



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 2, 2017 – 12:11 AM EDT

PDB ID : 3SY8  
Title : Crystal structure of the response regulator RocR  
Authors : Chen, M.W.; Kotaka, M.; Vonnrhein, C.; Bricogne, G.; Lescar, J.  
Deposited on : unknown  
Resolution : 2.50 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

<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



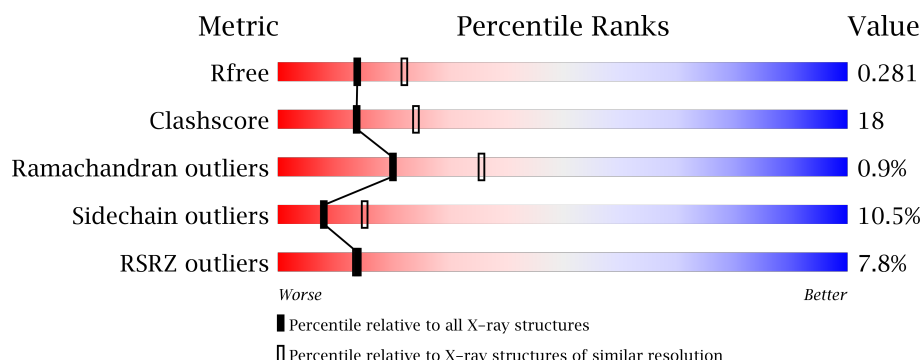
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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. 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	400	<div> <div>5%</div> <div> <div></div> <div>66%</div> <div>27%</div> <div>• • •</div> </div> </div>
1	B	400	<div> <div>5%</div> <div> <div></div> <div>72%</div> <div>22%</div> <div>• • •</div> </div> </div>
1	C	400	<div> <div>6%</div> <div> <div></div> <div>62%</div> <div>29%</div> <div>• 5%</div> </div> </div>
1	D	400	<div> <div>13%</div> <div> <div></div> <div>59%</div> <div>28%</div> <div>6% 7%</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 crite-



ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MG	B	401	-	-	-	X



## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11886 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 RocR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	389	Total	C	N	O	S	0	0	0
			2979	1899	509	554	17			
1	B	392	Total	C	N	O	S	0	0	0
			3007	1918	513	558	18			
1	C	380	Total	C	N	O	S	0	0	0
			2906	1856	493	540	17			
1	D	371	Total	C	N	O	S	0	0	0
			2807	1792	477	522	16			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	286	TRP	ARG	CONFLICT	UNP Q9HX69
A	393	LEU	-	EXPRESSION TAG	UNP Q9HX69
A	394	GLU	-	EXPRESSION TAG	UNP Q9HX69
A	395	HIS	-	EXPRESSION TAG	UNP Q9HX69
A	396	HIS	-	EXPRESSION TAG	UNP Q9HX69
A	397	HIS	-	EXPRESSION TAG	UNP Q9HX69
A	398	HIS	-	EXPRESSION TAG	UNP Q9HX69
A	399	HIS	-	EXPRESSION TAG	UNP Q9HX69
A	400	HIS	-	EXPRESSION TAG	UNP Q9HX69
B	286	TRP	ARG	CONFLICT	UNP Q9HX69
B	393	LEU	-	EXPRESSION TAG	UNP Q9HX69
B	394	GLU	-	EXPRESSION TAG	UNP Q9HX69
B	395	HIS	-	EXPRESSION TAG	UNP Q9HX69
B	396	HIS	-	EXPRESSION TAG	UNP Q9HX69
B	397	HIS	-	EXPRESSION TAG	UNP Q9HX69
B	398	HIS	-	EXPRESSION TAG	UNP Q9HX69
B	399	HIS	-	EXPRESSION TAG	UNP Q9HX69
B	400	HIS	-	EXPRESSION TAG	UNP Q9HX69
C	286	TRP	ARG	CONFLICT	UNP Q9HX69
C	393	LEU	-	EXPRESSION TAG	UNP Q9HX69
C	394	GLU	-	EXPRESSION TAG	UNP Q9HX69

