WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Sep 12, 2023 – 12:19 AM EDT

PDB ID	:	4MMO
Title	:	The crystal structure of a M20 family metallo-carboxypeptidase Sso-CP2 from Sulfolobus solfataricus
Authors	:	Dupuy, J.; Dutoit, R.; Durisotti, V.; Demarez, M.; Borel, F.; Van Elder, D.; Legrain, C.; Bauvois, C.
Deposited on	:	2013-09-09
Resolution	:	2.34 Å(reported)

The following versions of software and data (see references (i)) 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	:	2.35.1
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)	:	2.35.1

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 2.34 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.
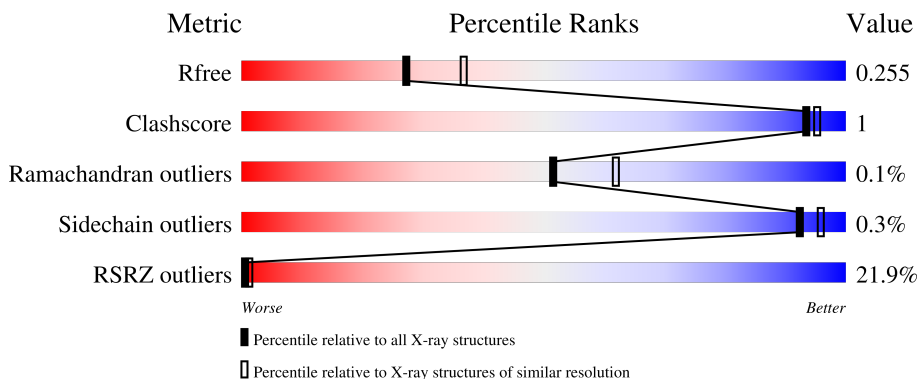
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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	437	7% / 96%
1	B	437	34% / 85% / 13%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12814 atoms, of which 6210 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 Sso-CP2 metallo-carboxypetidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	437	Total	C	H	N	O	S	0	0	0
			6845	2194	3431	561	650	9			
1	B	382	Total	C	H	N	O	S	0	0	0
			5611	1838	2771	465	530	7			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O$_4$S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total O S 5 4 1			0	0
3	A	1	Total O S 5 4 1			0	0
3	B	1	Total O S 5 4 1			0	0
3	B	1	Total O S 5 4 1			0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
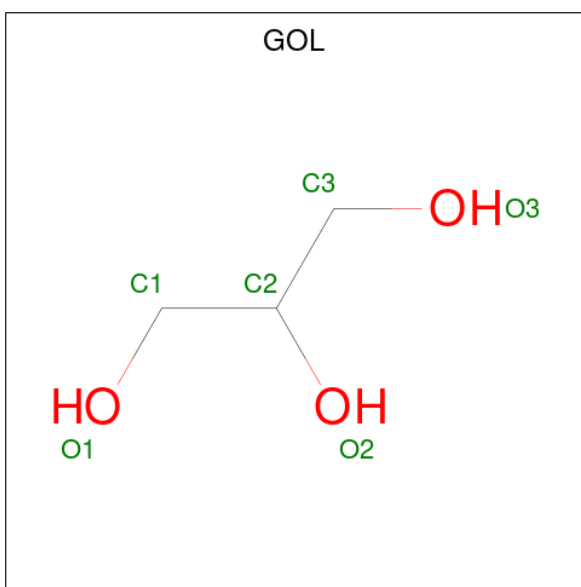
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total C H O 14 3 8 3				0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	188	Total O 188 188		0	0
5	B	130	Total O 130 130		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: Sso-CP2 metallo-carboxypetidase

Chain A:

