



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 01:18 AM GMT

PDB ID : 2A40  
Title : Ternary complex of the WH2 domain of WAVE with Actin-DNAse I  
Authors : Chereau, D.; Kerff, F.; Dominguez, R.  
Deposited on : 2005-06-27  
Resolution : 1.80 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>

---

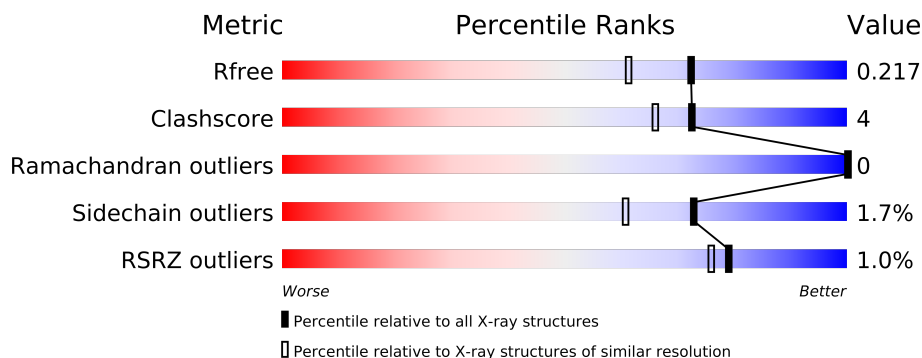
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.80 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	3513 (1.80-1.80)
Clashscore	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)
RSRZ outliers	66119	3515 (1.80-1.80)

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	375	
1	D	375	
2	B	260	
2	E	260	
3	C	32	
3	F	32	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
5	CA	A	1381	-	X
5	CA	E	1272	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
8	GOL	A	1383	-	X
8	GOL	B	1384	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 11357 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 Actin, alpha skeletal muscle.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	15	0
			2899	1843	481	552	23			
1	D	362	Total	C	N	O	S	0	12	0
			2878	1829	477	549	23			

- Molecule 2 is a protein called Deoxyribonuclease-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	260	Total	C	N	O	S	0	5	0
			2071	1312	345	406	8			
2	E	260	Total	C	N	O	S	0	12	0
			2109	1340	351	410	8			

- Molecule 3 is a protein called Wiskott-Aldrich syndrome protein family member 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	22	Total	C	N	O	0	1	0
			181	109	36	36			
3	F	22	Total	C	N	O	0	1	0
			181	109	36	36			

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	2	Total	C	N	O	0	0
			28	16	2	10		
4	E	2	Total	C	N	O	0	0
			28	16	2	10		

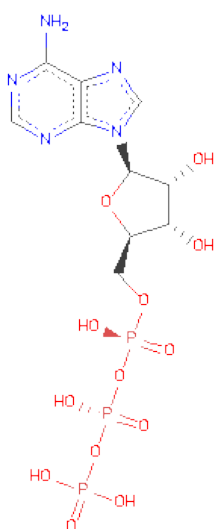
- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Ca	0	0
			1	1		
5	A	1	Total	Ca	0	0
			1	1		
5	D	1	Total	Ca	0	0
			1	1		
5	E	1	Total	Ca	0	0
			1	1		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

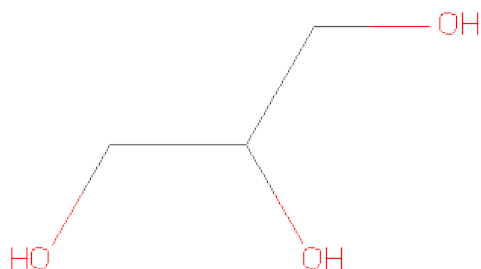
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Mg	0	0
			1	1		
6	E	1	Total	Mg	0	0
			1	1		

- Molecule 7 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
7	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			6	3	3		
8	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 9 is water.

