wwPDB X-ray Structure Validation Summary Report (i)

Jan 13, 2024 – 06:04 pm GMT

PDB ID	:	6Y9K
Title	:	Esterase EST8 with transacylase activity
Authors	:	Palm, G.J.; Lammers, M.; Berndt, L.
Deposited on	:	2020-03-09
Resolution	:	2.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

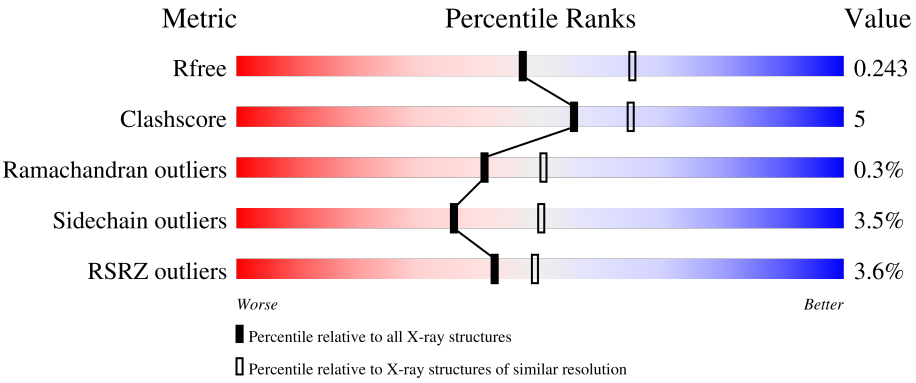
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	AAA	304	2% 88% 10%
1	BBB	304	4% 85% 12%
1	CCC	304	4% 83% 13%
1	DDD	304	4% 78% 18%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9012 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Esterase Est8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	297	Total	C	N	O	S	0	0	0
			2206	1401	380	409	16			
1	BBB	297	Total	C	N	O	S	0	1	0
			2209	1403	380	410	16			
1	CCC	297	Total	C	N	O	S	0	1	0
			2209	1403	380	410	16			
1	DDD	297	Total	C	N	O	S	0	1	0
			2209	1403	380	410	16			

- Molecule 2 is water.

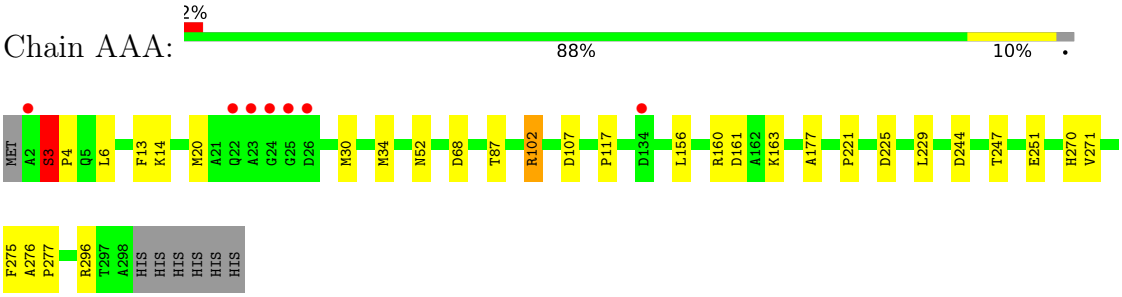
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	63	Total	O	0	0
			63	63		
2	BBB	32	Total	O	0	0
			32	32		
2	CCC	47	Total	O	0	0
			47	47		
2	DDD	37	Total	O	0	0
			37	37		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

