



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 5, 2023 – 09:44 PM EDT

PDB ID : 4EG2  
Title : 2.2 Angstrom Crystal Structure of Cytidine deaminase from *Vibrio cholerae* in Complex with Zinc and Uridine  
Authors : Minasov, G.; Wawrzak, Z.; Skarina, T.; Wang, Y.; Grimshaw, S.; Papazisi, L.; Savchenko, A.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2012-03-30  
Resolution : 2.20 Å (reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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) : 2.35

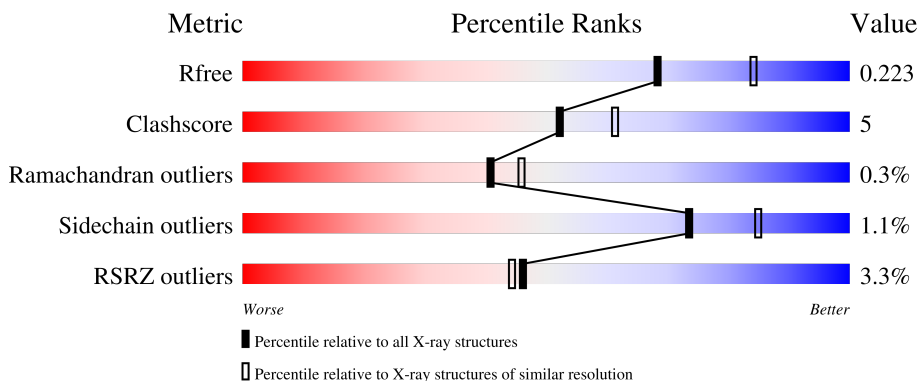
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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.



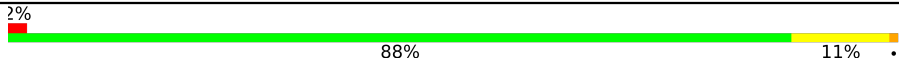

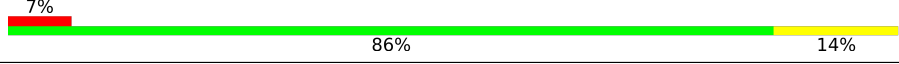
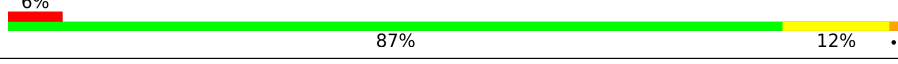
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 of 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	298	85% 14%
1	B	298	93% 6%
1	C	298	91% 9% .
1	D	298	88% 11% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	E	298	 2% 88% 11% .
1	F	298	 2% 88% 10% ..
1	G	298	 7% 86% 14%
1	H	298	 6% 87% 12% .

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 19771 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 Cytidine deaminase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	297	Total 2317	C 1462	N 390	O 451	S 3	Se 11	0	7	0
1	B	297	Total 2292	C 1449	N 384	O 445	S 3	Se 11	0	4	0
1	C	296	Total 2279	C 1437	N 381	O 447	S 3	Se 11	0	4	0
1	D	297	Total 2301	C 1454	N 385	O 448	S 3	Se 11	0	5	0
1	E	297	Total 2276	C 1436	N 384	O 442	S 3	Se 11	0	2	0
1	F	296	Total 2257	C 1426	N 377	O 440	S 3	Se 11	0	1	0
1	G	297	Total 2256	C 1425	N 378	O 439	S 3	Se 11	0	0	0
1	H	297	Total 2285	C 1441	N 384	O 446	S 3	Se 11	0	3	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q9KSM5
A	-1	ASN	-	expression tag	UNP Q9KSM5
A	0	ALA	-	expression tag	UNP Q9KSM5
B	-2	SER	-	expression tag	UNP Q9KSM5
B	-1	ASN	-	expression tag	UNP Q9KSM5
B	0	ALA	-	expression tag	UNP Q9KSM5
C	-2	SER	-	expression tag	UNP Q9KSM5
C	-1	ASN	-	expression tag	UNP Q9KSM5
C	0	ALA	-	expression tag	UNP Q9KSM5
D	-2	SER	-	expression tag	UNP Q9KSM5
D	-1	ASN	-	expression tag	UNP Q9KSM5
D	0	ALA	-	expression tag	UNP Q9KSM5
E	-2	SER	-	expression tag	UNP Q9KSM5

*Continued on next page...*

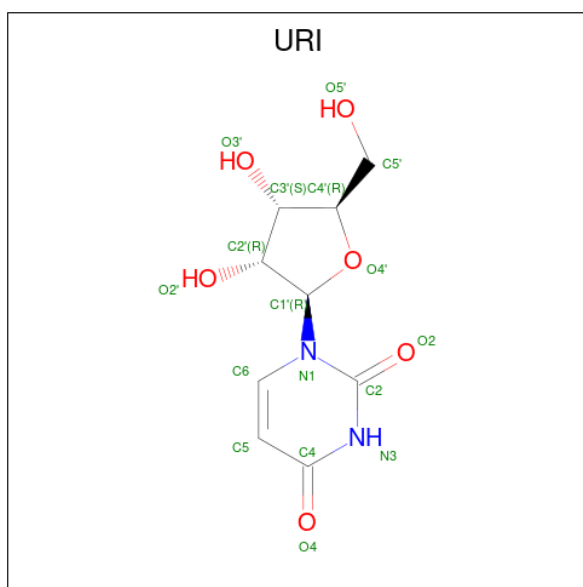
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	ASN	-	expression tag	UNP Q9KSM5
E	0	ALA	-	expression tag	UNP Q9KSM5
F	-2	SER	-	expression tag	UNP Q9KSM5
F	-1	ASN	-	expression tag	UNP Q9KSM5
F	0	ALA	-	expression tag	UNP Q9KSM5
G	-2	SER	-	expression tag	UNP Q9KSM5
G	-1	ASN	-	expression tag	UNP Q9KSM5
G	0	ALA	-	expression tag	UNP Q9KSM5
H	-2	SER	-	expression tag	UNP Q9KSM5
H	-1	ASN	-	expression tag	UNP Q9KSM5
H	0	ALA	-	expression tag	UNP Q9KSM5

