Full wwPDB X-ray Structure Validation Report ⓘ

May 11, 2021 – 04:02 pm BST

PDB ID	:	6Z0Y
Title	:	HtrA1 inactive protease domain S328A with CARASIL mutations D174R R274Q
Authors	:	Vetter, I.R.; Stege, P.; Ingendahl, L.; Ehrmann, M.
Deposited on	:	2020-05-11
Resolution	:	2.20 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 ①) were used in the production of this report:

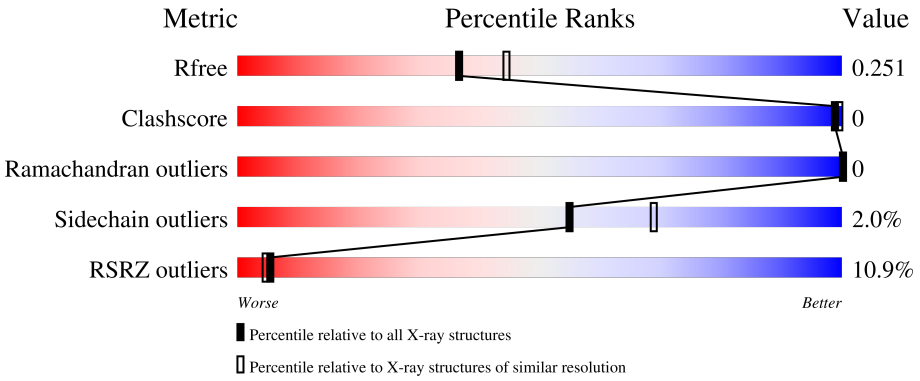
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.18
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.18

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	236	11% 84% · 15%
1	B	236	5% 79% · 17%
1	C	236	12% 81% · 17%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4583 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 Serine protease HTRA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	0	0
			1533	979	265	288	1			
1	B	197	Total	C	N	O	S	0	0	0
			1512	964	261	286	1			
1	C	197	Total	C	N	O	S	0	0	0
			1510	965	261	283	1			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	140	MET	-	initiating methionine	UNP Q92743
A	141	GLY	-	expression tag	UNP Q92743
A	142	SER	-	expression tag	UNP Q92743
A	143	SER	-	expression tag	UNP Q92743
A	144	HIS	-	expression tag	UNP Q92743
A	145	HIS	-	expression tag	UNP Q92743
A	146	HIS	-	expression tag	UNP Q92743
A	147	HIS	-	expression tag	UNP Q92743
A	148	HIS	-	expression tag	UNP Q92743
A	149	HIS	-	expression tag	UNP Q92743
A	150	SER	-	expression tag	UNP Q92743
A	151	SER	-	expression tag	UNP Q92743
A	152	GLY	-	expression tag	UNP Q92743
A	153	LEU	-	expression tag	UNP Q92743
A	154	VAL	-	expression tag	UNP Q92743
A	155	PRO	-	expression tag	UNP Q92743
A	156	ARG	-	expression tag	UNP Q92743
A	157	GLY	-	expression tag	UNP Q92743
A	158	SER	-	expression tag	UNP Q92743
A	159	HIS	-	expression tag	UNP Q92743
A	160	MET	-	expression tag	UNP Q92743
A	174	ARG	ASP	engineered mutation	UNP Q92743
A	274	GLN	ARG	engineered mutation	UNP Q92743

Continued on next page...

Continued from previous page...

