



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 1, 2018 – 11:47 AM EST

PDB ID : 5UFW  
Title : Estrogen Receptor Alpha Ligand Binding Domain in Complex with OP1154  
Authors : Fanning, S.W.; Hodges-Gallagher, L.; Myles, D.C.; Sun, R.; Fowler, C.E.;  
Green, B.D.; Harmon, C.L.; Greene, G.L.; Kushner, P.J.  
Deposited on : 2017-01-06  
Resolution : 1.58 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

<http://wwpdb.org/validation/2016/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  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030736  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

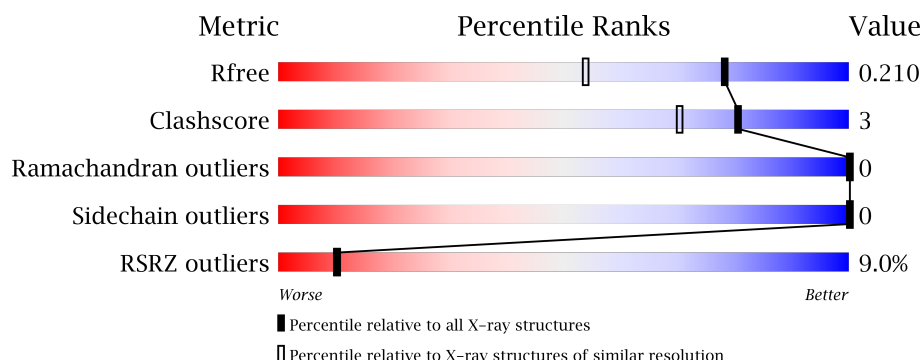
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.58 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))
$R_{free}$	100719	4211 (1.60-1.56)
Clashscore	112137	4539 (1.60-1.56)
Ramachandran outliers	110173	4423 (1.60-1.56)
Sidechain outliers	110143	4420 (1.60-1.56)
RSRZ outliers	101464	4232 (1.60-1.56)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	249	
1	B	249	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4490 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.

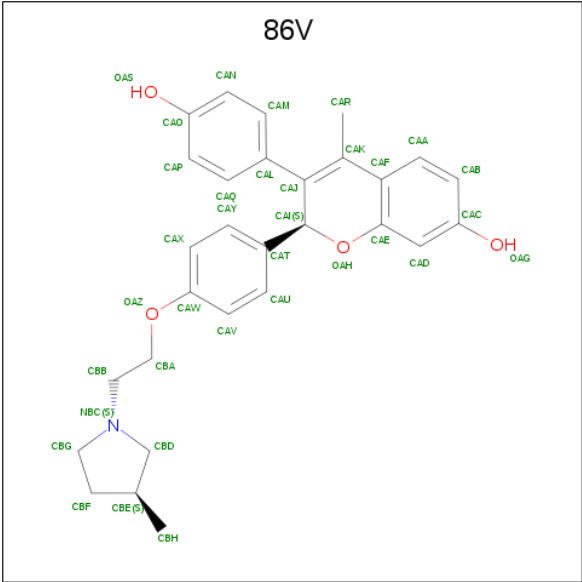
- Molecule 1 is a protein called Estrogen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	237	Total	C	N	O	S	0	10	0
			1946	1240	331	359	16			
1	B	229	Total	C	N	O	S	0	5	0
			1851	1186	315	334	16			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	381	SER	CYS	engineered mutation	UNP P03372
A	417	SER	CYS	engineered mutation	UNP P03372
A	530	SER	CYS	engineered mutation	UNP P03372
A	536	SER	LEU	engineered mutation	UNP P03372
B	381	SER	CYS	engineered mutation	UNP P03372
B	417	SER	CYS	engineered mutation	UNP P03372
B	530	SER	CYS	engineered mutation	UNP P03372
B	536	SER	LEU	engineered mutation	UNP P03372

- Molecule 2 is (2S)-3-(4-hydroxyphenyl)-4-methyl-2-(4-{2-[(3S)-3-methylpyrrolidin-1-yl]ethoxy}phenyl)-2H-1-benzopyran-7-ol (three-letter code: 86V) (formula: C<sub>29</sub>H<sub>31</sub>NO<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			34	29	1	4		
2	B	1	Total	C	N	O	0	0
			34	29	1	4		

