



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 16, 2023 – 10:39 AM EST

PDB ID : 3F1S  
Title : Crystal structure of Protein Z complexed with protein Z-dependent inhibitor  
Authors : Zhou, A.  
Deposited on : 2008-10-28  
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](mailto: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)  
Xtrriage (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

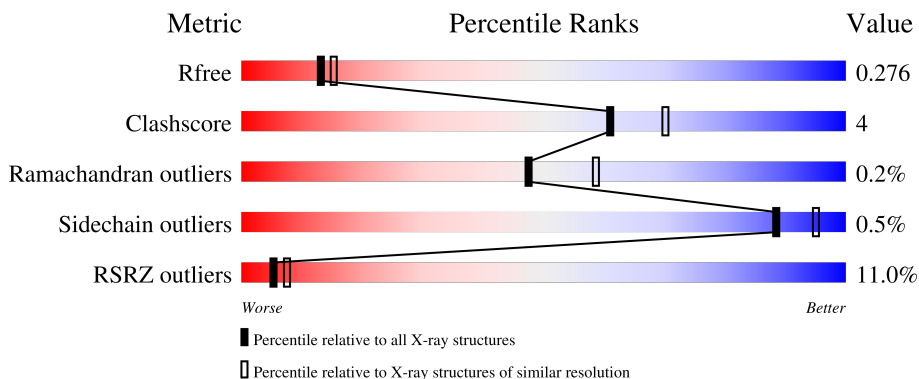
# 1 Overall quality at a glance

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	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  $\geq 3$ , 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  $\leq 5\%$ . 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	385	
2	B	283	

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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	A	424	-	-	-	X
6	NAG	B	1	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5241 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 Protein Z-dependent protease inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	385	3101	2002	523	555	21	0	0	0

- Molecule 2 is a protein called Vitamin K-dependent protein Z.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	259	2026	1266	364	376	20	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	361	ALA	-	expression tag	UNP P22891
B	362	HIS	-	expression tag	UNP P22891
B	363	HIS	-	expression tag	UNP P22891
B	364	HIS	-	expression tag	UNP P22891
B	365	HIS	-	expression tag	UNP P22891
B	366	HIS	-	expression tag	UNP P22891
B	367	HIS	-	expression tag	UNP P22891

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

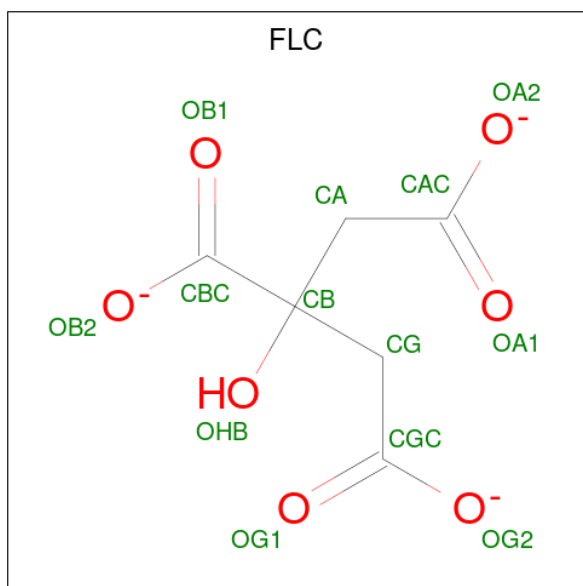
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Cl	0	0
			1	1		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

- Molecule 5 is CITRATE ANION (three-letter code: FLC) (formula:  $C_6H_5O_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			13	6	7		

