



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 17, 2018 – 11:21 PM EST

PDB ID : 6EOS
Title : DPP8 - Apo, space group 19
Authors : Ross, B.R.; Huber, R.
Deposited on : 2017-10-10
Resolution : 3.10 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030736
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030736

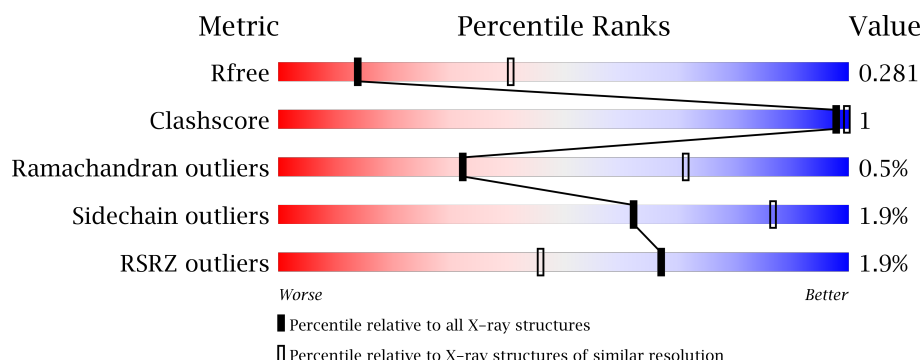
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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}	100719	1001 (3.12-3.08)
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)

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. 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 $\leq 5\%$. 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	A	898	<div> <div>0.1%</div> <div>87%</div> <div>9%</div> </div>
1	B	898	<div> <div>6%</div> <div>83%</div> <div>7%</div> <div>10%</div> </div>
1	C	898	<div> <div>0.1%</div> <div>88%</div> <div>8%</div> </div>
1	D	898	<div> <div>87%</div> <div>9%</div> </div>
1	E	898	<div> <div>0.1%</div> <div>89%</div> <div>8%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	898	<div><div></div><div>2%</div><div>86%</div><div>5%</div><div>9%</div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 39998 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 Dipeptidyl peptidase 8.

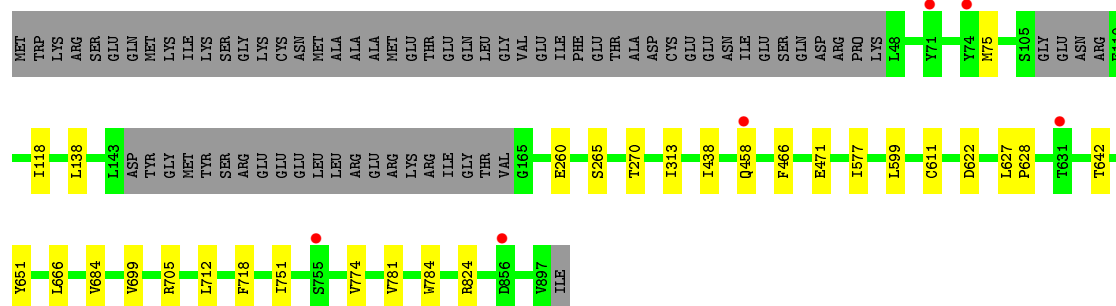
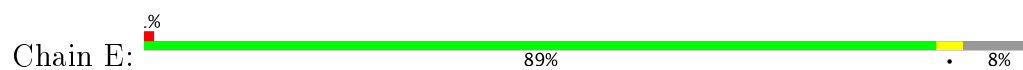
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	815	Total	C	N	O	S	0	0	0
			6620	4257	1107	1230	26			
1	C	829	Total	C	N	O	S	0	0	0
			6740	4334	1130	1250	26			
1	D	815	Total	C	N	O	S	0	0	0
			6613	4253	1104	1229	27			
1	E	825	Total	C	N	O	S	0	0	0
			6697	4308	1118	1243	28			
1	F	815	Total	C	N	O	S	0	0	0
			6620	4257	1107	1230	26			
1	B	811	Total	C	N	O	S	0	0	0
			6583	4233	1100	1224	26			

- Molecule 2 is water.