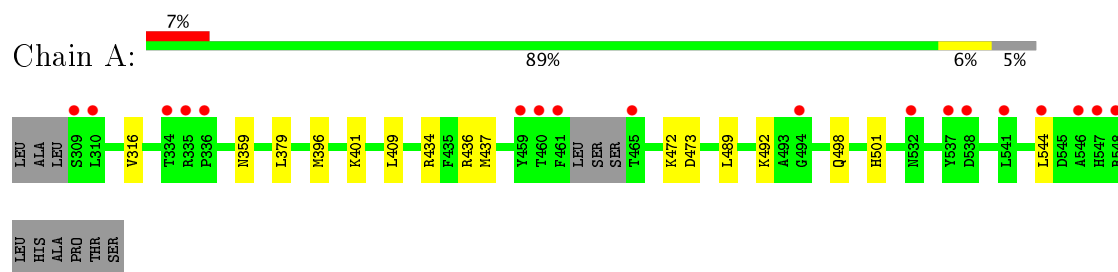
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	342	Total	O	0	2
			344	344		
3	B	281	Total	O	0	0
			281	281		

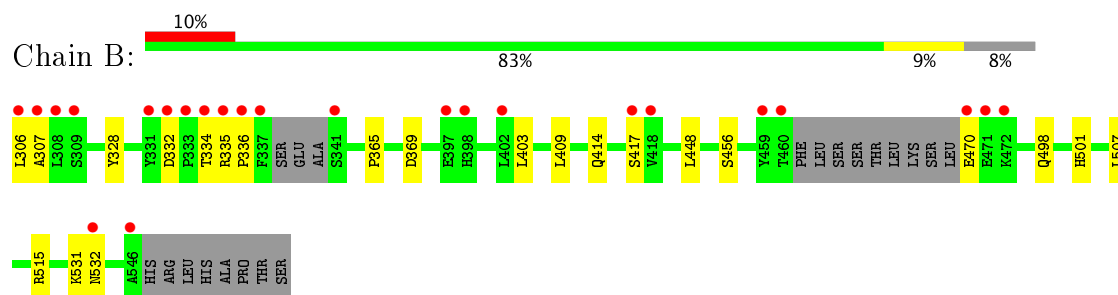
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Estrogen receptor



#### • Molecule 1: Estrogen receptor



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.57Å 57.88Å 87.70Å 90.00° 102.91° 90.00°	Depositor
Resolution (Å)	19.95 – 1.58 19.95 – 1.58	Depositor EDS
% Data completeness (in resolution range)	97.2 (19.95-1.58) 97.2 (19.95-1.58)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.99 (at 1.58Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.180 , 0.210 0.181 , 0.210	Depositor DCC
$R_{free}$ test set	3368 reflections (5.07%)	DCC
Wilson B-factor (Å <sup>2</sup> )	14.3	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 55.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4490	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.99% 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: 86V

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	0.33	0/1992	0.46	0/2693
1	B	0.38	0/1892	0.47	0/2559
All	All	0.36	0/3884	0.46	0/5252

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	1946	0	1956	10	1
1	B	1851	0	1869	14	1
2	A	34	0	0	0	0
2	B	34	0	0	0	0
3	A	344	0	0	3	2
3	B	281	0	0	3	1
All	All	4490	0	3825	24	3

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 3.