• Molecule 1: Sso-CP2 metallo-carboxypetidase

Chain B:
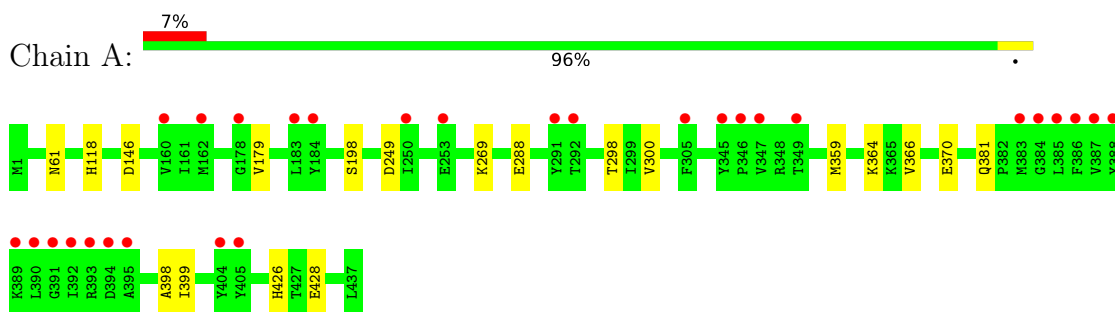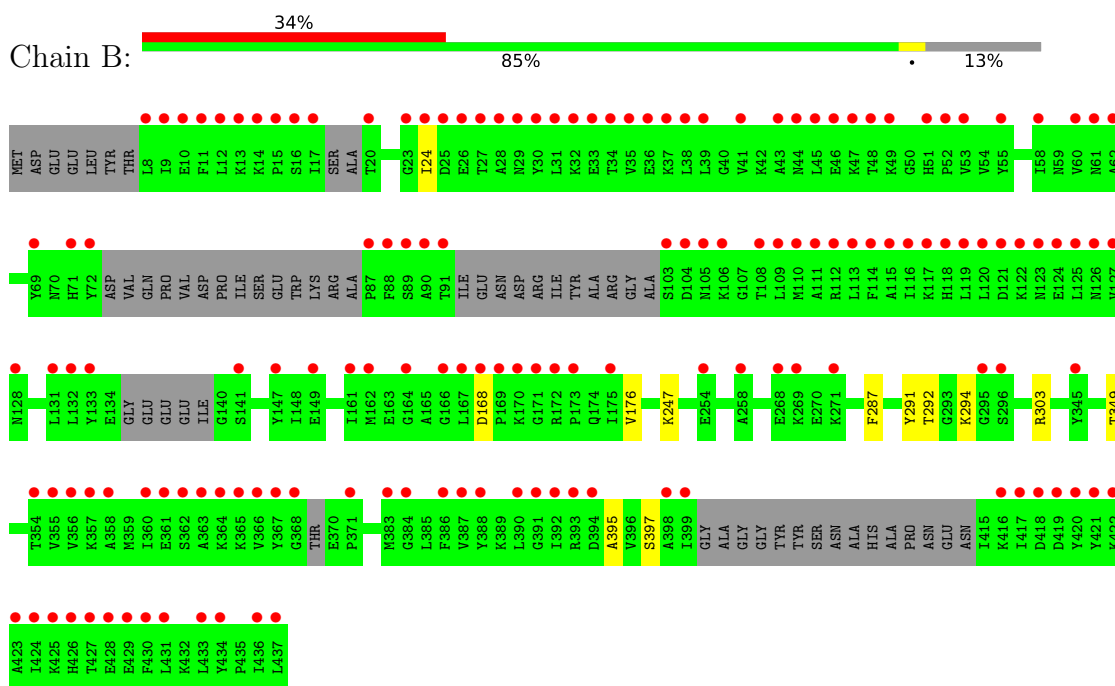
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	86.07Å 89.37Å 161.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.10 – 2.34 49.16 – 2.34	Depositor EDS
% Data completeness (in resolution range)	99.1 (46.10-2.34) 99.2 (49.16-2.34)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.00 (at 2.34Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.206 , 0.253 0.209 , 0.255	Depositor DCC
R_{free} test set	2654 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	46.7	Xtriage
Anisotropy	0.510	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 60.0	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	12814	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.37% 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: SO4, GOL, 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	0.25	0/3479	0.44	0/4706
1	B	0.23	0/2885	0.42	0/3909
All	All	0.24	0/6364	0.43	0/8615

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	3414	3431	3435	12	0
1	B	2840	2771	2761	7	0
2	A	1	0	0	0	0
3	A	15	0	0	0	0
3	B	10	0	0	0	0
4	B	6	8	8	0	0
5	A	188	0	0	2	0
5	B	130	0	0	3	0
All	All	6604	6210	6204	18	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.

All (18) 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:198:SER:OG	1:A:298:THR:OG1	2.09	0.69
1:A:61:ASN:ND2	5:A:705:HOH:O	2.32	0.63
1:B:291:TYR:OH	1:B:294:LYS:O	2.19	0.61
1:A:198:SER:HG	1:A:298:THR:HG1	1.49	0.60
1:A:364:LYS:NZ	1:A:370:GLU:OE1	2.38	0.56
1:A:249:ASP:OD1	1:A:269:LYS:NZ	2.40	0.55
1:B:303:ARG:NE	5:B:629:HOH:O	2.29	0.55
1:A:359:MET:HG2	1:A:398:ALA:HB3	1.93	0.51
1:A:146:ASP:OD1	5:A:677:HOH:O	2.20	0.46
1:A:198:SER:HA	1:A:300:VAL:HG23	1.97	0.46
1:B:24:ILE:N	5:B:627:HOH:O	2.46	0.44
1:A:179:VAL:HG11	1:A:381:GLN:HB3	2.00	0.43
1:B:247:LYS:NZ	5:B:673:HOH:O	2.45	0.42
1:A:288:GLU:HB2	1:B:292:THR:HG21	2.01	0.42
1:A:366:VAL:HG21	1:A:426:HIS:HA	2.02	0.41
1:A:118:HIS:ND1	1:A:428:GLU:OE2	2.48	0.41
1:B:176:VAL:HG13	1:B:397:SER:HB3	2.02	0.40
1:B:349:THR:HG21	1:B:395:ALA:H	1.85	0.40

