



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 10:20 AM GMT

PDB ID : 4C0A  
Title : Arf1(Delta1-17)in complex with BRAG2 Sec7-PH domain  
Authors : Aizel, K.; Biou, V.; Navaza, J.; Duarte, L.; Campanacci, V.; Cherfils, J.; Zeghouf, M.  
Deposited on : 2013-08-01  
Resolution : 3.30 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

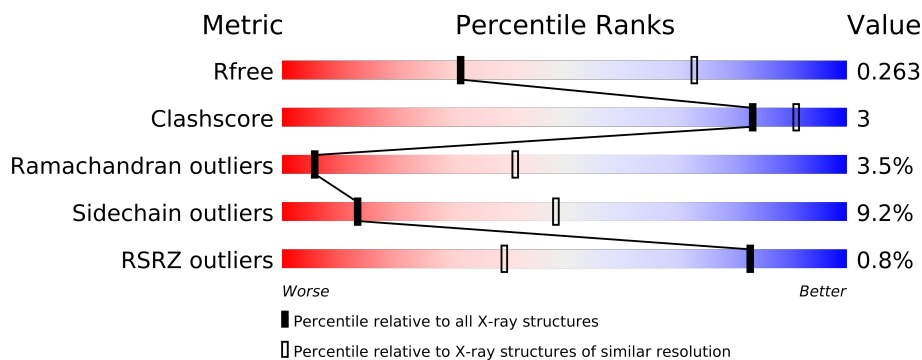
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.30 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))
$R_{free}$	66092	1341 (3.40-3.20)
Clashscore	79885	1696 (3.40-3.20)
Ramachandran outliers	78287	1664 (3.40-3.20)
Sidechain outliers	78261	1662 (3.40-3.20)
RSRZ outliers	66119	1342 (3.40-3.20)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	383	
1	B	383	
1	E	383	
1	F	383	
2	C	164	
2	D	164	
2	G	164	
2	H	164	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16930 atoms, of which 0 are hydrogen and 0 are deuterium.

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 IQ MOTIF AND SEC7 DOMAIN-CONTAINING PROTEIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	352	Total	C	N	O	S	0	0	0
			2897	1845	520	517	15			
1	B	357	Total	C	N	O	S	0	1	0
			2945	1875	531	524	15			
1	E	351	Total	C	N	O	S	0	0	0
			2893	1842	520	516	15			
1	F	351	Total	C	N	O	S	0	0	0
			2884	1834	519	516	15			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	381	GLY	-	EXPRESSION TAG	UNP Q6DN90
A	382	ALA	-	EXPRESSION TAG	UNP Q6DN90
A	383	MET	-	EXPRESSION TAG	UNP Q6DN90
A	384	GLY	-	EXPRESSION TAG	UNP Q6DN90
A	385	SER	-	EXPRESSION TAG	UNP Q6DN90
A	386	GLY	-	EXPRESSION TAG	UNP Q6DN90
A	387	ILE	-	EXPRESSION TAG	UNP Q6DN90
A	388	PRO	-	EXPRESSION TAG	UNP Q6DN90
A	389	MET	-	EXPRESSION TAG	UNP Q6DN90
A	498	LYS	GLU	ENGINEERED MUTATION	UNP Q6DN90
B	381	GLY	-	EXPRESSION TAG	UNP Q6DN90
B	382	ALA	-	EXPRESSION TAG	UNP Q6DN90
B	383	MET	-	EXPRESSION TAG	UNP Q6DN90
B	384	GLY	-	EXPRESSION TAG	UNP Q6DN90
B	385	SER	-	EXPRESSION TAG	UNP Q6DN90
B	386	GLY	-	EXPRESSION TAG	UNP Q6DN90
B	387	ILE	-	EXPRESSION TAG	UNP Q6DN90
B	388	PRO	-	EXPRESSION TAG	UNP Q6DN90
B	389	MET	-	EXPRESSION TAG	UNP Q6DN90
B	498	LYS	GLU	ENGINEERED MUTATION	UNP Q6DN90

