



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:03 PM GMT

PDB ID : 1DTQ
Title : CRYSTAL STRUCTURE OF HIV-1 REVERSE TRANSCRIPTASE IN COMPLEX WITH PETT-1 (PETT131A94)
Authors : Ren, J.; Diprose, J.; Warren, J.; Esnouf, R.M.; Bird, L.E.; Ikemizu, S.; Slater, M.; Milton, J.; Balzarini, J.; Stuart, D.I.; Stammers, D.K.
Deposited on : 2000-01-13
Resolution : 2.80 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

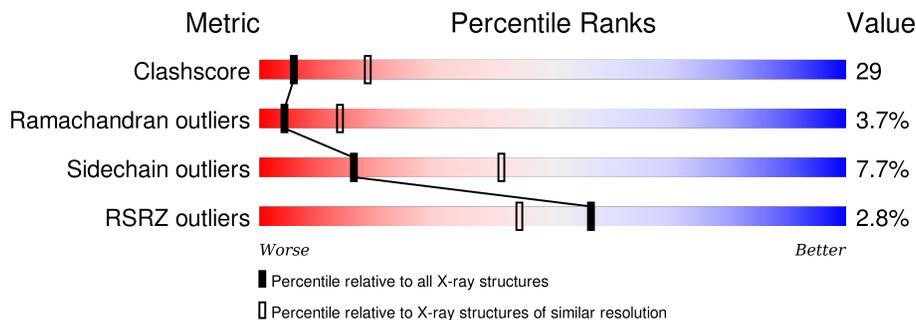
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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))
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	560	
2	B	440	

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
1	CSD	A	280	-	-	X	-

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 7687 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.

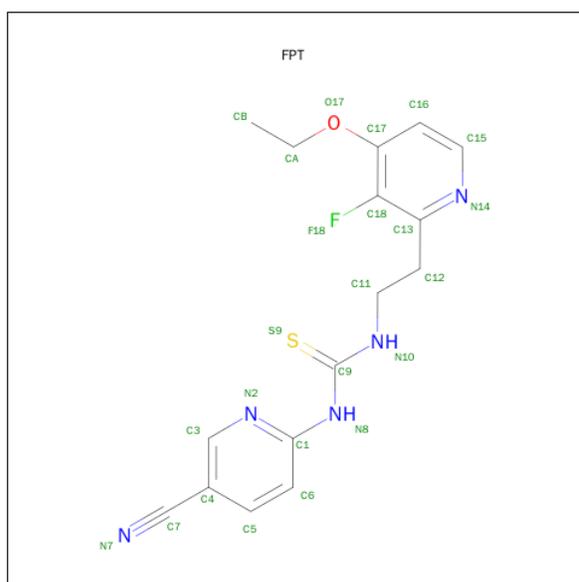
- Molecule 1 is a protein called HIV-1 RT A-CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	525	4310	2792	714	796	8	0	0	0

- Molecule 2 is a protein called HIV-1 RT B-CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	402	3327	2166	550	605	6	0	0	0

- Molecule 3 is N-[[3-FLUORO-4-ETHOXY-PYRID-2-YL]ETHYL]-N'-[5-NITRILOMETHYL-PYRIDYL]-THIOUREA (three-letter code: FPT) (formula: C₁₆H₁₆FN₅OS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	F	N	O			S
3	A	1	24	16	1	5	1	1	0	0

