



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2017 – 09:57 AM EDT

PDB ID : 2OBC
Title : The crystal structure of RibD from Escherichia coli in complex with a substrate analogue, ribose 5-phosphate (beta form), bound to the active site of the reductase domain
Authors : Moche, M.; Stenmark, P.; Gurmu, D.; Nordlund, P.; Structural Proteomics in Europe (SPINE)
Deposited on : unknown
Resolution : 3.00 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 : rb-20030345
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) : rb-20030345

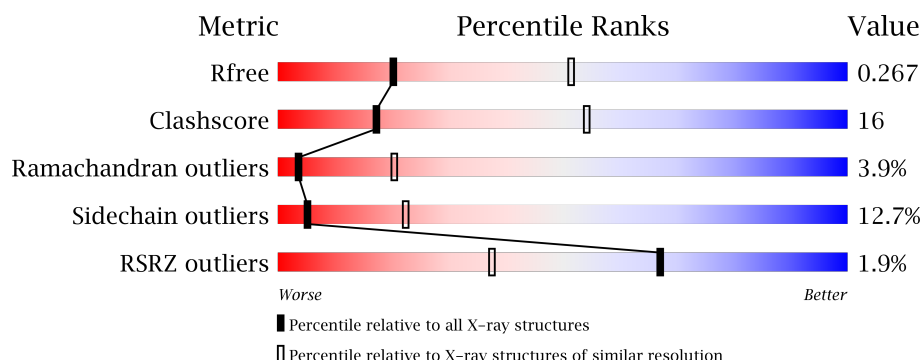
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (3.00-3.00)

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	402	<div> <div>0.5%</div> <div> <div></div> <div>56%</div> <div>26%</div> <div>6%</div> <div>12%</div> </div> </div>
1	B	402	<div> <div>2%</div> <div> <div></div> <div>58%</div> <div>26%</div> <div>6%</div> <div>10%</div> </div> </div>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5534 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 Riboflavin biosynthesis protein ribD.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	354	Total	C	N	O	S	Se	0	0	0
			2719	1711	494	499	6	9			
1	B	362	Total	C	N	O	S	Se	0	0	0
			2753	1729	501	511	3	9			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-26	MET	-	CLONING ARTIFACT	UNP P25539
A	-25	ASP	-	CLONING ARTIFACT	UNP P25539
A	-24	TYR	-	CLONING ARTIFACT	UNP P25539
A	-23	LYS	-	CLONING ARTIFACT	UNP P25539
A	-22	ASP	-	CLONING ARTIFACT	UNP P25539
A	-21	ASP	-	CLONING ARTIFACT	UNP P25539
A	-20	ASP	-	CLONING ARTIFACT	UNP P25539
A	-19	ASP	-	CLONING ARTIFACT	UNP P25539
A	-18	LYS	-	CLONING ARTIFACT	UNP P25539
A	-17	GLY	-	CLONING ARTIFACT	UNP P25539
A	-16	SER	-	CLONING ARTIFACT	UNP P25539
A	-15	SER	-	CLONING ARTIFACT	UNP P25539
A	-14	THR	-	CLONING ARTIFACT	UNP P25539
A	-13	SER	-	CLONING ARTIFACT	UNP P25539
A	-12	LEU	-	CLONING ARTIFACT	UNP P25539
A	-11	TYR	-	CLONING ARTIFACT	UNP P25539
A	-10	LYS	-	CLONING ARTIFACT	UNP P25539
A	-9	LYS	-	CLONING ARTIFACT	UNP P25539
A	-8	ALA	-	CLONING ARTIFACT	UNP P25539
A	-7	GLY	-	CLONING ARTIFACT	UNP P25539
A	-6	SER	-	CLONING ARTIFACT	UNP P25539
A	-5	GLU	-	CLONING ARTIFACT	UNP P25539
A	-4	THR	-	CLONING ARTIFACT	UNP P25539
A	-3	LEU	-	CLONING ARTIFACT	UNP P25539
A	-2	TYR	-	CLONING ARTIFACT	UNP P25539

Continued on next page...

Continued from previous page...