*Continued on next page...*



*Continued from previous page...*

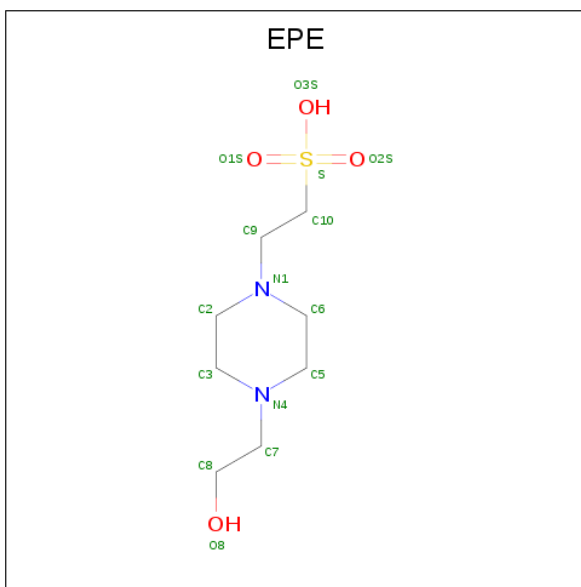
Chain	Residue	Modelled	Actual	Comment	Reference
C	395	HIS	-	EXPRESSION TAG	UNP Q9HX69
C	396	HIS	-	EXPRESSION TAG	UNP Q9HX69
C	397	HIS	-	EXPRESSION TAG	UNP Q9HX69
C	398	HIS	-	EXPRESSION TAG	UNP Q9HX69
C	399	HIS	-	EXPRESSION TAG	UNP Q9HX69
C	400	HIS	-	EXPRESSION TAG	UNP Q9HX69
D	286	TRP	ARG	CONFLICT	UNP Q9HX69
D	393	LEU	-	EXPRESSION TAG	UNP Q9HX69
D	394	GLU	-	EXPRESSION TAG	UNP Q9HX69
D	395	HIS	-	EXPRESSION TAG	UNP Q9HX69
D	396	HIS	-	EXPRESSION TAG	UNP Q9HX69
D	397	HIS	-	EXPRESSION TAG	UNP Q9HX69
D	398	HIS	-	EXPRESSION TAG	UNP Q9HX69
D	399	HIS	-	EXPRESSION TAG	UNP Q9HX69
D	400	HIS	-	EXPRESSION TAG	UNP Q9HX69

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	D	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 4 is water.

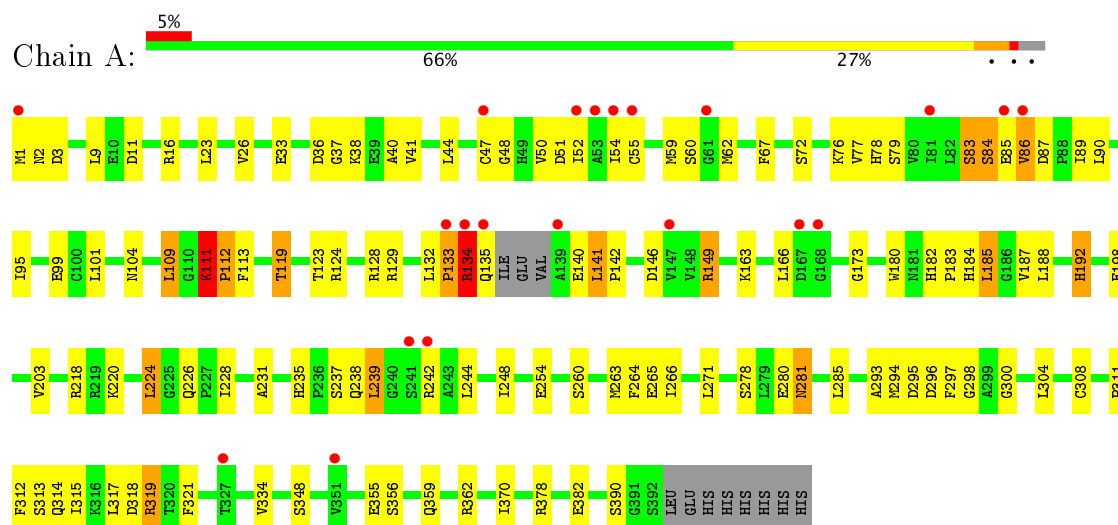
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	35	Total	O	0	0
			35	35		
4	B	53	Total	O	0	0
			53	53		
4	C	33	Total	O	0	0
			33	33		
4	D	47	Total	O	0	0
			47	47		



