Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 12:35 AM BST

PDB ID	:	5RE2
Title	:	PanDDA analysis group deposition – Endothiapepsin ground state model 59
Authors	:	Weiss, M.S.; Wollenhaupt, J.; Metz, A.; Barthel, T.; Lima, G.M.A.; Heine, A.; Mueller, U.; Klebe, G.
Deposited on	:	2020-03-24
Resolution	:	1.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

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

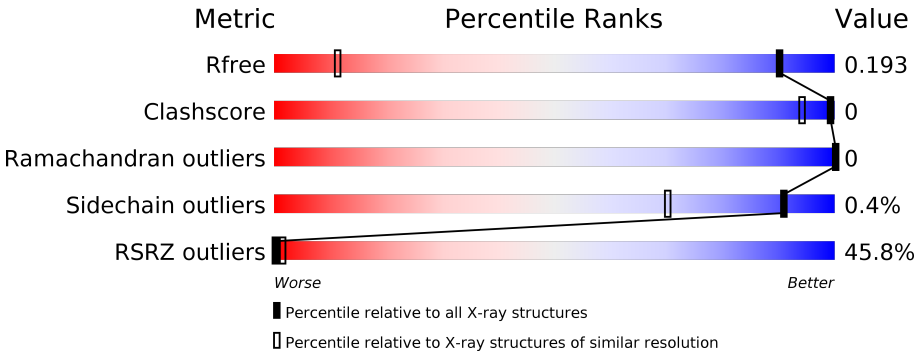
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 1.00 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	1050 (1.06-0.94)
Clashscore	141614	1117 (1.06-0.94)
Ramachandran outliers	138981	1043 (1.06-0.94)
Sidechain outliers	138945	1045 (1.06-0.94)
RSRZ outliers	127900	1023 (1.06-0.94)

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 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	A	419	36%, 77%, 21%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2717 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 Endothiapepsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	330	2462	1566	367	527	2	0	19	0

- Molecule 2 is water.

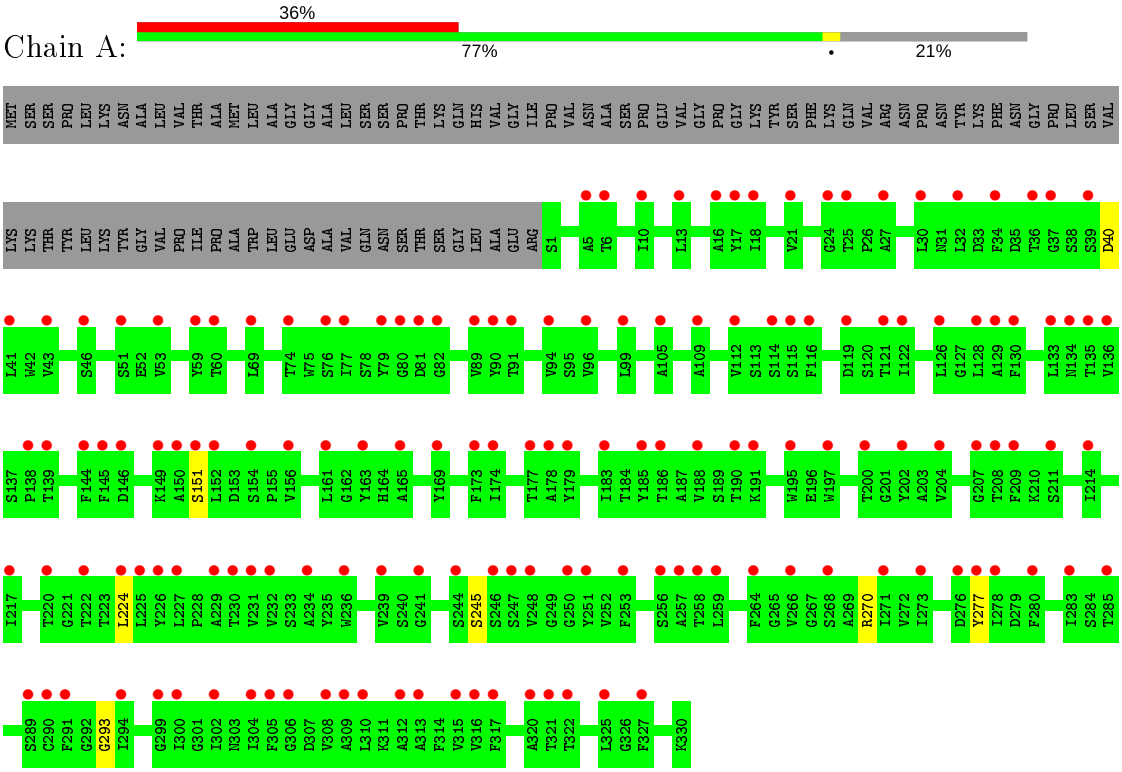
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	255	255	255	0	0