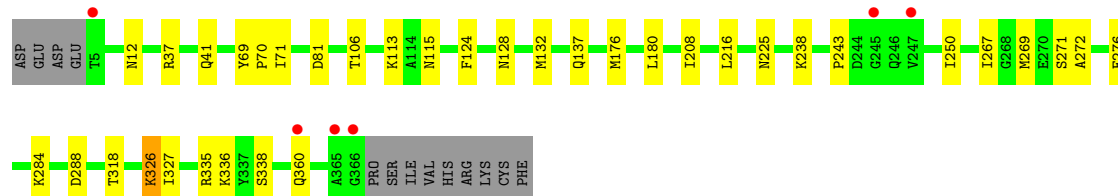
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	288	Total	O	0	0
			288	288		
9	B	167	Total	O	0	0
			167	167		
9	C	7	Total	O	0	0
			7	7		
9	D	253	Total	O	0	0
			253	253		
9	E	183	Total	O	0	0
			183	183		
9	F	4	Total	O	0	0
			4	4		

### 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.

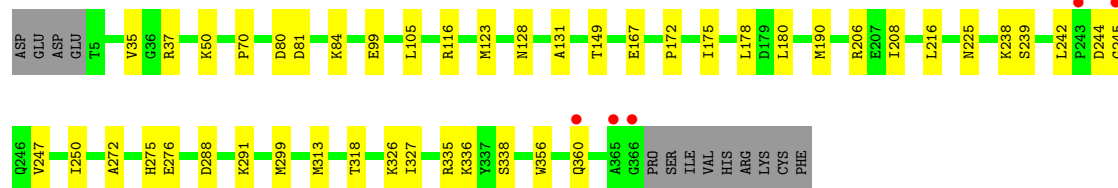
- Molecule 1: Actin, alpha skeletal muscle

Chain A: 



- Molecule 1: Actin, alpha skeletal muscle

Chain D: 



- Molecule 2: Deoxyribonuclease-1

Chain B: 



- Molecule 2: Deoxyribonuclease-1

Chain E: 



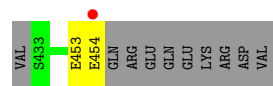
- Molecule 3: Wiskott-Aldrich syndrome protein family member 2

Chain C: 



- Molecule 3: Wiskott-Aldrich syndrome protein family member 2

Chain F:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.72Å 41.60Å 153.05Å 90.00° 108.79° 90.00°	Depositor
Resolution (Å)	47.51 – 1.80 47.49 – 1.80	Depositor EDS
% Data completeness (in resolution range)	80.6 (47.51-1.80) 80.6 (47.49-1.80)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.48 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.166 , 0.215 0.168 , 0.217	Depositor DCC
$R_{free}$ test set	5386 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtriage
Anisotropy	0.327	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 56.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	6 of 106912 reflections (0.006%)	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	11357	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	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 84.91 % 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 1.4638e-07. 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: GOL, MG, NAG, CA, HIC, ATP

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	0.63	0/2988	0.71	0/4043
1	D	0.57	0/2961	0.65	0/4007
2	B	0.65	2/2132 (0.1%)	0.73	1/2902 (0.0%)
2	E	0.60	0/2192	0.73	1/2981 (0.0%)
3	C	0.48	0/184	0.51	0/244
3	F	0.40	0/184	0.47	0/244
All	All	0.61	2/10641 (0.0%)	0.70	2/14421 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	122	SER	C-O	6.28	1.35	1.23
2	B	173	CYS	CB-SG	5.49	1.91	1.82

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	123	THR	CA-CB-CG2	-5.40	104.85	112.40
2	E	166	MET	CG-SD-CE	5.37	108.79	100.20

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	2899	0	2912	30	0
1	D	2878	0	2886	35	0
2	B	2071	0	2014	8	0
2	E	2109	0	2074	15	0
3	C	181	0	182	2	0
3	F	181	0	182	1	0
4	B	28	0	25	0	0
4	E	28	0	25	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
6	B	1	0	0	0	0
6	E	1	0	0	0	0
7	A	31	0	12	0	0
7	D	31	0	12	0	0
8	A	6	0	8	0	0
8	B	6	0	8	0	0
9	A	288	0	0	8	0
9	B	167	0	0	2	1
9	C	7	0	0	0	0
9	D	253	0	0	4	0
9	E	183	0	0	4	0
9	F	4	0	0	0	0
All	All	11357	0	10340	90	1

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 4.