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Zn 1 1	0	0
2	B	1	Total Zn 1 1	0	0
2	C	1	Total Zn 1 1	0	0
2	D	1	Total Zn 1 1	0	0
2	E	1	Total Zn 1 1	0	0
2	F	1	Total Zn 1 1	0	0
2	G	1	Total Zn 1 1	0	0
2	H	1	Total Zn 1 1	0	0

- Molecule 3 is URIDINE (three-letter code: URI) (formula: C<sub>9</sub>H<sub>12</sub>N<sub>2</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	Total	C	N	O	0	0
			17	9	2	6		
3	B	1	Total	C	N	O	0	0
			17	9	2	6		
3	C	1	Total	C	N	O	0	0
			17	9	2	6		
3	D	1	Total	C	N	O	0	0
			17	9	2	6		
3	E	1	Total	C	N	O	0	0
			17	9	2	6		
3	F	1	Total	C	N	O	0	0
			17	9	2	6		
3	G	1	Total	C	N	O	0	0
			17	9	2	6		
3	H	1	Total	C	N	O	0	0
			17	9	2	6		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0
4	E	1	Total C O 4 2 2	0	0
4	E	1	Total C O 4 2 2	0	0
4	E	1	Total C O 4 2 2	0	0
4	F	1	Total C O 4 2 2	0	0
4	G	1	Total C O 4 2 2	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Mg 1 1	0	0

- Molecule 6 is water.

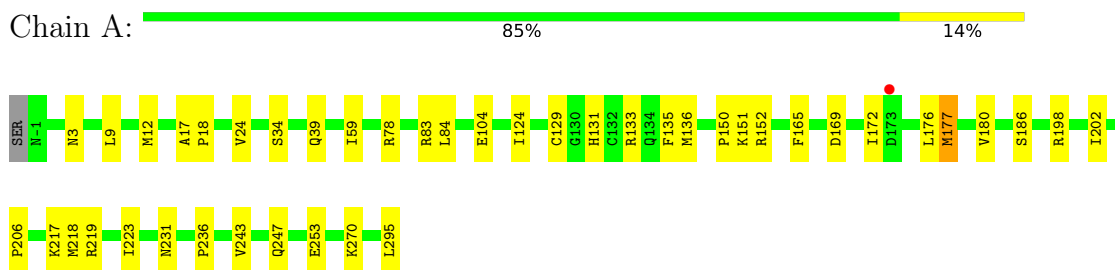
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	183	Total 185	O 185	0	4
6	B	223	Total 224	O 224	0	1
6	C	177	Total 178	O 178	0	1
6	D	182	Total 182	O 182	0	0
6	E	156	Total 157	O 157	0	1
6	F	163	Total 167	O 167	0	4
6	G	111	Total 111	O 111	0	0
6	H	122	Total 123	O 123	0	1



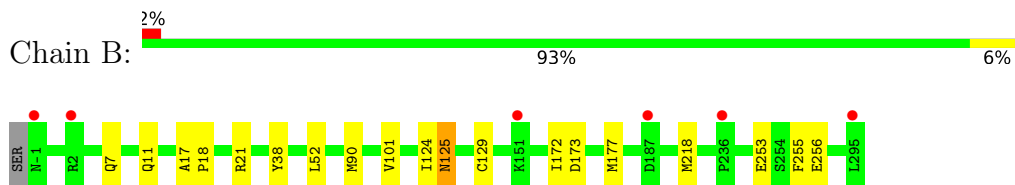
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