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	435/437 (100%)	428 (98%)	6 (1%)	1 (0%)	47	55
1	B	368/437 (84%)	355 (96%)	13 (4%)	0	100	100
All	All	803/874 (92%)	783 (98%)	19 (2%)	1 (0%)	51	62

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	399	ILE

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	363/374 (97%)	363 (100%)	0	100	100
1	B	277/374 (74%)	275 (99%)	2 (1%)	84	90
All	All	640/748 (86%)	638 (100%)	2 (0%)	92	96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	168	ASP
1	B	287	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	70	ASN
1	B	105	ASN

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 ⓘ

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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	#$|Z| > 2$
3	SO4	A	503	-	4,4,4	0.13	0	6,6,6	0.05	0
3	SO4	A	502	-	4,4,4	0.19	0	6,6,6	0.59	0
3	SO4	B	501	-	4,4,4	0.12	0	6,6,6	0.07	0
3	SO4	A	504	-	4,4,4	0.13	0	6,6,6	0.09	0
4	GOL	B	503	-	5,5,5	0.36	0	5,5,5	0.30	0
3	SO4	B	502	-	4,4,4	0.15	0	6,6,6	0.07	0

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
4	GOL	B	503	-	-	<mark>4/4/4/4</mark>	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	503	GOL	O1-C1-C2-C3
4	B	503	GOL	C1-C2-C3-O3
4	B	503	GOL	O2-C2-C3-O3
4	B	503	GOL	O1-C1-C2-O2

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.

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(Å^2)	Q<0.9
1	A	437/437 (100%)	0.58	29 (6%)	18	26	31, 47, 80, 131	0
1	B	382/437 (87%)	2.15	150 (39%)	0	0	32, 90, 172, 241	0
All	All	819/874 (93%)	1.31	179 (21%)	0	1	31, 53, 151, 241	0

