpkmpkt lpkslteeev rriinaaetl
    121 rdlrlilllly gaglrvselc nlrvedvnfe ygvivvrggk ggkdrvvpis esllseikry
    181 lesrnddsy 1fvmekrrk dklspktvwr lvkkygrkag veltpqlrh sfathmlerg
    241 idiriiqell ghsnlsttqi ytkvstkhk eavkkaklve siiggshhhh hh

//

name:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,208,209,210,211,212,213,214,215,216,217,218,219,220,221,222,223,224,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,243,244,245,246,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,287,288,289,290,291,292,293,294,295,296,297,298,299,300,301,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,338,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,371,372,373,374,375,376,377,378,379,380,381,382,383,384,385,386,387,388,389,390,391,392,393,394,395,396,397,398,399,400,401,402,403,404,405,406,407,408,409,410,411,412,413,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602,603,604,605,606,607,608,609,610,611,612,613,614,615,616,617,618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,637,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,656,657,658,659,660,661,662,663,664,665,666,667,668,669,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,738,739,740,741,742,743,744,745,746,747,748,749,750,751,752,753,754,755,756,757,758,759,760,761,762,763,764,765,766,767,768,769,770,771,772,773,774,775,776,777,778,779,780,781,782,783,784,785,786,787,788,789,790,791,792,793,794,795,796,797,798,799,800,801,802,803,804,805,806,807,808,809,810,811,812,813,814,815,816,817,818,819,820,821,822,823,824,825,826,827,828,829,830,831,832,833,834,835,836,837,838,839,840,841,842,843,844,845,846,847,848,849,850,851,852,853,854,855,856,857,858,859,860,861,862,863,864,865,866,867,868,869,870,871,872,873,874,875,876,877,878,879,880,881,882,883,884,885,886,887,888,889,890,891,892,893,894,895,896,897,898,899,900,901,902,903,904,905,906,907,908,909,910,911,912,913,914,915,916,917,918,919,920,921,922,923,924,925,926,927,928,929,930,931,932,933,934,935,936,937,938,939,940,941,942,943,944,945,946,947,948,949,950,951,952,953,954,955,956,957,958,959,960,961,962,963,964,965,966,967,968,969,970,971,972,973,974,975,976,977,978,979,980,981,982,983,984,985,986,987,988,989,990,991,992,993,994,995,996,997,998,999,1000,1001,1002,1003,1004,1005,1006,1007,1008,1009,1010,1011,1012,1013,1014,1015,1016,1017,1018,1019,1020,1021,1022,1023,1024,1025,1026,1027,1028,1029,1030,1031,1032,1033,1034,1035,1036,1037,1038,1039,1040,1041,1042,1043,1044,1045,1046,1047,1048,1049,1050,1051,1052,1053,1054,1055,1056,1057,1058,1059,1060,1061,1062,1063,1064,1065,1066,1067,1068,1069,1070,1071,1072,1073,1074,1075,1076,1077,1078,1079,1080,1081,1082,1083,1084,1085,1086,1087,1088,1089,1090,1091,1092,1093,1094,1095,1096,1097,1098,1099,1100,1101,1102,1103,1104,1105,1106,1107,1108,1109,1110,1111,1112,1113,1114,1115,1116,1117,1118,1119,1120,1121,1122,1123,1124,1125,1126,1127,1128,1129,1130,1131,1132,1133,1134,1135,1136,1137,1138,1139,1140,1141,1142,1143,1144,1145,1146,1147,1148,1149,1150,1151,1152,1153,1154,1155,1156,1157,1158,1159,1160,1161,1162,1163,1164,1165,1166,1167,1168,1169,1170,1171,1172,1173,1174,1175,1176,1177,1178,1179,1180,1181,1182,1183,1184,1185,1186,1187,1188,1189,1190,1191,1192,1193,1194,1195,1196,1197,1198,1199,1200,1201,1202,1203,1204,1205,1206,1207,1208,1209,1210,1211,1212,1213,1214,1215,1216,1217,1218,1219,1220,1221,1222,1223,1224,1225,1226,1227,1228,1229,1230,1231,1232,1233,1234,1235,1236,1237,1238,