



wwPDB X-ray Structure Validation Summary Report

Feb 15, 2017 – 12:33 am GMT

PDB ID : 3R9A
Title : Human alanine-glyoxylate aminotransferase in complex with the TPR domain of human PEX5P
Authors : Fodor, K.; Wilmanns, M.
Deposited on : 2011-03-25
Resolution : 2.35 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
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) : recalc28949

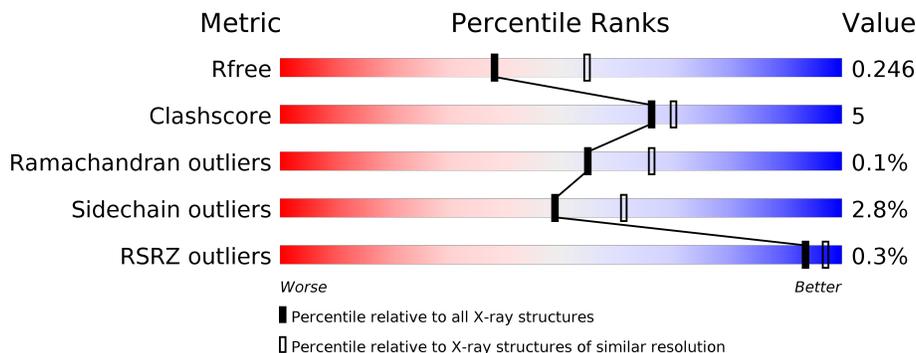
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.35 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	1522 (2.38-2.34)
Clashscore	112137	1626 (2.38-2.34)
Ramachandran outliers	110173	1605 (2.38-2.34)
Sidechain outliers	110143	1606 (2.38-2.34)
RSRZ outliers	101464	1528 (2.38-2.34)

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	394	
1	C	394	
2	B	328	
2	D	328	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	BTB	C	393	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11436 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 Serine–pyruvate aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	P				S
1	A	387	2999	1917	520	545	1	16	0	1	0
1	C	389	3014	1925	525	546	1	17	0	3	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP P21549
A	0	ALA	-	EXPRESSION TAG	UNP P21549
C	-1	GLY	-	EXPRESSION TAG	UNP P21549
C	0	ALA	-	EXPRESSION TAG	UNP P21549

- Molecule 2 is a protein called Peroxisomal targeting signal 1 receptor.

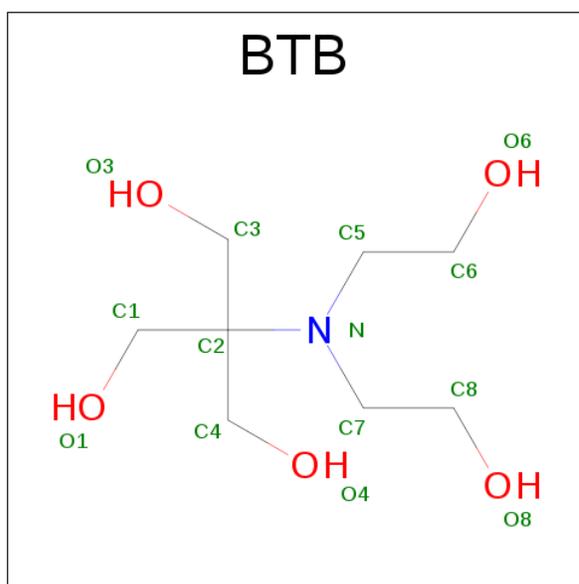
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	289	2274	1427	399	437	11	0	1	0
2	D	292	2304	1452	400	441	11	0	2	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	312	GLY	-	EXPRESSION TAG	UNP P50542
B	313	ALA	-	EXPRESSION TAG	UNP P50542
B	314	MET	-	EXPRESSION TAG	UNP P50542
D	312	GLY	-	EXPRESSION TAG	UNP P50542
D	313	ALA	-	EXPRESSION TAG	UNP P50542
D	314	MET	-	EXPRESSION TAG	UNP P50542

- Molecule 3 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPAN

E-1,3-DIOL (three-letter code: BTB) (formula: C₈H₁₉NO₅).