*Continued on next page...*

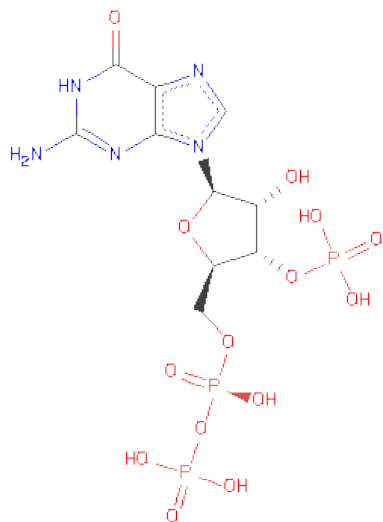
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	381	GLY	-	EXPRESSION TAG	UNP Q6DN90
E	382	ALA	-	EXPRESSION TAG	UNP Q6DN90
E	383	MET	-	EXPRESSION TAG	UNP Q6DN90
E	384	GLY	-	EXPRESSION TAG	UNP Q6DN90
E	385	SER	-	EXPRESSION TAG	UNP Q6DN90
E	386	GLY	-	EXPRESSION TAG	UNP Q6DN90
E	387	ILE	-	EXPRESSION TAG	UNP Q6DN90
E	388	PRO	-	EXPRESSION TAG	UNP Q6DN90
E	389	MET	-	EXPRESSION TAG	UNP Q6DN90
E	498	LYS	GLU	ENGINEERED MUTATION	UNP Q6DN90
F	381	GLY	-	EXPRESSION TAG	UNP Q6DN90
F	382	ALA	-	EXPRESSION TAG	UNP Q6DN90
F	383	MET	-	EXPRESSION TAG	UNP Q6DN90
F	384	GLY	-	EXPRESSION TAG	UNP Q6DN90
F	385	SER	-	EXPRESSION TAG	UNP Q6DN90
F	386	GLY	-	EXPRESSION TAG	UNP Q6DN90
F	387	ILE	-	EXPRESSION TAG	UNP Q6DN90
F	388	PRO	-	EXPRESSION TAG	UNP Q6DN90
F	389	MET	-	EXPRESSION TAG	UNP Q6DN90
F	498	LYS	GLU	ENGINEERED MUTATION	UNP Q6DN90

- Molecule 2 is a protein called ADP-RIBOSYLATION FACTOR 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	161	Total	C	N	O	S	0	0	0
			1293	819	227	242	5			
2	D	164	Total	C	N	O	S	0	0	0
			1315	831	232	247	5			
2	G	162	Total	C	N	O	S	0	0	0
			1298	822	228	243	5			
2	H	159	Total	C	N	O	S	0	0	0
			1277	810	222	240	5			

- Molecule 3 is GUANOSINE-3'-MONOPHOSPHATE-5'-DIPHOSPHATE (three-letter code: G3D) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>14</sub>P<sub>3</sub>).



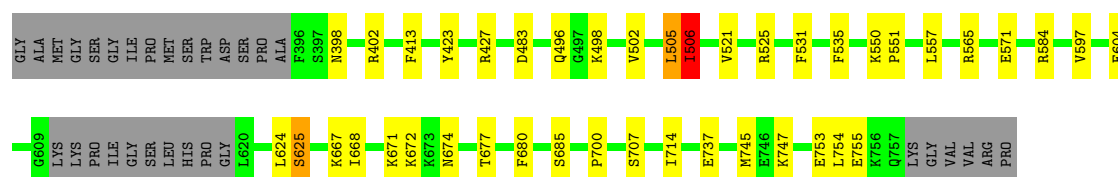
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	C	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
3	D	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
3	G	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
3	H	1	Total	C	N	O	P	0	0
			32	10	5	14	3		

### 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.

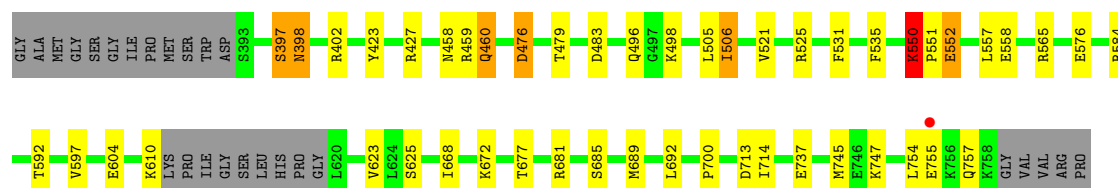
#### • Molecule 1: IQ MOTIF AND SEC7 DOMAIN-CONTAINING PROTEIN 1

