



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 10:28 am BST

PDB ID : 3QML  
Title : The structural analysis of Sil1-Bip complex reveals the mechanism for Sil1 to function as a novel nucleotide exchange factor  
Authors : Yan, M.; Li, J.Z.; Sha, B.D.  
Deposited on : 2011-02-04  
Resolution : 2.31 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

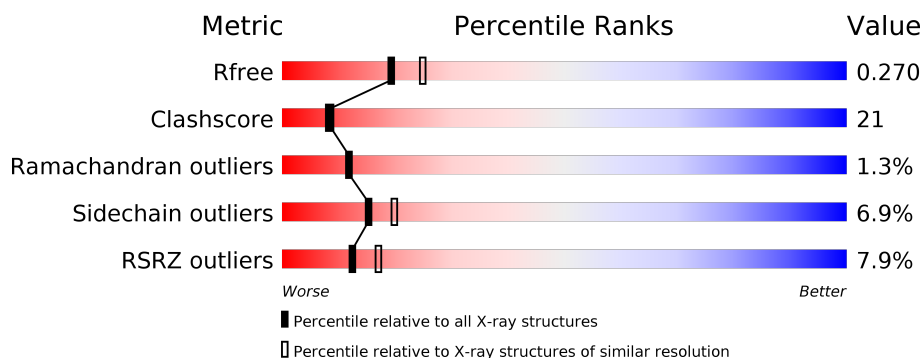
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.31 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))
$R_{free}$	130704	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	390	<div> <div>77%</div> <div>18%</div> <div>• •</div> </div>
1	B	390	<div> <div>20%</div> <div>53%</div> <div>37%</div> <div>5%</div> <div>5%</div> </div>
2	C	315	<div> <div>2%</div> <div>58%</div> <div>19%</div> <div>6%</div> <div>17%</div> </div>
2	D	315	<div> <div>5%</div> <div>49%</div> <div>34%</div> <div>•</div> <div>13%</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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PO4	A	4	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10378 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 78 kDa glucose-regulated protein homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	377	Total	C	N	O	S	0	0	0
			2893	1825	495	568	5			
1	B	370	Total	C	N	O	S	0	0	0
			2830	1788	483	554	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	GLY	-	EXPRESSION TAG	UNP P16474
A	38	MET	-	EXPRESSION TAG	UNP P16474
A	39	SER	-	EXPRESSION TAG	UNP P16474
A	40	HIS	-	EXPRESSION TAG	UNP P16474
A	41	ALA	-	EXPRESSION TAG	UNP P16474
A	42	SER	-	EXPRESSION TAG	UNP P16474
B	37	GLY	-	EXPRESSION TAG	UNP P16474
B	38	MET	-	EXPRESSION TAG	UNP P16474
B	39	SER	-	EXPRESSION TAG	UNP P16474
B	40	HIS	-	EXPRESSION TAG	UNP P16474
B	41	ALA	-	EXPRESSION TAG	UNP P16474
B	42	SER	-	EXPRESSION TAG	UNP P16474

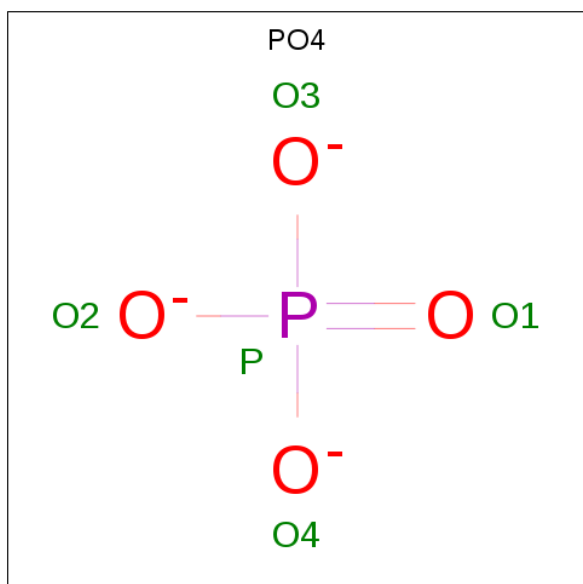
- Molecule 2 is a protein called Nucleotide exchange factor SIL1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	262	Total	C	N	O	S	0	0	0
			2160	1367	367	419	7			
2	D	273	Total	C	N	O	S	0	0	0
			2246	1423	384	432	7			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	107	GLY	-	EXPRESSION TAG	UNP Q08199
C	108	MET	-	EXPRESSION TAG	UNP Q08199
C	109	SER	-	EXPRESSION TAG	UNP Q08199
C	110	HIS	-	EXPRESSION TAG	UNP Q08199
C	111	ALA	-	EXPRESSION TAG	UNP Q08199
C	112	SER	-	EXPRESSION TAG	UNP Q08199
D	107	GLY	-	EXPRESSION TAG	UNP Q08199
D	108	MET	-	EXPRESSION TAG	UNP Q08199
D	109	SER	-	EXPRESSION TAG	UNP Q08199
D	110	HIS	-	EXPRESSION TAG	UNP Q08199
D	111	ALA	-	EXPRESSION TAG	UNP Q08199
D	112	SER	-	EXPRESSION TAG	UNP Q08199

