



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:09 AM GMT

PDB ID : 2PN5  
Title : Crystal Structure of TEP1r  
Authors : Baxter, R.H.G.  
Deposited on : 2007-04-23  
Resolution : 2.70 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  
<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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

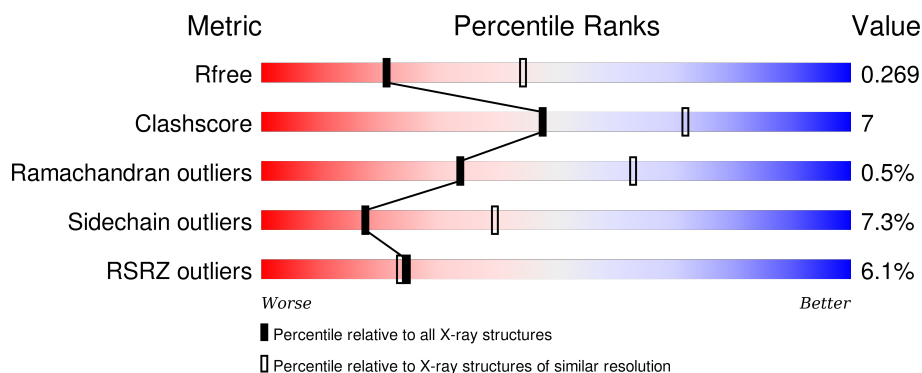
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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	1325	<div> <div>6%</div> <div>77%</div> <div>17%</div> <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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	A	1326	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	A	1328	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 10546 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 Thioester-containing protein I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1283	Total	C	N	O	S	0	7	0
			10333	6618	1724	1945	46			

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	Na	0	0
			3	3		

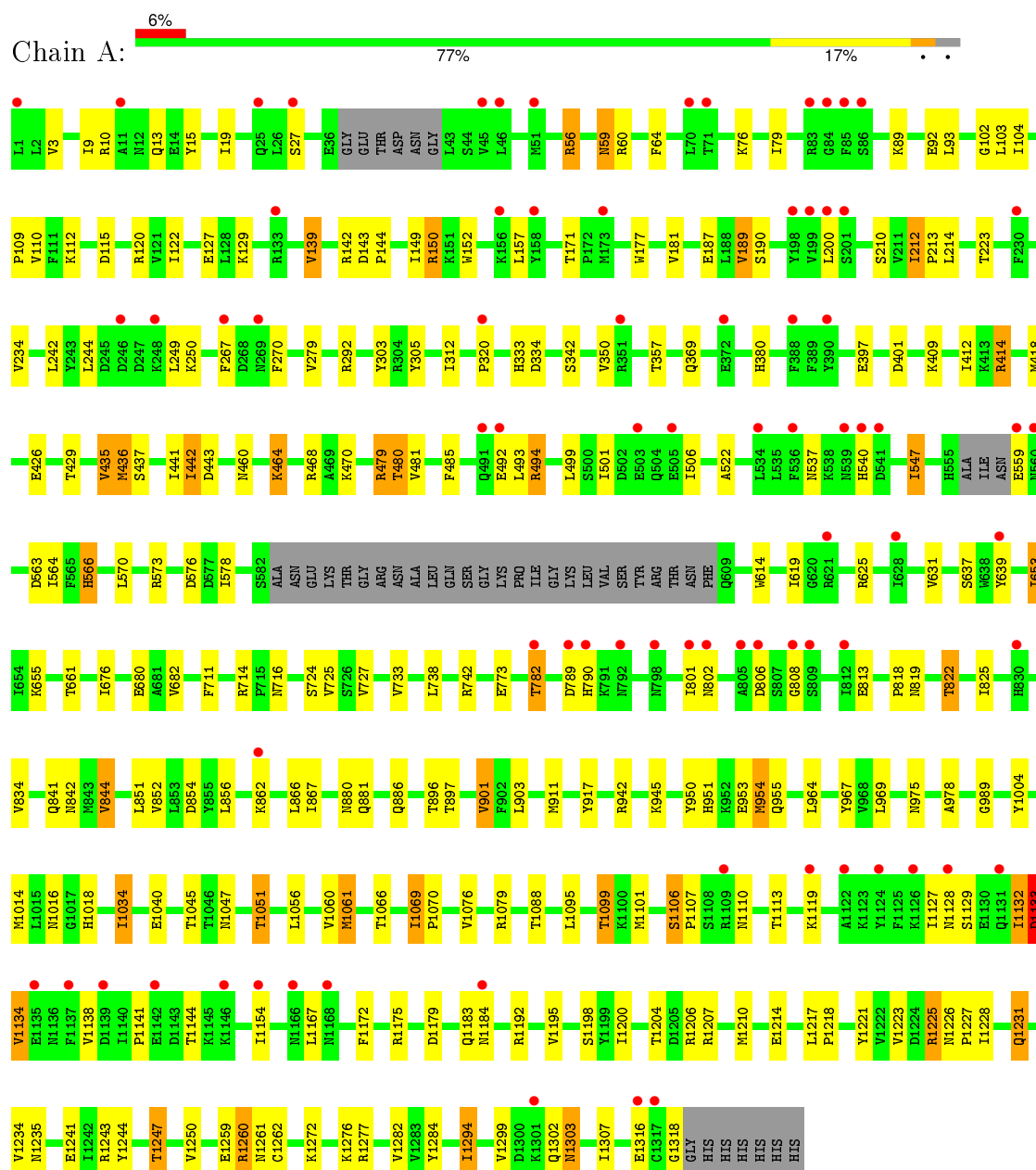
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	140	Total	O	0	0
			140	140		