Chain A:



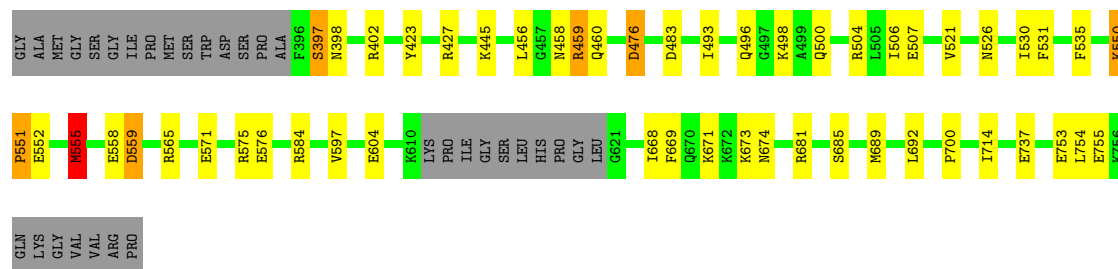
#### • Molecule 1: IQ MOTIF AND SEC7 DOMAIN-CONTAINING PROTEIN 1

Chain B:



#### • Molecule 1: IQ MOTIF AND SEC7 DOMAIN-CONTAINING PROTEIN 1

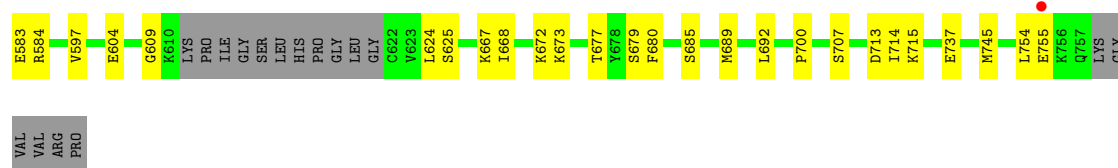
Chain E:



#### • Molecule 1: IQ MOTIF AND SEC7 DOMAIN-CONTAINING PROTEIN 1

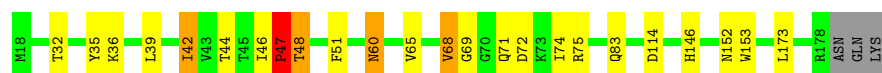
Chain F:





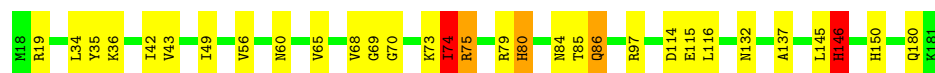
- Molecule 2: ADP-RIBOSYLATION FACTOR 1

Chain C:



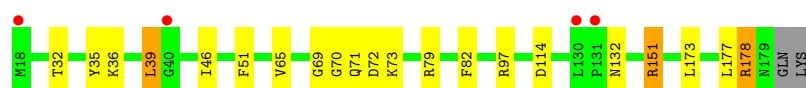
- Molecule 2: ADP-RIBOSYLATION FACTOR 1

Chain D:



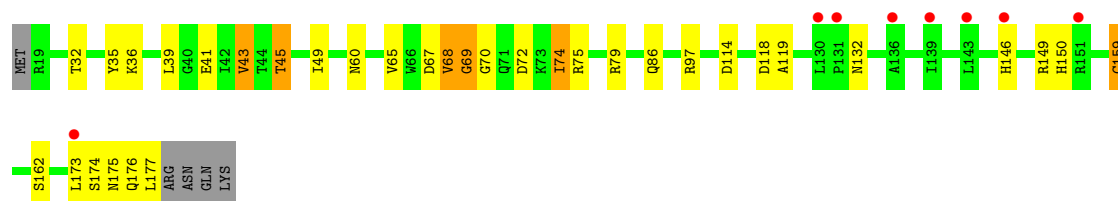
- Molecule 2: ADP-RIBOSYLATION FACTOR 1

Chain G:



- Molecule 2: ADP-RIBOSYLATION FACTOR 1

Chain H:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.98Å 65.78Å 196.86Å 90.00° 96.13° 90.00°	Depositor
Resolution (Å)	42.84 – 3.30 42.84 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.3 (42.84-3.30) 99.5 (42.84-3.30)	Depositor EDS
$R_{merge}$	0.03	Depositor
$R_{sym}$	20.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.70 (at 3.32Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.199 , 0.248 0.215 , 0.263	Depositor DCC
$R_{free}$ test set	1757 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	58.6	Xtriage
Anisotropy	0.565	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 44.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	1 of 35125 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	16930	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.53 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.1182e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: G3D

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	0.50	0/2950	0.78	4/3962 (0.1%)
1	B	0.54	0/3002	0.83	4/4030 (0.1%)
1	E	0.53	0/2946	0.82	3/3955 (0.1%)
1	F	0.54	0/2936	0.84	6/3944 (0.2%)
2	C	0.47	0/1316	0.81	2/1783 (0.1%)
2	D	0.51	0/1338	0.84	2/1813 (0.1%)
2	G	0.43	0/1321	0.74	0/1790
2	H	0.42	0/1300	0.80	0/1762
All	All	0.51	0/17109	0.81	21/23039 (0.1%)

There are no bond length outliers.

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	557	LEU	N-CA-C	-6.72	92.86	111.00
1	B	398	ASN	C-N-CA	6.20	137.20	121.70
1	A	506	ILE	N-CA-CB	6.10	124.82	110.80
1	F	474	GLU	C-N-CA	6.04	136.79	121.70
1	B	505	LEU	N-CA-C	-5.90	95.06	111.00
1	A	505	LEU	N-CA-C	-5.65	95.75	111.00
1	F	525	ARG	C-N-CA	5.55	135.57	121.70
2	C	60	ASN	CA-CB-CG	5.52	125.53	113.40
1	E	555	MET	N-CA-C	5.50	125.86	111.00
1	F	505	LEU	N-CA-C	-5.42	96.36	111.00
2	D	146	HIS	N-CA-CB	5.37	120.26	110.60
1	E	551	PRO	N-CA-C	5.32	125.94	112.10
2	D	86	GLN	N-CA-C	5.27	125.24	111.00
2	C	47	PRO	N-CA-C	5.27	125.80	112.10
1	E	397	SER	C-N-CA	5.26	134.85	121.70
1	B	557	LEU	N-CA-C	-5.21	96.92	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	506	ILE	N-CA-CB	5.18	122.71	110.80
1	A	525	ARG	C-N-CA	5.16	134.60	121.70
1	B	506	ILE	N-CA-CB	5.02	122.35	110.80
1	A	667	LYS	N-CA-C	-5.02	97.44	111.00
1	F	667	LYS	N-CA-C	-5.02	97.45	111.00

