WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 27, 2014 – 12:05 AM GMT

PDB ID   :   2VHH
Title   :   CRYSTAL STRUCTURE OF A PYRIMIDINE DEGRADING ENZYME
               FROM DROSOPHILA MELANOGASTER
Authors   :   Lundgren, S.; Lohkamp, B.; Andersen, B.; Piskur, J.; Dobritzsch, D.
Deposited on :   2007-11-21
Resolution :   2.80 Å(reported)

The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1   Overall quality at a glance ⓘ

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.256
Clashscore		8
Ramachandran outliers		0
Sidechain outliers		8.4%
RSRZ outliers		1.7%

Worse ← → Better

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
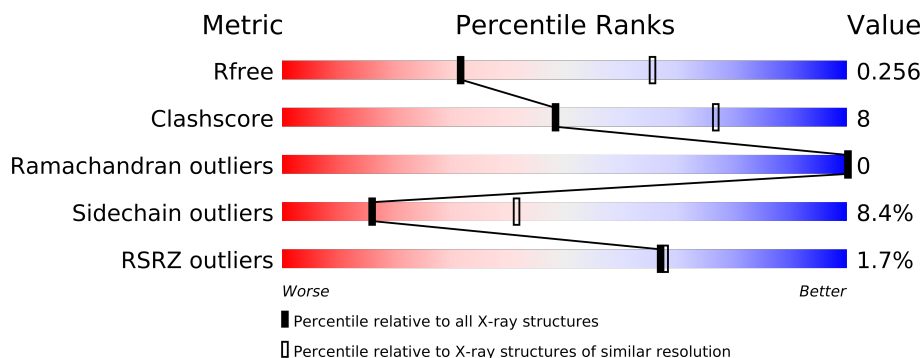
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1799 (2.80-2.80)
Clashscore	79885	2295 (2.80-2.80)
Ramachandran outliers	78287	2252 (2.80-2.80)
Sidechain outliers	78261	2254 (2.80-2.80)
RSRZ outliers	66119	1802 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	405	
1	B	405	
1	C	405	
1	D	405	

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 11865 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

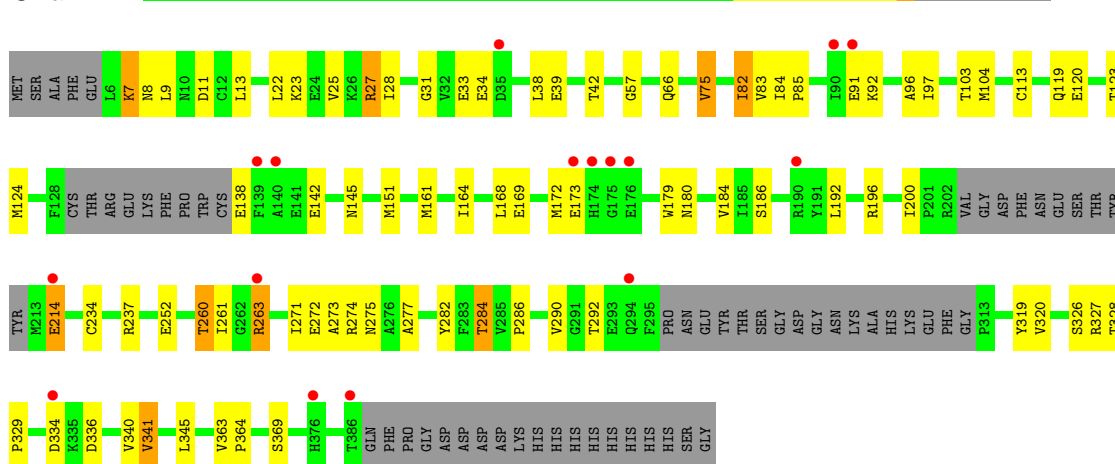- Molecule 1 is a protein called CG3027-PA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	345	Total	C	N	O	S	0	0	0
			2750	1743	485	508	14			
1	B	379	Total	C	N	O	S	0	0	0
			3027	1916	530	565	16			
1	C	382	Total	C	N	O	S	0	1	0
			3061	1937	536	572	16			
1	D	379	Total	C	N	O	S	0	0	0
			3027	1916	530	565	16			