### 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: Thioester-containing protein I



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	150.51Å 150.51Å 226.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.88 – 2.70 45.87 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (45.88-2.70) 99.8 (45.87-2.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.42 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, $R_{free}$	0.239 , 0.275 0.240 , 0.269	Depositor DCC
$R_{free}$ test set	3631 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	57.2	Xtriage
Anisotropy	0.160	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 53.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 71859 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	10546	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NA, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.46	1/10543 (0.0%)	0.59	1/14272 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1318	GLY	C-O	12.94	1.44	1.23

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1318	GLY	CA-C-O	-7.01	107.97	120.60

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	10333	0	10327	154	0
2	A	42	0	39	1	0
3	A	28	0	25	0	0
4	A	3	0	0	0	0
5	A	140	0	0	5	0
All	All	10546	0	10391	154	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 7.

All (154) 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:A:1303:ASN:ND2	1:A:1303:ASN:H	1.56	1.00
1:A:1303:ASN:N	1:A:1303:ASN:HD22	1.60	0.99
1:A:1231:GLN:HE21	1:A:1231:GLN:H	1.13	0.94
1:A:1244:TYR:O	1:A:1247:THR:HB	1.70	0.91
1:A:27:SER:HA	1:A:56:ARG:HD2	1.58	0.86
1:A:1066:THR:HG22	5:A:1440:HOH:O	1.77	0.82
1:A:637:SER:HB3	1:A:661:THR:HG22	1.62	0.81
1:A:1175:ARG:NH1	1:A:1200:ILE:O	2.17	0.78
1:A:89:LYS:HG2	1:A:573:ARG:HB3	1.65	0.78
1:A:950:TYR:H	1:A:955:GLN:HE22	1.28	0.78
1:A:1129:SER:OG	1:A:1132:ILE:HG23	1.86	0.76
1:A:1303:ASN:HD22	1:A:1303:ASN:H	0.78	0.75
1:A:1127:ILE:HG23	1:A:1132:ILE:HD13	1.68	0.74
1:A:187:GLU:OE2	2:A:1326:NAG:H62	1.89	0.73
1:A:1272:LYS:HB3	1:A:1307:ILE:HD12	1.71	0.73
1:A:9:ILE:HD13	1:A:93:LEU:HB3	1.71	0.71
1:A:1183:GLN:OE1	1:A:1192:ARG:NH1	2.24	0.70
1:A:187:GLU:OE1	1:A:190:SER:HB2	1.90	0.70
1:A:1221:TYR:OH	1:A:1299:VAL:HG11	1.91	0.69
1:A:1225:ARG:HG2	1:A:1226:ASN:HD22	1.57	0.69