All (90) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:12[B]:ASN:ND2	9:A:1461:HOH:O	1.62	1.30
2:E:121[B]:HIS:HD2	9:E:1433:HOH:O	1.19	1.21
1:D:299:MET:HE3	1:D:313[B]:MET:HG3	1.30	1.11
1:D:299:MET:CE	1:D:313[B]:MET:HG3	1.80	1.11
1:D:37:ARG:NH2	1:D:84:LYS:HE3	1.67	1.08
1:D:216:LEU:HD22	1:D:238:LYS:HE3	1.43	1.01
2:E:92[A]:LEU:HD11	2:E:119:SER:HB3	1.57	0.86
1:D:37:ARG:HH22	1:D:84:LYS:HE3	1.47	0.77
2:B:103:SER:OG	9:B:1502:HOH:O	1.96	0.76

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:226:VAL:CG2	2:B:260:THR:HG23	2.17	0.73
1:A:360:GLN:NE2	1:A:360:GLN:H	1.87	0.73
1:D:244:ASP:OD1	1:D:245:GLY:N	2.22	0.72
1:A:115[A]:ASN:ND2	9:A:1640:HOH:O	2.23	0.72
2:B:17[B]:SER:OG	9:B:1523:HOH:O	2.11	0.67
1:D:291:LYS:HE3	9:D:1622:HOH:O	1.94	0.67
1:A:176:MET:HE2	1:A:284[B]:LYS:HD2	1.77	0.67
1:D:216:LEU:HD22	1:D:238:LYS:CE	2.21	0.67
1:D:216:LEU:CD2	1:D:238:LYS:HE3	2.25	0.65
2:E:121[B]:HIS:ND1	9:E:1434:HOH:O	2.15	0.60
2:E:92[A]:LEU:HD11	2:E:119:SER:CB	2.31	0.59
1:D:299:MET:HE1	1:D:313[B]:MET:HG3	1.78	0.58
1:D:275:HIS:H	1:D:275:HIS:CD2	2.23	0.56
1:A:326:LYS:NZ	9:A:1571:HOH:O	2.39	0.54
1:D:116[B]:ARG:CZ	9:D:1499:HOH:O	2.55	0.54
1:A:360:GLN:H	1:A:360:GLN:HE21	1.56	0.53
2:B:226:VAL:HG22	2:B:260:THR:HG23	1.90	0.53
1:D:37:ARG:HH21	1:D:84:LYS:HE3	1.66	0.53
1:D:35:VAL:HG11	1:D:84:LYS:HD2	1.91	0.52
1:A:216:LEU:HD22	1:A:238[A]:LYS:HE3	1.91	0.52
1:A:106:THR:HB	1:A:137:GLN:HG3	1.91	0.52
3:C:453:GLU:HA	3:C:453:GLU:OE2	2.10	0.52
1:D:272:ALA:HB1	1:D:276:GLU:HB2	1.92	0.51
1:A:115[B]:ASN:ND2	9:A:1440:HOH:O	2.43	0.51
1:A:176:MET:CE	1:A:284[B]:LYS:HD2	2.41	0.50
1:D:299:MET:HE3	1:D:313[B]:MET:CG	2.22	0.50
1:A:70:PRO:HG3	1:A:81:ASP:HB3	1.94	0.50
1:D:80:ASP:O	1:D:84:LYS:HE2	2.12	0.49
1:A:12[B]:ASN:CG	9:A:1461:HOH:O	2.26	0.49
2:E:260:THR:HG22	9:E:1344:HOH:O	2.12	0.49
1:D:37:ARG:NH2	1:D:84:LYS:CE	2.58	0.49
1:A:176:MET:HE1	1:A:284[B]:LYS:HE3	1.95	0.49
1:A:272:ALA:HB1	1:A:276[B]:GLU:HB2	1.94	0.48
1:D:172:PRO:HA	1:D:175:ILE:HD12	1.95	0.48
2:E:121[B]:HIS:CD2	9:E:1433:HOH:O	2.11	0.48