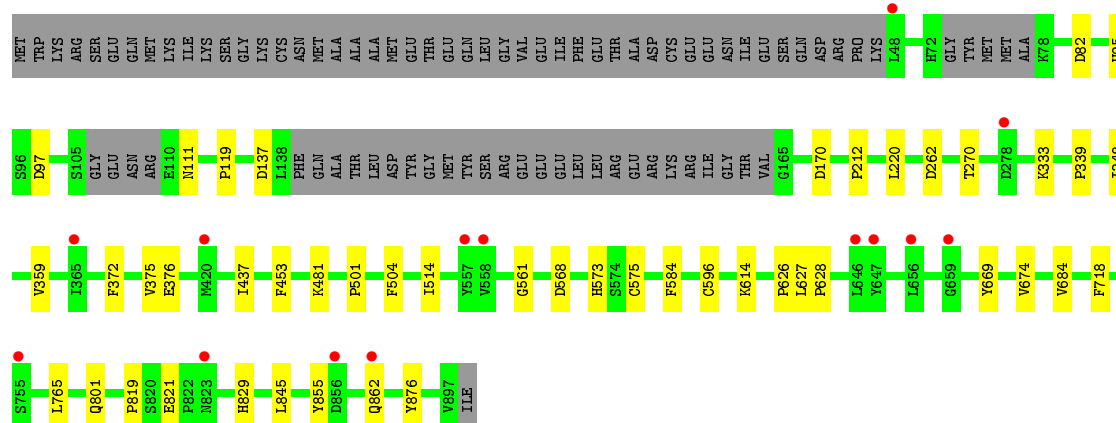
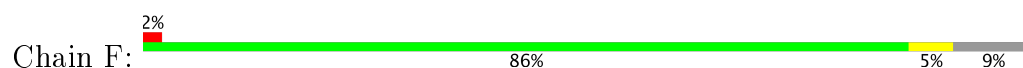
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	32	Total	O	0	0
			32	32		
2	C	23	Total	O	0	0
			23	23		
2	D	21	Total	O	0	0
			21	21		
2	E	21	Total	O	0	0
			21	21		
2	F	20	Total	O	0	0
			20	20		
2	B	8	Total	O	0	0
			8	8		



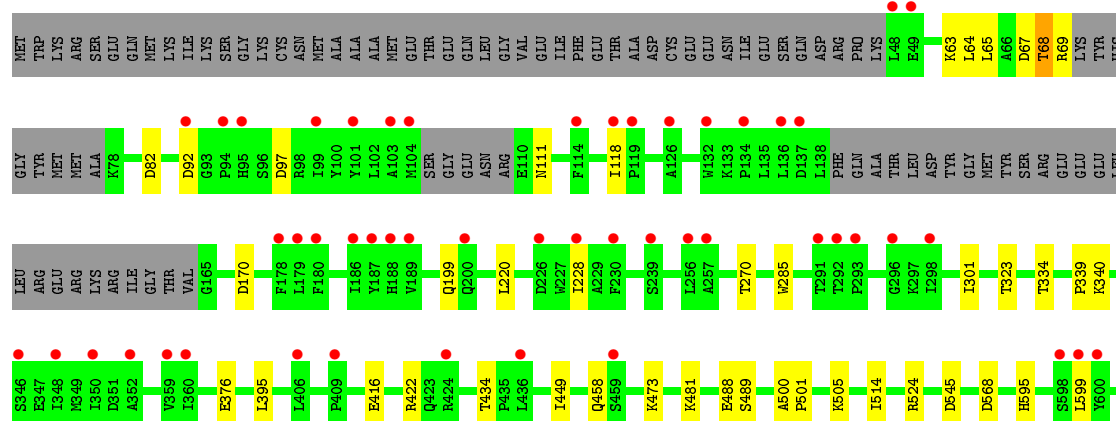
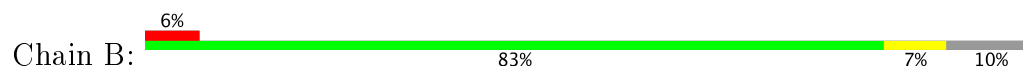
• Molecule 1: Dipeptidyl peptidase 8