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2897	0	0	5	0
1	B	2945	0	0	8	0
1	E	2893	0	0	8	0
1	F	2884	0	0	5	0
2	C	1293	0	0	6	0
2	D	1315	0	0	10	0
2	G	1298	0	0	4	0
2	H	1277	0	0	7	0
3	C	32	0	0	1	0
3	D	32	0	0	0	0
3	G	32	0	0	1	0
3	H	32	0	0	1	0
All	All	16930	0	0	50	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (50) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:H:86:GLN:O	2:H:119:ALA:CB	2.41	0.68
2:D:74:ILE:O	2:D:75:ARG:NH1	2.27	0.67
1:B:550:LYS:O	1:B:552:GLU:N	2.32	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:458:ASN:O	1:B:460:GLN:N	2.32	0.63
2:C:152:ASN:OD1	2:D:137:ALA:CB	2.47	0.62
1:A:502:VAL:O	1:A:505:LEU:O	2.17	0.62
2:D:79:ARG:O	2:D:80:HIS:CB	2.50	0.58
2:C:153:TRP:CZ2	2:D:146:HIS:CG	2.94	0.56
2:C:146:HIS:CD2	2:D:146:HIS:O	2.60	0.55
1:F:502:VAL:O	1:F:505:LEU:O	2.25	0.54
2:D:85:THR:CG2	2:D:116:LEU:CD2	2.86	0.54
2:C:32:THR:OG1	3:C:401:G3D:O2A	2.26	0.54
2:G:151:ARG:NH1	2:G:151:ARG:CG	2.71	0.53
2:C:46:ILE:O	2:C:47:PRO:O	2.26	0.53
1:F:580:GLY:O	1:F:584:ARG:CG	2.57	0.52
1:F:531:PHE:CZ	1:F:535:PHE:CE2	2.97	0.52
2:H:67:ASP:O	2:H:69:GLY:N	2.43	0.52
2:G:32:THR:N	3:G:401:G3D:O2A	2.43	0.52
1:E:476:ASP:OD1	1:E:476:ASP:N	2.42	0.52
2:H:32:THR:OG1	3:H:401:G3D:O2A	2.27	0.51
2:H:68:VAL:O	2:H:70:GLY:N	2.44	0.51
2:H:173:LEU:O	2:H:175:ASN:N	2.44	0.50
1:A:531:PHE:CZ	1:A:535:PHE:CE2	3.00	0.50
1:E:531:PHE:CZ	1:E:535:PHE:CE2	3.00	0.49
2:H:159:CYS:SG	2:H:162:SER:OG	2.70	0.49
1:B:531:PHE:CZ	1:B:535:PHE:CE2	3.00	0.49
2:C:68:VAL:CG1	2:C:69:GLY:N	2.76	0.49
2:D:68:VAL:O	2:D:70:GLY:N	2.45	0.48
1:A:483:ASP:OD1	1:A:584:ARG:NE	2.47	0.47
1:E:555:MET:O	1:E:559:ASP:CB	2.62	0.47
1:B:476:ASP:OD1	1:B:476:ASP:N	2.48	0.47
1:B:483:ASP:OD1	1:B:584:ARG:NE	2.48	0.47
1:B:479:THR:O	1:B:479:THR:CG2	2.63	0.47
1:E:483:ASP:OD1	1:E:584:ARG:NE	2.50	0.44
1:E:458:ASN:O	1:E:460:GLN:N	2.52	0.43
1:B:423:TYR:OH	1:B:427:ARG:NH1	2.52	0.43
2:D:19:ARG:O	2:D:85:THR:O	2.37	0.42
1:B:397:SER:OG	1:B:398:ASN:N	2.52	0.42
2:D:145:LEU:O	2:D:146:HIS:CB	2.67	0.42
1:E:531:PHE:CE1	1:E:535:PHE:CE2	3.08	0.42
2:D:97:ARG:NH1	2:D:132:ASN:O	2.53	0.42
1:E:423:TYR:OH	1:E:427:ARG:NH1	2.53	0.42
1:F:423:TYR:OH	1:F:427:ARG:NH1	2.53	0.42
1:A:413:PHE:C	1:A:413:PHE:CD1	2.94	0.42
2:H:97:ARG:NH1	2:H:132:ASN:O	2.53	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:557:LEU:O	1:F:558:GLU:CB	2.68	0.41
1:A:423:TYR:OH	1:A:427:ARG:NH1	2.54	0.41
2:G:177:LEU:O	2:G:178:ARG:C	2.59	0.41
1:E:459:ARG:NH1	1:E:500:GLN:O	2.53	0.41
2:G:97:ARG:NH1	2:G:132:ASN:O	2.53	0.41

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	348/383 (91%)	320 (92%)	18 (5%)	10 (3%)	7	47
1	B	354/383 (92%)	322 (91%)	21 (6%)	11 (3%)	7	45
1	E	347/383 (91%)	316 (91%)	21 (6%)	10 (3%)	7	47
1	F	347/383 (91%)	320 (92%)	18 (5%)	9 (3%)	8	50
2	C	159/164 (97%)	145 (91%)	6 (4%)	8 (5%)	3	28
2	D	162/164 (99%)	146 (90%)	6 (4%)	10 (6%)	2	21
2	G	160/164 (98%)	147 (92%)	8 (5%)	5 (3%)	7	45
2	H	157/164 (96%)	136 (87%)	13 (8%)	8 (5%)	3	28
All	All	2034/2188 (93%)	1852 (91%)	111 (6%)	71 (4%)	6	41