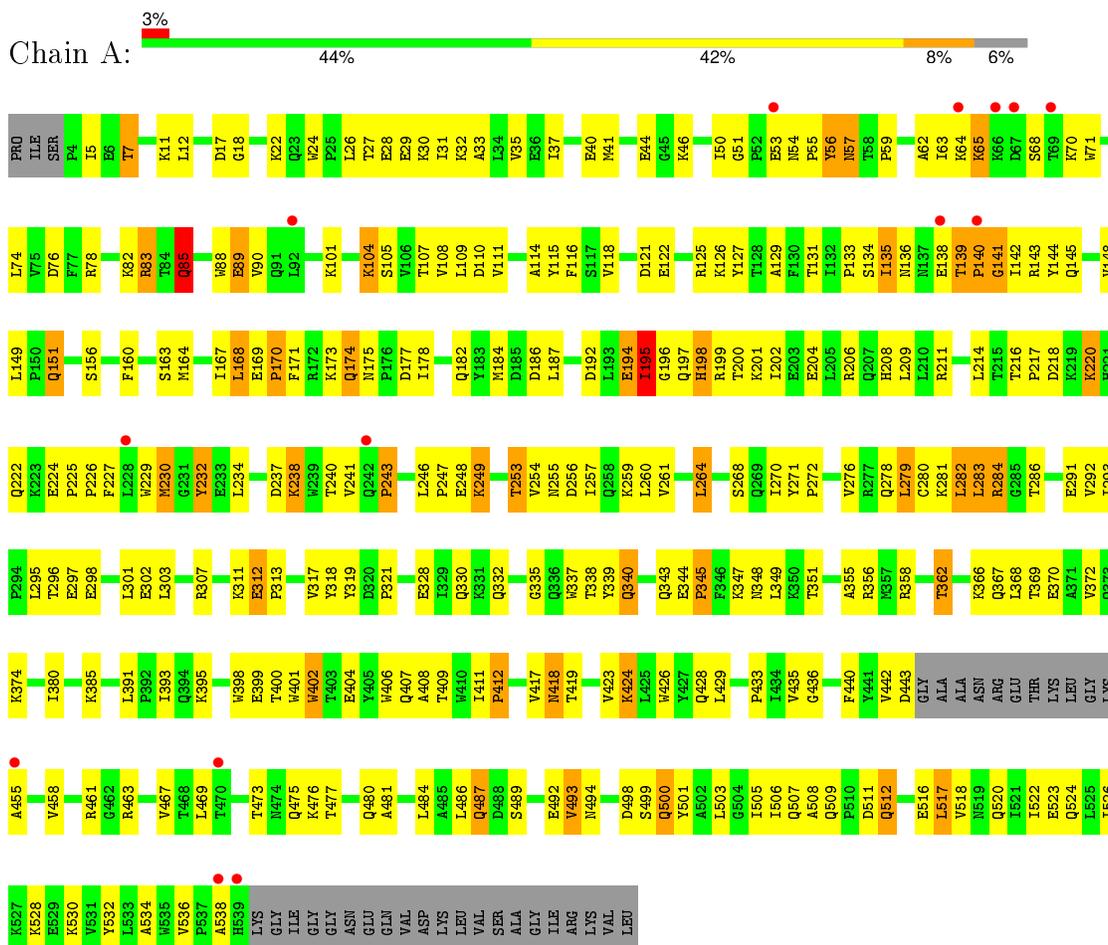
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	18	Total 18	O 18	0	0
4	B	8	Total 8	O 8	0	0

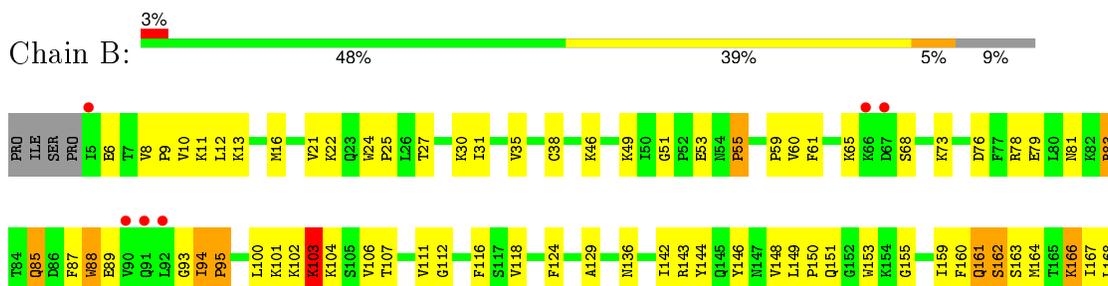
3 Residue-property plots

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: HIV-1 RT A-CHAIN



• Molecule 2: HIV-1 RT B-CHAIN



E169	P170	F171	R172	K173	Q174	N175	P176	D177	I178	V179	Y183	H184	D185	V189	G190	S191	D192	L193	E194	I195	G196	Q197	K201	L202	L205	R206	Q207	H208	L209	L210	R211	W212	G213	LEU	THR	THR	PRO	ASP	LYS	LYS	HIS	GLN	LYS	GLU	PRO	PRO	PHE	LEU	TRP	MET	G231	Y232	E233	L234	H235
P236	D237	K238	W239	T240	Y241	Q242	F243	K249	W252	T253	V254	M255	D256	I257	Q258	K259	L260	V261	G262	K263	W266	Q269	L270	Y271	P272	R277	Q278	K281	L282	L283	R284	K287	A288	L289	T290	E291	Y292	I293	T296	E297	E298	A299	E300	I301	LEU	E302	L303	V314	V317	Y318					
Y319	I325	I326	A327	E328	I329	Q330	K331	Q334	H337	T338	Y339	Y342	N348	L349	K350	K353	R356	MET	ARG	GLY	ALA	HIS	T362	S663	D664	V365	K366	Q367	L368	T369	E370	A371	V372	Q373	K374	I375	T376	T377	E378	S379	H383	G384	K385	T386	P387	K388	F389	I393	Q394						
H398	W401	W402	T403	E404	Y405	W406	W410	W414	E415	W418	T419	P420	P421	L422	W423	R425	L426	W426	Y427	Q428	LEU	GLU	LYS	GLU	PRO	PRO	ILE	VAL	GLY	ALA	GLU	THR	PHE																						