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	A	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total	O	P	0	0
			5	4	1		
3	D	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Mg	0	0
			2	2		
4	C	1	Total	Mg	0	0
			1	1		

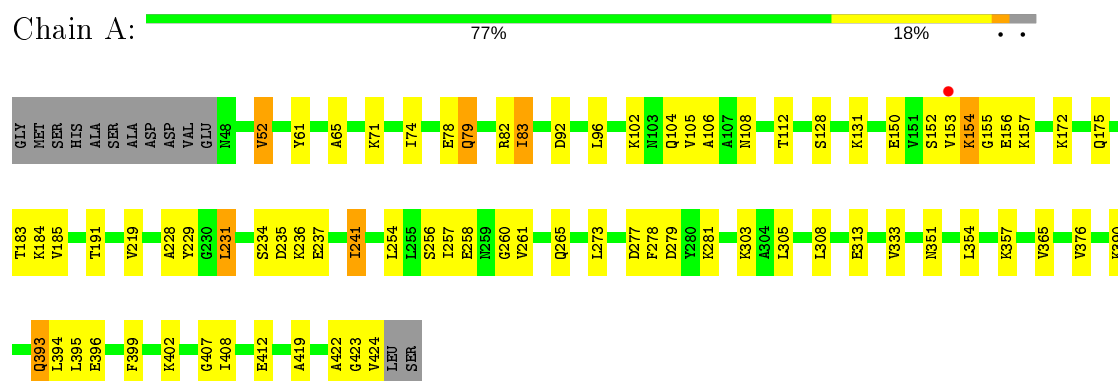
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	96	Total	O	0	0
			96	96		
5	B	48	Total	O	0	0
			48	48		
5	C	34	Total	O	0	0
			34	34		
5	D	28	Total	O	0	0
			28	28		

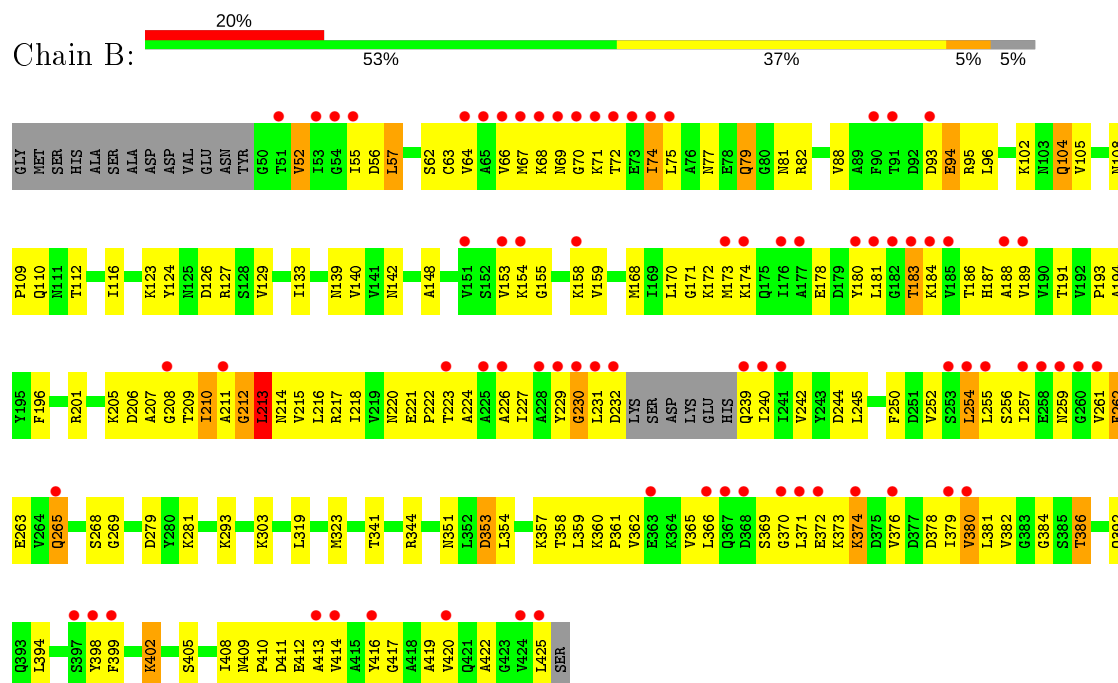
### 3 Residue-property plots

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: 78 kDa glucose-regulated protein homolog

