



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 09:07 PM GMT

PDB ID : 4UVC
Title : LSD1(KDM1A)-CoREST in complex with 1-Phenyl-Tranylcypromine
Authors : Vianello, P.; Botrugno, O.; Cappa, A.; Ciossani, G.; Dessanti, P.; Mai, A.;
Mattevi, A.; Meroni, G.; Minucci, S.; Thaler, F.; Tortorici, M.; Trifiro, P.;
Valente, S.; Villa, M.; Varasi, M.; Mercurio, C.
Deposited on : 2014-08-05
Resolution : 3.10 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

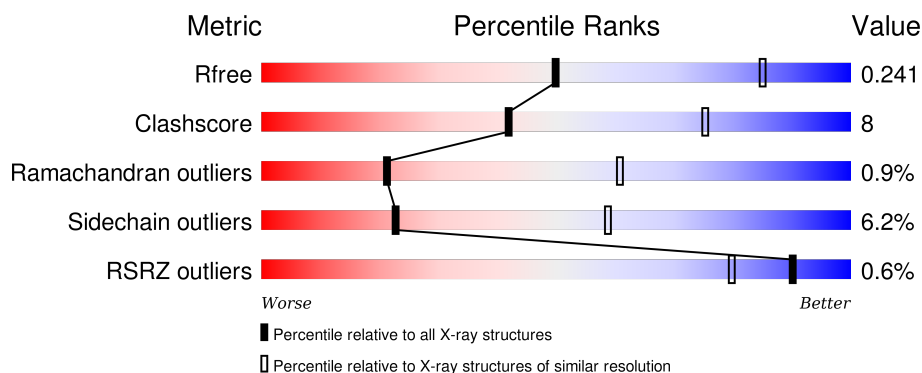
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	872	<div> <div> <div></div> <div>59%</div> <div>16%</div> <div>•</div> <div>24%</div> </div> </div>
2	B	482	<div> <div> <div>21%</div> <div>5%</div> <div>•</div> <div>72%</div> </div> </div>

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
3	D52	A	900	X	-	-	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6361 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 LYSINE-SPECIFIC HISTONE DEMETHYLASE 1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	666	5216	3323	906	967	20	0	0	0

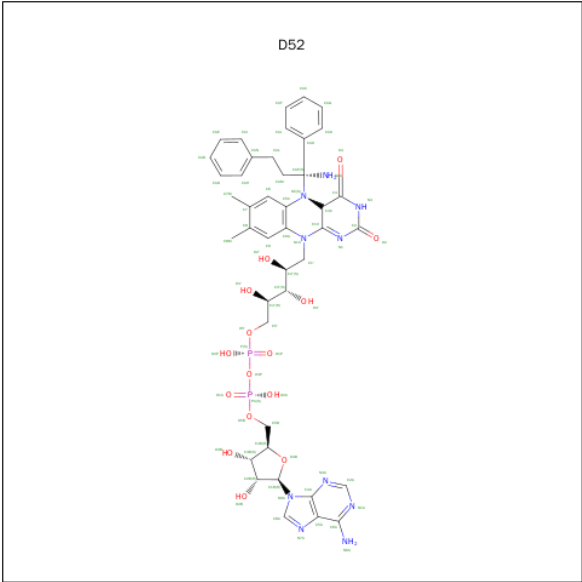
There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	171	PRO	ALA	CONFLICT	UNP O60341
A	.	-	ASP	DELETION	UNP O60341
A	.	-	THR	DELETION	UNP O60341
A	.	-	VAL	DELETION	UNP O60341
A	.	-	LYS	DELETION	UNP O60341

- Molecule 2 is a protein called REST COREPRESSOR 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	133	1076	676	194	203	3	0	0	0

- Molecule 3 is [(2R,3S,4R,5R)-5-(6-AMINOPURIN-9-YL)-3,4-BIS(OXIDANYL)OXOLAN-2-YL]METHOXY-OXIDANYL-PHOSPHORYL [(2R,3S,4S)-5-[5-[(1S)-1-AZANYL-1,3-DIPHENYL-PROPYL]-7,8-DIMETHYL-2,4-BIS(OXIDANYLIDENE)-4AH-BENZO[G]PTERIDIN-10-YL]-2,3,4-TRIS(OXIDANYL)PENTYL] HYDROGEN PHOSPHATE (three-letter code: D52) (formula: C₄₂H₅₀N₁₀O₁₅P₂).