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	137.10Å 115.00Å 65.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80 24.18 – 2.77	Depositor EDS
% Data completeness (in resolution range)	95.0 (30.00-2.80) 93.9 (24.18-2.77)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.43 (at 2.76Å)	Xtrriage
Refinement program	CNS	Depositor
R, R_{free}	0.224 , 0.295 0.211 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	54.4	Xtrriage
Anisotropy	0.108	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 73.9	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Outliers	0 of 25301 reflections	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7687	wwPDB-VP
Average B, all atoms (Å ²)	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% 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: FPT, CSD

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	0.45	0/4416	0.66	1/6002 (0.0%)
2	B	0.45	0/3420	0.67	0/4646
All	All	0.45	0/7836	0.67	1/10648 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	141	GLY	N-CA-C	-5.16	100.20	113.10

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	4310	0	4342	272	0
2	B	3327	0	3355	179	0
3	A	24	0	16	1	0
4	A	18	0	0	3	0
4	B	8	0	0	2	0
All	All	7687	0	7713	442	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 29.

The worst 5 of 442 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:139:THR:HB	1:A:140:PRO:HD2	1.21	1.11
1:A:408:ALA:HB1	2:B:364:ASP:HB3	1.36	1.04
1:A:337:TRP:HE1	1:A:367:GLN:HE21	1.06	1.02
1:A:139:THR:HB	1:A:140:PRO:CD	1.99	0.93
1:A:435:VAL:HG22	2:B:290:THR:HG21	1.55	0.87

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	519/560 (93%)	451 (87%)	47 (9%)	21 (4%)	4 12
2	B	396/440 (90%)	342 (86%)	41 (10%)	13 (3%)	5 16
All	All	915/1000 (92%)	793 (87%)	88 (10%)	34 (4%)	4 14

5 of 34 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	VAL
1	A	140	PRO
1	A	195	ILE
1	A	402	TRP
1	A	412	PRO

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	473/499 (95%)	433 (92%)	40 (8%)	13	36
2	B	366/400 (92%)	341 (93%)	25 (7%)	20	49
All	All	839/899 (93%)	774 (92%)	65 (8%)	16	41

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	340	GLN
1	A	493	VAL
2	B	394	GLN
1	A	345	PRO
1	A	424	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	509	GLN
2	B	57	ASN
2	B	269	GLN
1	A	512	GLN
2	B	91	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](#)

1 non-standard protein/DNA/RNA residue 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
1	CSD	A	280	1	3,7,8	0.66	0	3,8,10	4.03	1 (33%)

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
1	CSD	A	280	1	-	1/2/6/8	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	280	CSD	OD1-SG-CB	6.75	116.65	105.40

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	280	CSD	CA-CB-SG-OD1

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	280	CSD	5	0

5.5 Carbohydrates

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	FPT	A	999	-	24,25,25	1.63	5 (20%)	29,32,32	1.93	10 (34%)

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	FPT	A	999	-	-	1/15/15/15	0/2/2/2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	999	FPT	C1-N8	-2.67	1.35	1.40
3	A	999	FPT	C9-N8	-2.02	1.31	1.35
3	A	999	FPT	C1-N2	2.72	1.39	1.34
3	A	999	FPT	C18-C13	3.03	1.42	1.38
3	A	999	FPT	C13-N14	3.96	1.38	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	999	FPT	C16-C15-N14	-4.00	119.34	123.90
3	A	999	FPT	C4-C3-N2	-3.29	118.83	123.70
3	A	999	FPT	O17-C17-C16	-2.95	117.85	124.01
3	A	999	FPT	C6-C1-N8	-2.45	116.22	123.04
3	A	999	FPT	C16-C17-C18	2.12	121.29	117.84

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	999	FPT	N8-C9-N10-C11

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	999	FPT	1	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	524/560 (93%)	-0.17	14 (2%) 58 45	23, 62, 104, 140	0
2	B	402/440 (91%)	-0.29	12 (2%) 54 41	26, 57, 103, 144	0
All	All	926/1000 (92%)	-0.22	26 (2%) 56 44	23, 59, 104, 144	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	5	ILE	7.8
1	A	67	ASP	6.7
2	B	92	LEU	5.3
2	B	91	GLN	5.1
2	B	90	VAL	4.2

6.2 Non-standard residues in protein, DNA, RNA chains [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(Å ²)	Q < 0.9
1	CSD	A	280	8/9	0.94	0.17	-	39,49,63,63	0

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	FPT	A	999	24/24	0.94	0.19	0.12	36,45,66,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.