1239,1240,1241,1242,1243,1244,1245,1246,1247,1248,1249,1250,1251,1252,1253,1254,1255,1256,1257,1258,1259,1260,1261,1262,1263,1264,1265,1266,1267,1268,1269,1270,1271,1272,1273,1274,1275,1276,1277,1278,1279,1280,1281,1282,1283,1284,1285,1286,1287,1288,1289,1290,1291,1292,1293,1294,1295,1296,1297,1298,1299,1300,1301,1302,1303,1304,1305,1306,1307,1308,1309,1310,1311,1312,1313,1314,1315,1316,1317,1318,1319,1320,1321,1322,1323,1324,1325,1326,1327,1328,1329,1330,1331,1332,1333,1334,1335,1336,1337,1338,1339,1340,1341,1342,1343,1344,1345,1346,1347,1348,1349,1350,1351,1352,1353,1354,1355,1356,1357,1358,1359,1360,1361,1362,1363,1364,1365,1366,1367,1368,1369,1370,1371,1372,1373,1374,1375,1376,1377,1378,1379,1380,1381,1382,1383,1384,1385,1386,1387,1388,1389,1390,1391,1392,1393,1394,1395,1396,1397,1398,1399,1400,1401,1402,1403,1404,1405,1406,1407,1408,1409,1410,1411,1412,1413,1414,1415,1416,1417,1418,1419,1420,1421,1422,1423,1424,1425,1426,1427,1428,1429,1430,1431,1432,1433,1434,1435,1436,1437,1438,1439,1440,1441,1442,1443,1444,1445,1446,1447,1448,1449,1450,1451,1452,1453,1454,1455,1456,1457,1458,1459,1460,1461,1462,1463,1464,1465,1466,1467,1468,1469,1470,1471,1472,1473,1474,1475,1476,1477,1478,1479,1480,1481,1482,1483,1484,1485,1486,1487,1488,1489,1490,1491,1492,1493,1494,1495,1496,1497,1498,1499,1500,1501,1502,1503,1504,1505,1506,1507,1508,1509,1510,1511,1512,1513,1514,1515,1516,1517,1518,1519,1520,1521,1522,1523,1524,1525,1526,1527,1528,1529,1530,1531,1532,1533,1534,1535,1536,1537,1538,1539,1540,1541,1542,1543,1544,1545,1546,1547,1548,1549,1550,1551,1552,1553,1554,1555,1556,1557,1558,1559,1560,1561,1562,1563,1564,1565,1566,1567,1568,1569,1570,1571,1572,1573,1574,1575,1576,1577,1578,1579,1580,1581,1582,1583,1584,1585,1586,1587,1588,1589,1590,1591,1592,1593,1594,1595,1596,1597,1598,1599,1600,1601,1602,1603,1604,1605,1606,1607,1608,1609,1610,1611,1612,1613,1614,1615,1616,1617,1618,1619,1620,1621,1622,1623,1624,1625,1626,1627,1628,1629,1630,1631,1632,1633,1634,1635,1636,1637,1638,1639,1640,1641,1642,1643,1644,1645,1646,1647,1648,1649,1650,1651,1652,1653,1654,1655,1656,1657,1658,1659,1660,1661,1662,1663,1664,1665,1666,1667,1668,1669,1670,1671,1672,1673,1674,1675,1676,1677,1678,1679,1680,1681,1682,1683,1684,1685,1686,1687,1688,1689,1690,1691,1692,1693,1694,1695,1696,1697,1698,1699,1700,1701,1702,1703,1704,1705,1706,1707,1708,1709,1710,1711,1712,1713,1714,1715,1716,1717,1718,1719,1720,1721,1722,1723,1724,1725,1726,1727,1728,1729,1730,1731,1732,1733,1734,1735,1736,1737,1738,1739,1740,1741,1742,1743,1744,1745,1746,1747,1748,1749,1750,1751,1752,1753,1754,1755,1756,1757,1758,1759,1760,1761,1762,1763,1764,1765,1766,1767,1768,1769,1770,1771,1772,1773,1774,1775,1776,1777,1778,1779,1780,1781,1782,1783,1784,1785,1786,1787,1788,1789,1790,1791,1792,1793,1794,1795,1796,1797,1798,1799,1800,1801,1802,1803,1804,1805,1806,1807,1808,1809,1810,1811,1812,1813,1814,1815,1816,1817,1818,1819,1820,1821,1822,1823,1824,1825,1826,1827,1828,1829,1830,1831,1832,1833,1834,1835,1836,1837,1838,1839,1840,1841,1842,1843,1844,1845,1846,1847,1848,1849,1850,1851,1852,1853,1854,1855,1856,1857,1858,1859,1860,1861,1862,1863,1864,1865,1866,1867,1868,1869,1870,1871,1872,1873,1874,1875,1876,1877,1878,1879,1880,1881,1882,1883,1884,1885,1886,1887,1888,1889,1890,1891,1892,1893,1894,1895,1896,1897,1898,1899,1900,1901,1902,1903,1904,1905,1906,1907,1908,1909,1910,1911,1912,1913,1914,1915,1916,1917,1918,1919,1920,1921,1922,1923,