1:A:12[A]:ASN:HB3	1:A:71:ILE:CD1	2.43	0.48
1:A:37:ARG:NH2	1:A:81:ASP:OD1	2.36	0.48
3:C:437:SER:O	3:C:441[A]:SER:OG	2.30	0.48
2:E:24:TYR:HE1	2:E:27[B]:ARG:HH21	1.59	0.48
1:A:41:GLN:HB2	2:B:64:HIS:CD2	2.49	0.48
2:E:226:VAL:CG2	2:E:260:THR:HG23	2.45	0.47
1:D:318:THR:HA	1:D:327:ILE:HD13	1.96	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:326:LYS:HB3	1:D:326:LYS:HE3	1.69	0.47
1:D:149:THR:OG1	1:D:167:GLU:OE2	2.30	0.47
1:D:275:HIS:HD2	1:D:275:HIS:H	1.63	0.46
1:A:180[B]:LEU:HD13	1:A:267:ILE:HD11	1.97	0.46
1:A:69:TYR:HE1	9:A:1564:HOH:O	1.97	0.46
1:D:99:GLU:H	1:D:99:GLU:CD	2.19	0.46
2:E:106:ASN:HA	2:E:109:PHE:O	2.15	0.46
1:A:360:GLN:N	1:A:360:GLN:NE2	2.60	0.46
1:D:335:ARG:HA	1:D:338:SER:OG	2.17	0.45
1:D:105:LEU:HD11	1:D:123:MET:HG3	1.98	0.45
2:E:24:TYR:HE1	2:E:27[B]:ARG:NH2	2.14	0.45
1:D:336:LYS:HG2	9:D:1545:HOH:O	2.16	0.44
2:E:230:ALA:HA	2:E:256:GLU:O	2.17	0.44
1:A:318:THR:HA	1:A:327:ILE:CD1	2.47	0.44
1:D:208:ILE:HD13	1:D:242:LEU:HD22	1.99	0.44
1:A:176:MET:HE1	1:A:284[B]:LYS:CE	2.47	0.44
1:A:336:LYS:HE2	9:A:1632:HOH:O	2.18	0.44
3:F:454:GLU:H	3:F:454:GLU:HG3	1.56	0.44
1:D:70:PRO:HG3	1:D:81:ASP:HB3	2.00	0.43
1:A:113:LYS:HE3	1:A:113:LYS:HB3	1.91	0.43
1:A:124:PHE:O	1:A:128[A]:ASN:HA	2.19	0.43
1:A:124:PHE:O	1:A:128[B]:ASN:HA	2.19	0.43
1:D:190:MET:HE1	9:D:1512:HOH:O	2.18	0.42
2:B:220:LEU:HD12	2:B:220:LEU:HA	1.92	0.42
1:D:178:LEU:HG	1:D:180[B]:LEU:HB3	2.02	0.42
1:A:335:ARG:HA	1:A:338:SER:OG	2.20	0.42
2:E:115:VAL:HG12	2:E:158:TRP:CZ3	2.55	0.42
1:D:131:ALA:HB1	1:D:356:TRP:HB3	2.01	0.41
1:D:190:MET:CE	1:D:206:ARG:NH2	2.84	0.41
2:B:230:ALA:HA	2:B:256:GLU:O	2.19	0.41
2:B:92:LEU:HD11	2:B:119:SER:HB2	2.02	0.41
1:A:269[A]:MET:HG2	1:A:271:SER:HB3	2.02	0.41
2:E:30[B]:ARG:HH22	2:E:57:GLN:HE21	1.69	0.41
2:E:216:VAL:HG23	2:E:222:GLN:HG2	2.03	0.41
1:A:208:ILE:HD11	1:A:243:PRO:HG2	2.03	0.40
1:A:336:LYS:HG2	9:A:1441:HOH:O	2.21	0.40
1:D:275:HIS:CD2	1:D:275:HIS:N	2.89	0.40
1:D:239:SER:HB3	1:D:247:VAL:CG1	2.52	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
9:B:1464:HOH:O	9:B:1550:HOH:O[2.556]	2.18	0.02