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	B	1	14	8	1	5	0	0

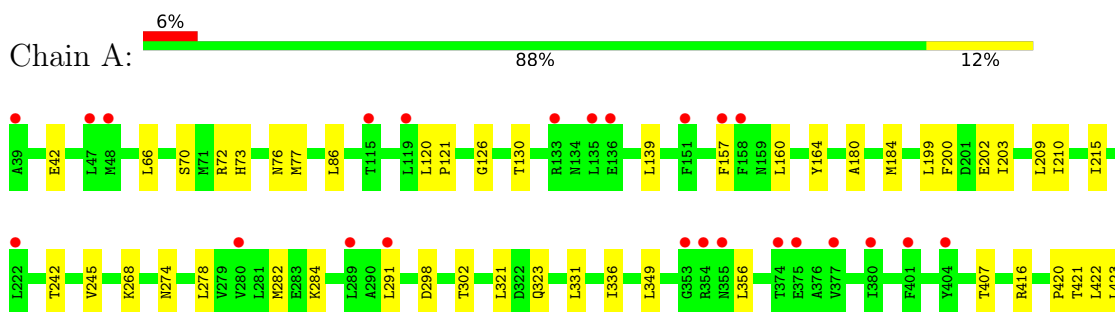
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	60	60	60	0	0
7	B	22	22	22	0	0

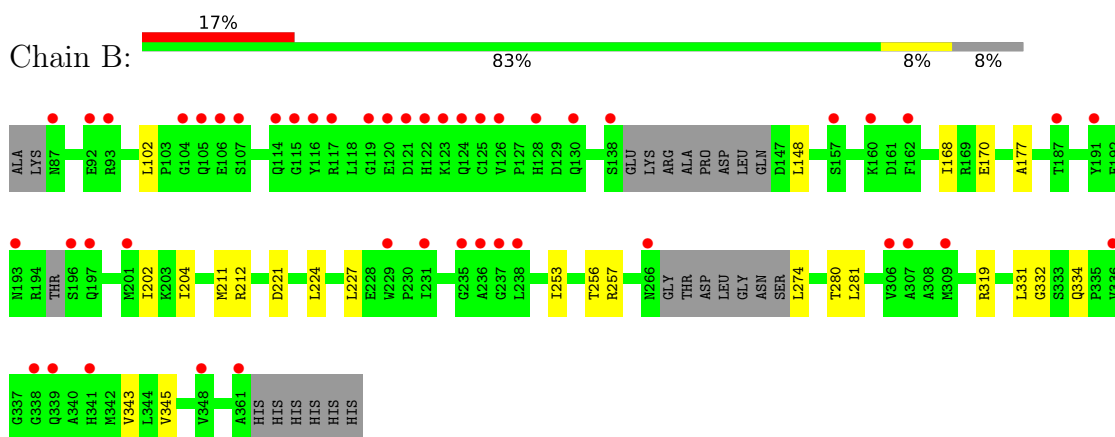
### 3 Residue-property plots [i](#)

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: Protein Z-dependent protease inhibitor



- Molecule 2: Vitamin K-dependent protein Z



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.56Å 96.51Å 117.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.63 – 2.30 44.62 – 2.30	Depositor EDS
% Data completeness (in resolution range)	92.4 (44.63-2.30) 92.4 (44.62-2.30)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.50 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.5.0042	Depositor
R, $R_{free}$	0.224 , 0.271 0.229 , 0.276	Depositor DCC
$R_{free}$ test set	1640 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.7	Xtrriage
Anisotropy	0.264	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 44.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5241	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.58% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, NAG, CL, FLC

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	0.34	0/3171	0.52	0/4272
2	B	0.32	0/2071	0.50	0/2809
All	All	0.33	0/5242	0.51	0/7081

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	3101	0	3146	29	0
2	B	2026	0	1958	17	0
3	A	1	0	0	1	0
4	A	4	0	6	0	0
5	A	13	0	5	0	0
6	B	14	0	13	0	0
7	A	60	0	0	3	0
7	B	22	0	0	0	0
All	All	5241	0	5128	46	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 4.