1924,1925,1926,1927,1928,1929,1930,1931,1932,1933,1934,1935,1936,1937,1938,1939,1940,1941,1942,1943,1944,1945,1946,1947,1948,1949,1950,1951,1952,1953,1954,1955,1956,1957,1958,1959,1960,1961,1962,1963,1964,1965,1966,1967,1968,1969,1970,1971,1972,1973,1974,1975,1976,1977,1978,1979,1980,1981,1982,1983,1984,1985,1986,1987,1988,1989,1990,1991,1992,1993,1994,1995,1996,1997,1998,1999,2000,2001,2002,2003,2004,2005,2006,2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021,2022,2023,2024,2025,2026,2027,2028,2029,2030,2031,2032,2033,2034,2035,2036,2037,2038,2039,2040,2041,2042,2043,2044,2045,2046,2047,2048,2049,2050,2051,2052,2053,2054,2055,2056,2057,2058,2059,2060,2061,2062,2063,2064,2065,2066,2067,2068,2069,2070,2071,2072,2073,2074,2075,2076,2077,2078,2079,2080,2081,2082,2083,2084,2085,2086,2087,2088,2089,2090,2091,2092,2093,2094,2095,2096,2097,2098,2099,2100,2101,2102,2103,2104,2105,2106,2107,2108,2109,2110,2111,2112,2113,2114,2115,2116,2117,2118,2119,2120,2121,2122,2123,2124,2125,2126,2127,2128,2129,2130,2131,2132,2133,2134,2135,2136,2137,2138,2139,2140,2141,2142,2143,2144,2145,2146,2147,2148,2149,2150,2151,2152,2153,2154,2155,2156,2157,2158,2159,2160,2161,2162,2163,2164,2165,2166,2167,2168,2169,2170,2171,2172,2173,2174,2175,2176,2177,2178,2179,2180,2181,2182,2183,2184,2185,2186,2187,2188,2189,2190,2191,2192,2193,2194,2195,2196,2197,2198,2199,2200,2201,2202,2203,2204,2205,2206,2207,2208,2209,2210,2211,2212,2213,2214,2215,2216,2217,2218,2219,2220,2221,2222,2223,2224,2225,2226,2227,2228,2229,2230,2231,2232,2233,2234,2235,2236,2237,2238,2239,2240,2241,2242,2243,2244,2245,2246,2247,2248,2249,2250,2251,2252,2253,2254,2255,2256,2257,2258,2259,2260,2261,2262,2263,2264,2265,2266,2267,2268,2269,2270,2271,2272,2273,2274,2275,2276,2277,2278,2279,2280,2281,2282,2283,2284,2285,2286,2287,2288,2289,2290,2291,2292,2293,2294,2295,2296,2297,2298,2299,2300,2301,2302,2303,2304,2305,2306,2307,2308,2309,2310,2311,2312,2313,2314,2315,2316,2317,2318,2319,2320,2321,2322,2323,2324,2325,2326,2327,2328,2329,2330,2331,2332,2333,2334,2335,2336,2337,2338,2339,2340,2341,2342,2343,2344,2345,2346,2347,2348,2349,2350,2351,2352,2353,2354,2355,2356,2357,2358,2359,2360,2361,2362,2363,2364,2365,2366,2367,2368,2369,2370,2371,2372,2373,2374,2375,2376,2377,2378,2379,2380,2381,2382,2383,2384,2385,2386,2387,2388,2389,2390,2391,2392,2393,2394,2395,2396,2397,2398,2399,2400,2401,2402,2403,2404,2405,2406,2407,2408,2409,2410,2411,2412,2413,2414,2415,2416,2417,2418,2419,2420,2421,2422,2423,2424,2425,2426,2427,2428,2429,2430,2431,2432,2433,2434,2435,2436,2437,2438,2439,2440,2441,2442,2443,2444,2445,2446,2447,2448,2449,2450,2451,2452,2453,2454,2455,2456,2457,2458,2459,2460,2461,2462,2463,2464,2465,2466,2467,2468,2469,2470,2471,2472,2473,2474,2475,2476,2477,2478,2479,2480,2481,2482,2483,2484,2485,2486,2487,2488,2489,2490,2491,2492,2493,2494,2495,2496,2497,2498,2499,2500,2501,2502,2503,2504,2505,2506,2507,2508,2509,2510,2511,2512,2513,2514,2515,2516,2517,2518,2519,2520,2521,2522,2523,2524,2525,2526,2527,2528,2529,2530,2531,2532,2533,2534,2535,2536,2537,2538,2539,2540,2541,2542,2543,2544,2545,2546,2547,2548,2549,2550,2551,2552,2553,2554,2555,2556,2557,2558,2559,2560,2561,2562,2563,2564,2565,2566,2567,2568,2569,2570,2571,2572,2573,2574,2575,2576,2577,2578,2579,2580,2581,2582
```

f) The protein is 4A8E: The Structure of A Dimeric Xer Recombinase From Archaea

DEFINITION Chain A, Probable Tyrosine Recombinase Xerc-like.

