



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:12 AM GMT

PDB ID : 3DRP
Title : HIV reverse transcriptase in complex with inhibitor R8e
Authors : Yan, Y.; Prasad, S.
Deposited on : 2008-07-11
Resolution : 2.60 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
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

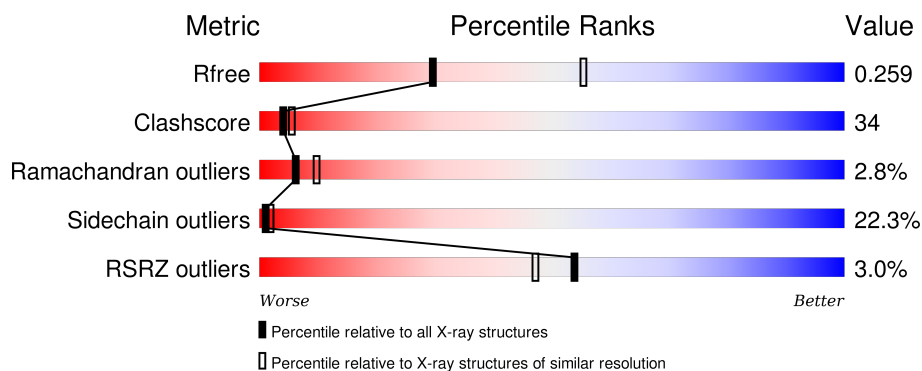
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

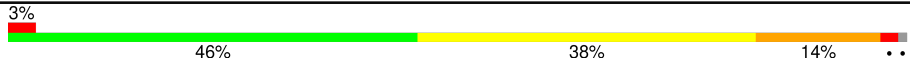
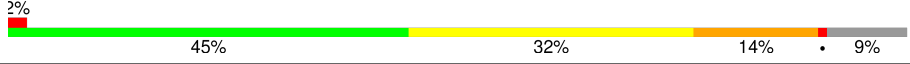
The reported resolution of this entry is 2.60 Å.

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}	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	 3% 46% 38% 14% ..
2	B	443	 2% 45% 32% 14% • 9%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8340 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
			4542	2934	760	840	8			

There are 3 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

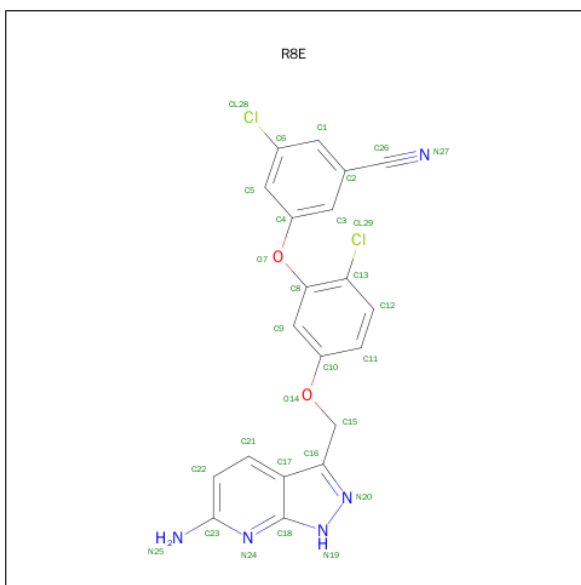
- Molecule 2 is a protein called p51 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	404	Total	C	N	O	S	0	0	0
			3344	2176	554	608	6			