• Molecule 1: Dipeptidyl peptidase 8



• Molecule 1: Dipeptidyl peptidase 8





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	148.12Å 266.78Å 268.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.40 – 3.10 44.46 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.40-3.10) 99.9 (44.46-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.18	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.65 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.247 , 0.283 0.248 , 0.281	Depositor DCC
R_{free} test set	9593 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	83.4	Xtriage
Anisotropy	0.549	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 59.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.003 for -h,l,k	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	39998	wwPDB-VP
Average B, all atoms (Å ²)	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.28% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	A	0.36	0/6805	0.58	0/9234
1	B	0.39	0/6766	0.57	0/9182
1	C	0.36	0/6925	0.58	0/9393
1	D	0.36	0/6797	0.57	0/9223
1	E	0.36	0/6885	0.59	0/9343
1	F	0.37	0/6805	0.58	0/9234
All	All	0.37	0/40983	0.58	0/55609

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	A	6620	0	6439	12	0
1	B	6583	0	6405	17	0
1	C	6740	0	6576	14	0
1	D	6613	0	6433	7	0
1	E	6697	0	6515	13	0
1	F	6620	0	6439	11	0
2	A	32	0	0	0	0
2	B	8	0	0	0	0
2	C	23	0	0	0	0
2	D	21	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	21	0	0	0	0
2	F	20	0	0	0	0
All	All	39998	0	38807	74	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 1.

The worst 5 of 74 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:A:64:LEU:O	1:A:68:THR:HG23	1.93	0.67
1:A:333:LYS:HG2	1:A:336:THR:HG23	1.90	0.53
1:E:666:LEU:HD23	1:E:751:ILE:HD12	1.91	0.52
1:E:774:VAL:HG12	1:E:824:ARG:HA	1.90	0.52
1:C:651:TYR:HB2	1:C:699:VAL:HB	1.92	0.52

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	A	807/898 (90%)	767 (95%)	35 (4%)	5 (1%)	28	67
1	B	803/898 (89%)	728 (91%)	70 (9%)	5 (1%)	28	67
1	C	821/898 (91%)	777 (95%)	42 (5%)	2 (0%)	51	84
1	D	807/898 (90%)	754 (93%)	47 (6%)	6 (1%)	25	64
1	E	819/898 (91%)	765 (93%)	54 (7%)	0	100	100
1	F	807/898 (90%)	750 (93%)	49 (6%)	8 (1%)	18	57
All	All	4864/5388 (90%)	4541 (93%)	297 (6%)	26 (0%)	32	71

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	68	THR
1	C	158	ARG
1	F	561	GLY
1	A	70	LYS
1	D	138	LEU

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	724/795 (91%)	714 (99%)	10 (1%)	71	90
1	B	720/795 (91%)	694 (96%)	26 (4%)	40	75
1	C	737/795 (93%)	727 (99%)	10 (1%)	71	90
1	D	723/795 (91%)	712 (98%)	11 (2%)	70	89
1	E	731/795 (92%)	724 (99%)	7 (1%)	80	92
1	F	724/795 (91%)	707 (98%)	17 (2%)	56	83
All	All	4359/4770 (91%)	4278 (98%)	81 (2%)	62	87

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	622	ASP
1	F	568	ASP
1	B	629	ASP
1	E	718	PHE
1	F	170	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	882	HIS
1	D	730	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	274	GLN
1	D	200	GLN
1	D	591	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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	815/898 (90%)	-0.03	7 (0%) 84 69	57, 86, 125, 160	0
1	B	811/898 (90%)	0.49	56 (6%) 18 7	69, 123, 160, 189	0
1	C	829/898 (92%)	0.04	7 (0%) 86 71	64, 91, 127, 174	0
1	D	815/898 (90%)	-0.03	1 (0%) 95 90	65, 95, 126, 152	0
1	E	825/898 (91%)	0.07	6 (0%) 87 75	66, 94, 135, 185	0
1	F	815/898 (90%)	0.29	14 (1%) 70 49	76, 112, 141, 168	0
All	All	4910/5388 (91%)	0.14	91 (1%) 67 46	57, 98, 141, 189	0

The worst 5 of 91 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	189	VAL	5.8
1	B	187	TYR	5.7
1	B	94	PRO	5.3
1	B	239	SER	4.7
1	E	71	TYR	4.5

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.