



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 09:45 pm GMT

PDB ID : 3DRR
Title : HIV reverse transcriptase Y181C mutant in complex with inhibitor R8e
Authors : Yan, Y.
Deposited on : 2008-07-11
Resolution : 2.89 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

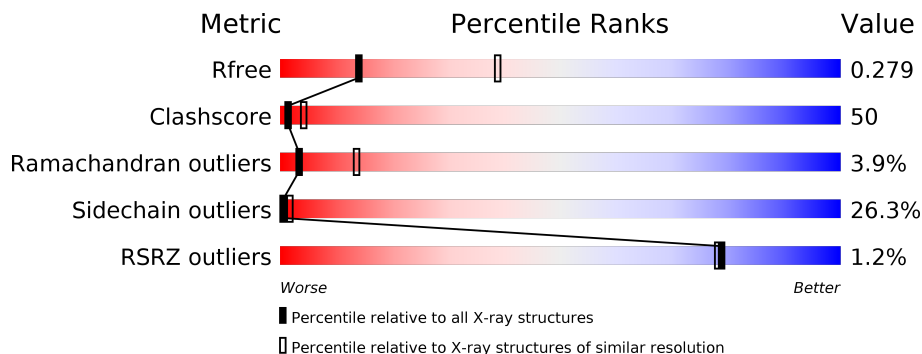
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.89 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}	111664	1716 (2.90-2.90)
Clashscore	122126	1924 (2.90-2.90)
Ramachandran outliers	120053	1884 (2.90-2.90)
Sidechain outliers	120020	1886 (2.90-2.90)
RSRZ outliers	108989	1669 (2.90-2.90)

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 ≥ 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	563	<div> <div>%</div> <div> <div></div> <div>30%</div> <div>50%</div> <div>18%</div> <div>..</div> </div> </div>
2	B	443	<div> <div>%</div> <div> <div></div> <div>31%</div> <div>43%</div> <div>16%</div> <div>• 9%</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8249 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 Reverse transcriptase/ribonuclease H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	558	Total	C	N	O	S	0	0	0
			4536	2928	760	839	9			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	EXPRESSION TAG	UNP P04585
A	-1	ASN	-	EXPRESSION TAG	UNP P04585
A	0	SER	-	EXPRESSION TAG	UNP P04585
A	181	CYS	TYR	ENGINEERED	UNP P04585

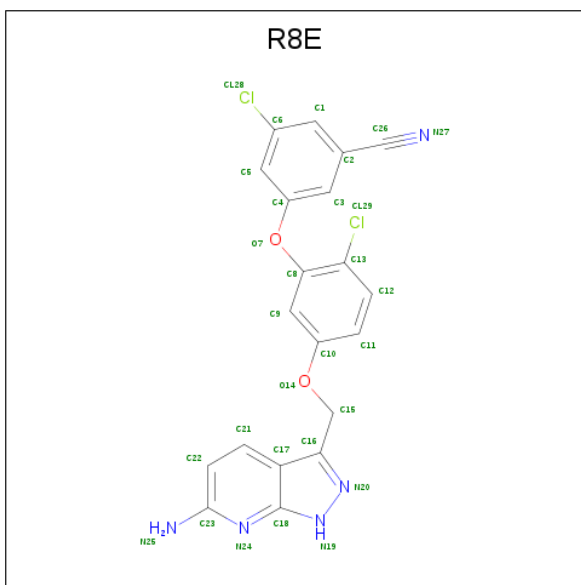
- Molecule 2 is a protein called p51 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	404	Total	C	N	O	S	0	0	0
			3338	2170	554	607	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	MET	-	EXPRESSION TAG	UNP P04585
B	-1	ASN	-	EXPRESSION TAG	UNP P04585
B	0	SER	-	EXPRESSION TAG	UNP P04585
B	181	CYS	TYR	ENGINEERED	UNP P04585