There are 3 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

- 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		

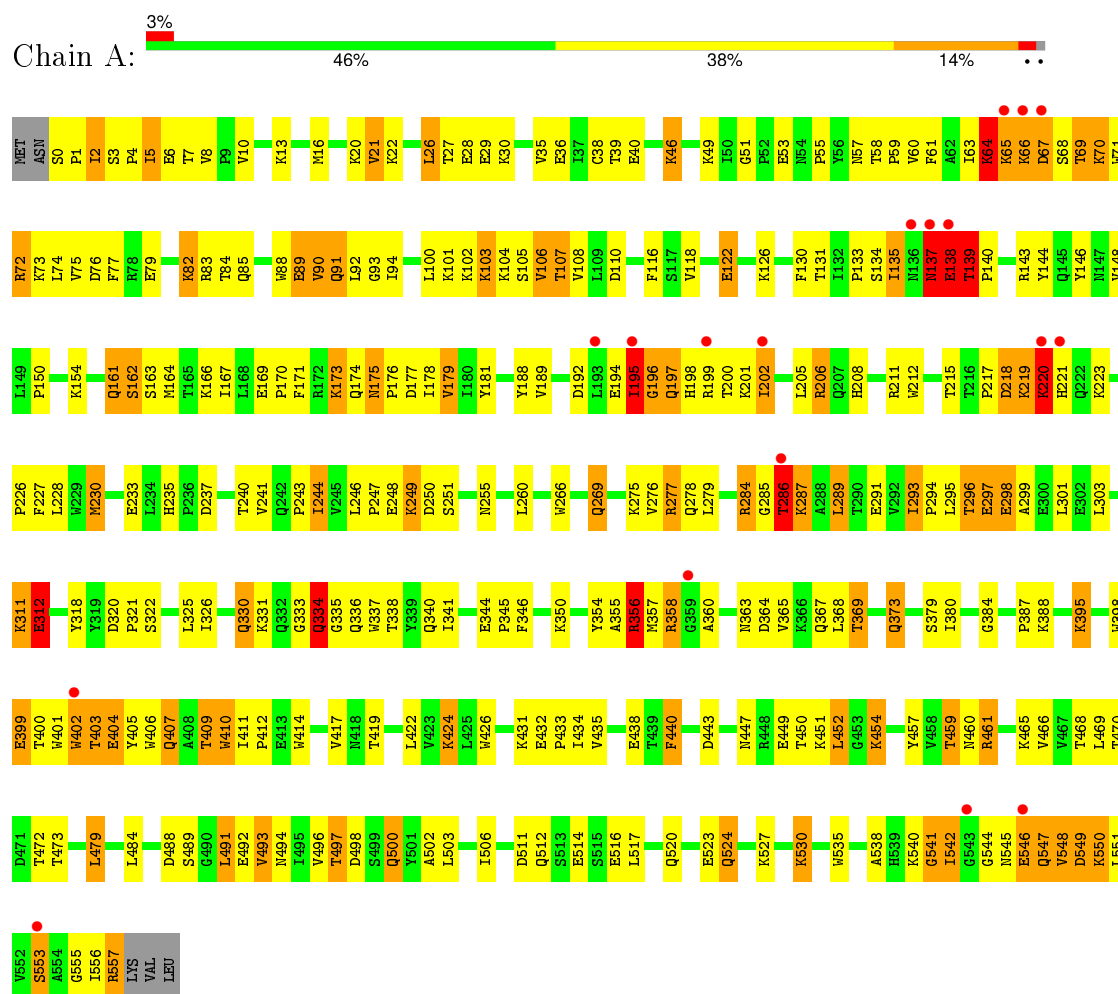
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	232	Total O 232 232	0	0
4	B	193	Total O 193 193	0	0

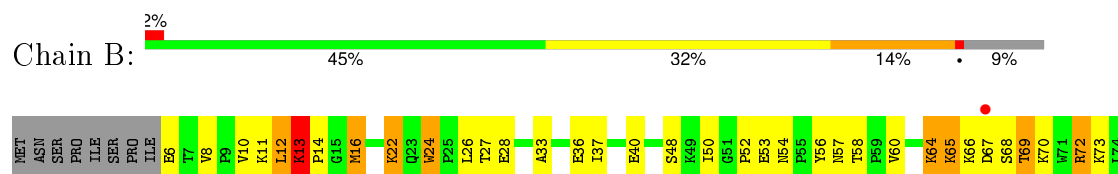
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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