All (71) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	672	LYS
1	B	459	ARG
1	B	625	SER
1	B	713	ASP
2	C	39	LEU
2	C	47	PRO
2	C	71	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	D	42	ILE
2	D	43	VAL
2	D	74	ILE
2	D	80	HIS
2	D	86	GLN
1	E	398	ASN
1	E	526	ASN
1	E	671	LYS
1	F	558	GLU
1	F	625	SER
1	F	673	LYS
1	F	713	ASP
2	G	70	GLY
2	G	178	ARG
2	H	43	VAL
2	H	45	THR
2	H	68	VAL
2	H	69	GLY
1	A	625	SER
1	B	397	SER
1	B	558	GLU
2	C	44	THR
2	C	68	VAL
2	C	72	ASP
2	D	69	GLY
2	D	146	HIS
2	D	150	HIS
1	E	550	LYS
1	E	673	LYS
1	F	609	GLY
1	F	680	PHE
2	G	39	LEU
2	G	82	PHE
2	H	72	ASP
2	H	174	SER
1	A	714	ILE
2	D	73	LYS
1	E	459	ARG
1	E	551	PRO
1	E	674	ASN
2	G	69	GLY
1	A	506	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	550	LYS
1	A	671	LYS
1	A	674	ASN
1	A	680	PHE
1	B	714	ILE
2	D	49	ILE
1	E	397	SER
1	F	714	ILE
1	B	592	THR
2	H	49	ILE
1	A	700	PRO
1	B	700	PRO
2	C	48	THR
1	E	700	PRO
1	F	700	PRO
1	A	551	PRO
1	B	550	LYS
1	B	623	VAL
2	C	42	ILE
1	F	526	ASN
1	B	551	PRO
2	H	74	ILE

### 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	323/347 (93%)	300 (93%)	23 (7%)	21	65
1	B	329/347 (95%)	301 (92%)	28 (8%)	15	54
1	E	323/347 (93%)	289 (90%)	34 (10%)	10	40
1	F	321/347 (92%)	292 (91%)	29 (9%)	14	50
2	C	138/142 (97%)	126 (91%)	12 (9%)	15	53
2	D	140/142 (99%)	128 (91%)	12 (9%)	15	54
2	G	138/142 (97%)	125 (91%)	13 (9%)	13	47
2	H	137/142 (96%)	118 (86%)	19 (14%)	5	25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	1849/1956 (94%)	1679 (91%)	170 (9%)	13	48