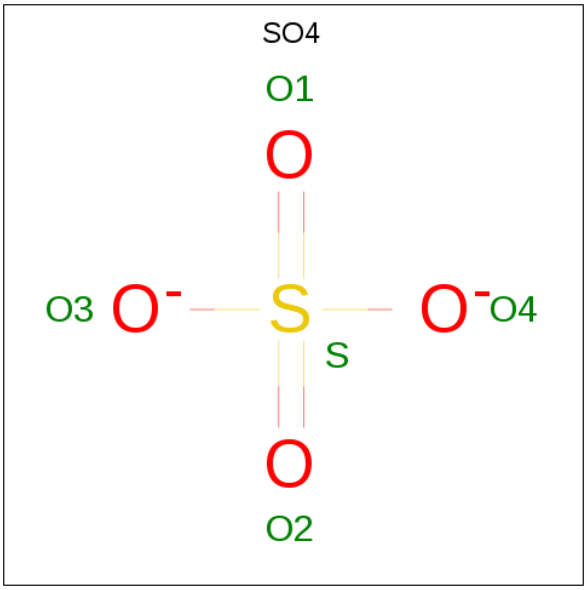
Chain	Residue	Modelled	Actual	Comment	Reference
A	328	ALA	SER	engineered mutation	UNP Q92743
B	140	MET	-	initiating methionine	UNP Q92743
B	141	GLY	-	expression tag	UNP Q92743
B	142	SER	-	expression tag	UNP Q92743
B	143	SER	-	expression tag	UNP Q92743
B	144	HIS	-	expression tag	UNP Q92743
B	145	HIS	-	expression tag	UNP Q92743
B	146	HIS	-	expression tag	UNP Q92743
B	147	HIS	-	expression tag	UNP Q92743
B	148	HIS	-	expression tag	UNP Q92743
B	149	HIS	-	expression tag	UNP Q92743
B	150	SER	-	expression tag	UNP Q92743
B	151	SER	-	expression tag	UNP Q92743
B	152	GLY	-	expression tag	UNP Q92743
B	153	LEU	-	expression tag	UNP Q92743
B	154	VAL	-	expression tag	UNP Q92743
B	155	PRO	-	expression tag	UNP Q92743
B	156	ARG	-	expression tag	UNP Q92743
B	157	GLY	-	expression tag	UNP Q92743
B	158	SER	-	expression tag	UNP Q92743
B	159	HIS	-	expression tag	UNP Q92743
B	160	MET	-	expression tag	UNP Q92743
B	174	ARG	ASP	engineered mutation	UNP Q92743
B	274	GLN	ARG	engineered mutation	UNP Q92743
B	328	ALA	SER	engineered mutation	UNP Q92743
C	140	MET	-	initiating methionine	UNP Q92743
C	141	GLY	-	expression tag	UNP Q92743
C	142	SER	-	expression tag	UNP Q92743
C	143	SER	-	expression tag	UNP Q92743
C	144	HIS	-	expression tag	UNP Q92743
C	145	HIS	-	expression tag	UNP Q92743
C	146	HIS	-	expression tag	UNP Q92743
C	147	HIS	-	expression tag	UNP Q92743
C	148	HIS	-	expression tag	UNP Q92743
C	149	HIS	-	expression tag	UNP Q92743
C	150	SER	-	expression tag	UNP Q92743
C	151	SER	-	expression tag	UNP Q92743
C	152	GLY	-	expression tag	UNP Q92743
C	153	LEU	-	expression tag	UNP Q92743
C	154	VAL	-	expression tag	UNP Q92743
C	155	PRO	-	expression tag	UNP Q92743
C	156	ARG	-	expression tag	UNP Q92743

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	157	GLY	-	expression tag	UNP Q92743
C	158	SER	-	expression tag	UNP Q92743
C	159	HIS	-	expression tag	UNP Q92743
C	160	MET	-	expression tag	UNP Q92743
C	174	ARG	ASP	engineered mutation	UNP Q92743
C	274	GLN	ARG	engineered mutation	UNP Q92743
C	328	ALA	SER	engineered mutation	UNP Q92743

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is water.

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

Continued on next page...

Continued from previous page...

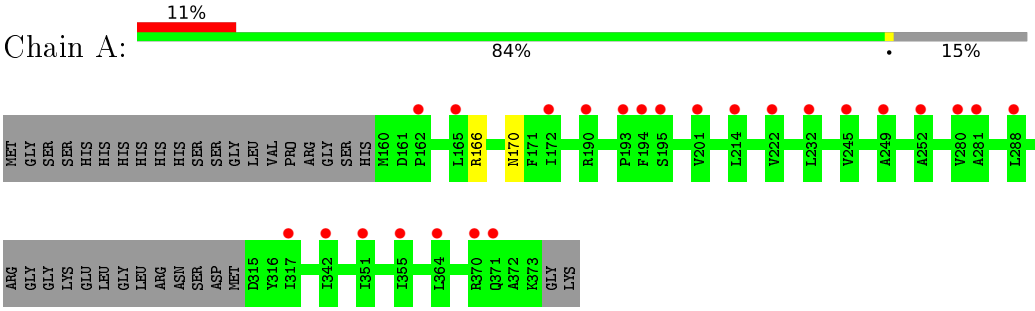
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	4	Total O 4 4	0	0

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.

