Full wwPDB X-ray Structure Validation Report ⓘ

Nov 30, 2023 – 02:03 PM EST

PDB ID	:	8EHN
Title	:	PRRSV-1 PLP2 domain
Authors	:	Bailey-Elkin, B.A.; Mark, B.L.
Deposited on	:	2022-09-14
Resolution	:	2.30 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

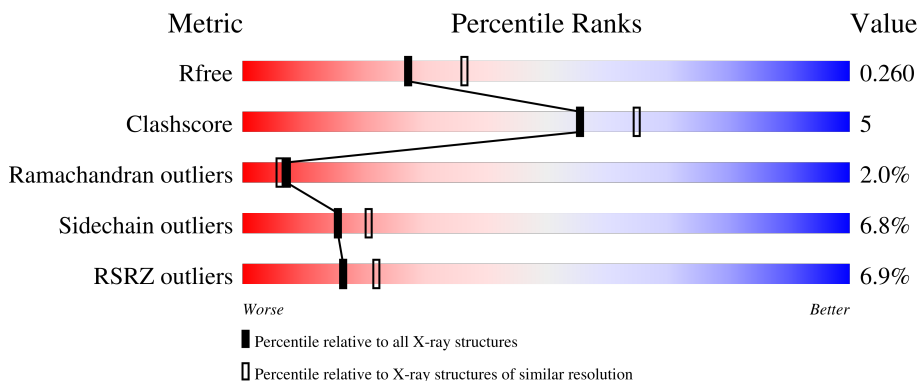|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

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	Percentile Ranks	Value
Rfree		0.260
Clashscore		5
Ramachandran outliers		2.0%
Sidechain outliers		6.8%
RSRZ outliers		6.9%

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$_{free}$	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	A	204	4% / 62% / 12% / 26%
2	B	204	6% / 63% / 11% · / 24%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	ACT	A	602	-	-	-	X

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 2339 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 Papain-like protease 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	150	1131	710	197	215	9	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	397	THR	ALA	conflict	UNP W0NX70
A	400	GLY	GLU	conflict	UNP W0NX70
A	405	LEU	PRO	conflict	UNP W0NX70
A	406	ALA	THR	conflict	UNP W0NX70
A	409	ILE	VAL	conflict	UNP W0NX70
A	411	PRO	LEU	conflict	UNP W0NX70
A	438	VAL	MET	conflict	UNP W0NX70
A	451	PRO	THR	conflict	UNP W0NX70
A	467	ALA	VAL	conflict	UNP W0NX70
A	513	SER	PRO	conflict	UNP W0NX70
A	564	ASP	ASN	conflict	UNP W0NX70
A	579	GLY	-	expression tag	UNP W0NX70
A	580	SER	-	expression tag	UNP W0NX70
A	581	SER	-	expression tag	UNP W0NX70
A	582	GLY	-	expression tag	UNP W0NX70
A	583	HIS	-	expression tag	UNP W0NX70
A	584	HIS	-	expression tag	UNP W0NX70
A	585	HIS	-	expression tag	UNP W0NX70
A	586	HIS	-	expression tag	UNP W0NX70
A	587	HIS	-	expression tag	UNP W0NX70
A	588	HIS	-	expression tag	UNP W0NX70

- Molecule 2 is a protein called Papain-like protease 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	155	Total 1176	C 737	N 203	O 226	S 10	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	397	THR	ALA	conflict	UNP W0NX70
B	400	GLY	GLU	conflict	UNP W0NX70
B	405	LEU	PRO	conflict	UNP W0NX70
B	406	ALA	THR	conflict	UNP W0NX70
B	409	ILE	VAL	conflict	UNP W0NX70
B	411	PRO	LEU	conflict	UNP W0NX70
B	438	VAL	MET	conflict	UNP W0NX70
B	451	PRO	THR	conflict	UNP W0NX70
B	467	ALA	VAL	conflict	UNP W0NX70
B	513	SER	PRO	conflict	UNP W0NX70
B	564	ASP	ASN	conflict	UNP W0NX70
B	579	GLY	-	expression tag	UNP W0NX70
B	580	SER	-	expression tag	UNP W0NX70
B	581	SER	-	expression tag	UNP W0NX70
B	582	GLY	-	expression tag	UNP W0NX70
B	583	HIS	-	expression tag	UNP W0NX70
B	584	HIS	-	expression tag	UNP W0NX70
B	585	HIS	-	expression tag	UNP W0NX70
B	586	HIS	-	expression tag	UNP W0NX70
B	587	HIS	-	expression tag	UNP W0NX70
B	588	HIS	-	expression tag	UNP W0NX70