ACCESSION 4A8E\_A

VERSION 4A8E\_A

DBSOURCE pdb: molecule 4A8E, chain A, release Jan 15, 2014;

deposition: Nov 21, 2011.

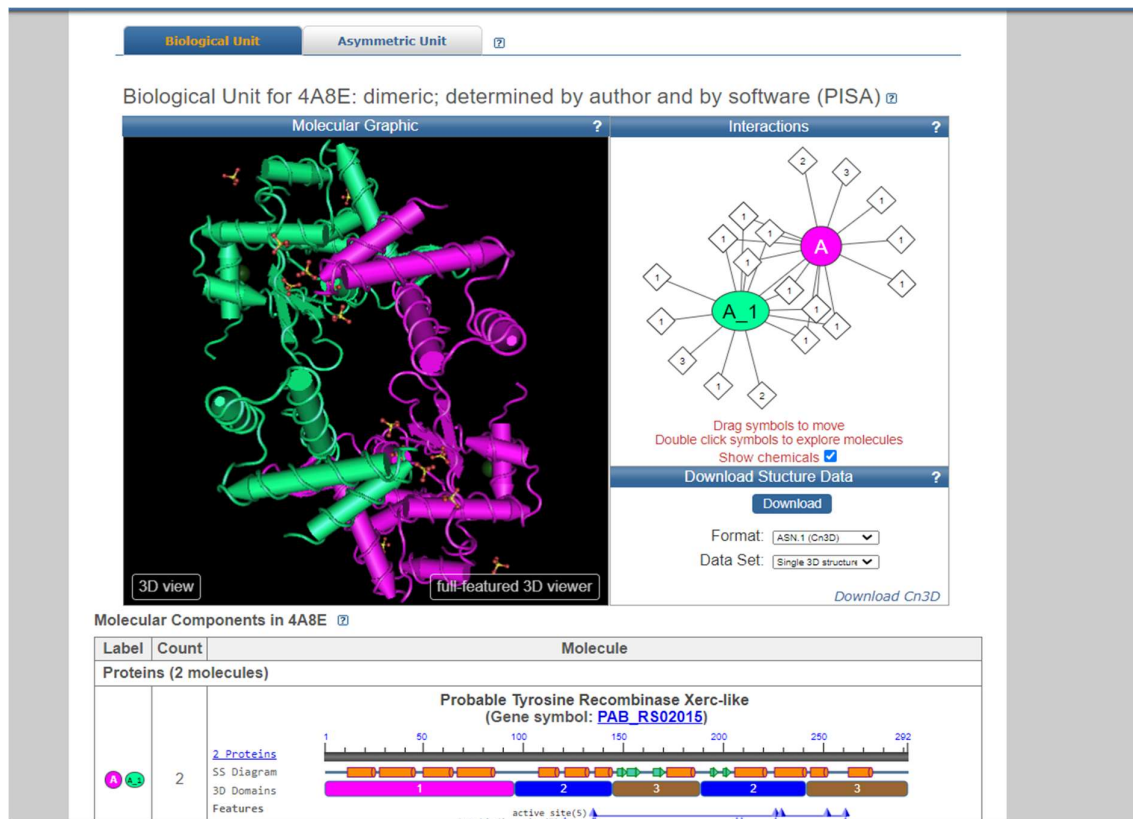
class: Cell Cycle;

SOURCE Pyrococcus abyssi

ORGANISM Pyrococcus abyssi

Archaea; Euryarchaeota; Thermococci; Thermococcales;

Thermococcaceae; Pyrococcus.



3D view

full-featured 3D viewer

Data Set: Single 3D structure

[Download Cn3D](#)

**Molecular Components in 4A8E** [?](#)

Label	Count	Molecule
<b>Proteins (2 molecules)</b>		
	2	<div style="text-align: center;"> <b>Probable Tyrosine Recombinase Xerc-like</b>            (Gene symbol: <a href="#">PAB_RS02015</a>)         </div> <div style="margin-top: 10px;"> </div> <div style="margin-top: 10px;"> <a href="#">2 Proteins</a>  <a href="#">SS Diagram</a>  <a href="#">3D Domains</a>  <a href="#">Features</a>  <a href="#">Domain Families</a>  <a href="#">Super Families</a> </div>
<b>Chemicals and Non-standard biopolymers (18 molecules)</b>		
	14	Sulfate Ion <div style="float: right; text-align: center;"> </div>
	2	Chloride Ion <div style="float: right; text-align: center;"> </div>
	2	1,2-Ethanediol <div style="float: right; text-align: center;"> </div>

\* Click molecule labels to explore molecular sequence information.