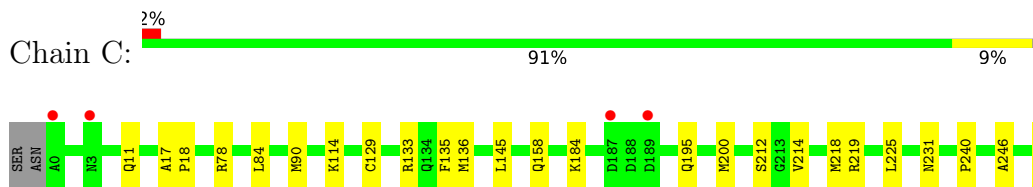
- Molecule 1: Cytidine deaminase



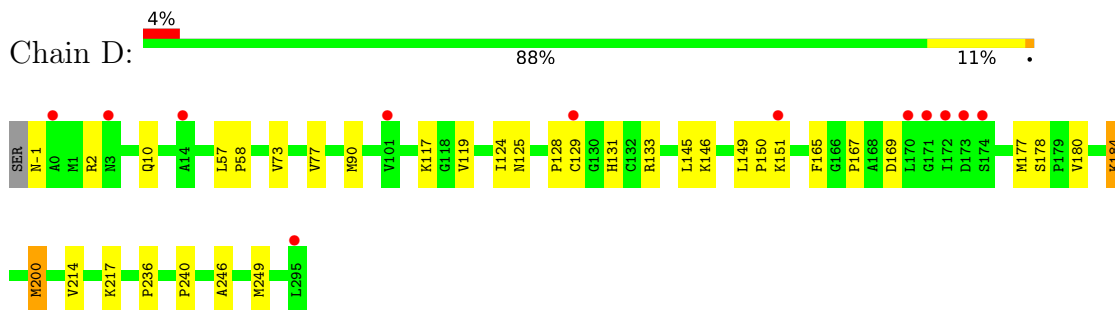
- Molecule 1: Cytidine deaminase



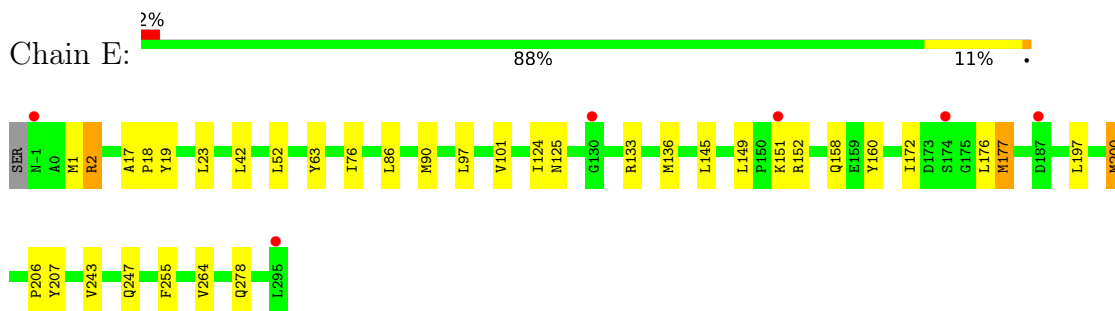
- Molecule 1: Cytidine deaminase



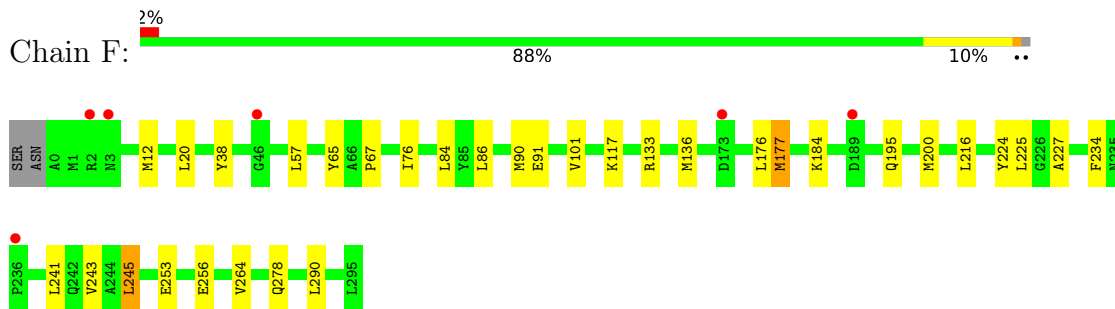
- Molecule 1: Cytidine deaminase



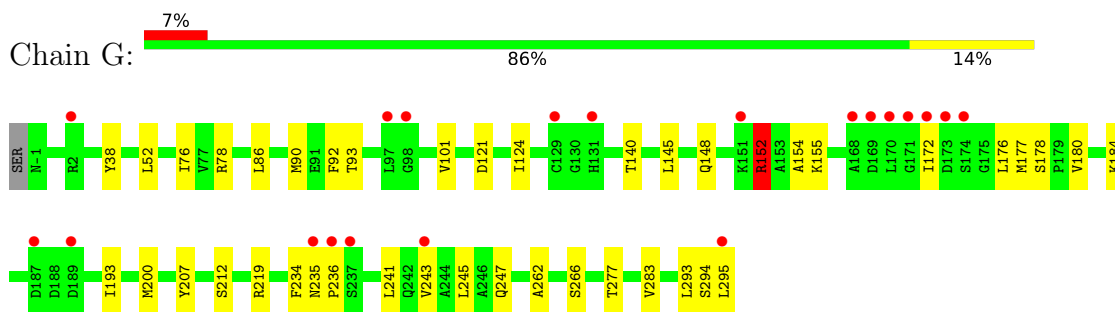
- Molecule 1: Cytidine deaminase



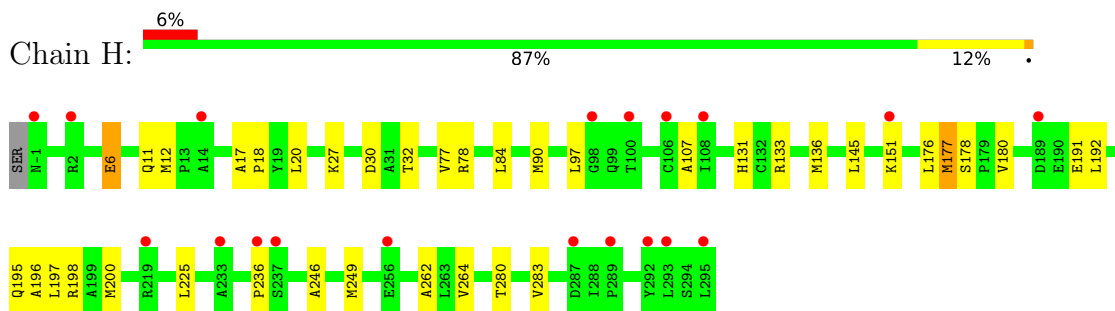
• Molecule 1: Cytidine deaminase



• Molecule 1: Cytidine deaminase



• Molecule 1: Cytidine deaminase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.81Å 163.73Å 111.95Å 90.00° 97.70° 90.00°	Depositor
Resolution (Å)	40.21 – 2.20 40.21 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (40.21-2.20) 98.9 (40.21-2.20)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.07 (at 2.20Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.159 , 0.210 0.173 , 0.223	Depositor DCC
$R_{free}$ test set	6173 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.0	Xtrriage
Anisotropy	0.566	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 44.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	19771	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.40% 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: URI, MG, ACT, ZN

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	0.55	0/2351	0.68	0/3170
1	B	0.54	1/2326 (0.0%)	0.65	0/3137
1	C	0.53	0/2312	0.65	0/3118
1	D	0.53	0/2335	0.63	0/3149
1	E	0.48	0/2309	0.63	0/3113
1	F	0.51	0/2290	0.65	0/3088
1	G	0.44	0/2289	0.61	1/3087 (0.0%)
1	H	0.45	0/2318	0.61	0/3125
All	All	0.51	1/18530 (0.0%)	0.64	1/24987 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	129	CYS	CB-SG	-5.17	1.73	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	G	152	ARG	NE-CZ-NH1	5.67	123.14	120.30