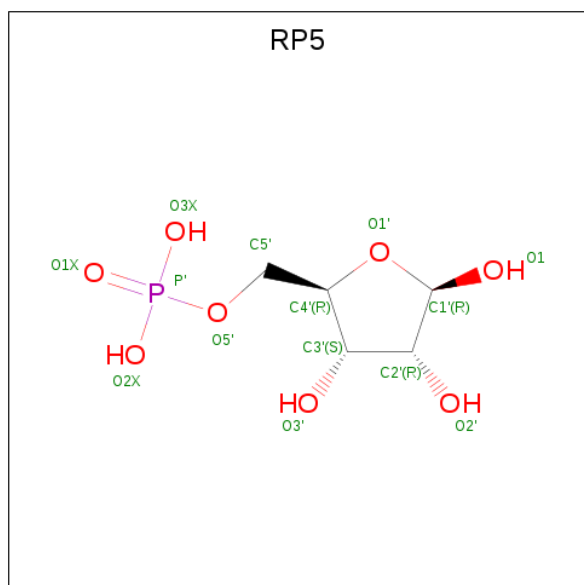
Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ILE	-	CLONING ARTIFACT	UNP P25539
A	0	GLN	-	CLONING ARTIFACT	UNP P25539
A	1	GLY	-	CLONING ARTIFACT	UNP P25539
A	368	SER	-	EXPRESSION TAG	UNP P25539
A	369	THR	-	EXPRESSION TAG	UNP P25539
A	370	HIS	-	EXPRESSION TAG	UNP P25539
A	371	HIS	-	EXPRESSION TAG	UNP P25539
A	372	HIS	-	EXPRESSION TAG	UNP P25539
A	373	HIS	-	EXPRESSION TAG	UNP P25539
A	374	HIS	-	EXPRESSION TAG	UNP P25539
A	375	HIS	-	EXPRESSION TAG	UNP P25539
B	-26	MET	-	CLONING ARTIFACT	UNP P25539
B	-25	ASP	-	CLONING ARTIFACT	UNP P25539
B	-24	TYR	-	CLONING ARTIFACT	UNP P25539
B	-23	LYS	-	CLONING ARTIFACT	UNP P25539
B	-22	ASP	-	CLONING ARTIFACT	UNP P25539
B	-21	ASP	-	CLONING ARTIFACT	UNP P25539
B	-20	ASP	-	CLONING ARTIFACT	UNP P25539
B	-19	ASP	-	CLONING ARTIFACT	UNP P25539
B	-18	LYS	-	CLONING ARTIFACT	UNP P25539
B	-17	GLY	-	CLONING ARTIFACT	UNP P25539
B	-16	SER	-	CLONING ARTIFACT	UNP P25539
B	-15	SER	-	CLONING ARTIFACT	UNP P25539
B	-14	THR	-	CLONING ARTIFACT	UNP P25539
B	-13	SER	-	CLONING ARTIFACT	UNP P25539
B	-12	LEU	-	CLONING ARTIFACT	UNP P25539
B	-11	TYR	-	CLONING ARTIFACT	UNP P25539
B	-10	LYS	-	CLONING ARTIFACT	UNP P25539
B	-9	LYS	-	CLONING ARTIFACT	UNP P25539
B	-8	ALA	-	CLONING ARTIFACT	UNP P25539
B	-7	GLY	-	CLONING ARTIFACT	UNP P25539
B	-6	SER	-	CLONING ARTIFACT	UNP P25539
B	-5	GLU	-	CLONING ARTIFACT	UNP P25539
B	-4	THR	-	CLONING ARTIFACT	UNP P25539
B	-3	LEU	-	CLONING ARTIFACT	UNP P25539
B	-2	TYR	-	CLONING ARTIFACT	UNP P25539
B	-1	ILE	-	CLONING ARTIFACT	UNP P25539
B	0	GLN	-	CLONING ARTIFACT	UNP P25539
B	1	GLY	-	CLONING ARTIFACT	UNP P25539
B	368	SER	-	EXPRESSION TAG	UNP P25539
B	369	THR	-	EXPRESSION TAG	UNP P25539
B	370	HIS	-	EXPRESSION TAG	UNP P25539

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	371	HIS	-	EXPRESSION TAG	UNP P25539
B	372	HIS	-	EXPRESSION TAG	UNP P25539
B	373	HIS	-	EXPRESSION TAG	UNP P25539
B	374	HIS	-	EXPRESSION TAG	UNP P25539
B	375	HIS	-	EXPRESSION TAG	UNP P25539