All (24) 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:335:ARG:HB3	1:B:336:PRO:HD2	1.49	0.93
1:B:470:GLU:N	3:B:703:HOH:O	2.27	0.67
1:A:434:ARG:NH1	1:A:437:MET:SD	2.67	0.66
1:B:306:LEU:N	3:B:706:HOH:O	2.34	0.59
1:B:403[A]:LEU:HD23	1:B:409:LEU:HD13	1.85	0.58
1:B:335:ARG:HB3	1:B:336:PRO:CD	2.31	0.57
1:B:332:ASP:OD2	1:B:334:THR:OG1	2.22	0.57
1:B:369:ASP:OD1	3:B:701:HOH:O	2.18	0.53
1:A:434:ARG:HA	1:A:437:MET:HE3	1.92	0.52
1:A:396:MET:O	1:A:436:ARG:HD3	2.12	0.50
1:B:456:SER:HA	1:B:515:ARG:NH2	2.26	0.50
1:B:332:ASP:OD2	1:B:335:ARG:HG3	2.12	0.49
1:A:492:LYS:HE3	3:A:937:HOH:O	2.14	0.46
1:B:414:GLN:O	1:B:417:SER:HB3	2.16	0.46
1:A:379:LEU:HD12	1:A:544:LEU:HD11	1.98	0.46
1:A:316:VAL:HG21	1:A:489:LEU:HD21	1.98	0.45
1:A:472:LYS:NZ	3:A:725:HOH:O	2.49	0.45
1:B:498:GLN:HA	1:B:501[A]:HIS:CE1	2.52	0.45
1:A:473[B]:ASP:OD2	3:A:701:HOH:O	2.22	0.42
1:A:498:GLN:HA	1:A:501[B]:HIS:CE1	2.54	0.42
1:A:401:LYS:HD3	1:A:409:LEU:HG	2.01	0.41
1:B:531:LYS:C	1:B:532:ASN:HD22	2.22	0.41
1:B:307:ALA:HB1	1:B:365:PRO:HB3	2.03	0.40
1:B:448:LEU:HD21	1:B:507:LEU:HB3	2.03	0.40

All (3) 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)
1:A:359[A]:ASN:ND2	1:B:328:TYR:OH[3_455]	2.09	0.11
3:A:844:HOH:O	3:B:930:HOH:O[3_455]	2.11	0.09
3:A:707:HOH:O	3:A:902:HOH:O[2_555]	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	243/249 (98%)	241 (99%)	2 (1%)	0	100	100
1	B	229/249 (92%)	226 (99%)	3 (1%)	0	100	100
All	All	472/498 (95%)	467 (99%)	5 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	220/224 (98%)	220 (100%)	0	100	100
1	B	206/224 (92%)	206 (100%)	0	100	100
All	All	426/448 (95%)	426 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
1	B	413	ASN
1	B	532	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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

2 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	86V	A	601	-	38,38,38	4.19	21 (55%)	48,54,54	1.54	5 (10%)
2	86V	B	601	-	38,38,38	4.19	18 (47%)	48,54,54	1.67	6 (12%)

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	86V	A	601	-	-	0/14/39/39	0/5/5/5
2	86V	B	601	-	-	0/14/39/39	0/5/5/5

All (39) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	601	86V	CAT-CAI	-7.90	1.39	1.51
2	A	601	86V	CAT-CAI	-7.85	1.39	1.51
2	B	601	86V	OAH-CAI	-5.36	1.34	1.45
2	A	601	86V	OAH-CAI	-5.17	1.34	1.45
2	A	601	86V	CAB-CAC	-4.96	1.29	1.38
2	B	601	86V	CAB-CAC	-4.96	1.29	1.38
2	A	601	86V	CAF-CAK	-4.77	1.38	1.45
2	B	601	86V	CAF-CAK	-4.62	1.38	1.45
2	B	601	86V	OAH-CAE	-3.75	1.32	1.37
2	A	601	86V	OAH-CAE	-3.34	1.33	1.37
2	B	601	86V	CBB-NBC	-2.83	1.40	1.47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	601	86V	CAF-CAE	-2.74	1.35	1.40
2	A	601	86V	CBB-NBC	-2.53	1.41	1.47
2	A	601	86V	CAF-CAE	-2.49	1.36	1.40
2	A	601	86V	CAM-CAL	-2.26	1.35	1.39
2	A	601	86V	CAN-CAO	-2.15	1.34	1.38
2	A	601	86V	OAS-CAO	2.03	1.41	1.37
2	A	601	86V	CAL-CAJ	2.05	1.52	1.48
2	B	601	86V	OAS-CAO	2.08	1.41	1.37
2	B	601	86V	OAZ-CAW	2.62	1.43	1.37
2	A	601	86V	OAZ-CAW	2.92	1.44	1.37
2	A	601	86V	CAD-CAC	5.41	1.47	1.39
2	B	601	86V	CAD-CAC	5.52	1.47	1.39
2	A	601	86V	CAB-CAA	5.82	1.49	1.38
2	B	601	86V	CAB-CAA	6.29	1.50	1.38
2	B	601	86V	CAJ-CAK	6.56	1.44	1.34
2	B	601	86V	CAX-CAW	6.64	1.52	1.38
2	A	601	86V	CAX-CAW	6.67	1.52	1.38
2	B	601	86V	CAP-CAO	6.70	1.52	1.38
2	A	601	86V	CAJ-CAK	6.71	1.44	1.34
2	A	601	86V	CAP-CAO	6.86	1.52	1.38
2	B	601	86V	CAQ-CAL	7.32	1.51	1.39
2	B	601	86V	CAM-CAN	7.58	1.52	1.38
2	A	601	86V	CAM-CAN	7.71	1.52	1.38
2	A	601	86V	CAQ-CAL	7.82	1.52	1.39
2	A	601	86V	CAV-CAU	8.12	1.53	1.38
2	B	601	86V	CAV-CAU	8.41	1.53	1.38
2	A	601	86V	CAY-CAT	8.49	1.52	1.39
2	B	601	86V	CAY-CAT	8.55	1.52	1.39