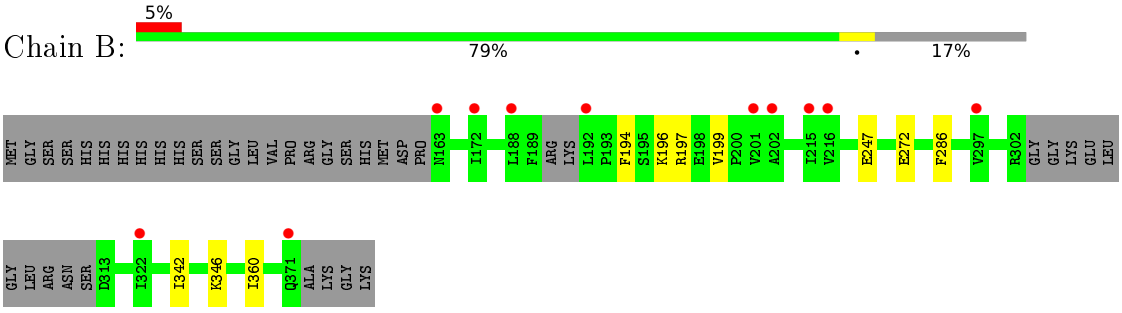
- Molecule 1: Serine protease HTRA1

Chain A:



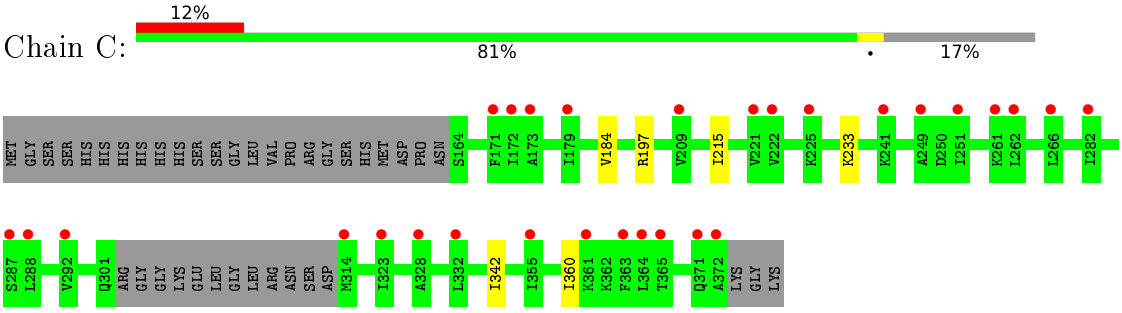
- Molecule 1: Serine protease HTRA1

Chain B:



- Molecule 1: Serine protease HTRA1

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α, β, γ	101.42Å 101.42Å 144.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.87 – 2.20 47.87 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.87-2.20) 100.0 (47.87-2.20)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.247 , 0.253 0.246 , 0.251	Depositor DCC
R_{free} test set	2141 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	67.0	Xtriage
Anisotropy	0.270	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 31.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.469 for h,-h-k,-l	Xtriage
Reported twinning fraction	0.517 for H, K, L 0.483 for -K, -H, -L	Depositor
Outliers	0 of 42814 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4583	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 3.29% 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

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	0.26	0/1558	0.41	0/2109
1	B	0.26	0/1535	0.42	0/2077
1	C	0.26	0/1534	0.41	0/2076
All	All	0.26	0/4627	0.41	0/6262

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	1533	0	1585	1	0
1	B	1512	0	1554	1	0
1	C	1510	0	1563	2	0
2	A	5	0	0	0	0
2	B	15	0	0	0	0
3	A	3	0	0	0	0
3	B	1	0	0	0	0
3	C	4	0	0	0	1
All	All	4583	0	4702	4	1

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 (4) 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:B:342:ILE:HG13	1:B:360:ILE:HD11	1.93	0.51
1:C:342:ILE:HG13	1:C:360:ILE:HD11	1.96	0.47
1:A:166:ARG:O	1:A:170:ASN:ND2	2.48	0.47
1:C:184:VAL:HG21	1:C:215:ILE:HD13	2.00	0.42

All (1) 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)
3:C:404:HOH:O	3:C:404:HOH:O[2_775]	1.74	0.46