All (46) 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)
2:B:148:LEU:HD22	2:B:274:LEU:HD21	1.67	0.77
2:B:256:THR:O	2:B:319:ARG:NH1	2.22	0.72
3:A:1:CL:CL	7:A:441:HOH:O	2.44	0.72
1:A:298:ASP:O	1:A:302:THR:HG23	1.99	0.63
1:A:199:LEU:HD23	1:A:200:PHE:CE2	2.37	0.60
2:B:204:ILE:HD13	2:B:227:LEU:HD13	1.85	0.57
1:A:157:PHE:CG	1:A:349:LEU:HD13	2.40	0.57
2:B:331:LEU:HD23	2:B:334:GLN:HE21	1.70	0.57
1:A:139:LEU:HD11	1:A:215:ILE:HD11	1.86	0.56
1:A:120:LEU:HB3	1:A:121:PRO:HD3	1.88	0.55
2:B:253:ILE:HG22	2:B:281:LEU:HD22	1.92	0.52
1:A:284:LYS:HB3	2:B:102:LEU:HD11	1.92	0.52
1:A:86:LEU:C	1:A:86:LEU:HD23	2.30	0.51
1:A:73:HIS:NE2	7:A:441:HOH:O	2.34	0.51
1:A:274:ASN:OD1	1:A:407:THR:HG23	2.11	0.51
2:B:204:ILE:CD1	2:B:227:LEU:HD13	2.41	0.50
2:B:202:ILE:CG2	2:B:227:LEU:HD12	2.42	0.50
1:A:76:ASN:HB3	1:A:321:LEU:HD22	1.95	0.49
1:A:282:MET:HE1	1:A:291:LEU:HB2	1.95	0.49
1:A:66:LEU:HD13	1:A:77:MET:HE1	1.95	0.48
2:B:257:ARG:HG3	2:B:280:THR:HG22	1.94	0.48
1:A:323:GLN:NE2	1:A:423:LEU:O	2.43	0.47
1:A:160:LEU:HD22	1:A:164:TYR:HE2	1.78	0.47
2:B:343:VAL:HG12	2:B:345:VAL:HG13	1.96	0.47
2:B:331:LEU:CD2	2:B:334:GLN:HE21	2.26	0.47
1:A:282:MET:HE1	1:A:291:LEU:HD22	1.96	0.47
1:A:157:PHE:CD2	1:A:349:LEU:HD13	2.50	0.46
1:A:202:GLU:O	1:A:203:ILE:HD13	2.16	0.46
2:B:168:ILE:HD11	2:B:224:LEU:HD21	1.97	0.46
1:A:157:PHE:CD1	1:A:349:LEU:HD13	2.52	0.45
1:A:76:ASN:ND2	1:A:422:LEU:HD12	2.32	0.45
1:A:126:GLY:O	1:A:130:THR:HG23	2.17	0.44
1:A:180:ALA:O	1:A:184:MET:HE2	2.18	0.44
1:A:245:VAL:HG12	1:A:421:THR:CG2	2.47	0.44
1:A:282:MET:HE3	1:A:282:MET:HB2	1.89	0.44
1:A:331:LEU:HD22	1:A:336:ILE:HD12	2.00	0.43
2:B:211:MET:HG3	2:B:212:ARG:HG2	2.01	0.42
1:A:70:SER:OG	1:A:416:ARG:NH2	2.52	0.42
1:A:268:LYS:HG2	1:A:278:LEU:CD2	2.49	0.42
2:B:177:ALA:HB3	2:B:221:ASP:OD1	2.20	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:256:THR:HG22	2:B:257:ARG:O	2.21	0.41
1:A:72:ARG:NH1	7:A:441:HOH:O	2.52	0.41
1:A:209:LEU:HD23	1:A:210:ILE:N	2.35	0.41
1:A:245:VAL:HG11	1:A:420:PRO:HG2	2.03	0.41
2:B:331:LEU:HD23	2:B:332:GLY:N	2.36	0.40
2:B:168:ILE:HD11	2:B:224:LEU:CD2	2.52	0.40

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	383/385 (100%)	373 (97%)	10 (3%)	0	100	100
2	B	251/283 (89%)	243 (97%)	7 (3%)	1 (0%)	34	42
All	All	634/668 (95%)	616 (97%)	17 (3%)	1 (0%)	47	58

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	170	GLU

### 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	340/343 (99%)	337 (99%)	3 (1%)	78	89
2	B	226/246 (92%)	226 (100%)	0	100	100
All	All	566/589 (96%)	563 (100%)	3 (0%)	88	95