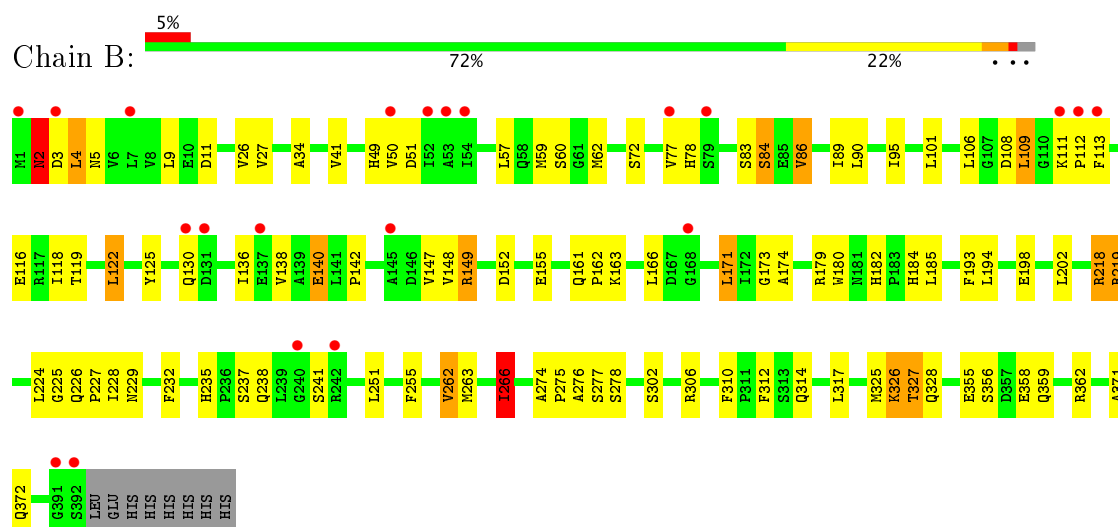
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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: RocR