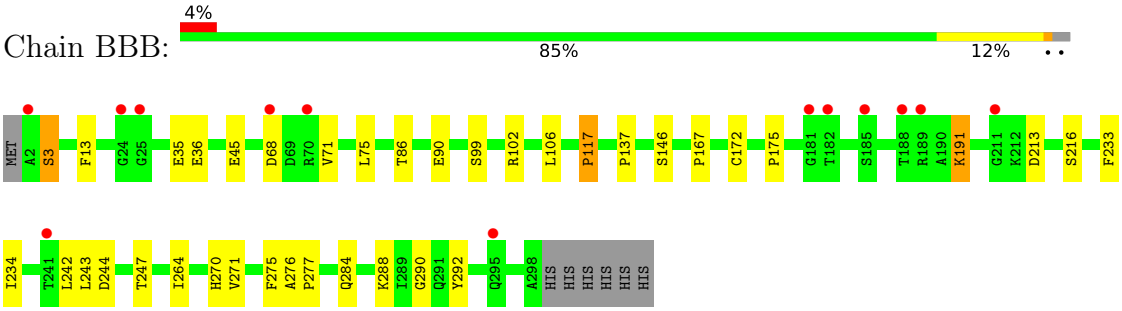
• Molecule 1: Esterase Est8

Chain AAA:



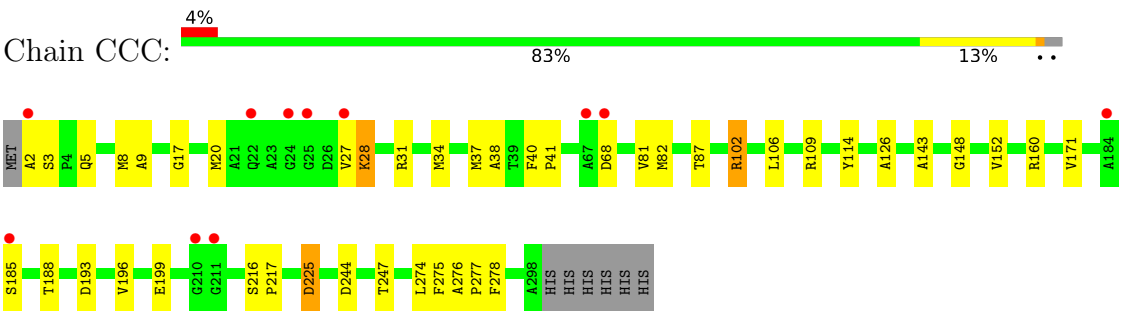
• Molecule 1: Esterase Est8

Chain BBB:



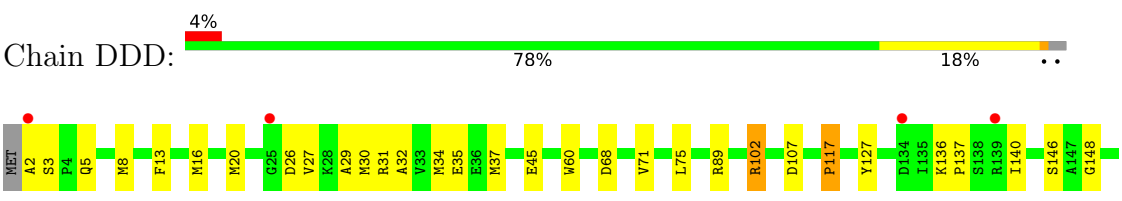
• Molecule 1: Esterase Est8

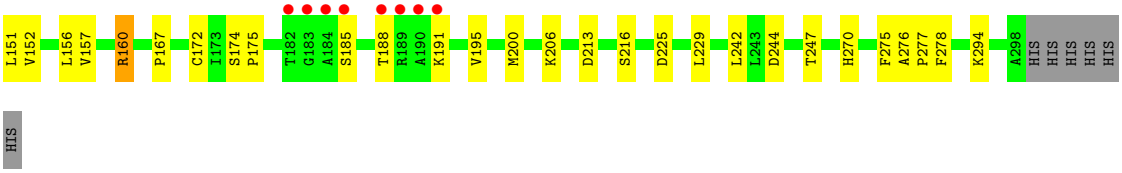
Chain CCC:



• Molecule 1: Esterase Est8

Chain DDD:




L151
V152
L156
V157
R160
P167
C172
I173
S174
P175
T182
G183
A184
S185
T188
R189
A190
K191
V195
M200
K206
D213
S216
D225
L229
L242
L243
D244
T247
H270
F275
A276
P277
F278
K294
A298
HIS
HIS
HIS
HIS
HIS

HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	78.17Å 78.18Å 112.66Å 90.00° 100.68° 90.00°	Depositor
Resolution (Å)	49.50 – 2.30 49.47 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.1 (49.50-2.30) 98.8 (49.47-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.19	Depositor
$< I/\sigma(I) >$ [1]	1.31 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.205 , 0.239 0.212 , 0.243	Depositor DCC
R_{free} test set	1492 reflections (2.52%)	wwPDB-VP
Wilson B-factor (Å^2)	37.7	Xtriage
Anisotropy	0.586	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 35.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9012	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.36% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	AAA	0.72	0/2255	0.89	0/3063
1	BBB	0.74	0/2261	0.89	1/3071 (0.0%)
1	CCC	0.71	0/2261	0.89	0/3071
1	DDD	0.74	0/2261	0.89	2/3071 (0.1%)
All	All	0.73	0/9038	0.89	3/12276 (0.0%)

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	DDD	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	DDD	160	ARG	NE-CZ-NH2	-6.38	117.11	120.30
1	BBB	117	PRO	N-CA-CB	-6.26	95.71	102.60
1	DDD	117	PRO	N-CA-CB	-5.41	96.65	102.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	DDD	2	ALA	Peptide

5.2 Too-close contacts (i)

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 hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	AAA	2206	0	2224	16	0
1	BBB	2209	0	2229	25	0
1	CCC	2209	0	2229	26	0
1	DDD	2209	0	2229	27	0
2	AAA	63	0	0	2	0
2	BBB	32	0	0	6	0
2	CCC	47	0	0	1	0
2	DDD	37	0	0	0	0
All	All	9012	0	8911	91	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 91 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:BBB:146[B]:SER:OG	2:BBB:401:HOH:O	1.87	0.89
1:BBB:213:ASP:O	1:BBB:216:SER:HB2	1.78	0.83
1:BBB:146[A]:SER:OG	2:BBB:401:HOH:O	1.94	0.83
1:DDD:160:ARG:NH2	1:DDD:225:ASP:O	2.16	0.75
1:DDD:244:ASP:HA	1:DDD:247:THR:OG1	1.88	0.74

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	AAA	295/304 (97%)	283 (96%)	11 (4%)	1 (0%)	41	50
1	BBB	296/304 (97%)	277 (94%)	18 (6%)	1 (0%)	41	50
1	CCC	296/304 (97%)	282 (95%)	13 (4%)	1 (0%)	41	50
1	DDD	296/304 (97%)	277 (94%)	18 (6%)	1 (0%)	41	50
All	All	1183/1216 (97%)	1119 (95%)	60 (5%)	4 (0%)	41	50

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	3	SER
1	BBB	3	SER
1	CCC	3	SER
1	DDD	157	VAL

5.3.2 Protein sidechains (i)

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	AAA	222/229 (97%)	215 (97%)	7 (3%)	39	54
1	BBB	223/229 (97%)	218 (98%)	5 (2%)	52	69
1	CCC	223/229 (97%)	219 (98%)	4 (2%)	59	75
1	DDD	223/229 (97%)	208 (93%)	15 (7%)	16	21
All	All	891/916 (97%)	860 (96%)	31 (4%)	36	50

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	199	GLU
1	DDD	195	VAL
1	DDD	20	MET
1	DDD	206	LYS
1	DDD	117	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

5.3.3 RNA

There are no RNA molecules 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 monosaccharides 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

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	AAA	297/304 (97%)	0.27	7 (2%)	59	66	27, 41, 67, 84	0
1	BBB	297/304 (97%)	0.27	13 (4%)	34	41	29, 44, 79, 104	0
1	CCC	297/304 (97%)	0.21	11 (3%)	41	48	27, 41, 70, 88	0
1	DDD	297/304 (97%)	0.28	12 (4%)	38	45	25, 42, 73, 95	0
All	All	1188/1216 (97%)	0.26	43 (3%)	42	49	25, 42, 72, 104	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	CCC	25	GLY	5.3
1	BBB	25	GLY	5.2
1	BBB	188	THR	5.1
1	AAA	25	GLY	4.9
1	DDD	2	ALA	4.1

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.