- Molecule 2 is 5-O-phosphono-beta-D-ribofuranose (three-letter code: RP5) (formula: $C_5H_{11}O_8P$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	P	0	0
			14	5	8	1		

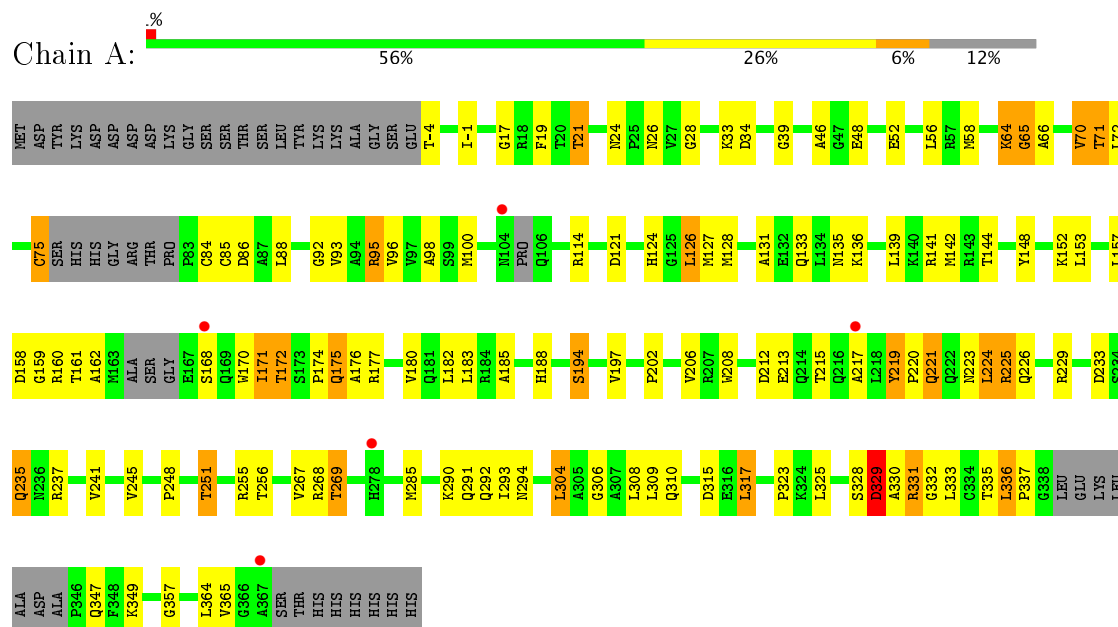
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	31	Total	O	0	0
			31	31		
3	B	17	Total	O	0	0
			17	17		

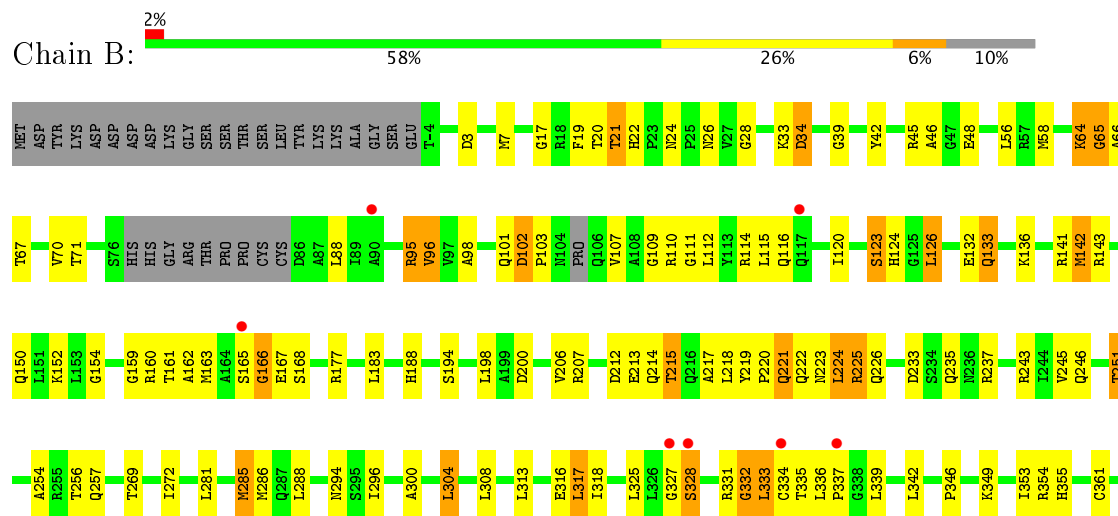
3 Residue-property plots

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: Riboflavin biosynthesis protein ribD