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	42	GLU
1	A	242	THR
1	A	356	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 1 is 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	EDO	A	424	-	3,3,3	0.51	0	2,2,2	0.14	0
6	NAG	B	1	2	14,14,15	0.57	0	17,19,21	0.90	0
5	FLC	A	425	-	12,12,12	0.97	0	17,17,17	1.21	1 (5%)

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	EDO	A	424	-	-	0/1/1/1	-
6	NAG	B	1	2	-	0/6/23/26	0/1/1/1
5	FLC	A	425	-	-	8/16/16/16	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	425	FLC	OB2-CBC-CB	2.82	117.94	113.05

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	425	FLC	CG-CB-CBC-OB1
5	A	425	FLC	CG-CB-CBC-OB2
5	A	425	FLC	OHB-CB-CBC-OB1
5	A	425	FLC	OHB-CB-CBC-OB2
5	A	425	FLC	CAC-CA-CB-CBC
5	A	425	FLC	CAC-CA-CB-CG
5	A	425	FLC	CAC-CA-CB-OHB
5	A	425	FLC	OHB-CB-CG-CGC

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	385/385 (100%)	0.52	24 (6%) 20 26	25, 32, 47, 53	4 (1%)
2	B	259/283 (91%)	0.99	47 (18%) 1 1	31, 37, 48, 52	1 (0%)
All	All	644/668 (96%)	0.71	71 (11%) 5 7	25, 35, 47, 53	5 (0%)

All (71) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	197	GLN	6.1
2	B	120	GLU	5.7
2	B	105	GLN	5.2
2	B	115	GLY	5.1
2	B	119	GLY	4.7
1	A	135	LEU	4.5
2	B	229	TRP	4.5
2	B	126	VAL	4.5
2	B	125	CYS	4.4
2	B	339	GLN	4.4
2	B	117	ARG	4.3
2	B	236	ALA	4.2
2	B	138	SER	4.1
1	A	353	GLY	4.1
1	A	115	THR	4.1
2	B	121	ASP	4.0
2	B	106	GLU	3.9
2	B	237	GLY	3.8
2	B	122	HIS	3.8
2	B	114	GLN	3.7
2	B	235	GLY	3.7
1	A	354	ARG	3.6
1	A	380	ILE	3.6
1	A	136	GLU	3.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	B	266	ASN	3.5
2	B	307	ALA	3.4
2	B	93	ARG	3.4
1	A	48	MET	3.4
2	B	124	GLN	3.3
1	A	157	PHE	3.2
2	B	196	SER	3.1
2	B	123	LYS	3.0
1	A	39	ALA	3.0
2	B	130	GLN	3.0
1	A	119	LEU	3.0
1	A	151	PHE	2.9
1	A	355	ASN	2.9
2	B	306	VAL	2.9
2	B	191	TYR	2.8
2	B	107	SER	2.8
2	B	336	VAL	2.8
1	A	374	THR	2.7
2	B	348	VAL	2.7
2	B	338	GLY	2.7
1	A	280	VAL	2.6
2	B	361	ALA	2.5
2	B	341	HIS	2.5
1	A	404	TYR	2.5
2	B	201	MET	2.5
2	B	92	GLU	2.5
2	B	104	GLY	2.5
2	B	116	TYR	2.4
2	B	128	HIS	2.4
1	A	289	LEU	2.4
1	A	401	PHE	2.4
2	B	193	ASN	2.4
1	A	291	LEU	2.4
2	B	238	LEU	2.3
1	A	133	ARG	2.3
1	A	158	PHE	2.2
2	B	187	THR	2.2
2	B	87	ASN	2.2
1	A	47	LEU	2.1
2	B	162	PHE	2.1
2	B	309	MET	2.1
1	A	375	GLU	2.0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	160	LYS	2.0
1	A	377	VAL	2.0
1	A	222	LEU	2.0
2	B	231	ILE	2.0
2	B	157	SER	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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
5	FLC	A	425	13/13	0.59	0.32	83,84,85,86	0
4	EDO	A	424	4/4	0.63	0.47	70,70,70,70	0
6	NAG	B	1	14/15	0.72	0.40	61,65,67,68	0
3	CL	A	1	1/1	0.95	0.08	68,68,68,68	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.