**Citing MMDB**  
 Madej T, Lanczycki CJ, Zhang D, Thiessen PA, Geer RC, Marchler-Bauer A, Bryant SH. " *MMDB and VAST+: tracking structural similarities between macromolecular complexes*. *Nucleic Acids Res.* **2014 Jan; 42(Database issue):D297-303**

2.

a) Here is the logo sequence

From the below we need to go to the blast2logo server at <https://services.healthtech.dtu.dk/services/Blast2logo-1.1/>.

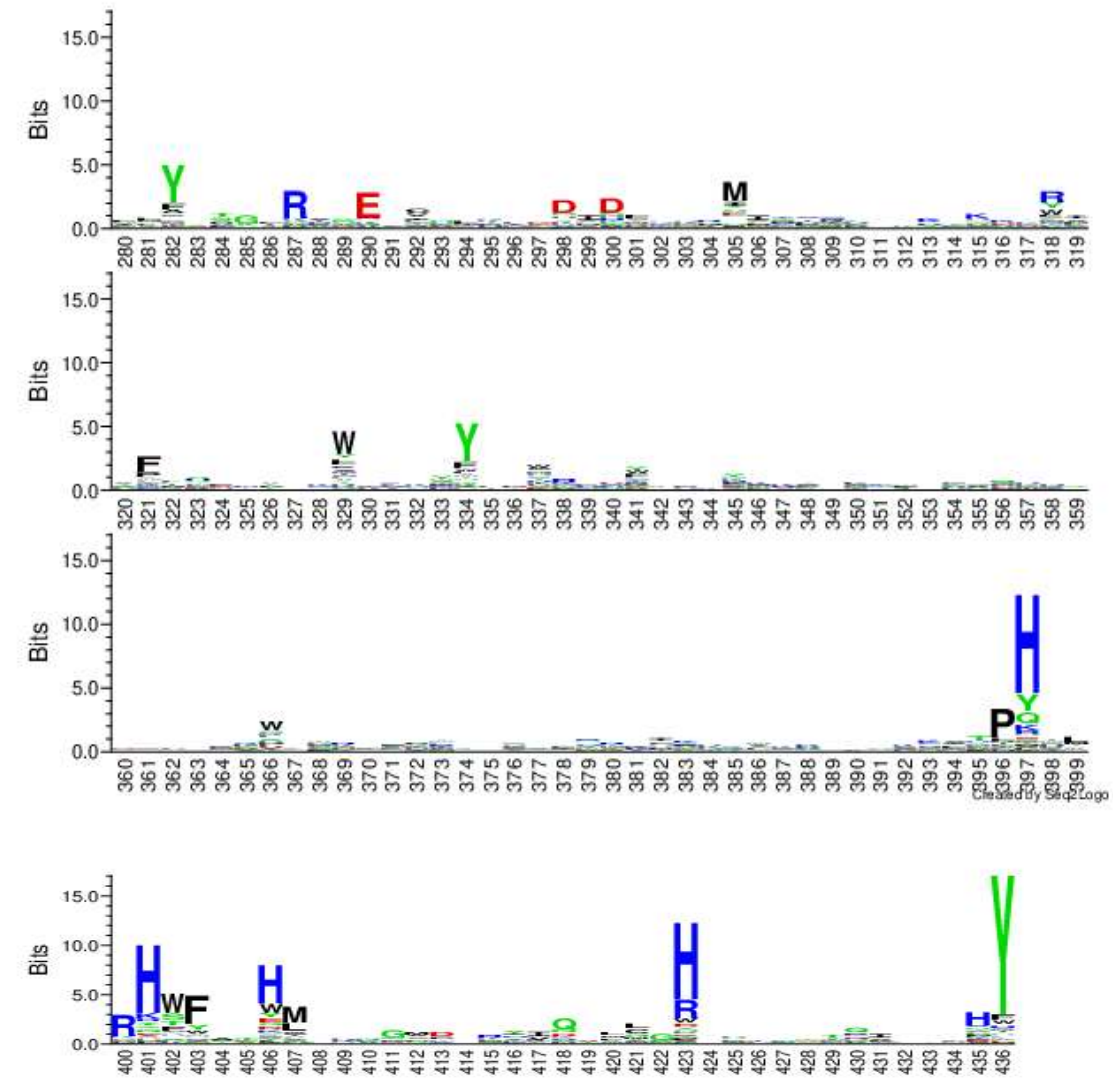
Now, to create a logo which shows the conservation pattern along the sequence. The amino acid letter codes are depicted in sizes according to their conservation. Paste the query sequence to the window, change the database to NR, the iterations to 3 and the Evalve to 0.001 and hit "submit".