E399	T400	T403	E404	Y405	W406	Q407	A408	T409	E413	M418	T419	P420	P421	L422	W423	K424	L425	W426	Y427	Q428	LEU	GLU	LYS	GLU	PRO	ILE	VAL	GLY	ALA	GLU	THR	PHE	V314	V317	Y318	Y319	D320	P321	I326	K331	Q332	G335	Q336	W337	I341	E344	P345	F346	K347	N348	L349	Y354	A355	R356	MET	ARG	GLY	ALA	H361	T362	N363	K366	Q367	E370	Q373	K374	I375	T376	T377	E378	S379	K385	T386	P387	K388	I393	Q394	K395	E396	K249	D250	S251	W252	V254	N255	D256	I257	L260	K263	L264	W265	A267	S268	Q269	I270	Y271	G272	G273	I274	K275	W276	Q277	Q278	L279	C280	K281	L282	T286	K287	A288	L289	T290	E291	V292	I293	P294	L295	T296	E297	E298	A299	E300	L301	E302	L303	A304	E305	N306	R307	E308	I309	L310	K311	D185	D186	L187	G190	L193	E194	I195	G196	Q197	H198	R199	T200	K201	I202	E203	E204	L205	R206	Q207	H208	L209	L210	R211	W212	G213	L214	T215	THR	PRO	ASP	LYS	LYS	HIS	GLN	LYS	GLU	PRO	PRO	PHE	LEU	TRP	MET	G231	Y232	P236	D237	K238	W239	T240	V241	Q242	P243	I244	V245	L246	P247	E248	V75	L80	N81	K82	R83	T84	Q85	D86	F87	W88	E89	V90	Q91	L92	K101	K102	K103	K104	S105	V111	A114	L120	F124	R125	K126	Y127	T131	I135	E138	T139	P140	Q145	Y146	S162	K166	R172	K173	Q174	N175	P176	V179	I180	Y181	Q182	Y183	M184
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	118.86Å 155.03Å 155.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	94.49 – 2.60 47.16 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.6 (94.49-2.60) 99.8 (47.16-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.85 (at 2.61Å)	Xtriage
Refinement program	BUSTER-TNT 2.1.1	Depositor
R, R_{free}	0.187 , 0.251 0.194 , 0.259	Depositor DCC
R_{free} test set	2235 reflections (5.30%)	DCC
Wilson B-factor (Å ²)	55.1	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 83.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.54$, $\langle L^2 \rangle = 0.38$	Xtriage
Outliers	0 of 44453 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8340	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.66	0/4659	0.83	2/6330 (0.0%)
2	B	0.66	0/3438	0.81	3/4671 (0.1%)
All	All	0.66	0/8097	0.82	5/11001 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	175	ASN	C-N-CD	-9.03	100.73	120.60
2	B	13	LYS	C-N-CD	-7.11	104.95	120.60
1	A	312	GLU	C-N-CD	-6.80	105.65	120.60
2	B	205	LEU	CA-CB-CG	5.19	127.23	115.30
2	B	320	ASP	C-N-CD	-5.13	109.32	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	4542	0	4595	349	0
2	B	3344	0	3369	220	0
3	A	29	0	13	2	0
4	A	232	0	0	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	193	0	0	7	0
All	All	8340	0	7977	541	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 34.

The worst 5 of 541 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:296:THR:HG22	1:A:299:ALA:H	1.09	1.18
1:A:255:ASN:HB2	1:A:289:LEU:HD22	1.20	1.14
2:B:13:LYS:HE3	2:B:85:GLN:HB3	1.35	1.05
1:A:2:ILE:HD11	1:A:46:LYS:HZ1	1.20	1.04
2:B:209:LEU:HG	2:B:214:LEU:HD12	1.38	1.01

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%)	486 (87%)	51 (9%)	19 (3%)	5	7
2	B	398/443 (90%)	369 (93%)	21 (5%)	8 (2%)	9	18
All	All	954/1006 (95%)	855 (90%)	72 (8%)	27 (3%)	6	10

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	VAL
1	A	138	GLU
1	A	195	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	286	THR
1	A	334	GLN

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	498/503 (99%)	384 (77%)	114 (23%)	1	2
2	B	368/403 (91%)	289 (78%)	79 (22%)	1	2
All	All	866/906 (96%)	673 (78%)	193 (22%)	1	2

5 of 193 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	422	LEU
1	A	547	GLN
2	B	309	ILE
1	A	431	LYS
1	A	493	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	480	GLN
1	A	519	ASN
2	B	373	GLN
1	A	509	GLN
1	A	520	GLN

5.3.3 RNA ⓘ

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	-	31,32,32	2.20	9 (29%)	33,45,45	2.35	10 (30%)

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 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	R8E	C2-C26	-6.20	1.29	1.44
3	A	601	R8E	C8-C13	-4.48	1.31	1.39
3	A	601	R8E	C11-C10	-2.44	1.33	1.38
3	A	601	R8E	C18-N19	-2.31	1.30	1.34
3	A	601	R8E	C9-C10	-2.26	1.34	1.38

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	R8E	C8-C13-CL29	-5.74	112.33	119.42
3	A	601	R8E	C1-C6-CL28	-4.51	113.54	119.14
3	A	601	R8E	C11-C12-C13	-4.47	113.29	119.95
3	A	601	R8E	O14-C15-C16	-4.13	99.26	109.89
3	A	601	R8E	O7-C8-C13	-3.47	111.73	119.84

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	R8E	2	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%)	-0.05	18 (3%)	51 44	28, 55, 93, 125	0
2	B	404/443 (91%)	-0.11	11 (2%)	58 51	27, 50, 108, 126	0
All	All	962/1006 (95%)	-0.08	29 (3%)	54 47	27, 54, 102, 126	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	193	LEU	4.0
1	A	67	ASP	3.9
1	A	66	LYS	3.8
2	B	362	THR	3.6
1	A	286	THR	3.3

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, 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	LLDF	B-factors(\AA^2)	Q<0.9
3	R8E	A	601	29/29	0.98	0.14	-0.67	43,50,59,61	0

6.5 Other polymers [i](#)

There are no such residues in this entry.