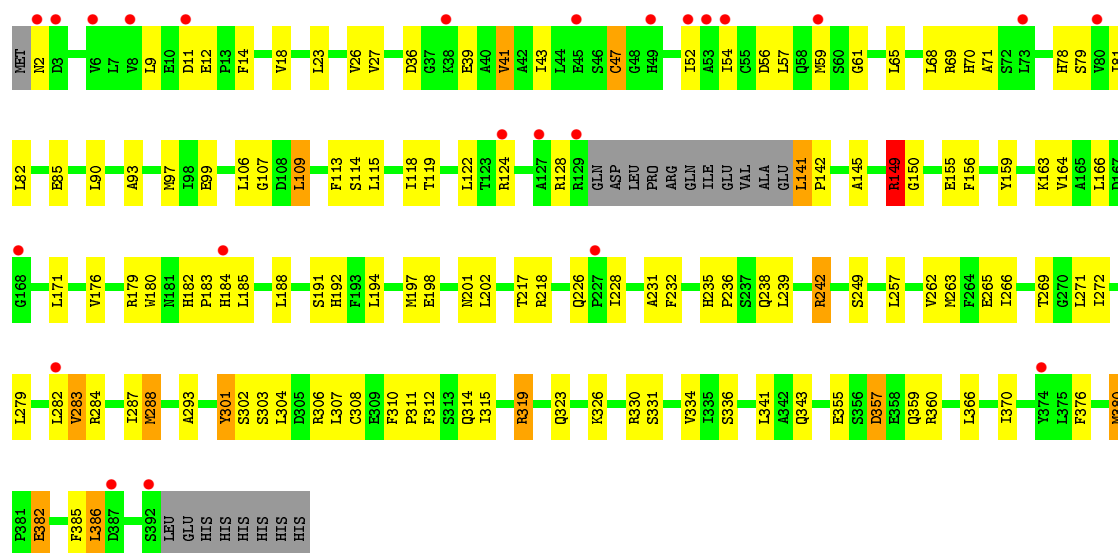
#### • Molecule 1: RocR



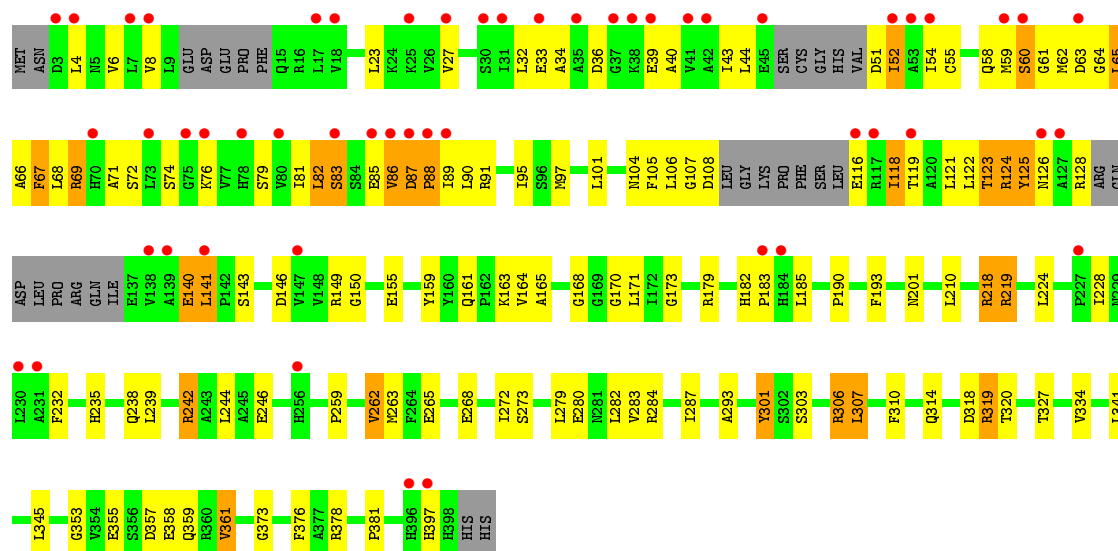
#### • Molecule 1: RocR







• Molecule 1: RocR





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.80 Å   118.80 Å   495.10 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	19.96 – 2.50 19.96 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.96-2.50) 83.3 (19.96-2.50)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.73 (at 2.50 Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.216   ,   0.285 0.211   ,   0.281	Depositor DCC
$R_{free}$ test set	3044 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	65.5	Xtriage
Anisotropy	0.095	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 44.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11886	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% 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: MG, EPE

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	0.71	2/3036 (0.1%)	0.79	1/4118 (0.0%)
1	B	0.81	0/3065	0.85	2/4158 (0.0%)
1	C	0.68	2/2962 (0.1%)	0.75	3/4018 (0.1%)
1	D	0.73	0/2855	0.86	3/3871 (0.1%)
All	All	0.73	4/11918 (0.0%)	0.81	9/16165 (0.1%)

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	1
1	C	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	308	CYS	CB-SG	-7.81	1.69	1.82
1	A	55	CYS	CB-SG	-7.13	1.70	1.82
1	A	280	GLU	CG-CD	5.48	1.60	1.51
1	C	319	ARG	CG-CD	5.46	1.65	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	179	ARG	NE-CZ-NH1	9.19	124.89	120.30
1	A	111	LYS	C-N-CD	-7.68	103.70	120.60
1	C	315	ILE	CB-CA-C	-6.98	97.64	111.60

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	306	ARG	NE-CZ-NH2	-5.90	117.35	120.30
1	C	61	GLY	N-CA-C	-5.73	98.78	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	111	LYS	Peptide
1	C	141	LEU	Peptide

## 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	2979	0	3002	91	0
1	B	3007	0	3040	88	0
1	C	2906	0	2930	106	0
1	D	2807	0	2811	153	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	D	15	0	17	2	0
4	A	35	0	0	7	0
4	B	53	0	0	5	0
4	C	33	0	0	2	0
4	D	47	0	0	4	0
All	All	11886	0	11800	417	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 18.