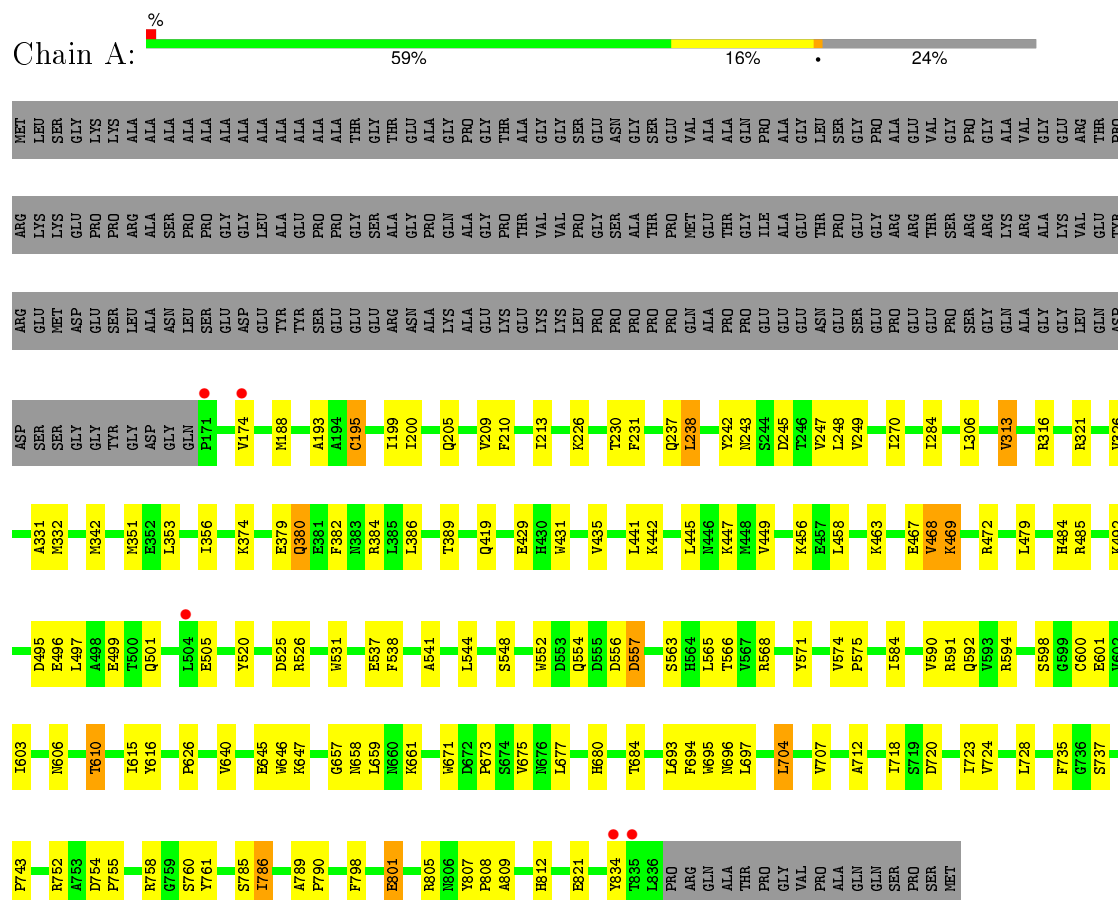


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			69	42	10	15	2		

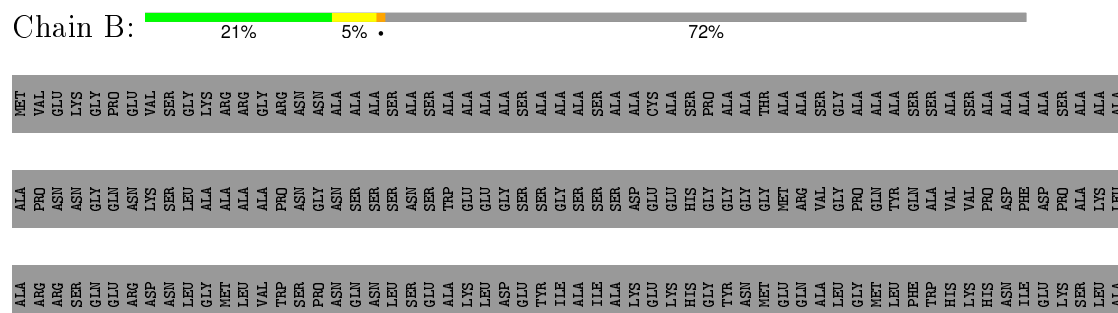
3 Residue-property plots

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: LYSINE-SPECIFIC HISTONE DEMETHYLASE 1A