• Molecule 1: Riboflavin biosynthesis protein ribD



L364

A367

SER

THR

HIS

HIS

HIS

HIS

HIS

HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	173.10 Å 173.10 Å 76.70 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.86 – 3.00 28.85 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (28.86-3.00) 99.9 (28.85-3.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.22 (at 3.00 Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.222 , 0.276 0.214 , 0.267	Depositor DCC
R_{free} test set	1346 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	70.9	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 65.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.028 for -h,-k,l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5534	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% 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, 0.333 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: RP5

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	0/2761	0.67	1/3729 (0.0%)
1	B	0.45	1/2795 (0.0%)	0.64	0/3781
All	All	0.45	1/5556 (0.0%)	0.65	1/7510 (0.0%)

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
1	A	0	2
1	B	0	2
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	116	GLN	CD-OE1	5.33	1.35	1.24

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	317	LEU	CA-CB-CG	5.27	127.42	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	217	ALA	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	66	ALA	Peptide
1	B	217	ALA	Peptide
1	B	66	ALA	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	2719	0	2721	88	0
1	B	2753	0	2739	88	0
2	A	14	0	9	1	0
3	A	31	0	0	7	0
3	B	17	0	0	5	0
All	All	5534	0	5469	170	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 16.

The worst 5 of 170 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:B:194:SER:HB3	1:B:233:ASP:HB3	1.23	1.14
1:B:206:VAL:H	1:B:226:GLN:NE2	1.53	1.06
1:A:75:CYS:HA	1:A:84:CYS:HB3	1.40	1.01
1:B:213:GLU:HA	3:B:390:HOH:O	1.62	0.95
1:A:235:GLN:HG3	1:A:237:ARG:HD2	1.54	0.90

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	344/402 (86%)	308 (90%)	26 (8%)	10 (3%)	5	28
1	B	356/402 (89%)	310 (87%)	29 (8%)	17 (5%)	2	16
All	All	700/804 (87%)	618 (88%)	55 (8%)	27 (4%)	3	20

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	172	THR
1	A	329	ASP
1	B	107	VAL
1	B	162	ALA
1	B	167	GLU

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	284/321 (88%)	242 (85%)	42 (15%)	3	16
1	B	283/321 (88%)	253 (89%)	30 (11%)	8	30
All	All	567/642 (88%)	495 (87%)	72 (13%)	5	22

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	269	THR
1	A	349	LYS
1	B	285	MSE
1	A	285	MSE
1	A	329	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	292	GLN
1	B	133	GLN
1	B	235	GLN
1	A	291	GLN
1	B	246	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	$\# Z > 2$
2	RP5	A	401	-	14,14,14	0.57	0	20,21,21	1.30	4 (20%)

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
2	RP5	A	401	-	-	0/6/22/22	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	A	401	RP5	O1-C1'-O1'	-2.26	108.10	111.14
2	A	401	RP5	C1'-C2'-C3'	2.01	104.85	102.30
2	A	401	RP5	C2'-C3'-C4'	2.14	106.79	102.62
2	A	401	RP5	O1'-C1'-C2'	2.71	107.88	104.46

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	RP5	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	345/402 (85%)	-0.23	5 (1%) 75 49	43, 56, 75, 94	0
1	B	353/402 (87%)	-0.17	8 (2%) 61 31	44, 58, 73, 80	0
All	All	698/804 (86%)	-0.20	13 (1%) 67 37	43, 58, 75, 94	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	334	CYS	4.0
1	B	337	PRO	3.6
1	B	328	SER	2.8
1	A	278	HIS	2.4
1	B	327	GLY	2.4

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
2	RP5	A	401	14/14	0.91	0.19	-0.55	75,81,82,82	0

6.5 Other polymers [i](#)

There are no such residues in this entry.