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

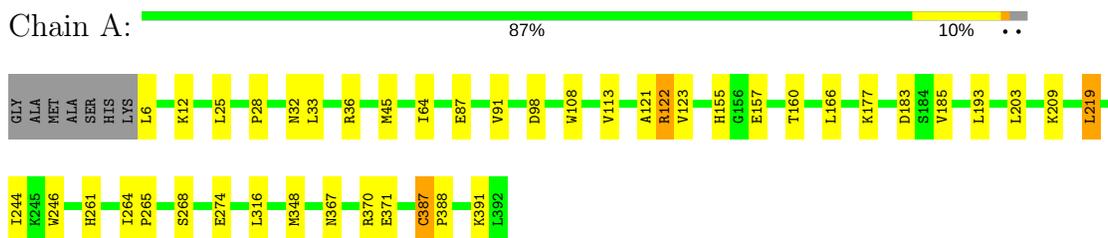
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	307	Total	O	0	0
			307	307		
4	B	186	Total	O	0	0
			186	186		
4	C	204	Total	O	0	0
			204	204		
4	D	134	Total	O	0	0
			134	134		

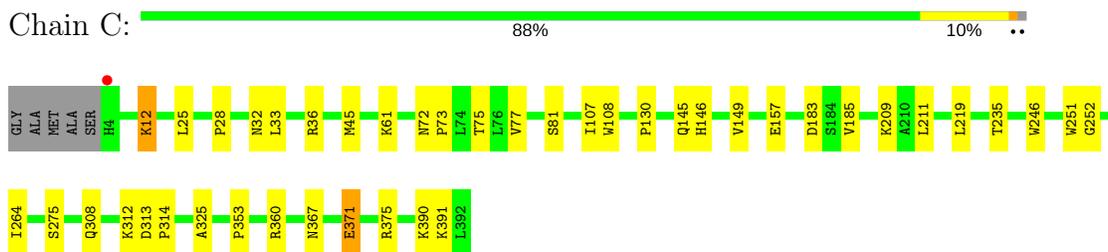
3 Residue-property plots [i](#)

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.

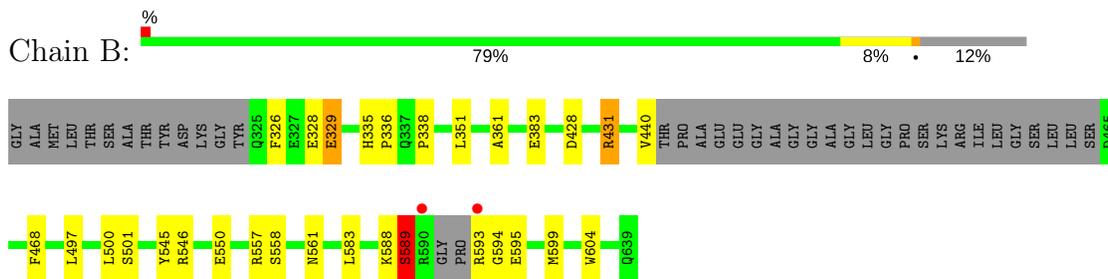
- Molecule 1: Serine–pyruvate aminotransferase



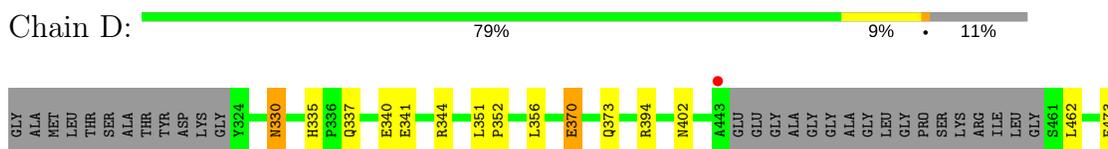
- Molecule 1: Serine–pyruvate aminotransferase



- Molecule 2: Peroxisomal targeting signal 1 receptor



- Molecule 2: Peroxisomal targeting signal 1 receptor





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	59.37Å 99.16Å 127.70Å 90.00° 96.66° 90.00°	Depositor
Resolution (Å)	78.09 – 2.35 78.12 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.7 (78.09-2.35) 99.6 (78.12-2.35)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.77 (at 2.34Å)	Xtrriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.184 , 0.239 0.191 , 0.246	Depositor DCC
R_{free} test set	3096 reflections (5.34%)	DCC
Wilson B-factor (Å ²)	29.3	Xtrriage
Anisotropy	0.058	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 33.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	11436	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.46% 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

5.1 Standard geometry

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

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	0.42	0/3045	0.57	0/4130
1	C	0.38	0/3065	0.54	0/4161
2	B	0.44	0/2319	0.54	0/3141
2	D	0.41	0/2354	0.53	0/3193
All	All	0.41	0/10783	0.55	0/14625

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	329	GLU	Peptide

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	2999	0	3036	28	0
1	C	3014	0	3026	30	0
2	B	2274	0	2217	23	0
2	D	2304	0	2261	22	0
3	C	14	0	19	2	0
4	A	307	0	0	4	0
4	B	186	0	0	0	0
4	C	204	0	0	1	0
4	D	134	0	0	4	0
All	All	11436	0	10559	97	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 5.