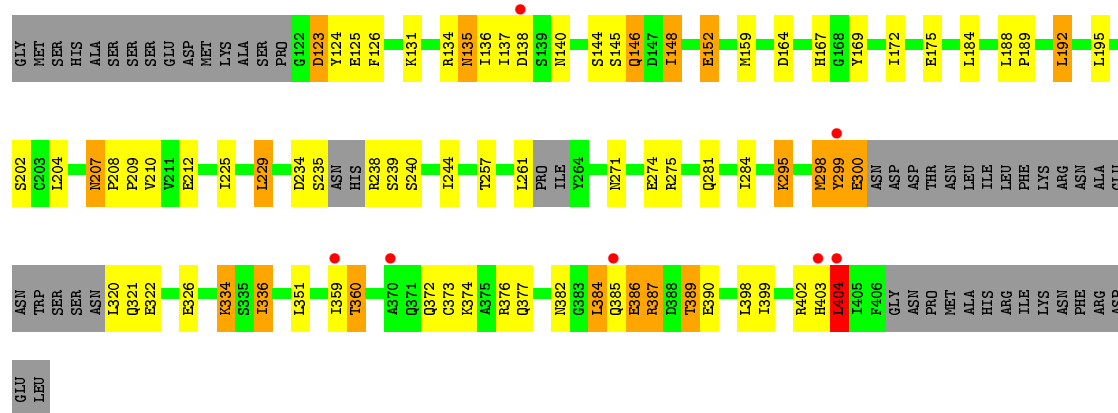


- Molecule 1: 78 kDa glucose-regulated protein homolog

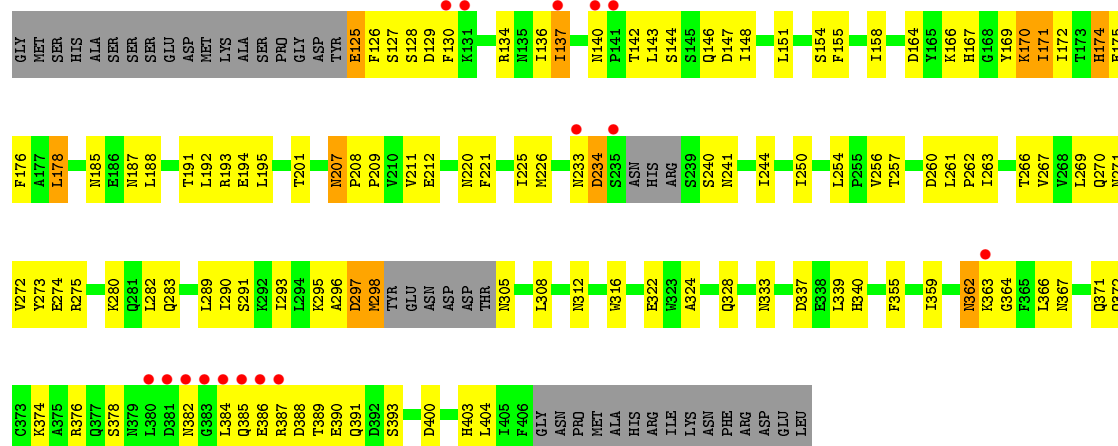


- Molecule 2: Nucleotide exchange factor SIL1





### • Molecule 2: Nucleotide exchange factor SIL1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	226.35Å 116.59Å 55.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.13 – 2.31 46.13 – 2.31	Depositor EDS
% Data completeness (in resolution range)	91.5 (46.13-2.31) 91.5 (46.13-2.31)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.69 (at 2.32Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.1_351)	Depositor
R, $R_{free}$	0.210 , 0.271 0.210 , 0.270	Depositor DCC
$R_{free}$ test set	3052 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	46.1	Xtriage
Anisotropy	0.408	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 57.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10378	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, MG

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	0/2931	0.60	0/3952
1	B	0.36	0/2865	0.53	0/3863
2	C	0.41	0/2194	0.57	1/2954 (0.0%)
2	D	0.37	0/2283	0.54	1/3078 (0.0%)
All	All	0.40	0/10273	0.56	2/13847 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	234	ASP	CB-CG-OD2	5.21	122.99	118.30
2	D	234	ASP	CB-CG-OD2	5.16	122.94	118.30

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	2893	0	2950	69	0
1	B	2830	0	2897	184	0
2	C	2160	0	2142	72	0
2	D	2246	0	2240	102	0
3	A	10	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	10	0	0	1	0
3	C	10	0	0	0	0
3	D	10	0	0	1	0
4	A	2	0	0	0	0
4	C	1	0	0	0	0
5	A	96	0	0	3	0
5	B	48	0	0	9	0
5	C	34	0	0	4	0
5	D	28	0	0	2	0
All	All	10378	0	10229	422	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 21.