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	196/236 (83%)	192 (98%)	4 (2%)	0	100	100
1	B	191/236 (81%)	183 (96%)	8 (4%)	0	100	100
1	C	193/236 (82%)	180 (93%)	13 (7%)	0	100	100
All	All	580/708 (82%)	555 (96%)	25 (4%)	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	170/199 (85%)	170 (100%)	0	100	100
1	B	168/199 (84%)	160 (95%)	8 (5%)	25	32
1	C	167/199 (84%)	165 (99%)	2 (1%)	71	83
All	All	505/597 (85%)	495 (98%)	10 (2%)	55	69

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	194	PHE
1	B	196	LYS
1	B	197	ARG
1	B	199	VAL
1	B	247	GLU
1	B	272	GLU
1	B	286	PHE
1	B	346	LYS
1	C	197	ARG
1	C	233	LYS

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

4 ligands are 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	SO4	B	401	-	4,4,4	0.32	0	6,6,6	0.05	0
2	SO4	B	402	-	4,4,4	0.30	0	6,6,6	0.10	0
2	SO4	A	401	-	4,4,4	0.33	0	6,6,6	0.05	0
2	SO4	B	403	-	4,4,4	0.32	0	6,6,6	0.05	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.

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	200/236 (84%)	1.01	25 (12%) 3 3	60, 69, 82, 88	0
1	B	197/236 (83%)	0.85	11 (5%) 24 23	62, 71, 82, 91	0
1	C	197/236 (83%)	1.05	29 (14%) 2 2	72, 79, 90, 93	0
All	All	594/708 (83%)	0.97	65 (10%) 5 5	60, 74, 86, 93	0

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	194	PHE	5.4
1	A	193	PRO	5.2
1	B	163	ASN	3.9
1	A	249	ALA	3.8
1	A	371	GLN	3.7
1	C	171	PHE	3.5
1	C	221	VAL	3.5
1	B	201	VAL	3.4
1	A	245	VAL	3.2
1	C	288	LEU	3.1
1	A	162	PRO	3.1
1	C	266	LEU	2.8
1	A	222	VAL	2.8
1	A	281	ALA	2.8
1	C	372	ALA	2.7
1	C	355	ILE	2.7
1	A	165	LEU	2.7
1	B	371	GLN	2.6
1	B	192	LEU	2.6
1	C	222	VAL	2.6
1	A	288	LEU	2.5
1	C	314	MET	2.5
1	A	293	THR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	195	SER	2.5
1	A	364	LEU	2.4
1	C	282	ILE	2.4
1	A	232	LEU	2.4
1	C	179	ILE	2.4
1	C	371	GLN	2.4
1	C	249	ALA	2.3
1	B	188	LEU	2.3
1	C	287	SER	2.3
1	A	201	VAL	2.3
1	C	364	LEU	2.3
1	C	365	THR	2.3
1	A	190	ARG	2.3
1	B	216	VAL	2.3
1	A	317	ILE	2.3
1	A	280	VAL	2.2
1	C	323	ILE	2.2
1	C	225	LYS	2.2
1	A	342	ILE	2.2
1	C	262	LEU	2.2
1	A	252	ALA	2.2
1	C	261	LYS	2.2
1	B	297	VAL	2.1
1	C	292	VAL	2.1
1	A	355	ILE	2.1
1	C	251	ILE	2.1
1	C	363	PHE	2.1
1	B	322	ILE	2.1
1	C	241	LYS	2.1
1	C	328	ALA	2.1
1	A	172	ILE	2.1
1	B	172	ILE	2.1
1	B	215	ILE	2.1
1	A	214	LEU	2.1
1	B	202	ALA	2.1
1	C	209	VAL	2.1
1	C	332	LEU	2.1
1	C	173	ALA	2.1
1	C	361	LYS	2.1
1	C	172	ILE	2.0
1	A	370	ARG	2.0
1	A	351	ILE	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 (i)

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, 95^{th} 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	SO4	B	401	5/5	0.90	0.17	70,70,70,70	0
2	SO4	B	402	5/5	0.92	0.15	72,72,72,72	0
2	SO4	A	401	5/5	0.95	0.10	72,72,72,72	0
2	SO4	B	403	5/5	0.95	0.12	60,60,60,60	0

6.5 Other polymers (i)

There are no such residues in this entry.