## 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	374/375 (100%)	369 (99%)	5 (1%)	0	100	100
1	D	371/375 (99%)	365 (98%)	6 (2%)	0	100	100
2	B	263/260 (101%)	255 (97%)	8 (3%)	0	100	100
2	E	270/260 (104%)	266 (98%)	4 (2%)	0	100	100
3	C	21/32 (66%)	21 (100%)	0	0	100	100
3	F	21/32 (66%)	21 (100%)	0	0	100	100
All	All	1320/1334 (99%)	1297 (98%)	23 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	319/317 (101%)	314 (98%)	5 (2%)	75	63
1	D	316/317 (100%)	310 (98%)	6 (2%)	69	56
2	B	234/229 (102%)	232 (99%)	2 (1%)	87	83
2	E	241/229 (105%)	238 (99%)	3 (1%)	82	74
3	C	20/29 (69%)	17 (85%)	3 (15%)	4	0
3	F	20/29 (69%)	19 (95%)	1 (5%)	34	14
All	All	1150/1150 (100%)	1130 (98%)	20 (2%)	73	61

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	132	MET
1	A	225	ASN
1	A	250	ILE
1	A	288	ASP
1	A	326	LYS
2	B	2	LYS
2	B	101	CYS
3	C	441[A]	SER
3	C	441[B]	SER
3	C	454	GLU
1	D	50	LYS
1	D	128	ASN
1	D	225	ASN
1	D	250	ILE
1	D	288	ASP
1	D	360	GLN
2	E	101	CYS
2	E	119	SER
2	E	156	GLN
3	F	453	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	92	ASN
1	A	225	ASN
1	A	280	ASN
1	A	360	GLN
2	B	61	ASN
2	B	155	GLN
2	B	161	ASN
2	B	236	GLN
2	B	243	ASN
1	D	92	ASN
1	D	128	ASN
1	D	275	HIS
1	D	280	ASN
2	E	57	GLN
2	E	61	ASN
2	E	74	ASN
2	E	155	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	E	161	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

2 non-standard protein/DNA/RNA residues 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
1	HIC	A	73	1	11,11,12	4.82	3 (27%)	12,14,16	1.89	3 (25%)
1	HIC	D	73	1	11,11,12	5.30	2 (18%)	12,14,16	1.76	4 (33%)

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
1	HIC	A	73	1	-	0/4/6/8	0/1/1/1
1	HIC	D	73	1	-	0/4/6/8	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	73	HIC	O-C	17.22	1.23	1.11
1	A	73	HIC	O-C	15.36	1.21	1.11
1	A	73	HIC	CD2-CG	3.13	1.39	1.36
1	D	73	HIC	CD2-CG	2.43	1.38	1.36
1	A	73	HIC	CA-C	2.13	1.52	1.48

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	73	HIC	C-CA-N	-4.51	109.32	113.83
1	D	73	HIC	CB-CG-ND1	3.31	127.93	122.17
1	A	73	HIC	CB-CG-ND1	3.07	127.52	122.17
1	D	73	HIC	C-CA-N	-2.51	111.33	113.83
1	D	73	HIC	CB-CG-CD2	-2.44	125.28	130.79
1	A	73	HIC	CB-CG-CD2	-2.27	125.67	130.79
1	D	73	HIC	CG-CB-CA	-2.20	110.84	113.85

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

4 carbohydrates 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	$\# Z  > 2$
4	NAG	B	270	2,4	12,14,15	0.53	0	15,19,21	1.31	2 (13%)
4	NAG	B	271	4	12,14,15	0.71	1 (8%)	15,19,21	0.95	0
4	NAG	E	270	2,4	12,14,15	0.68	1 (8%)	15,19,21	0.89	0
4	NAG	E	271	4	12,14,15	0.62	0	15,19,21	1.53	2 (13%)

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
4	NAG	B	270	2,4	-	0/6/23/26	0/1/1/1
4	NAG	B	271	4	-	0/6/23/26	0/1/1/1
4	NAG	E	270	2,4	-	0/6/23/26	0/1/1/1

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	E	271	4	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	271	NAG	O5-C5	-2.20	1.41	1.45
4	E	270	NAG	O5-C5	-2.10	1.41	1.45

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	271	NAG	O5-C5-C6	3.76	110.92	106.98
4	B	270	NAG	O5-C5-C6	3.22	110.36	106.98
4	B	270	NAG	C3-C2-N2	-2.75	107.57	111.76
4	E	271	NAG	C3-C2-N2	-2.66	107.71	111.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 6 are monoatomic - leaving 4 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	$\# Z  > 2$
7	ATP	A	1379	5	33,33,33	1.39	5 (15%)	52,52,52	1.57	7 (13%)
8	GOL	A	1383	-	5,5,5	0.32	0	5,5,5	0.88	0
8	GOL	B	1384	-	5,5,5	0.40	0	5,5,5	0.69	0
7	ATP	D	1380	5	33,33,33	1.16	2 (6%)	52,52,52	1.76	9 (17%)

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
7	ATP	A	1379	5	-	0/22/38/38	0/1/3/3
8	GOL	A	1383	-	-	0/4/4/4	0/0/0/0
8	GOL	B	1384	-	-	0/4/4/4	0/0/0/0
7	ATP	D	1380	5	-	0/22/38/38	0/1/3/3