- Molecule 3 is 3-{5-[(6-amino-1H-pyrazolo[3,4-b]pyridin-3-yl)methoxy]-2-chlorophenoxy}-5-chlorobenzonitrile (three-letter code: R8E) (formula: C₂₀H₁₃Cl₂N₅O₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	Cl	N	O	0	0
			29	20	2	5	2		

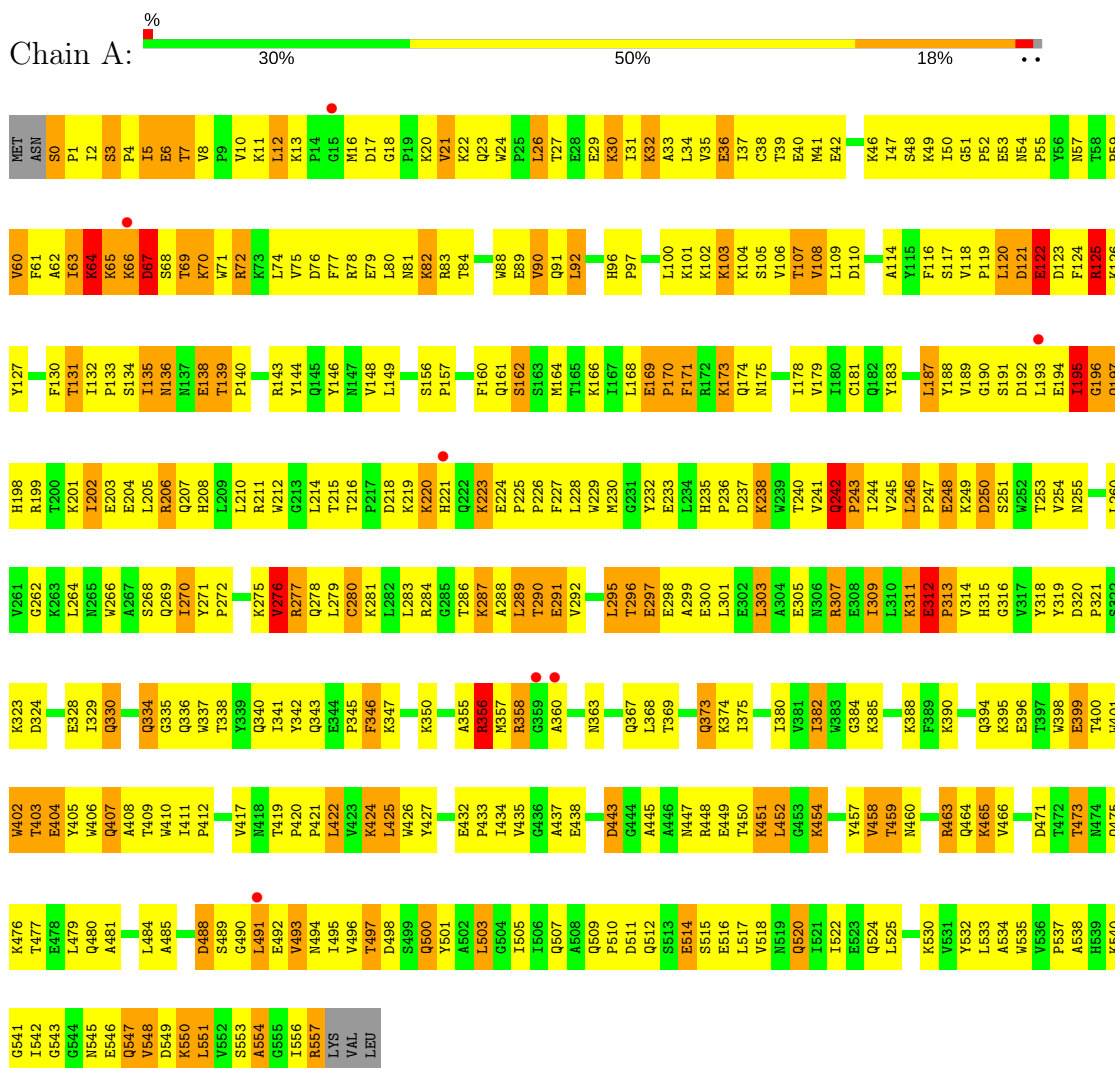
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	195	Total O 195 195	0	0
4	B	151	Total O 151 151	0	0