All (179) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	39	LEU	15.1
1	B	38	LEU	11.4
1	B	120	LEU	10.8
1	B	17	ILE	10.4
1	B	31	LEU	10.2
1	B	114	PHE	10.0
1	B	398	ALA	9.5
1	B	418	ASP	8.9
1	B	111	ALA	8.8
1	B	52	PRO	8.3
1	B	53	VAL	8.1
1	B	16	SER	8.1
1	B	421	TYR	7.9
1	B	417	ILE	7.8
1	B	9	ILE	7.8
1	B	115	ALA	7.6
1	B	11	PHE	7.2
1	B	364	LYS	6.8
1	B	430	PHE	6.7
1	B	113	LEU	6.7
1	B	48	THR	6.4
1	B	29	ASN	6.3
1	B	366	VAL	6.3
1	B	124	GLU	6.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	362	SER	6.2
1	B	45	LEU	6.1
1	B	125	LEU	6.1
1	B	360	ILE	6.0
1	B	399	ILE	5.9
1	B	43	ALA	5.9
1	B	167	LEU	5.9
1	B	126	ASN	5.8
1	B	30	TYR	5.7
1	B	426	HIS	5.6
1	B	169	PRO	5.5
1	B	10	GLU	5.5
1	B	87	PRO	5.5
1	A	405	TYR	5.4
1	B	387	VAL	5.4
1	B	41	VAL	5.3
1	B	27	THR	5.3
1	B	354	THR	5.3
1	B	371	PRO	5.3
1	B	51	HIS	5.2
1	B	35	VAL	5.2
1	B	71	HIS	5.2
1	B	431	LEU	5.2
1	B	131	LEU	5.1
1	B	26	GLU	5.1
1	B	388	TYR	5.0
1	B	173	PRO	5.0
1	A	404	TYR	4.9
1	B	425	LYS	4.9
1	A	388	TYR	4.9
1	B	24	ILE	4.9
1	B	429	GLU	4.8
1	B	88	PHE	4.8
1	B	12	LEU	4.8
1	B	422	LYS	4.7
1	B	72	TYR	4.7
1	B	13	LYS	4.7
1	B	109	LEU	4.7
1	B	437	LEU	4.6
1	B	44	ASN	4.6
1	B	28	ALA	4.6
1	B	34	THR	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	365	LYS	4.6
1	B	367	TYR	4.6
1	B	36	GLU	4.5
1	B	58	ILE	4.5
1	B	15	PRO	4.4
1	B	25	ASP	4.4
1	B	123	ASN	4.4
1	B	108	THR	4.4
1	B	424	ILE	4.3
1	B	119	LEU	4.3
1	B	393	ARG	4.3
1	B	14	LYS	4.2
1	B	132	LEU	4.2
1	B	37	LYS	4.1
1	B	436	ILE	4.1
1	B	103	SER	3.9
1	A	392	ILE	3.9
1	B	61	ASN	3.9
1	B	166	GLY	3.9
1	B	368	GLY	3.9
1	B	121	ASP	3.8
1	B	127	VAL	3.8
1	B	32	LYS	3.8
1	B	112	ARG	3.8
1	B	8	LEU	3.7
1	B	110	MET	3.6
1	B	116	ILE	3.5
1	A	387	VAL	3.5
1	B	427	THR	3.5
1	B	55	TYR	3.5
1	B	122	LYS	3.5
1	B	164	GLY	3.4
1	B	141	SER	3.4
1	B	394	ASP	3.4
1	B	392	ILE	3.4
1	A	345	TYR	3.4
1	B	420	TYR	3.4
1	B	91	THR	3.4
1	B	105	ASN	3.4
1	A	395	ALA	3.4
1	B	433	LEU	3.3
1	B	33	GLU	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	356	VAL	3.3
1	B	357	LYS	3.3
1	A	386	PHE	3.3
1	B	296	SER	3.3
1	B	62	ALA	3.2
1	B	391	GLY	3.2
1	B	133	TYR	3.2
1	B	423	ALA	3.2
1	A	305	PHE	3.2
1	B	358	ALA	3.2
1	B	254	GLU	3.1
1	B	168	ASP	3.1
1	B	361	GLU	3.1
1	B	90	ALA	3.1
1	B	69	TYR	3.0
1	B	106	LYS	3.0
1	A	391	GLY	3.0
1	A	349	THR	2.9
1	B	118	HIS	2.9
1	B	171	GLY	2.9
1	B	172	ARG	2.8
1	B	161	ILE	2.8
1	B	390	LEU	2.8
1	A	390	LEU	2.8
1	B	23	GLY	2.8
1	B	434	TYR	2.7
1	A	347	VAL	2.7
1	B	104	ASP	2.7
1	B	46	GLU	2.7
1	B	175	ILE	2.7
1	B	386	PHE	2.7
1	A	383	MET	2.7
1	B	128	ASN	2.6
1	B	47	LYS	2.6
1	B	170	LYS	2.6
1	A	393	ARG	2.6
1	B	258	ALA	2.6
1	A	160	VAL	2.6
1	A	385	LEU	2.6
1	B	20	THR	2.5
1	B	89	SER	2.5
1	B	345	TYR	2.5

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	149	GLU	2.5
1	A	291	TYR	2.4
1	A	384	GLY	2.4
1	A	184	TYR	2.4
1	A	178	GLY	2.4
1	B	363	ALA	2.4
1	B	60	VAL	2.4
1	B	355	VAL	2.4
1	A	250	ILE	2.3
1	B	428	GLU	2.3
1	B	117	LYS	2.3
1	A	162	MET	2.3
1	A	346	PRO	2.2
1	B	419	ASP	2.2
1	B	384	GLY	2.2
1	B	295	GLY	2.2
1	B	269	LYS	2.2
1	A	292	THR	2.2
1	A	183	LEU	2.2
1	B	268	GLU	2.2
1	B	416	LYS	2.1
1	A	389	LYS	2.1
1	B	271	LYS	2.1
1	B	383	MET	2.0
1	B	49	LYS	2.0
1	B	147	TYR	2.0
1	A	253	GLU	2.0
1	B	162	MET	2.0
1	A	394	ASP	2.0

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 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($\mathring{A}^2$)	Q<0.9
3	SO4	A	502	5/5	0.71	0.34	74,99,111,113	0
2	ZN	A	501	1/1	0.85	0.14	31,31,31,31	1
3	SO4	B	502	5/5	0.85	0.23	75,84,94,94	0
4	GOL	B	503	6/6	0.86	0.20	36,62,84,90	0
3	SO4	A	504	5/5	0.93	0.14	58,77,89,91	0
3	SO4	A	503	5/5	0.95	0.20	64,78,96,98	0
3	SO4	B	501	5/5	0.95	0.15	51,63,76,78	0

6.5 Other polymers ⓘ

There are no such residues in this entry.