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	2317	0	2276	26	0
1	B	2292	0	2257	13	0
1	C	2279	0	2238	20	0
1	D	2301	0	2262	25	0
1	E	2276	0	2244	29	0
1	F	2257	0	2224	23	0
1	G	2256	0	2225	35	0
1	H	2285	0	2247	31	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
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	17	0	12	0	0
3	B	17	0	12	0	0
3	C	17	0	12	0	0
3	D	17	0	12	0	0
3	E	17	0	12	0	0
3	F	17	0	12	0	0
3	G	17	0	12	0	0
3	H	17	0	12	0	0
4	A	4	0	3	0	0
4	B	4	0	3	1	0
4	D	8	0	6	0	0
4	E	12	0	9	1	0
4	F	4	0	3	1	0
4	G	4	0	3	0	0
5	B	1	0	0	0	0
6	A	185	0	0	5	0
6	B	224	0	0	2	0
6	C	178	0	0	4	0
6	D	182	0	0	5	0
6	E	157	0	0	1	0
6	F	167	0	0	1	0
6	G	111	0	0	2	0
6	H	123	0	0	1	0
All	All	19771	0	18096	200	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.

All (200) 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:H:200:MSE:CE	1:H:264:VAL:HG12	1.62	1.29
1:H:200:MSE:HE3	1:H:264:VAL:HG12	1.18	1.10
1:H:200:MSE:HE3	1:H:264:VAL:CG1	1.87	1.04
1:B:125[A]:ASN:HD22	1:B:125[A]:ASN:N	1.54	0.99
1:B:125[A]:ASN:ND2	1:B:125[A]:ASN:H	1.49	0.97
1:G:93:THR:O	1:G:283:VAL:HG11	1.66	0.95
1:E:172:ILE:HD13	1:E:206:PRO:HB2	1.53	0.89
1:E:172:ILE:HD12	1:E:207:TYR:CE1	2.10	0.86
1:F:216:LEU:HD11	1:F:245:LEU:HD13	1.59	0.84
1:E:90:MSE:HE2	1:E:101:VAL:HG21	1.58	0.84
1:G:93:THR:O	1:G:283:VAL:CG1	2.26	0.82
1:B:125[A]:ASN:HD22	1:B:125[A]:ASN:H	0.80	0.78
1:G:200:MSE:HG3	1:G:212:SER:HA	1.64	0.78
1:E:90:MSE:CE	1:E:101:VAL:HG21	2.14	0.77
1:E:90:MSE:HE2	1:E:101:VAL:CG2	2.14	0.77
1:H:90:MSE:HE2	1:H:249:MSE:CE	2.17	0.75
1:H:90:MSE:HE1	1:H:246:ALA:HA	1.73	0.71
1:G:184:LYS:HG3	1:G:200:MSE:HE1	1.73	0.70
1:G:152:ARG:HG2	1:G:152:ARG:HH11	1.57	0.70
1:G:152:ARG:HH11	1:G:152:ARG:CG	2.07	0.67
1:H:90:MSE:HE2	1:H:249:MSE:HE2	1.77	0.66
1:E:1:MSE:HE3	1:E:63:TYR:CZ	2.31	0.66
1:G:293:LEU:C	1:G:293:LEU:HD23	2.18	0.64
1:D:124[A]:ILE:HD11	1:D:128:PRO:HD3	1.79	0.63
1:G:293:LEU:HD23	1:G:294:SER:N	2.14	0.63
1:B:7:GLN:HG2	6:B:448:HOH:O	1.99	0.63
1:C:184:LYS:HG3	1:C:200:MSE:HE1	1.81	0.63
1:C:214:VAL:HG23	1:C:240:PRO:HB2	1.80	0.62
1:F:133:ARG:HA	1:F:136:MSE:HE3	1.82	0.61
1:C:200:MSE:HG3	1:C:212:SER:HA	1.81	0.61
1:H:200:MSE:HE1	1:H:264:VAL:HG12	1.73	0.61
1:D:90:MSE:HE1	1:D:246:ALA:HA	1.82	0.61
1:F:76:ILE:HG12	1:F:86:LEU:HD22	1.81	0.61
1:E:172:ILE:HD13	1:E:206:PRO:CB	2.28	0.60
1:A:152:ARG:HG3	6:A:443:HOH:O	2.01	0.59
1:H:200:MSE:CE	1:H:264:VAL:CG1	2.54	0.59
1:C:90:MSE:HE2	1:C:249:MSE:CE	2.33	0.58
1:F:256:GLU:HG3	6:F:486:HOH:O	2.03	0.58
1:G:145:LEU:C	1:G:145:LEU:HD12	2.23	0.58
1:D:90:MSE:HE2	1:D:249:MSE:CE	2.33	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:38:TYR:CD1	1:F:84:LEU:HD11	2.39	0.58
1:F:184:LYS:HG3	1:F:200:MSE:HE2	1.86	0.58
1:D:200:MSE:HE2	1:D:200:MSE:C	2.25	0.57
1:D:131:HIS:CE1	1:D:236:PRO:HA	2.40	0.56
1:A:131:HIS:CE1	1:A:236:PRO:HA	2.40	0.56
1:B:255:PHE:CE2	4:B:304:ACT:H3	2.40	0.56
1:D:90:MSE:HE2	1:D:249:MSE:SE	2.55	0.56
1:A:78:ARG:HA	1:A:84:LEU:HD23	1.86	0.56
1:D:217:LYS:HE2	6:D:477:HOH:O	2.05	0.56
1:G:78:ARG:HB3	1:G:121:ASP:OD1	2.06	0.56
1:G:193:ILE:HD13	1:G:262:ALA:HB1	1.88	0.56