The below screenshot is the resulting logo plot.





L280



Created by Seq2Logo

d) The PBD entry is 5VFZ\_A and the below is the image we depicts.

Welcome to the MPI Bioinformatics Toolkit! Would you like a tour?

Skip tour

Start



Search Alignment Sequence Analysis 2ary Structure 3ary Structure Classification Utils

HHblits HHpred HMMER PatternSearch ProtBLAST/PSI-BLAST pLM-BLAST

Sign In

ID	Date	Tool
2157623		HHPR

## HHpred

Job ID: 2157623, Created: 8 minutes ago

Input Parameters Results Raw Output Probability Plot Query Template MSA Query MSA

Vis Hits Aln Select All Forward Forward Query A3M Model using selection Download HHR Color Seqs Wrap Seqs

Number of Hits: 50

Query MSA diversity (Neff): 10.8467

Detected sequence features: ■ Coiled coil segment(s) ■ Transmembrane segment(s)

## Visualization

Resubmit Section

ID	Date	Tool
2157623		HHPR

Vis Hits Aln Select All Forward Forward Query A3M Model using selection Download HHR Color Seqs Wrap Seqs

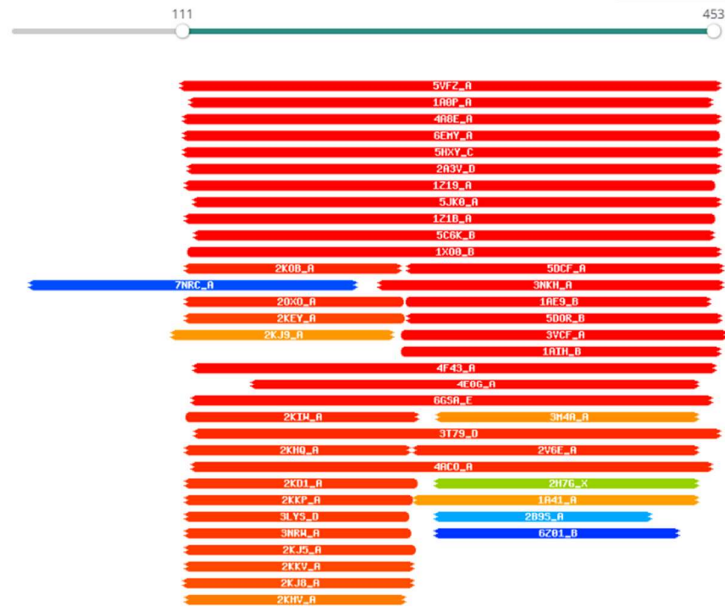
Number of Hits: 50

Query MSA diversity (Neff): 10.8467

Detected sequence features: ■ Coiled coil segment(s) ■ Transmembrane segment(s)

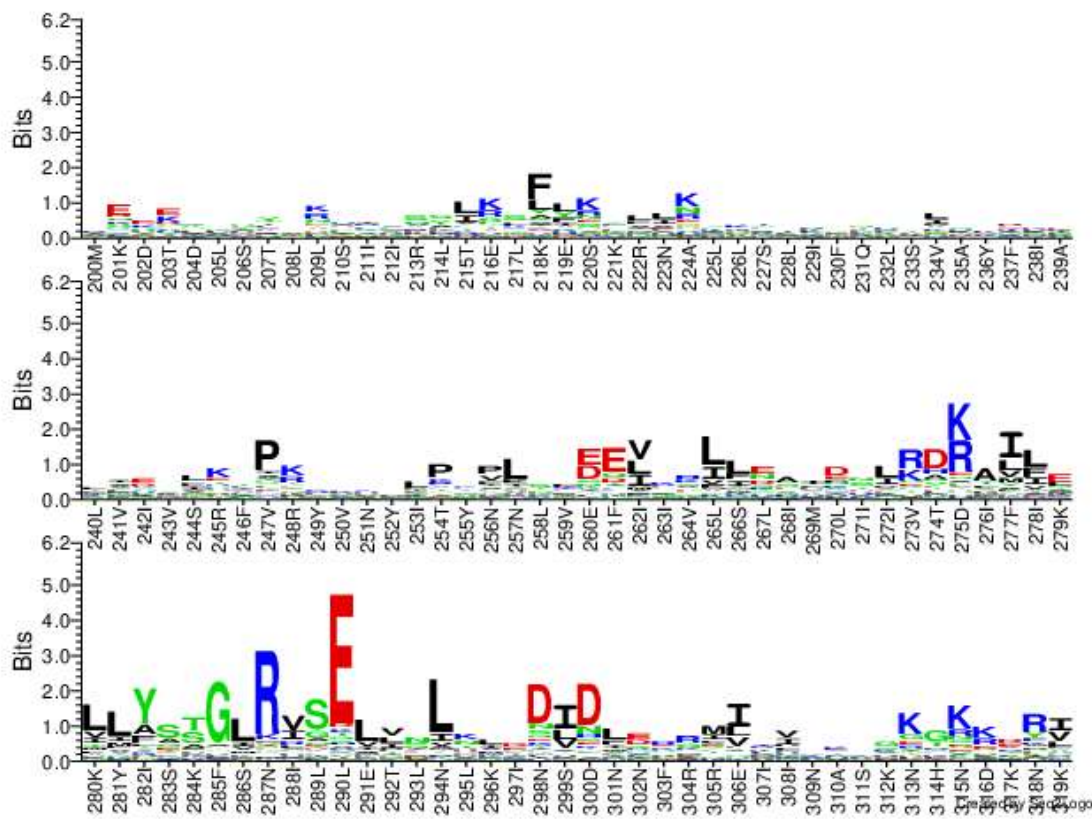
## Visualization

Resubmit Section

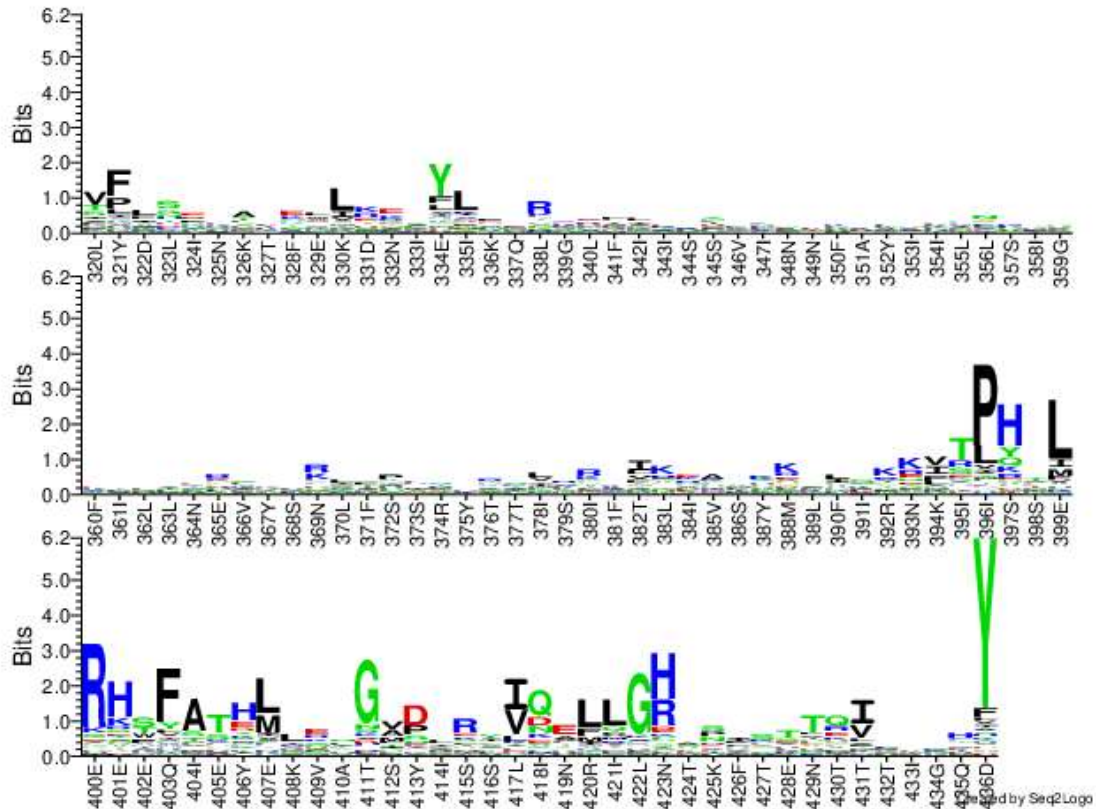


## Hitlist









3)

**a.** There are 1142 atoms in the PDB with the entry 3CHY.

In summary, there are 1141 atoms and one chain molecule so that will be 1142 atoms there in the protein.

**b.** Oxygen is the element (O) present in the atom 289.

**c.** In atom 289 I found the amino acid as GLU, and its name is Glutamic acid and the three-letter code for that acid is GLU.