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Zn 1	0	0
3	B	1	Total 1	Zn 1	0	0

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

ACT

OXT

O⁻

C

CH3

O

O

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0

• Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	10	Total 10	O 10	0	0
5	B	8	Total 8	O 8	0	0

3 Residue-property plots ⓘ

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: Papain-like protease 2

Chain A:
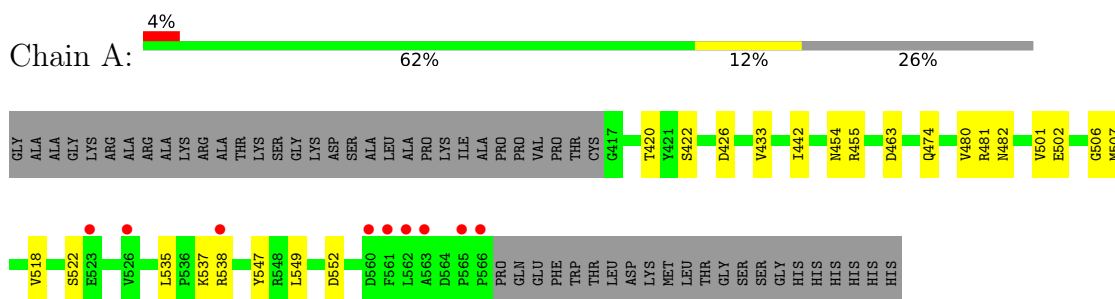
• Molecule 2: Papain-like protease 2

Chain B:
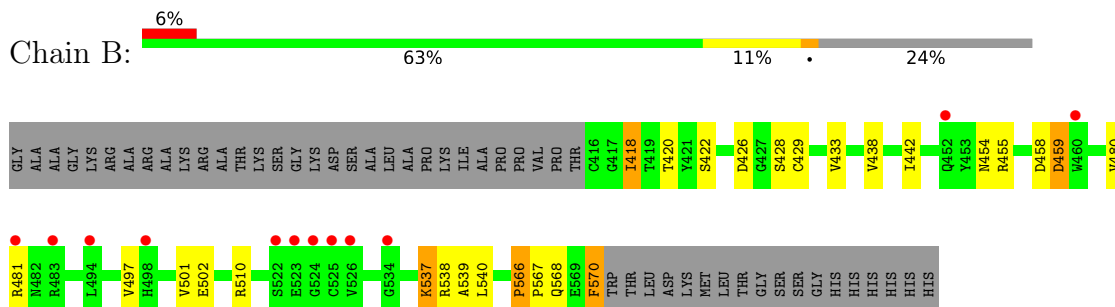
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	67.50Å 102.35Å 100.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.58 – 2.30 37.55 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.3 (37.58-2.30) 90.2 (37.55-2.30)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.73 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.200 , 0.253 0.208 , 0.260	Depositor DCC
R_{free} test set	1568 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	58.1	Xtriage
Anisotropy	0.608	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 75.8	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2339	wwPDB-VP
Average B, all atoms (Å^2)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ACT, OCS, ZN

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	0.47	0/1161	0.70	0/1591
2	B	0.47	0/1198	0.74	0/1640
All	All	0.47	0/2359	0.72	0/3231

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	1131	0	1096	12	0
2	B	1176	0	1129	13	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	12	0	9	0	0
5	A	10	0	0	0	0
5	B	8	0	0	0	0
All	All	2339	0	2234	23	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.

All (23) 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:420:THR:HG22	1:A:502:GLU:HG2	1.38	1.04
1:A:442:ILE:HD11	1:A:535:LEU:HD23	1.67	0.76
2:B:567:PRO:HA	2:B:570:PHE:CD1	2.22	0.75
1:A:420:THR:HG22	1:A:502:GLU:CG	2.16	0.73
1:A:422:SER:HB2	2:B:418:ILE:HD13	1.72	0.69
2:B:567:PRO:HA	2:B:570:PHE:CE1	2.28	0.68
2:B:455:ARG:HD3	2:B:459:ASP:O	1.94	0.68
2:B:433:VAL:HG13	2:B:501:VAL:HB	1.80	0.62
1:A:481:ARG:NH1	1:A:518:VAL:O	2.37	0.58
1:A:433:VAL:HG13	1:A:501:VAL:HB	1.86	0.56
2:B:438:VAL:HG11	2:B:540:LEU:HB3	1.88	0.55
1:A:422:SER:HB2	2:B:418:ILE:CD1	2.37	0.54
2:B:497:VAL:HG22	2:B:497:VAL:O	2.08	0.53
1:A:463:ASP:OD2	1:A:481:ARG:NH2	2.44	0.51
1:A:474:GLN:HE21	1:A:474:GLN:HA	1.77	0.49
1:A:455:ARG:NH2	1:A:549:LEU:O	2.46	0.49
2:B:567:PRO:CA	2:B:570:PHE:CE1	2.94	0.49
1:A:426:ASP:O	1:A:454:ASN:HA	2.14	0.47
2:B:538:ARG:O	2:B:540:LEU:N	2.48	0.45
2:B:420:THR:OG1	2:B:502:GLU:HG2	2.18	0.43
1:A:455:ARG:NH2	1:A:547:TYR:O	2.49	0.42
2:B:566:PRO:HA	2:B:567:PRO:HD3	1.87	0.42
2:B:426:ASP:O	2:B:454:ASN:HA	2.20	0.42