1:F:216:LEU:CD1	1:F:245:LEU:HD13	2.34	0.56
1:A:186:SER:HB2	1:A:295:LEU:HD11	1.88	0.55
1:H:195:GLN:O	1:H:198:ARG:HB2	2.07	0.54
1:H:280:THR:O	1:H:283:VAL:HG22	2.07	0.54
1:C:184:LYS:CG	1:C:200:MSE:HE1	2.36	0.54
1:E:76:ILE:HG12	1:E:86:LEU:HD22	1.89	0.54
1:G:155:LYS:NZ	6:G:435:HOH:O	2.39	0.54
1:G:184:LYS:CG	1:G:200:MSE:HE1	2.38	0.54
1:G:172:ILE:HD12	1:G:207:TYR:CE1	2.43	0.53
1:A:270:LYS:NZ	6:A:524:HOH:O	2.42	0.53
1:C:195:GLN:HB3	1:C:225:LEU:HD21	1.91	0.53
1:G:38:TYR:CE1	1:G:52:LEU:HD23	2.43	0.53
1:H:195:GLN:HB3	1:H:225:LEU:HD21	1.90	0.53
1:H:6[A]:GLU:N	1:H:6[A]:GLU:OE2	2.42	0.53
1:A:135:PHE:HA	1:A:231:ASN:ND2	2.24	0.52
1:C:90:MSE:HE2	1:C:249:MSE:HE2	1.92	0.52
1:H:17:ALA:HB3	1:H:18:PRO:HD3	1.90	0.52
1:A:295:LEU:HD23	1:A:295:LEU:O	2.10	0.52
1:D:165:PHE:HA	1:D:169:ASP:OD2	2.10	0.52
1:B:90:MSE:HE2	1:B:101:VAL:HG21	1.92	0.51
1:H:11:GLN:NE2	6:H:513:HOH:O	2.44	0.51
1:A:218:MSE:HE2	1:A:253:GLU:HB3	1.90	0.51
1:D:125[B]:ASN:OD1	1:D:125[B]:ASN:N	2.45	0.50
1:E:200:MSE:HE3	1:E:264:VAL:HG12	1.92	0.50
1:H:30:ASP:O	1:H:32:THR:HG23	2.11	0.50
1:G:243:VAL:O	1:G:247:GLN:HG2	2.10	0.50
1:A:3[B]:ASN:HB3	6:A:581[B]:HOH:O	2.11	0.50
1:E:172:ILE:CD1	1:E:206:PRO:HB2	2.33	0.49
1:F:57:LEU:HD23	1:F:76:ILE:HG13	1.93	0.49
1:G:200:MSE:HG3	1:G:212:SER:CA	2.40	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:12:MSE:HG3	1:H:20:LEU:HD23	1.94	0.49
1:A:34:SER:HA	1:A:83[B]:ARG:HG3	1.94	0.49
1:G:234:PHE:CD2	1:H:97:LEU:HD22	2.47	0.49
1:F:241:LEU:HG	1:F:245:LEU:HD22	1.95	0.49
1:D:-1:ASN:OD1	1:D:2:ARG:HG2	2.13	0.49
1:F:90:MSE:HE2	1:F:101:VAL:HG21	1.94	0.48
1:A:133:ARG:HA	1:A:136:MSE:HE3	1.95	0.48
1:B:90:MSE:HE2	1:B:101:VAL:CG2	2.44	0.48
1:E:145:LEU:HD12	1:E:145:LEU:C	2.32	0.48
1:H:178:SER:O	1:H:180:VAL:HG23	2.14	0.48
1:E:145:LEU:HD12	1:E:145:LEU:O	2.14	0.48
1:E:97:LEU:HD22	1:F:234:PHE:CD2	2.49	0.48
1:C:114:LYS:NZ	6:C:562:HOH:O	2.38	0.48
1:E:124:ILE:HG22	1:E:125:ASN:N	2.29	0.48
1:E:243:VAL:O	1:E:247:GLN:HG2	2.14	0.48
1:G:266:SER:HA	1:G:295:LEU:O	2.14	0.48
1:B:172:ILE:HG22	1:B:173:ASP:O	2.14	0.47
1:C:17:ALA:HB3	1:C:18:PRO:HD3	1.95	0.47
1:E:278:GLN:HG3	6:E:513:HOH:O	2.12	0.47
1:D:145:LEU:HD12	1:D:145:LEU:C	2.35	0.47
1:A:172:ILE:HD13	1:A:206:PRO:HB2	1.96	0.47
1:G:90:MSE:HE2	1:G:101:VAL:HG21	1.96	0.47
1:A:129:CYS:O	1:A:133:ARG:HG3	2.15	0.47
1:D:150:PRO:O	1:D:151:LYS:HG2	2.14	0.47
1:G:90:MSE:HE3	1:G:92:PHE:CZ	2.50	0.47
1:H:133:ARG:HA	1:H:136:MSE:HE3	1.97	0.47
1:A:12:MSE:HE3	1:A:59:ILE:HD11	1.97	0.47
1:F:216:LEU:CD1	1:F:245:LEU:CD1	2.93	0.46
1:D:146:LYS:NZ	6:D:491:HOH:O	2.48	0.46
1:G:90:MSE:HE2	1:G:101:VAL:CG2	2.45	0.46
1:F:216:LEU:HD11	1:F:245:LEU:CD1	2.39	0.46
1:F:227:ALA:O	1:F:243:VAL:HG21	2.16	0.46
1:F:278:GLN:HG2	1:F:290:LEU:HD23	1.98	0.46
1:G:76:ILE:HG12	1:G:86:LEU:HD22	1.97	0.46
1:H:6[A]:GLU:OE2	1:H:6[A]:GLU:CA	2.64	0.46
1:A:150:PRO:C	1:A:152:ARG:H	2.19	0.46
1:H:78:ARG:HA	1:H:84:LEU:HD23	1.97	0.46
1:B:125[A]:ASN:N	1:B:125[A]:ASN:ND2	2.26	0.45
1:B:17:ALA:HB3	1:B:18:PRO:HD3	1.99	0.45
1:D:10:GLN:HG2	6:D:468:HOH:O	2.15	0.45
1:G:90:MSE:HE3	1:G:92:PHE:HZ	1.81	0.45