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: Endothiapepsin

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	45.29Å 72.85Å 52.46Å 90.00° 109.22° 90.00°	Depositor
Resolution (Å)	42.80 – 1.00 42.77 – 1.00	Depositor EDS
% Data completeness (in resolution range)	97.8 (42.80-1.00) 97.8 (42.77-1.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.96 (at 1.00Å)	Xtriage
Refinement program	REFMAC 5.8.0238, PHENIX 1.16.3549	Depositor
R, R_{free}	0.184 , 0.184 0.191 , 0.193	Depositor DCC
R_{free} test set	8532 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	10.6	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.45 , 56.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2717	wwPDB-VP
Average B, all atoms (Å^2)	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.57% 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	A	0.87	2/2552 (0.1%)	0.88	2/3496 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	151	SER	CA-CB	-5.41	1.44	1.52
1	A	245	SER	CA-CB	-5.04	1.45	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	270	ARG	NE-CZ-NH1	6.83	123.72	120.30
1	A	277	TYR	CB-CG-CD1	5.43	124.26	121.00

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	2462	0	2329	1	0
2	A	255	0	0	0	2
All	All	2717	0	2329	1	2

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 0.

All (1) 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:224:LEU:HD22	1:A:293:GLY:HA2	2.02	0.41

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:641:HOH:O	2:A:645:HOH:O[2_747]	1.96	0.24
2:A:422:HOH:O	2:A:622:HOH:O[1_556]	2.12	0.08

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	347/419 (83%)	344 (99%)	3 (1%)	0	100 100

There are no Ramachandran outliers to report.

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	A	270/336 (80%)	269 (100%)	1 (0%)	91 70

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	ASP

Some 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	A	330/419 (78%)	1.98	151 (45%) 0 1	8, 11, 18, 25	0