All (170) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	398	ASN
1	A	402	ARG
1	A	496	GLN
1	A	498	LYS
1	A	506	ILE
1	A	521	VAL
1	A	557	LEU
1	A	565	ARG
1	A	571	GLU
1	A	597	VAL
1	A	604	GLU
1	A	624	LEU
1	A	625	SER
1	A	668	ILE
1	A	677	THR
1	A	685	SER
1	A	707	SER
1	A	737	GLU
1	A	745	MET
1	A	747	LYS
1	A	753	GLU
1	A	754	LEU
1	A	755	GLU
1	B	402	ARG
1	B	460	GLN
1	B	476	ASP
1	B	496	GLN
1	B	498	LYS
1	B	506	ILE
1	B	521	VAL
1	B	525	ARG
1	B	550	LYS
1	B	552	GLU
1	B	565	ARG
1	B	576	GLU
1	B	597	VAL
1	B	604	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	610	LYS
1	B	668	ILE
1	B	672	LYS
1	B	677	THR
1	B	681	ARG
1	B	685	SER
1	B	689	MET
1	B	692	LEU
1	B	737	GLU
1	B	745	MET
1	B	747	LYS
1	B	754	LEU
1	B	755	GLU
1	B	757	GLN
2	C	35	TYR
2	C	36	LYS
2	C	42	ILE
2	C	48	THR
2	C	51	PHE
2	C	60	ASN
2	C	65	VAL
2	C	74	ILE
2	C	75	ARG
2	C	83	GLN
2	C	114	ASP
2	C	173	LEU
2	D	34	LEU
2	D	35	TYR
2	D	36	LYS
2	D	56	VAL
2	D	60	ASN
2	D	65	VAL
2	D	74	ILE
2	D	75	ARG
2	D	84	ASN
2	D	114	ASP
2	D	115	GLU
2	D	180	GLN
1	E	402	ARG
1	E	445	LYS
1	E	456	LEU
1	E	476	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	493	ILE
1	E	496	GLN
1	E	498	LYS
1	E	504	ARG
1	E	506	ILE
1	E	507	GLU
1	E	521	VAL
1	E	530	ILE
1	E	550	LYS
1	E	552	GLU
1	E	555	MET
1	E	558	GLU
1	E	559	ASP
1	E	565	ARG
1	E	571	GLU
1	E	575	ARG
1	E	576	GLU
1	E	597	VAL
1	E	604	GLU
1	E	668	ILE
1	E	669	PHE
1	E	681	ARG
1	E	685	SER
1	E	689	MET
1	E	692	LEU
1	E	714	ILE
1	E	737	GLU
1	E	753	GLU
1	E	754	LEU
1	E	755	GLU
1	F	398	ASN
1	F	412	LEU
1	F	459	ARG
1	F	476	ASP
1	F	496	GLN
1	F	498	LYS
1	F	506	ILE
1	F	512	ARG
1	F	521	VAL
1	F	558	GLU
1	F	571	GLU
1	F	576	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	583	GLU
1	F	597	VAL
1	F	604	GLU
1	F	624	LEU
1	F	668	ILE
1	F	672	LYS
1	F	677	THR
1	F	679	SER
1	F	685	SER
1	F	689	MET
1	F	692	LEU
1	F	707	SER
1	F	715	LYS
1	F	737	GLU
1	F	745	MET
1	F	754	LEU
1	F	755	GLU
2	G	35	TYR
2	G	36	LYS
2	G	39	LEU
2	G	46	ILE
2	G	51	PHE
2	G	65	VAL
2	G	71	GLN
2	G	72	ASP
2	G	73	LYS
2	G	79	ARG
2	G	114	ASP
2	G	151	ARG
2	G	173	LEU
2	H	35	TYR
2	H	36	LYS
2	H	39	LEU
2	H	41	GLU
2	H	43	VAL
2	H	45	THR
2	H	60	ASN
2	H	65	VAL
2	H	74	ILE
2	H	75	ARG
2	H	79	ARG
2	H	114	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	H	118	ASP
2	H	146	HIS
2	H	149	ARG
2	H	150	HIS
2	H	159	CYS
2	H	176	GLN
2	H	177	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

4 ligands are 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	G3D	C	401	-	34,34,34	1.28	2 (5%)	51,54,54	3.80	6 (11%)
3	G3D	D	401	-	34,34,34	1.11	3 (8%)	51,54,54	3.41	6 (11%)
3	G3D	G	401	-	34,34,34	1.12	2 (5%)	51,54,54	3.83	8 (15%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	G3D	H	401	-	34,34,34	1.04	3 (8%)	51,54,54	3.46	6 (11%)

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	G3D	C	401	-	-	0/21/37/37	0/1/3/3
3	G3D	D	401	-	-	0/21/37/37	0/1/3/3
3	G3D	G	401	-	-	0/21/37/37	0/1/3/3
3	G3D	H	401	-	-	0/21/37/37	0/1/3/3

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	401	G3D	PA-O3A	-4.91	1.51	1.59
3	G	401	G3D	PA-O3A	-4.06	1.52	1.59
3	C	401	G3D	PB-O3A	-3.72	1.53	1.60
3	D	401	G3D	C6-C5	3.44	1.46	1.41
3	H	401	G3D	PA-O3A	-3.04	1.54	1.59
3	G	401	G3D	PB-O3A	-2.90	1.55	1.60
3	D	401	G3D	PA-O3A	-2.89	1.54	1.59
3	H	401	G3D	PB-O3A	-2.56	1.55	1.60
3	D	401	G3D	PB-O3A	-2.48	1.55	1.60
3	H	401	G3D	C6-C5	2.34	1.45	1.41