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1379	ATP	C4-N9	-4.14	1.31	1.37
7	A	1379	ATP	C5-C4	2.80	1.46	1.40
7	D	1380	ATP	C4-N9	-2.54	1.34	1.37
7	D	1380	ATP	C5-C4	2.46	1.46	1.40
7	A	1379	ATP	PG-O3B	2.43	1.64	1.60
7	A	1379	ATP	PA-O3A	2.23	1.63	1.59
7	A	1379	ATP	PB-O3B	2.15	1.63	1.59

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	1380	ATP	N3-C2-N1	-6.85	122.98	128.71
7	A	1379	ATP	N3-C2-N1	-5.25	124.32	128.71
7	D	1380	ATP	N3-C4-N9	4.93	134.34	125.43
7	A	1379	ATP	N3-C4-N9	4.69	133.90	125.43
7	D	1380	ATP	C8-N9-C4	3.41	109.50	106.90
7	A	1379	ATP	O4'-C1'-C2'	-3.11	102.00	106.77
7	A	1379	ATP	O3G-PG-O2G	2.77	118.38	107.61
7	D	1380	ATP	O3G-PG-O2G	2.73	118.24	107.61
7	A	1379	ATP	C5-C4-N3	-2.72	119.78	125.70
7	D	1380	ATP	C5-C4-N3	-2.64	119.94	125.70
7	A	1379	ATP	C8-N9-C4	2.34	108.69	106.90
7	A	1379	ATP	C4-C5-N7	-2.34	107.52	109.52
7	D	1380	ATP	C2-N3-C4	2.19	120.25	114.01
7	D	1380	ATP	C4-C5-N7	-2.13	107.70	109.52
7	D	1380	ATP	C4'-O4'-C1'	-2.05	107.52	109.75
7	D	1380	ATP	O3A-PB-O3B	-2.05	97.49	101.66

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	362/375 (96%)	-0.38	6 (1%) 67 62	19, 26, 40, 61	0
1	D	362/375 (96%)	-0.32	5 (1%) 72 68	21, 30, 44, 61	0
2	B	260/260 (100%)	-0.50	0 100 100	23, 28, 41, 55	0
2	E	260/260 (100%)	-0.60	0 100 100	22, 27, 37, 50	0
3	C	22/32 (68%)	0.21	1 (4%) 32 25	34, 40, 55, 59	0
3	F	22/32 (68%)	0.47	1 (4%) 32 25	33, 43, 53, 61	0
All	All	1288/1334 (96%)	-0.41	13 (1%) 79 76	19, 28, 43, 61	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	366	GLY	5.9
1	A	366	GLY	5.5
1	A	365	ALA	4.7
1	D	365	ALA	4.6
1	D	360	GLN	3.4
1	A	5	THR	3.2
1	D	245	GLY	2.7
3	F	454	GLU	2.6
1	A	360	GLN	2.5
1	A	245	GLY	2.4
3	C	433	SER	2.4
1	A	247	VAL	2.2
1	D	243	PRO	2.2

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

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
1	HIC	D	73	11/12	0.08	-1.24	26,34,38,39	0
1	HIC	A	73	11/12	0.07	-1.34	23,30,34,35	0

### 6.3 Carbohydrates

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
4	NAG	B	270	14/15	0.21	6.22	34,35,42,43	0
4	NAG	E	270	14/15	0.19	6.21	33,37,43,44	0
4	NAG	B	271	14/15	0.32	4.12	40,43,45,45	0
4	NAG	E	271	14/15	0.26	3.44	42,44,46,47	0

### 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
8	GOL	A	1383	6/6	0.17	4.42	52,58,59,60	0
8	GOL	B	1384	6/6	0.15	4.34	29,32,35,37	0
5	CA	A	1381	1/1	0.11	3.31	20,20,20,20	0
5	CA	E	1272	1/1	0.09	2.55	23,23,23,23	0
7	ATP	A	1379	31/31	0.12	1.67	18,21,25,25	0
7	ATP	D	1380	31/31	0.11	1.27	21,24,26,27	0
5	CA	D	1382	1/1	0.10	1.24	23,23,23,23	0
5	CA	B	1273	1/1	0.09	1.00	24,24,24,24	0
6	MG	E	1275	1/1	0.08	0.41	27,27,27,27	0
6	MG	B	1274	1/1	0.06	-1.72	32,32,32,32	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.