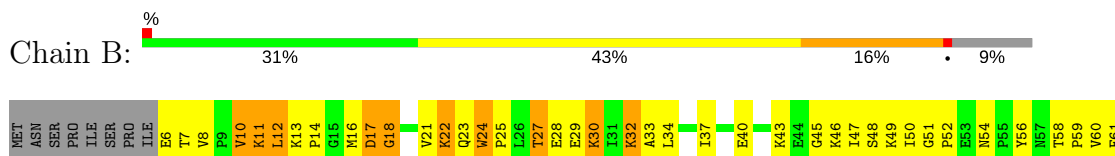
3 Residue-property plots

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: Reverse transcriptase/ribonuclease H



• Molecule 2: p51 RT



A62	I135	T215	R277	Q343	E413
I63	Y144	THR	Q278	E344	
K64	Y146	PRO	L279	P345	F416
K65	Y149	ASP	C280	F346	V417
K66	Y150	LYS	K281	K347	N418
D67	L149	LYS	L282	N348	T419
S68	P150	HIS	L283	L349	P420
T69	W153	GLN	R284		P421
K70	W154	LYS	G285	Y354	L422
W71	PRU	GLU	T286	A355	V423
R72	K154	PRO	T287	R356	K424
K73	G155	PRU	A288	MET	L425
L74	S156	PHE	L289	ARG	H426
V75	P157	LEU	T290	GLY	V427
D76	A158	TRP	E291	ALA	Q428
F77	F160	MET	V292	H361	LEU
R78	I159	G231	I293	T362	GLU
E79	Y232	Y232	P294	N363	LYS
L80	S163	E233	L295	D364	GLU
N81	K166	L234	T296	V365	PRO
K82	I167	D237	E297	K366	ILE
R83	L168	K238	E298	Q367	VAL
T84	E169		A299	L368	GLY
Q85	R172	V241	E300	T369	ALA
D86	K173	Q242	L301	E370	GLU
F87	E88	P243	E302	A371	THR
W88	Q174	I244	L303	V372	PHE
E89	M175	V245	A304	Q373	
V90	P176	L246	E305	K374	
Q91	D177	P247	N306	I375	
L92	I178	E248	R307	T376	
P95	I179	K249	E308	T377	
H96	V179	D250	L309	I382	
P97	I180	S251	L310	W383	
A98	G181	W252	K311	G384	
G99	Q182	T253	E312	K385	
L100	L187	V254	P313	T386	
K101	Y188	N255	V314	P387	
K102	V189	D256	Y319	K388	
K103	I195	I257	D320	F389	
V108	R199	Q258	P321	K390	
V111	T200	K259	S322	L391	
Y115	K201	L260	K323	P392	
F116	I202	V261	D324	I393	
S117	E203	G262	L325	Q394	
V118	E204	N265	I326	E395	
P119	L205	V266	Q330	T397	
L120	R206	A267	K331	W398	
D121	Q207	S268	Q332	E399	
E122	H208	Q269	I335	T400	
	L209	I270	G336	W401	
	L210	Y271	W337	W402	
R125	K211	P272	T338	T403	
K126	W212	G273	Y339	E404	
Y127	G213	I274	Q340	Y405	
T128	L214	K275	I341	W406	
A129		V276	Y342	Q407	

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	118.93Å 154.68Å 154.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.89 47.14 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.00-2.89) 99.6 (47.14-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.91 (at 2.91Å)	Xtriage
Refinement program	BUSTER-TNT 2.1.1	Depositor
R, R_{free}	0.183 , 0.269 0.192 , 0.279	Depositor DCC
R_{free} test set	1617 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	58.9	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 81.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.53$, $\langle L^2 \rangle = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	8249	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: R8E