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
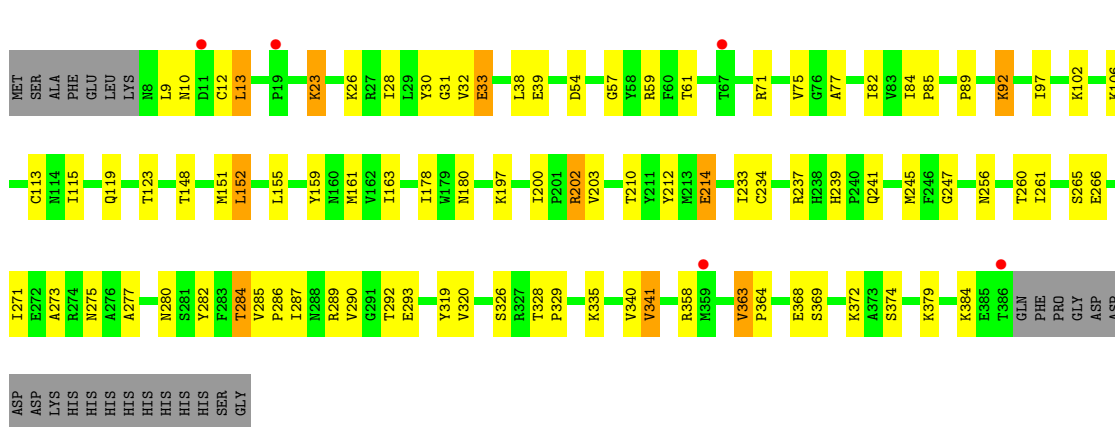
- Molecule 1: CG3027-PA

Chain A:

- Molecule 1: CG3027-PA

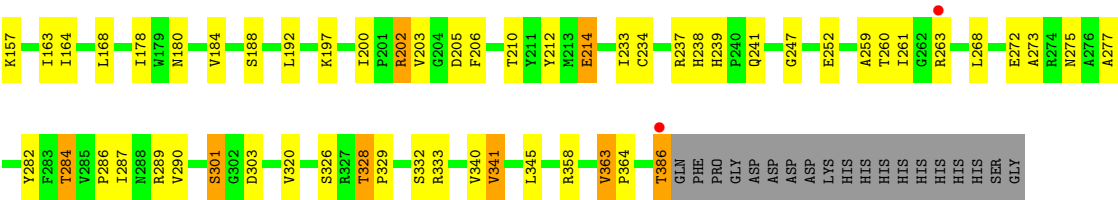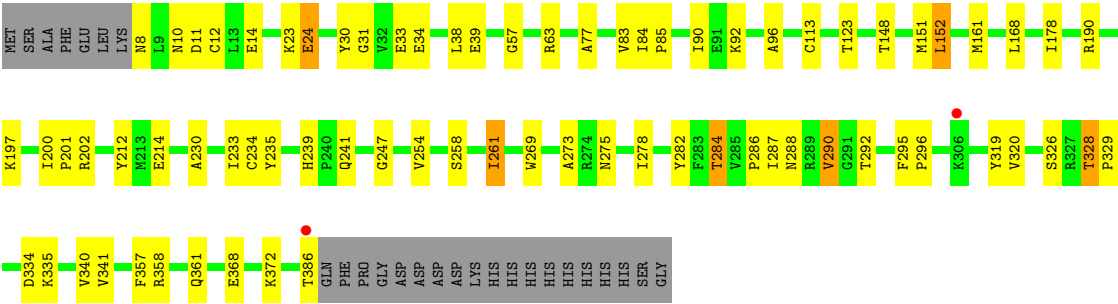Chain B:

- Molecule 1: CG3027-PA

Chain C:

K157 I163 I164 L168 I178 W179 N180 V184 S188 L192 K197 I200 P201 R202 V203 G204 D205 F206 T210 Y211 Y212 M213 E214 I233 C234 R237 H238 H239 P240 Q241 G247 E252 A259 T260 I261 G262 R263 L268 E272 A273 R274 N275 A276 A277

Y282 F283 T284 V285 P286 I287 N288 R289 V290 S301 G302 D303 V320 S326 R327 T328 P329 S332 R333 V340 V341 L345 R358 V363 P364 T386 GLN PHE PRO GLY ASP ASP ASP LYS HIS HIS HIS HIS HIS HIS HIS SER GLY

- Molecule 1: CG3027-PA

Chain D:

MET SER ALA PHE GLU LEU LYS N8 L9 N10 D11 C12 L13 E14 K23 E24 Y30 G31 V32 E33 E34 L38 E39 G57 R63 A77 V83 I84 P85 I90 E91 K92 A96 C113 T123 T148 M151 L152 M161 L168 I178 R190

K197 I200 P201 R202 Y212 M213 E214 A230 I233 C234 Y235 H239 P240 Q241 G247 V254 S258 I261 W269 A273 R274 N275 I278 Y282 F283 T284 V285 P286 I287 N288 R289 V290 G291 T292 F295 P296 K306 Y319 V320 S326 R327 T328 P329

D334 K335 V340 V341 F357 R358 Q361 E368 K372 T386 GLN PHE PRO GLY ASP ASP ASP LYS HIS HIS HIS HIS HIS HIS HIS SER GLY

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	95.04Å 197.88Å 97.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.00 – 2.80 56.16 – 2.80	Depositor EDS
% Data completeness (in resolution range)	94.8 (35.00-2.80) 94.8 (56.16-2.80)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.16 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.215 , 0.255 0.218 , 0.256	Depositor DCC
R_{free} test set	2237 reflections (5.37%)	DCC
Wilson B-factor (Å^2)	41.1	Xtriage
Anisotropy	0.159	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 7.8	EDS
Estimated twinning fraction	0.025 for l,-k,h	Xtriage
L-test for twinning	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Outliers	0 of 43921 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	11865	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	A	0.49	0/2810	0.59	0/3796
1	B	0.45	0/3101	0.62	0/4197
1	C	0.43	0/3135	0.62	0/4242
1	D	0.46	0/3101	0.59	0/4197
All	All	0.46	0/12147	0.61	0/16432

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	1
1	C	0	2
1	D	0	1
All	All	0	6

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	261	ILE	Peptide
1	A	82	ILE	Peptide
1	B	261	ILE	Peptide
1	C	261	ILE	Peptide
1	C	82	ILE	Peptide

5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2750	0	2726	53	0
1	B	3027	0	2958	56	0
1	C	3061	0	2993	51	0
1	D	3027	0	2958	41	0
All	All	11865	0	11635	183	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

The worst 5 of 183 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:82:ILE:HG12	1:A:83:VAL:H	1.32	0.95
1:C:205:ASP:OD1	1:C:301:SER:HB3	1.70	0.90
1:B:320:VAL:HG21	1:B:341:VAL:HG11	1.60	0.83
1:C:84:ILE:HB	1:C:85:PRO:HD2	1.64	0.79
1:D:84:ILE:HB	1:D:85:PRO:HD2	1.66	0.78

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	337/405 (83%)	322 (96%)	15 (4%)	0	100	100
1	B	377/405 (93%)	364 (97%)	13 (3%)	0	100	100
1	C	381/405 (94%)	370 (97%)	11 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	377/405 (93%)	366 (97%)	11 (3%)	0	100	100
All	All	1472/1620 (91%)	1422 (97%)	50 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	294/346 (85%)	267 (91%)	27 (9%)	13	36
1	B	323/346 (93%)	296 (92%)	27 (8%)	16	41
1	C	327/346 (94%)	297 (91%)	30 (9%)	13	36
1	D	323/346 (93%)	301 (93%)	22 (7%)	22	54
All	All	1267/1384 (92%)	1161 (92%)	106 (8%)	16	41

5 of 106 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	290	VAL
1	C	51	ASN
1	D	258	SER
1	B	341	VAL
1	B	384	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	275	ASN
1	C	51	ASN
1	D	8	ASN
1	B	256	ASN
1	C	275	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	345/405 (85%)	0.07	16 (4%)	31	31	20, 33, 55, 61	0
1	B	379/405 (93%)	-0.15	5 (1%)	74	75	20, 33, 55, 61	0
1	C	382/405 (94%)	-0.18	2 (0%)	88	90	20, 33, 55, 62	0
1	D	379/405 (93%)	-0.12	2 (0%)	88	90	20, 33, 54, 61	0
All	All	1485/1620 (91%)	-0.10	25 (1%)	67	68	20, 33, 55, 62	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	386	THR	5.6
1	B	67	THR	4.4
1	A	376	HIS	4.1
1	A	90	ILE	4.0
1	B	386	THR	3.8

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.