• Molecule 2: REST COREPRESSOR 1



ASP	SER	GLN	HIS
LEU	VAL	ALA	GLY
PRO	MET	LYS	LYS
ASN	ASN	ASN	GLU
PHE	ARG	ARG	GLU
THR	HIS	ALA	THR
PRO	ALA	LYS	ASN
PHE	ARG	R308	GLY
PRO	LYS	K312	PRO
ASP	GLN	G313	SER
GLU	LYS	M314	ASN
TRP	ARG	F315	GLN
THR	GLU	L316	LYS
VAL	ARG	S317	PRO
GLU	GLU	V321	VAL
ASP	GLU	V324	LYS
LYS	VAL	N327	SER
LEU	ASP	A330	ASP
PHE	GLU	A331	ASN
ALA	ALA	Q337	GLU
PHE	ASN	V343	GLU
PHE	GLY	R347	GLU
HIS	ASN	Q348	ALA
GLY	ASN	I349	PRO
LYS	PRO	T355	VAL
THR	ILE	N356	LEU
PHE	ASP	I367	ASP
HIS	ILE	Y370	VAL
ARG	GLU	R371	ARG
ILE	VAL	L372	TYR
GLN	GLN	P373	ALA
ASP	ASN	C379	SER
GLN	LYS	A404	GLU
MET	LEU	V417	PRO
LEU	PRO	F421	THR
PRO	SER	R425	VAL
ASP	LEU	N429	THR
LYS	LYS	L434	VAL
SER	LYS	W437	THR
ILE	GLU	E440	THR
ALA	VAL		
SER	PRO		
LEU	PRO		
VAL	THR		
LYS	GLU		
PHE	THR		
TYR	VAL		
TYR	PRO		
SER	GLN		
LYS	VAL		
LYS	TRP		
THR	LYS		
ARG	GLU		
THR	LYS		
LYS	HIS		
THR	SER		
THR	THR		

4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	118.96Å 179.09Å 234.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.53 – 3.10 65.38 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (68.53-3.10) 100.0 (65.38-3.10)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.64 (at 3.13Å)	Xtriage
Refinement program	REFMAC 5.7.0027	Depositor
R, R_{free}	0.210 , 0.241 0.212 , 0.241	Depositor DCC
R_{free} test set	886 reflections (1.97%)	DCC
Wilson B-factor (Å ²)	78.0	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 57.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	0 of 45852 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	6361	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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: D52

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	0.46	3/5329 (0.1%)	0.55	0/7229
2	B	0.40	1/1091 (0.1%)	0.52	0/1471
All	All	0.45	4/6420 (0.1%)	0.55	0/8700

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	437	TRP	CD2-CE2	5.20	1.47	1.41
1	A	531	TRP	CD2-CE2	5.17	1.47	1.41
1	A	646	TRP	CD2-CE2	5.07	1.47	1.41
1	A	431	TRP	CD2-CE2	5.04	1.47	1.41

There are no bond angle outliers.

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	5216	0	5246	94	0
2	B	1076	0	1091	22	0
3	A	69	0	47	7	0
All	All	6361	0	6384	107	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 8.

The worst 5 of 107 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:199:ILE:HD11	1:A:248:LEU:HD11	1.50	0.93
1:A:356:ILE:HD11	1:A:566:THR:HG23	1.66	0.76
1:A:456:LYS:HA	2:B:370:TYR:CE2	2.26	0.69
1:A:693:LEU:HD12	1:A:694:PHE:H	1.59	0.66
1:A:374:LYS:HE3	1:A:525:ASP:OD1	1.95	0.66

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	664/872 (76%)	613 (92%)	45 (7%)	6 (1%)	21	61
2	B	131/482 (27%)	117 (89%)	13 (10%)	1 (1%)	24	63
All	All	795/1354 (59%)	730 (92%)	58 (7%)	7 (1%)	21	61

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	313	VAL
2	B	429	ASN
1	A	801	GLU
1	A	554	GLN
1	A	805	ARG

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	565/711 (80%)	532 (94%)	33 (6%)	25	61
2	B	117/395 (30%)	108 (92%)	9 (8%)	16	50
All	All	682/1106 (62%)	640 (94%)	42 (6%)	23	59

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	568	ARG
1	A	610	THR
2	B	348	GLN
1	A	571	TYR
1	A	598	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	461	GLN
2	B	337	GLN
1	A	633	GLN
1	A	380	GLN
1	A	484	HIS

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](#)

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	D52	A	900	-	64,76,76	1.98	7 (10%)	72,115,115	1.29	3 (4%)

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	D52	A	900	-	1/1/16/17	0/39/100/100	0/7/8/8

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	900	D52	CAP-CAO	-10.59	1.42	1.52
3	A	900	D52	CAL-CAN	-4.43	1.39	1.51
3	A	900	D52	C10-N1	2.43	1.37	1.31
3	A	900	D52	C8-C7	2.95	1.48	1.41
3	A	900	D52	C5A-C4A	3.60	1.48	1.40

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	900	D52	N3A-C2A-N1A	-7.18	123.39	128.89
3	A	900	D52	C1B-N9A-C4A	-3.16	122.17	126.94
3	A	900	D52	PA-O3P-P	-2.89	124.61	132.73

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	900	D52	C4X

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	900	D52	7	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	666/872 (76%)	0.03	5 (0%) 87 75	42, 79, 120, 191	0
2	B	133/482 (27%)	0.25	0 100 100	72, 108, 133, 154	0
All	All	799/1354 (59%)	0.07	5 (0%) 90 80	42, 85, 126, 191	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	835	THR	4.9
1	A	834	TYR	3.1
1	A	504	LEU	2.4
1	A	171	PRO	2.3
1	A	174	VAL	2.2

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	D52	A	900	69/69	0.98	0.24	-0.07	49,68,111,119	0

6.5 Other polymers ⓘ

There are no such residues in this entry.