**d.** In this atom we will find the below axis from the PDB file.

X= 40.782

Y = -0.025

Z= 6.322

**e.** Searching for 3CHY in SCOPe (at scop.berkeley.edu) reveals that this protein belongs to the following classification:

Class: Alpha and beta

According to the Scope (Structural Classification of Proteins - extended) database, 3CHY corresponds to a protein called "Nucleotide-binding domain of the cystic fibrosis transmembrane conductance regulator (CFTR)". Therefore, 3CHY belongs to the class of proteins that have both alpha helices and beta sheets in their structure, the NBP fold,

which is characterized by a central parallel beta-sheet flanked by alpha helices, and the ABC transporter-like family, which includes ATP-binding cassette transporters and other proteins with similar domains involved in ATP binding and hydrolysis.

From the below image from the scope we got the result as:

CLASS: Alpha and beta proteins(a/b)

FOLD: Flavodoxin-like 3 layers, a/b/a; parallel beta sheet of 5 strand order 21345

SUPER FAMILY: CheY-like

FAMILY: CheY-related

Protein: CheY protein.

SCOPe Browse Stats & History Downloads Help

### Lineage for d3chya\_ (3chy A:)

1. Root: SCOPe 2.08
2. Class c: Alpha and beta proteins (a/b) [51349] (148 folds)
3. Fold c.23: Flavodoxin-like [52171] (15 superfamilies)  
*3 layers, a/b/a; parallel beta-sheet of 5 strand, order 21345*
4. Superfamily c.23.1: CheY-like [52172] (8 families) S
5. Family c.23.1.1: CheY-related [52173] (26 proteins)
6. Protein CheY protein [52174] (6 species)
7. Species *Escherichia coli* [TaxId:562] [52175] (42 PDB entries)  
*Uniprot P06143*
8. Domain d3chya\_ : 3chy A: [31038]  
*complexed with so4*

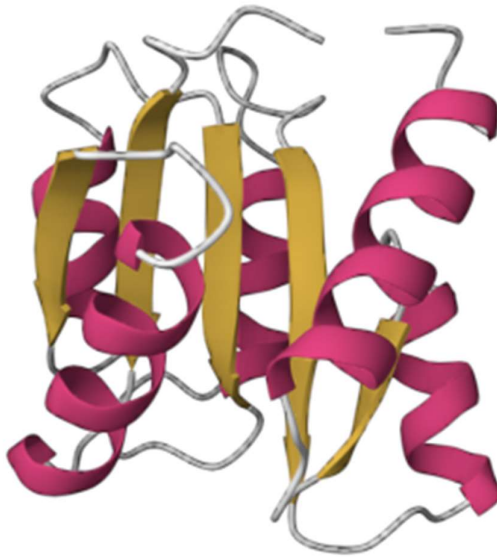
4)

a.

We found 5 helices (refer only to those coloured in red).



We found 5 beta strands (refer only to those coloured in yellow).



b. Hide the default cartoon representation and create a ball and stick representation. And from that we will be having the below image.

RCSB PDB

Deposit

Search

Visualize

Analyze

Download

Learn

About

Documentation

Careers

MyPDB

Contact us

Polymer

3 reprs

Line

Molecular Surface

Orientation

Point

Putty

Spacefill

Non-covalent Interactions

Validation Clashes

Set Coloring

Modify by Selection

Select This

Edit Label

Cartoon Representation

Cartoon Representation

Ball & Stick Representation

Water

Ball & Stick

Ion

Ball & Stick

Unit Cell P 21 21 21

Density

Quality Assessment

Assembly Symmetry

Export Models

3CHY | Model 1 | Instance ASM\_1 | A | LEU 23 [auth 24]

CHEV

**Programming Objects:**

This is the program for Part 3 in R language: using bio3d package.

```
library(bio3d)
pdb.id <- "3chy"
pdb.file <- get.pdb(pdb.id)
atom.count <- length(atom.select(pdb.file, all=TRUE)$xyz)
cat(paste("Total number of atoms: ", atom.count, "\n"))
atom.289 <- atom.select(pdb.file, atom=289)
atom.type <- atom.289$aname
cat(paste("Atom type of atom 289: ", atom.type, "\n"))
residue.289 <- atom.289$resno
residue.type <- pdb.file$atom$pdb$resnam[residue.289]
cat(paste("Residue type of atom 289: ", residue.type, "\n"))
coords <- atom.289$xyz
x <- coords[1]
y <- coords[2]
z <- coords[3]
cat(paste("Coordinates of atom 289: x=", x, ", y=", y, ", z=", z, "\n"))
```

**Temporary output:**

Total number of atoms: 1142

Atom type of atom 289: OE2

Residue type of atom 289: GLU

Coordinates of atom 289: x=40.782, y=-0.025, z=6.32