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:152:ARG:CG	1:G:152:ARG:NH1	2.74	0.45
1:H:197:LEU:HD12	1:H:200:MSE:HG2	1.97	0.45
1:D:125[A]:ASN:ND2	6:D:473:HOH:O	2.49	0.45
1:C:129:CYS:O	1:C:133:ARG:HG3	2.17	0.45
1:D:214:VAL:HG23	1:D:240:PRO:HB2	1.98	0.45
1:G:184:LYS:CB	1:G:200:MSE:HE1	2.46	0.45
1:A:180:VAL:HG23	1:A:202:ILE:HA	1.99	0.45
1:D:129:CYS:O	1:D:133:ARG:HG3	2.16	0.45
1:A:150:PRO:O	1:A:152:ARG:N	2.49	0.45
1:B:38:TYR:CE1	1:B:52:LEU:HD23	2.52	0.44
1:D:124[A]:ILE:HD11	1:D:149:LEU:HD21	1.99	0.44
1:E:176:LEU:O	1:E:177:MSE:HB2	2.17	0.44
1:A:9:LEU:HD11	1:A:24:VAL:CG2	2.47	0.44
1:A:198:ARG:HD2	6:A:452:HOH:O	2.16	0.44
1:B:218:MSE:HE2	1:B:253:GLU:HB3	2.00	0.44
1:C:11:GLN:HB2	6:C:439:HOH:O	2.18	0.44
1:E:90:MSE:CE	1:E:101:VAL:CG2	2.87	0.44
1:D:167:PRO:HB2	6:D:426:HOH:O	2.18	0.44
1:C:90:MSE:HE1	1:C:246:ALA:HA	2.00	0.43
1:H:192:LEU:HG	1:H:262:ALA:CB	2.48	0.43
1:G:140:THR:HA	6:G:503:HOH:O	2.18	0.43
1:G:176:LEU:O	1:G:177:MSE:HB2	2.18	0.43
1:A:17:ALA:HB3	1:A:18:PRO:HD3	2.01	0.43
1:F:67:PRO:HD2	1:F:91:GLU:HB2	2.01	0.43
1:F:65:TYR:HB2	4:F:303:ACT:H2	2.00	0.43
1:F:200:MSE:SE	1:F:264:VAL:HG12	2.69	0.43
1:C:218:MSE:HE2	1:C:253:GLU:HB3	2.01	0.43
1:D:77:VAL:HG13	1:D:119:VAL:HG13	2.01	0.43
1:E:149:LEU:HD22	1:E:160:TYR:OH	2.19	0.43
1:C:133:ARG:HA	1:C:136:MSE:HE3	2.00	0.43
1:E:42:LEU:HA	1:E:52:LEU:HD22	2.01	0.43
1:F:176:LEU:O	1:F:177:MSE:HB2	2.19	0.43
1:D:57:LEU:N	1:D:58:PRO:CD	2.82	0.42
1:D:184:LYS:HG3	1:D:200:MSE:HE1	2.01	0.42
1:G:241:LEU:HD22	1:G:277:THR:HG23	2.00	0.42
1:A:104:GLU:CG	1:A:124[B]:ILE:HD13	2.48	0.42
1:B:256:GLU:HG2	6:B:543:HOH:O	2.18	0.42
1:E:19:TYR:CE2	1:E:23:LEU:HD11	2.54	0.42
1:E:255:PHE:H	4:E:303:ACT:H1	1.84	0.42
1:F:195:GLN:HB3	1:F:225:LEU:HD21	2.00	0.42
1:H:145:LEU:HD12	1:H:145:LEU:C	2.40	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:224:TYR:CE2	1:F:253:GLU:HG3	2.55	0.42
1:G:293:LEU:C	1:G:293:LEU:CD2	2.87	0.42
1:H:176:LEU:O	1:H:177:MSE:HB2	2.20	0.42
1:A:165:PHE:HA	1:A:169:ASP:OD2	2.19	0.42
1:E:197:LEU:HD12	1:E:200:MSE:HG2	2.01	0.42
1:C:78:ARG:HB2	1:C:84:LEU:CD2	2.50	0.42
1:C:90:MSE:HE3	1:C:90:MSE:HB2	1.88	0.42
1:D:178:SER:O	1:D:180:VAL:HG23	2.20	0.42
1:E:2:ARG:NE	1:E:2:ARG:HA	2.34	0.42
1:A:243:VAL:O	1:A:247:GLN:HG2	2.20	0.41
1:E:17:ALA:HB3	1:E:18:PRO:HD3	2.01	0.41
1:G:178:SER:O	1:G:180:VAL:HG23	2.20	0.41
1:H:191[A]:GLU:HA	1:H:191[A]:GLU:OE1	2.20	0.41
1:C:135:PHE:HA	1:C:231:ASN:ND2	2.36	0.41
1:G:145:LEU:HD12	1:G:145:LEU:O	2.20	0.41
1:G:235:ASN:N	1:G:236:PRO:CD	2.84	0.41
1:H:196:ALA:HB3	1:H:264:VAL:HG23	2.03	0.41
1:C:158:GLN:HG3	6:C:425:HOH:O	2.20	0.41
1:C:145:LEU:C	1:C:145:LEU:HD12	2.41	0.41
1:E:158[B]:GLN:H	1:E:158[B]:GLN:CD	2.24	0.41
1:F:12:MSE:HG3	1:F:20:LEU:HD23	2.03	0.41
1:H:131:HIS:CE1	1:H:236:PRO:HA	2.56	0.41
1:A:217:LYS:HD2	6:A:504:HOH:O	2.19	0.41
1:H:77:VAL:HG23	1:H:107:ALA:HB1	2.02	0.41
1:C:114:LYS:HE3	6:C:442:HOH:O	2.20	0.41
1:E:1:MSE:CE	1:E:63:TYR:CZ	3.02	0.41
1:D:150:PRO:C	1:D:151:LYS:HG2	2.41	0.41
1:G:92:PHE:CZ	1:G:245:LEU:HD13	2.56	0.41
1:H:200:MSE:HE3	1:H:264:VAL:HG11	1.91	0.41
1:A:217:LYS:HA	1:A:223:ILE:HD13	2.01	0.40
1:H:90:MSE:HE3	1:H:90:MSE:HB2	1.96	0.40
1:E:133:ARG:HA	1:E:136:MSE:HE3	2.02	0.40
1:E:151:LYS:O	1:E:152:ARG:HG2	2.21	0.40
1:D:73:VAL:HA	1:D:125[B]:ASN:HD21	1.87	0.40
1:G:148:GLN:OE1	1:G:154:ALA:HB2	2.22	0.40
1:A:176:LEU:O	1:A:177:MSE:HB2	2.21	0.40
1:F:76:ILE:HG12	1:F:86:LEU:CD2	2.51	0.40