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	0.64	0/4652	0.88	4/6320 (0.1%)
2	B	0.63	0/3431	0.84	4/4661 (0.1%)
All	All	0.64	0/8083	0.86	8/10981 (0.1%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	312	GLU	C-N-CD	-13.44	91.04	120.60
2	B	420	PRO	C-N-CD	-11.79	94.67	120.60
2	B	51	GLY	C-N-CD	-6.55	106.19	120.60
1	A	216	THR	C-N-CD	6.06	141.12	128.40
1	A	242	GLN	C-N-CD	-6.05	107.30	120.60

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	4536	0	4591	499	0
2	B	3338	0	3365	325	0
3	A	29	0	13	5	0
4	A	195	0	0	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	151	0	0	8	0
All	All	8249	0	7969	796	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 50.

The worst 5 of 796 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:65:LYS:HE3	1:A:70:LYS:HG2	1.19	1.16
1:A:255:ASN:HB2	1:A:289:LEU:HD22	1.16	1.14
1:A:63:ILE:HD11	1:A:74:LEU:HD21	1.19	1.13
2:B:260:LEU:HD22	2:B:264:LEU:HD12	1.30	1.13
1:A:473:THR:HG23	1:A:476:LYS:HD2	1.20	1.11

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	556/563 (99%)	466 (84%)	62 (11%)	28 (5%)	2	8
2	B	398/443 (90%)	350 (88%)	39 (10%)	9 (2%)	7	26
All	All	954/1006 (95%)	816 (86%)	101 (11%)	37 (4%)	3	13

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	SER
1	A	90	VAL
1	A	136	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	170	PRO
1	A	286	THR

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	498/503 (99%)	368 (74%)	130 (26%)	0	2
2	B	368/403 (91%)	270 (73%)	98 (27%)	0	2
All	All	866/906 (96%)	638 (74%)	228 (26%)	0	2

5 of 228 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	451	LYS
1	A	550	LYS
2	B	305	GLU
1	A	458	VAL
1	A	493	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	494	ASN
2	B	54	ASN
2	B	367	GLN
1	A	519	ASN
1	A	524	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand 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$
3	R8E	A	601	-	30,32,32	2.35	8 (26%)	37,45,45	3.08	13 (35%)

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
3	R8E	A	601	-	-	0/9/11/11	0/4/4/4

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	R8E	C2-C26	-9.55	1.22	1.44
3	A	601	R8E	C5-C4	-2.91	1.33	1.38
3	A	601	R8E	C18-N24	-2.46	1.31	1.36
3	A	601	R8E	C26-N27	-2.29	1.09	1.14
3	A	601	R8E	C17-C18	-2.13	1.37	1.43

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	R8E	C2-C1-C6	-10.15	110.92	119.18
3	A	601	R8E	C5-C6-CL28	-6.49	110.86	119.15
3	A	601	R8E	C8-C13-CL29	-5.47	112.98	119.42
3	A	601	R8E	C3-C2-C26	-4.56	113.69	119.52
3	A	601	R8E	N25-C23-N24	-4.23	114.71	118.28

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	R8E	5	0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	558/563 (99%)	-0.08	7 (1%) 77 76	21, 50, 83, 112	0
2	B	404/443 (91%)	-0.14	5 (1%) 79 78	24, 49, 100, 118	0
All	All	962/1006 (95%)	-0.10	12 (1%) 79 78	21, 50, 94, 118	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	193	LEU	3.0
2	B	90	VAL	3.0
2	B	67	ASP	2.7
2	B	270	ILE	2.7
1	A	66	LYS	2.5

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. 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(\AA^2)	Q<0.9
3	R8E	A	601	29/29	0.96	0.23	42,52,75,78	0

6.5 Other polymers [i](#)

There are no such residues in this entry.