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	401	G3D	C6-C5-N7	-21.14	131.29	134.14
3	G	401	G3D	C6-C5-N7	-21.08	131.30	134.14
3	D	401	G3D	C6-C5-N7	-18.76	131.61	134.14
3	H	401	G3D	C6-C5-N7	-18.05	131.71	134.14
3	G	401	G3D	PA-O3A-PB	14.17	173.24	131.68
3	C	401	G3D	PA-O3A-PB	14.08	172.97	131.68
3	H	401	G3D	PA-O3A-PB	13.85	172.31	131.68
3	D	401	G3D	PA-O3A-PB	13.05	169.96	131.68
3	C	401	G3D	P1-O3'-C3'	5.10	132.70	121.96
3	G	401	G3D	P1-O3'-C3'	5.09	132.68	121.96
3	H	401	G3D	P1-O3'-C3'	4.93	132.34	121.96
3	D	401	G3D	C6-N1-C2	3.35	125.37	119.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	401	G3D	C6-N1-C2	3.25	125.19	119.51
3	H	401	G3D	C6-N1-C2	3.24	125.17	119.51
3	H	401	G3D	O2B-PB-O3A	3.22	120.43	105.14
3	C	401	G3D	C6-N1-C2	3.20	125.10	119.51
3	C	401	G3D	O2B-PB-O3A	3.09	119.80	105.14
3	G	401	G3D	O2B-PB-O3A	3.09	119.78	105.14
3	D	401	G3D	P1-O3'-C3'	2.97	128.20	121.96
3	D	401	G3D	O4P-P1-O3'	2.94	115.56	107.09
3	D	401	G3D	O2B-PB-O3A	2.59	117.43	105.14
3	H	401	G3D	O4P-P1-O3'	2.37	113.92	107.09
3	G	401	G3D	O5'-PA-O1A	-2.22	100.66	109.37
3	G	401	G3D	O3B-PB-O2B	2.11	115.82	107.61
3	G	401	G3D	O3'-P1-O6P	-2.10	100.92	106.79
3	C	401	G3D	O3B-PB-O2B	2.05	115.59	107.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	352/383 (91%)	-0.06	0 100 100	25, 59, 103, 132	0
1	B	357/383 (93%)	-0.13	1 (0%) 91 63	16, 52, 104, 126	0
1	E	351/383 (91%)	-0.15	0 100 100	27, 53, 94, 134	0
1	F	351/383 (91%)	-0.12	1 (0%) 91 63	25, 52, 92, 120	0
2	C	161/164 (98%)	-0.08	0 100 100	34, 70, 114, 141	0
2	D	164/164 (100%)	-0.09	0 100 100	31, 65, 102, 127	0
2	G	162/164 (98%)	0.23	4 (2%) 54 14	54, 100, 139, 155	0
2	H	159/164 (96%)	0.51	8 (5%) 28 6	69, 135, 196, 215	0
All	All	2057/2188 (94%)	-0.03	14 (0%) 83 42	16, 61, 132, 215	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	173	LEU	2.9
2	H	139	ILE	2.7
2	H	143	LEU	2.5
2	G	18	MET	2.4
2	H	146	HIS	2.3
2	H	136	ALA	2.3
2	H	130	LEU	2.3
1	F	755	GLU	2.3
2	H	131	PRO	2.2
2	G	130	LEU	2.2
2	G	131	PRO	2.2
2	G	40	GLY	2.1
2	H	151	ARG	2.1
1	B	755	GLU	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	G3D	H	401	32/32	0.23	-0.37	97,113,124,129	0
3	G3D	C	401	32/32	0.18	-0.47	42,69,88,93	0
3	G3D	G	401	32/32	0.21	-0.53	66,89,99,105	0
3	G3D	D	401	32/32	0.15	-0.81	47,53,83,91	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.