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	302/298 (101%)	298 (99%)	2 (1%)	2 (1%)	22	22
1	B	299/298 (100%)	293 (98%)	5 (2%)	1 (0%)	41	46
1	C	298/298 (100%)	292 (98%)	6 (2%)	0	100	100
1	D	300/298 (101%)	294 (98%)	5 (2%)	1 (0%)	41	46
1	E	297/298 (100%)	293 (99%)	3 (1%)	1 (0%)	41	46
1	F	295/298 (99%)	292 (99%)	2 (1%)	1 (0%)	41	46
1	G	295/298 (99%)	290 (98%)	5 (2%)	0	100	100
1	H	298/298 (100%)	291 (98%)	6 (2%)	1 (0%)	41	46
All	All	2384/2384 (100%)	2343 (98%)	34 (1%)	7 (0%)	41	46

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	151	LYS
1	F	177	MSE
1	A	177	MSE
1	D	177	MSE
1	B	177	MSE
1	H	177	MSE
1	E	177	MSE

### 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	244/227 (108%)	242 (99%)	2 (1%)	81	90
1	B	241/227 (106%)	235 (98%)	6 (2%)	47	60
1	C	240/227 (106%)	239 (100%)	1 (0%)	91	96
1	D	242/227 (107%)	239 (99%)	3 (1%)	71	83
1	E	239/227 (105%)	237 (99%)	2 (1%)	81	90
1	F	237/227 (104%)	235 (99%)	2 (1%)	81	90
1	G	237/227 (104%)	234 (99%)	3 (1%)	69	81
1	H	240/227 (106%)	236 (98%)	4 (2%)	60	74
All	All	1920/1816 (106%)	1897 (99%)	23 (1%)	73	83

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	39	GLN
1	A	219	ARG
1	B	11	GLN
1	B	21	ARG
1	B	124[A]	ILE
1	B	124[B]	ILE
1	B	125[A]	ASN
1	B	125[B]	ASN
1	C	219	ARG
1	D	117	LYS
1	D	184	LYS
1	D	200	MSE
1	E	2	ARG
1	E	200	MSE
1	F	117	LYS
1	F	245	LEU
1	G	124	ILE
1	G	152	ARG
1	G	219	ARG
1	H	6[A]	GLU
1	H	6[B]	GLU
1	H	27	LYS
1	H	151	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	39	GLN
1	C	11	GLN
1	C	278	GLN
1	F	11	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 26 ligands modelled in this entry, 9 are monoatomic - leaving 17 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	$\# Z  > 2$
4	ACT	E	303	-	3,3,3	0.72	0	3,3,3	1.35	0
4	ACT	F	303	-	3,3,3	0.72	0	3,3,3	1.27	0
4	ACT	D	304	-	3,3,3	0.85	0	3,3,3	1.46	0
3	URI	F	302	-	18,18,18	0.35	0	26,26,26	0.58	0
3	URI	C	302	-	18,18,18	0.50	0	26,26,26	0.69	0
3	URI	D	302	-	18,18,18	0.44	0	26,26,26	0.67	0
3	URI	B	303	-	18,18,18	0.39	0	26,26,26	0.48	0
3	URI	A	302	-	18,18,18	0.49	0	26,26,26	0.64	0
4	ACT	A	303	-	3,3,3	0.74	0	3,3,3	1.47	0
4	ACT	D	303	-	3,3,3	0.72	0	3,3,3	1.58	1 (33%)
4	ACT	E	304	-	3,3,3	0.75	0	3,3,3	1.31	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	URI	E	302	-	18,18,18	0.54	0	26,26,26	0.66	0
3	URI	H	302	2	18,18,18	0.65	0	26,26,26	0.57	0
4	ACT	B	304	-	3,3,3	0.79	0	3,3,3	1.47	0
4	ACT	G	303	-	3,3,3	0.75	0	3,3,3	1.64	1 (33%)
3	URI	G	302	-	18,18,18	0.55	0	26,26,26	0.59	0
4	ACT	E	305	-	3,3,3	0.71	0	3,3,3	1.47	0

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	URI	F	302	-	-	0/6/22/22	0/2/2/2
3	URI	C	302	-	-	0/6/22/22	0/2/2/2
3	URI	D	302	-	-	0/6/22/22	0/2/2/2
3	URI	B	303	-	-	0/6/22/22	0/2/2/2
3	URI	A	302	-	-	0/6/22/22	0/2/2/2
3	URI	E	302	-	-	0/6/22/22	0/2/2/2
3	URI	H	302	2	-	2/6/22/22	0/2/2/2
3	URI	G	302	-	-	0/6/22/22	0/2/2/2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	303	ACT	OXT-C-O	-2.03	114.56	122.05
4	D	303	ACT	OXT-C-CH3	2.03	123.57	115.18