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	148/204 (72%)	139 (94%)	7 (5%)	2 (1%)	11	11

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	152/204 (74%)	141 (93%)	7 (5%)	4 (3%)	5	4
All	All	300/408 (74%)	280 (93%)	14 (5%)	6 (2%)	7	6

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	507	MET
2	B	539	ALA
2	B	459	ASP
2	B	537	LYS
1	A	506	GLY
2	B	566	PRO

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	123/164 (75%)	117 (95%)	6 (5%)	25	35
2	B	127/163 (78%)	116 (91%)	11 (9%)	10	12
All	All	250/327 (76%)	233 (93%)	17 (7%)	16	21

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	480	VAL
1	A	482	ASN
1	A	522	SER
1	A	537	LYS
1	A	538	ARG
1	A	552	ASP
2	B	418	ILE
2	B	422	SER
2	B	428	SER
2	B	442	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	458	ASP
2	B	480	VAL
2	B	481	ARG
2	B	510	ARG
2	B	537	LYS
2	B	568	GLN
2	B	570	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	474	GLN
2	B	452	GLN
2	B	498	HIS
2	B	568	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	OCS	B	429	2	7,8,9	1.47	1 (14%)	6,11,13	3.23	2 (33%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	OCS	B	429	2	-	3/4/7/9	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	429	OCS	OD1-SG	3.47	1.55	1.45

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	429	OCS	OD1-SG-CB	-6.98	98.64	106.94
2	B	429	OCS	OD2-SG-OD3	3.28	119.29	111.27

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	429	OCS	CA-CB-SG-OD1
2	B	429	OCS	CA-CB-SG-OD2
2	B	429	OCS	CA-CB-SG-OD3

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z|>2$	Counts	RMSZ	$\#|Z|>2$
4	ACT	A	604	-	3,3,3	0.98	0	3,3,3	0.78	0
4	ACT	A	603	-	3,3,3	0.96	0	3,3,3	0.79	0
4	ACT	A	602	-	3,3,3	0.95	0	3,3,3	0.83	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

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, 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(Å²)	Q<0.9
1	A	150/204 (73%)	0.49	9 (6%)	21	28	52, 83, 125, 151	0
2	B	154/204 (75%)	0.74	12 (7%)	13	17	54, 88, 140, 163	0
All	All	304/408 (74%)	0.62	21 (6%)	16	22	52, 85, 138, 163	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	565	PRO	5.7
2	B	522	SER	5.6
2	B	523	GLU	5.2
1	A	566	PRO	4.5
1	A	523	GLU	4.2
1	A	561	PHE	4.2
2	B	526	VAL	4.1
2	B	460	TRP	4.1
2	B	498	HIS	3.5
2	B	481	ARG	3.2
1	A	563	ALA	3.1
1	A	526	VAL	3.1
2	B	524	GLY	2.8
2	B	483	ARG	2.8
1	A	538	ARG	2.8
2	B	525	CYS	2.6
2	B	452	GLN	2.6
2	B	534	GLY	2.5
1	A	560	ASP	2.2
2	B	494	LEU	2.1
1	A	562	LEU	2.1

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	OCS	B	429	9/10	0.92	0.13	72,91,124,125	0

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	ACT	A	602	4/4	0.35	0.71	126,140,146,157	0
4	ACT	A	604	4/4	0.67	0.40	107,109,136,140	0
4	ACT	A	603	4/4	0.88	0.14	84,99,104,121	0
3	ZN	B	601	1/1	0.95	0.08	89,89,89,89	0
3	ZN	A	601	1/1	0.98	0.09	78,78,78,78	0

6.5 Other polymers ⓘ

There are no such residues in this entry.