The worst 5 of 422 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:213:LEU:HD12	1:B:214:ASN:H	1.14	1.10
1:A:234:SER:HA	1:A:235:ASP:HB2	1.32	1.05
1:A:52:VAL:HG22	1:A:422:ALA:HB1	1.38	1.00
2:D:233:ASN:ND2	2:D:271:ASN:ND2	2.12	0.97
2:D:382:ASN:HB2	2:D:387:ARG:HH22	1.29	0.95

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	375/390 (96%)	361 (96%)	13 (4%)	1 (0%)	41 50
1	B	366/390 (94%)	327 (89%)	32 (9%)	7 (2%)	8 6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	254/315 (81%)	242 (95%)	8 (3%)	4 (2%)	9	8
2	D	267/315 (85%)	241 (90%)	21 (8%)	5 (2%)	8	6
All	All	1262/1410 (90%)	1171 (93%)	74 (6%)	17 (1%)	12	12

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	79	GLN
2	C	404	LEU
2	D	389	THR
1	B	213	LEU
1	B	230	GLY

### 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	313/323 (97%)	297 (95%)	16 (5%)	24	33
1	B	306/323 (95%)	285 (93%)	21 (7%)	15	20
2	C	248/295 (84%)	225 (91%)	23 (9%)	9	10
2	D	259/295 (88%)	241 (93%)	18 (7%)	15	20
All	All	1126/1236 (91%)	1048 (93%)	78 (7%)	15	20

5 of 78 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	386	THR
2	C	192	LEU
2	D	289	LEU
1	B	402	LYS
2	C	146	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	285	GLN
2	C	167	HIS
2	D	362	ASN
1	B	351	ASN
2	C	135	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](#)

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](#)

Of 11 ligands modelled in this entry, 3 are monoatomic - leaving 8 for Mogul analysis.

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	PO4	A	4	-	4,4,4	1.29	0	6,6,6	0.79	0
3	PO4	D	8	-	4,4,4	0.83	0	6,6,6	0.63	0
3	PO4	B	3	-	4,4,4	1.22	0	6,6,6	0.75	0
3	PO4	B	1	-	4,4,4	0.89	0	6,6,6	0.44	0
3	PO4	C	6	-	4,4,4	0.82	0	6,6,6	0.50	0
3	PO4	C	5	-	4,4,4	0.85	0	6,6,6	0.47	0
3	PO4	A	2	-	4,4,4	1.07	0	6,6,6	0.60	0
3	PO4	D	7	-	4,4,4	0.80	0	6,6,6	0.56	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	4	PO4	3	0
3	B	3	PO4	1	0
3	D	7	PO4	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	377/390 (96%)	0.08	1 (0%) 94 96	27, 43, 78, 115	0
1	B	370/390 (94%)	1.04	77 (20%) 1 1	35, 80, 165, 252	0
2	C	262/315 (83%)	0.19	7 (2%) 54 62	35, 55, 95, 131	0
2	D	273/315 (86%)	0.41	16 (5%) 22 29	36, 67, 128, 181	0
All	All	1282/1410 (90%)	0.45	101 (7%) 12 17	27, 59, 134, 252	0

The worst 5 of 101 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	183	THR	19.6
1	B	366	LEU	8.9
1	B	69	ASN	8.6
1	B	71	LYS	8.0
1	B	260	GLY	7.7

### 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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PO4	C	6	5/5	0.81	0.14	109,109,115,116	0
4	MG	A	427	1/1	0.88	0.32	64,64,64,64	0
3	PO4	D	8	5/5	0.91	0.14	98,100,106,107	0
3	PO4	D	7	5/5	0.91	0.10	96,100,107,108	0
4	MG	A	1	1/1	0.93	0.27	62,62,62,62	0
3	PO4	A	4	5/5	0.94	0.23	29,42,56,66	0
3	PO4	B	3	5/5	0.94	0.22	33,41,60,70	0
3	PO4	C	5	5/5	0.96	0.11	69,70,83,84	0
3	PO4	B	1	5/5	0.97	0.19	53,78,85,88	0
4	MG	C	1	1/1	0.98	0.11	58,58,58,58	0
3	PO4	A	2	5/5	0.99	0.20	46,46,49,70	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.