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	H	302	URI	C3'-C4'-C5'-O5'
3	H	302	URI	O4'-C4'-C5'-O5'

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	303	ACT	1	0
4	F	303	ACT	1	0
4	B	304	ACT	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	286/298 (95%)	-0.22	1 (0%) 94 93	17, 25, 45, 58	0
1	B	286/298 (95%)	-0.14	6 (2%) 63 61	17, 26, 43, 63	0
1	C	285/298 (95%)	-0.12	5 (1%) 68 66	18, 28, 46, 62	0
1	D	286/298 (95%)	0.11	12 (4%) 36 34	19, 30, 52, 75	0
1	E	286/298 (95%)	-0.16	6 (2%) 63 61	21, 32, 49, 62	0
1	F	285/298 (95%)	-0.08	6 (2%) 63 61	20, 30, 51, 91	0
1	G	286/298 (95%)	0.32	20 (6%) 16 15	24, 41, 68, 110	0
1	H	286/298 (95%)	0.29	19 (6%) 18 17	24, 37, 62, 86	0
All	All	2286/2384 (95%)	-0.00	75 (3%) 46 44	17, 31, 55, 110	0

All (75) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	0	ALA	6.1
1	G	171	GLY	5.4
1	H	219[A]	ARG	5.2
1	H	2	ARG	4.9
1	G	295	LEU	4.8
1	D	172	ILE	4.4
1	D	173	ASP	4.4
1	H	-1	ASN	4.0
1	E	151	LYS	3.7
1	G	170	LEU	3.7
1	D	170	LEU	3.6
1	G	169	ASP	3.5
1	H	236	PRO	3.4
1	G	187	ASP	3.3
1	D	101	VAL	3.3
1	G	173	ASP	3.3

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	H	189	ASP	3.2
1	E	187	ASP	3.0
1	D	174	SER	3.0
1	D	0	ALA	3.0
1	C	295	LEU	2.9
1	G	174	SER	2.9
1	C	189	ASP	2.9
1	B	2	ARG	2.9
1	F	189	ASP	2.8
1	F	173	ASP	2.8
1	H	151	LYS	2.8
1	H	14	ALA	2.7
1	B	-1	ASN	2.7
1	F	3	ASN	2.7
1	G	129	CYS	2.7
1	E	-1	ASN	2.7
1	G	98	GLY	2.6
1	H	289	PRO	2.5
1	G	2	ARG	2.5
1	H	256	GLU	2.4
1	H	292	TYR	2.4
1	G	189	ASP	2.4
1	G	243	VAL	2.4
1	H	98	GLY	2.4
1	G	236	PRO	2.3
1	F	2	ARG	2.3
1	H	237	SER	2.3
1	D	171	GLY	2.3
1	G	172	ILE	2.3
1	D	14	ALA	2.3
1	D	151	LYS	2.3
1	E	295	LEU	2.3
1	G	131	HIS	2.3
1	G	168	ALA	2.3
1	H	100	THR	2.2
1	A	173	ASP	2.2
1	G	237	SER	2.2
1	E	130	GLY	2.2
1	F	46	GLY	2.2
1	G	235	ASN	2.2
1	H	295	LEU	2.2
1	E	174	SER	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	151	LYS	2.2
1	G	151	LYS	2.2
1	F	236	PRO	2.2
1	D	129	CYS	2.1
1	C	3	ASN	2.1
1	H	108	ILE	2.1
1	B	236	PRO	2.1
1	B	295	LEU	2.1
1	D	3	ASN	2.1
1	D	295	LEU	2.1
1	G	97	LEU	2.1
1	H	287	ASP	2.1
1	H	106	CYS	2.1
1	H	233	ALA	2.1
1	H	293	LEU	2.0
1	C	187[A]	ASP	2.0
1	B	187	ASP	2.0

## 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 monosaccharides 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(Å <sup>2</sup> )	Q<0.9
4	ACT	D	303	4/4	0.66	0.27	58,58,58,59	0
4	ACT	E	304	4/4	0.73	0.25	59,60,60,60	0
4	ACT	G	303	4/4	0.76	0.19	61,62,62,62	0
3	URI	H	302	17/17	0.83	0.23	56,62,64,65	0
4	ACT	E	305	4/4	0.85	0.20	60,61,61,62	0
4	ACT	F	303	4/4	0.88	0.21	60,61,61,61	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ACT	D	304	4/4	0.90	0.24	52,53,53,53	0
4	ACT	E	303	4/4	0.90	0.15	55,56,57,57	0
3	URI	C	302	17/17	0.92	0.15	33,42,44,46	0
4	ACT	A	303	4/4	0.93	0.23	36,37,39,39	0
3	URI	E	302	17/17	0.93	0.16	31,37,40,44	0
5	MG	B	301	1/1	0.93	0.07	38,38,38,38	0
4	ACT	B	304	4/4	0.94	0.10	53,53,53,54	0
2	ZN	A	301	1/1	0.94	0.20	36,36,36,36	1
2	ZN	B	302	1/1	0.94	0.33	34,34,34,34	1
3	URI	A	302	17/17	0.95	0.14	25,31,35,36	0
3	URI	G	302	17/17	0.96	0.16	33,38,45,45	0
2	ZN	F	301	1/1	0.96	0.19	40,40,40,40	1
3	URI	D	302	17/17	0.96	0.18	27,33,35,37	0
2	ZN	D	301	1/1	0.96	0.30	32,32,32,32	1
3	URI	B	303	17/17	0.97	0.14	29,33,37,37	0
3	URI	F	302	17/17	0.97	0.13	23,30,32,33	0
2	ZN	E	301	1/1	0.98	0.15	36,36,36,36	1
2	ZN	G	301	1/1	0.98	0.27	40,40,40,40	1
2	ZN	H	301	1/1	0.98	0.13	50,50,50,50	1
2	ZN	C	301	1/1	0.99	0.08	32,32,32,32	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.