The worst 5 of 417 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:C:93:ALA:HB1	1:D:283:VAL:HG21	1.23	1.15

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:68:LEU:HG	1:D:69:ARG:NH1	1.70	1.07
1:D:66:ALA:N	1:D:69:ARG:HE	1.53	1.06
4:A:421:HOH:O	1:B:277:SER:HB3	1.50	1.06
1:D:125:TYR:HA	1:D:128:ARG:HB2	1.43	1.00

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	385/400 (96%)	356 (92%)	24 (6%)	5 (1%)	14	25
1	B	390/400 (98%)	370 (95%)	18 (5%)	2 (0%)	32	53
1	C	376/400 (94%)	349 (93%)	25 (7%)	2 (0%)	32	53
1	D	361/400 (90%)	338 (94%)	18 (5%)	5 (1%)	13	23
All	All	1512/1600 (94%)	1413 (94%)	85 (6%)	14 (1%)	20	36

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	111	LYS
1	A	112	PRO
1	A	133	PRO
1	A	134	ARG
1	D	67	PHE

### 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	324/337 (96%)	286 (88%)	38 (12%)	6	12
1	B	328/337 (97%)	298 (91%)	30 (9%)	11	21
1	C	317/337 (94%)	289 (91%)	28 (9%)	12	22
1	D	300/337 (89%)	263 (88%)	37 (12%)	5	10
All	All	1269/1348 (94%)	1136 (90%)	133 (10%)	8	15

5 of 133 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	224	LEU
1	C	65	LEU
1	D	244	LEU
1	B	241	SER
1	B	358	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	209	GLN
1	B	359	GLN
1	D	235	HIS
1	B	229	ASN
1	B	235	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 ⓘ

There are no carbohydrates in this entry.



## 5.6 Ligand geometry

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 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	EPE	D	3380	2	15,15,15	0.71	1 (6%)	18,20,20	2.82	7 (38%)

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	EPE	D	3380	2	-	0/9/19/19	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	3380	EPE	C10-S	2.10	1.80	1.77

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	3380	EPE	C9-N1-C6	-3.87	101.33	111.26
3	D	3380	EPE	C9-N1-C2	-2.58	104.65	111.26
3	D	3380	EPE	C5-C6-N1	2.06	114.80	110.63
3	D	3380	EPE	C6-N1-C2	3.56	116.93	108.87
3	D	3380	EPE	C6-C5-N4	3.73	118.18	110.63

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	D	3380	EPE	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, 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	389/400 (97%)	0.06	21 (5%) 26 27	37, 63, 85, 98	0
1	B	392/400 (98%)	-0.05	21 (5%) 26 27	34, 54, 79, 96	0
1	C	380/400 (95%)	0.21	24 (6%) 21 21	53, 74, 100, 109	0
1	D	371/400 (92%)	0.36	53 (14%) 3 2	35, 60, 141, 146	0
All	All	1532/1600 (95%)	0.14	119 (7%) 14 14	34, 63, 104, 146	0

The worst 5 of 119 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	73	LEU	7.5
1	D	139	ALA	7.1
1	B	1	MET	6.1
1	C	227	PRO	5.9
1	C	392	SER	5.5

### 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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	MG	B	401	1/1	0.98	0.19	4.96	49,49,49,49	0
3	EPE	D	3380	15/15	0.94	0.14	0.05	56,61,69,70	0
2	MG	C	401	1/1	0.94	0.06	-2.16	90,90,90,90	0
2	MG	A	401	1/1	0.95	0.05	-2.32	49,49,49,49	0
2	MG	D	401	1/1	0.91	0.06	-4.17	50,50,50,50	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.