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	601	86V	OAH-CAI-CAJ	-5.84	102.98	111.48
2	A	601	86V	OAH-CAI-CAJ	-4.84	104.44	111.48
2	B	601	86V	CAT-CAI-CAJ	-4.47	106.42	114.81
2	A	601	86V	CAT-CAI-CAJ	-3.47	108.28	114.81
2	A	601	86V	CAM-CAL-CAJ	-2.94	117.11	120.91
2	B	601	86V	CAQ-CAL-CAJ	-2.89	117.17	120.91
2	B	601	86V	CAR-CAK-CAJ	-2.59	121.31	124.27
2	B	601	86V	CBG-NBC-CBD	2.83	106.86	104.05
2	A	601	86V	CBG-NBC-CBD	3.57	107.59	104.05
2	A	601	86V	CAE-OAH-CAI	3.79	127.99	116.50

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	601	86V	CAE-OAH-CAI	4.86	131.20	116.50

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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

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	237/249 (95%)	0.32	18 (7%) 15 15	8, 18, 41, 55	2 (0%)
1	B	229/249 (91%)	0.54	24 (10%) 7 7	10, 21, 46, 60	1 (0%)
All	All	466/498 (93%)	0.43	42 (9%) 10 10	8, 20, 44, 60	3 (0%)

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	417	SER	7.2
1	B	336	PRO	6.8
1	B	337	PHE	6.2
1	B	418	VAL	5.0
1	B	459	TYR	4.3
1	B	307	ALA	4.3
1	A	334	THR	4.2
1	A	548	ARG	4.2
1	B	335	ARG	3.8
1	B	306	LEU	3.6
1	B	331	TYR	3.4
1	B	532	ASN	3.2
1	B	470	GLU	3.2
1	A	336	PRO	3.2
1	B	308	LEU	3.1
1	A	459	TYR	3.0
1	A	465	THR	3.0
1	A	460	THR	2.9
1	A	547	HIS	2.9
1	A	310	LEU	2.7
1	A	461	PHE	2.7
1	B	402	LEU	2.7
1	B	332	ASP	2.6
1	B	309	SER	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	335	ARG	2.5
1	B	333	PRO	2.5
1	A	537	TYR	2.5
1	B	471	GLU	2.4
1	B	334	THR	2.4
1	A	494	GLY	2.4
1	B	398	HIS	2.3
1	A	544	LEU	2.2
1	B	472	LYS	2.2
1	A	546	ALA	2.2
1	B	460	THR	2.1
1	A	541	LEU	2.1
1	B	341	SER	2.1
1	B	397	GLU	2.1
1	A	538	ASP	2.0
1	B	546	ALA	2.0
1	A	532	ASN	2.0
1	A	309	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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	86V	A	601	34/34	0.94	0.09	-0.03	7,12,20,36	0
2	86V	B	601	34/34	0.92	0.10	-0.07	9,13,19,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.