The worst 5 of 97 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:25:LEU:H	1:A:32:ASN:HD21	1.24	0.86
2:B:583:LEU:HB3	2:B:599:MET:HE3	1.57	0.85
2:B:588:LYS:HA	2:B:589:SER:HB2	1.58	0.84
2:D:552:GLN:HE21	2:D:554:GLY:H	1.25	0.81
1:C:183:ASP:OD1	1:C:209:LLP:H2'2	1.86	0.76

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	385/394 (98%)	377 (98%)	7 (2%)	1 (0%)	44	53
1	C	389/394 (99%)	380 (98%)	9 (2%)	0	100	100
2	B	284/328 (87%)	279 (98%)	4 (1%)	1 (0%)	38	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	288/328 (88%)	282 (98%)	6 (2%)	0	100	100
All	All	1346/1444 (93%)	1318 (98%)	26 (2%)	2 (0%)	55	66

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	589	SER
1	A	387	CYS

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	319/325 (98%)	311 (98%)	8 (2%)	53	66
1	C	317/325 (98%)	308 (97%)	9 (3%)	49	61
2	B	238/264 (90%)	233 (98%)	5 (2%)	59	72
2	D	244/264 (92%)	234 (96%)	10 (4%)	35	44
All	All	1118/1178 (95%)	1086 (97%)	32 (3%)	49	60

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	36	ARG
1	C	219	LEU
2	D	539	GLU
1	C	157	GLU
1	C	235	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	72	ASN
1	C	261	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	419	GLN
1	C	32	ASN
2	D	536	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 non-standard protein/DNA/RNA residues 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
1	LLP	A	209	1	24,24,25	1.85	5 (20%)	28,32,34	1.68	7 (25%)
1	LLP	C	209	1	24,24,25	2.00	6 (25%)	28,32,34	2.29	9 (32%)

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	LLP	A	209	1	-	0/15/17/19	0/1/1/1
1	LLP	C	209	1	-	0/15/17/19	0/1/1/1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	209	LLP	O3-C3	-5.91	1.23	1.37
1	C	209	LLP	O3-C3	-5.75	1.23	1.37
1	A	209	LLP	C6-N1	2.25	1.39	1.34
1	C	209	LLP	C6-N1	2.33	1.39	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	209	LLP	C4'-NZ	2.38	1.34	1.27

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	209	LLP	C5-C6-N1	-3.38	118.15	123.87
1	A	209	LLP	CE-NZ-C4'	-3.26	109.57	119.03
1	C	209	LLP	C5'-C5-C4	-3.15	116.16	121.66
1	C	209	LLP	CE-NZ-C4'	-2.88	110.66	119.03
1	A	209	LLP	C5-C6-N1	-2.61	119.45	123.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	209	LLP	3	0
1	C	209	LLP	4	0

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	BTB	C	393	-	13,13,13	0.43	0	9,16,16	1.19	2 (22%)

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	BTB	C	393	-	-	0/21/21/21	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	393	BTB	C7-N-C2	2.13	120.00	113.70
3	C	393	BTB	C5-N-C2	2.69	121.66	113.70

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	C	393	BTB	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	386/394 (97%)	-0.34	0 100 100	16, 26, 39, 52	0
1	C	388/394 (98%)	-0.29	1 (0%) 93 97	20, 32, 47, 57	0
2	B	289/328 (88%)	-0.29	2 (0%) 87 93	18, 28, 42, 63	0
2	D	292/328 (89%)	-0.15	1 (0%) 93 97	21, 36, 57, 73	0
All	All	1355/1444 (93%)	-0.27	4 (0%) 93 97	16, 30, 49, 73	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	443	ALA	3.2
1	C	4	HIS	2.7
2	B	590	ARG	2.3
2	B	593	ARG	2.1

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	LLP	A	209	24/25	0.94	0.18	-	21,35,39,40	0
1	LLP	C	209	24/25	0.92	0.20	-	24,34,42,43	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	BTB	C	393	14/14	0.85	0.21	3.53	59,61,61,63	0

6.5 Other polymers [i](#)

There are no such residues in this entry.