All (151) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	150	ALA	6.1
1	A	82	GLY	5.8
1	A	246[A]	SER	5.6
1	A	300[A]	ILE	5.2
1	A	321	THR	5.1
1	A	10	ILE	4.5
1	A	294	ILE	4.1
1	A	299[A]	GLY	3.9
1	A	304	ILE	3.8
1	A	325[A]	LEU	3.8
1	A	208	THR	3.8
1	A	322	THR	3.7
1	A	74	THR	3.6
1	A	122	ILE	3.6
1	A	139	THR	3.5
1	A	128	LEU	3.5
1	A	145	PHE	3.5
1	A	77	ILE	3.4
1	A	273	ILE	3.4
1	A	226	TYR	3.3
1	A	69	LEU	3.2
1	A	320	ALA	3.2
1	A	135	THR	3.2
1	A	53	VAL	3.2
1	A	315	VAL	3.2
1	A	18	ILE	3.1
1	A	214	ILE	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	34	PHE	3.0
1	A	188	VAL	3.0
1	A	126	LEU	3.0
1	A	305	PHE	2.9
1	A	136	VAL	2.9
1	A	178	ALA	2.9
1	A	183	ILE	2.9
1	A	138	PRO	2.8
1	A	121	THR	2.8
1	A	186	THR	2.8
1	A	236	TRP	2.8
1	A	89	VAL	2.8
1	A	105	ALA	2.8
1	A	217	ILE	2.8
1	A	144	PHE	2.8
1	A	225	LEU	2.8
1	A	197	TRP	2.7
1	A	152	LEU	2.7
1	A	191	LYS	2.7
1	A	247	SER	2.7
1	A	239	VAL	2.7
1	A	310	LEU	2.7
1	A	37	GLY	2.7
1	A	25	THR	2.7
1	A	119	ASP	2.6
1	A	177	THR	2.6
1	A	115	SER	2.6
1	A	316	VAL	2.6
1	A	36	THR	2.6
1	A	165	ALA	2.6
1	A	278	ILE	2.6
1	A	43	VAL	2.6
1	A	46	SER	2.6
1	A	312	ALA	2.6
1	A	134	ASN	2.6
1	A	264	PHE	2.6
1	A	231	VAL	2.5
1	A	146	ASP	2.5
1	A	30	LEU	2.5
1	A	161	LEU	2.5
1	A	224	LEU	2.5
1	A	174[A]	ILE	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	283	ILE	2.5
1	A	302	ILE	2.5
1	A	27	ALA	2.5
1	A	41	LEU	2.5
1	A	154[A]	SER	2.5
1	A	227	LEU	2.5
1	A	163	TYR	2.5
1	A	290	CYS	2.5
1	A	280	PHE	2.5
1	A	109	ALA	2.5
1	A	59	TYR	2.5
1	A	91	THR	2.4
1	A	51	SER	2.4
1	A	17	TYR	2.4
1	A	90	TYR	2.4
1	A	130	PHE	2.4
1	A	114	SER	2.4
1	A	151	SER	2.4
1	A	276[A]	ASP	2.4
1	A	291	PHE	2.4
1	A	207	GLY	2.4
1	A	32	LEU	2.3
1	A	169	TYR	2.3
1	A	258	THR	2.3
1	A	195	TRP	2.3
1	A	241	GLY	2.3
1	A	266	VAL	2.3
1	A	16	ALA	2.3
1	A	129	ALA	2.3
1	A	257	ALA	2.3
1	A	13	LEU	2.3
1	A	244	SER	2.3
1	A	79	TYR	2.3
1	A	259	LEU	2.3
1	A	268[A]	SER	2.3
1	A	190	THR	2.3
1	A	285	THR	2.3
1	A	253	PHE	2.3
1	A	313	ALA	2.3
1	A	96	VAL	2.3
1	A	256	SER	2.3
1	A	173	PHE	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	229	ALA	2.2
1	A	24	GLY	2.2
1	A	80	GLY	2.2
1	A	248	VAL	2.2
1	A	185	TYR	2.2
1	A	202	TYR	2.2
1	A	133	LEU	2.2
1	A	116	PHE	2.2
1	A	327	PHE	2.2
1	A	251	TYR	2.2
1	A	277	TYR	2.2
1	A	306	GLY	2.2
1	A	271	ILE	2.2
1	A	309	ALA	2.2
1	A	317	PHE	2.2
1	A	6	THR	2.2
1	A	60	THR	2.2
1	A	200	THR	2.2
1	A	149[A]	LYS	2.1
1	A	94	VAL	2.1
1	A	230	THR	2.1
1	A	99	LEU	2.1
1	A	209	PHE	2.1
1	A	76	SER	2.1
1	A	211	SER	2.1
1	A	156	VAL	2.1
1	A	204	VAL	2.1
1	A	232	VAL	2.1
1	A	5	ALA	2.1
1	A	220	THR	2.1
1	A	289[A]	SER	2.1
1	A	250	GLY	2.1
1	A	234	ALA	2.1
1	A	39	SER	2.1
1	A	222	THR	2.1
1	A	21	VAL	2.0
1	A	112	VAL	2.0
1	A	308	VAL	2.0
1	A	179	TYR	2.0
1	A	81	ASP	2.0

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

There are no such residues in this entry.