1:A:1056:LEU:HD11	1:A:1101:MET:CE	2.23	0.68
1:A:1227:PRO:HG2	1:A:1228:ILE:HD12	1.75	0.68
1:A:414:ARG:HD3	1:A:493:LEU:HD11	1.75	0.68
1:A:436:MET:HE2	1:A:441:ILE:HG12	1.77	0.67
1:A:143:ASP:HB2	1:A:144:PRO:HD2	1.76	0.66
1:A:866:LEU:HD21	1:A:1134:VAL:HG22	1.78	0.66
1:A:412:ILE:HG12	1:A:418:MET:HG2	1.78	0.66
1:A:1056:LEU:HD11	1:A:1101:MET:HE3	1.78	0.65
1:A:951:HIS:H	1:A:955:GLN:HE21	1.45	0.65
1:A:342:SER:HB3	1:A:357:THR:HG22	1.81	0.63
1:A:1235:ASN:HB3	1:A:1261:ASN:HD21	1.64	0.62
1:A:1095:LEU:O	1:A:1099:THR:HB	1.99	0.62
1:A:818:PRO:HB3	1:A:1132:ILE:HD12	1.81	0.61
1:A:342:SER:CB	1:A:357:THR:HG22	2.31	0.60
1:A:1235:ASN:HB3	1:A:1261:ASN:ND2	2.17	0.60
1:A:813:GLU:HA	1:A:1141:PRO:HD2	1.84	0.60
1:A:350:VAL:HG12	1:A:350:VAL:O	2.02	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:109:PRO:HB2	1:A:537:ASN:ND2	2.16	0.59
1:A:818:PRO:CB	1:A:1132:ILE:HD12	2.32	0.59
1:A:819:ASN:HB3	1:A:822:THR:HG22	1.85	0.59
1:A:1045:THR:OG1	1:A:1051:THR:HG21	2.03	0.58
1:A:292:ARG:NH2	1:A:1225:ARG:HD2	2.18	0.58
1:A:171:THR:HB	1:A:880:ASN:HD21	1.68	0.58
1:A:436:MET:CE	1:A:441:ILE:HG12	2.33	0.58
1:A:856:LEU:HD13	1:A:867:ILE:HG12	1.85	0.58
1:A:547:ILE:HD12	1:A:653:ILE:HD11	1.86	0.58
1:A:822:THR:HB	1:A:1133:ASP:O	2.04	0.57
1:A:1231:GLN:H	1:A:1231:GLN:NE2	1.94	0.57
1:A:1217:LEU:HD11	1:A:1223:VAL:HG13	1.87	0.57
1:A:143:ASP:HB2	1:A:144:PRO:CD	2.34	0.57
1:A:563:ASP:HB3	1:A:566:HIS:HB2	1.87	0.57
1:A:841:GLN:HA	1:A:1088:THR:HG21	1.86	0.57
1:A:244:LEU:HB2	1:A:249:LEU:HB2	1.87	0.56
1:A:818:PRO:HD2	1:A:822:THR:HG21	1.88	0.56
1:A:841:GLN:O	1:A:844:VAL:HG22	2.05	0.56
1:A:782:THR:HG21	1:A:1079:ARG:NH1	2.20	0.56
1:A:480:THR:HG22	5:A:1366:HOH:O	2.04	0.56
1:A:825:ILE:HG21	1:A:866:LEU:HD22	1.87	0.56
1:A:676:ILE:HG21	1:A:682:VAL:HG21	1.88	0.56
1:A:1195:VAL:O	1:A:1262:CYS:HA	2.06	0.55
1:A:856:LEU:CD1	1:A:867:ILE:HG12	2.37	0.54
1:A:103:LEU:HB2	1:A:122:ILE:HB	1.90	0.54
1:A:120:ARG:HG3	1:A:614:TRP:CZ2	2.43	0.54
1:A:773:GLU:HG2	1:A:1277:ARG:HD3	1.90	0.54
1:A:242:LEU:HD21	1:A:279:VAL:HG11	1.90	0.54
1:A:1223:VAL:HB	1:A:1227:PRO:HB3	1.91	0.53
1:A:3:VAL:HG11	1:A:79:ILE:HD13	1.90	0.53
1:A:950:TYR:N	1:A:955:GLN:HE22	2.03	0.53
1:A:292:ARG:HH22	1:A:1225:ARG:HD2	1.72	0.53
1:A:485:PHE:CD1	1:A:564:ILE:HD11	2.44	0.53
1:A:969:LEU:HD21	1:A:989:GLY:HA3	1.90	0.53
1:A:437:SER:HB3	1:A:442:ILE:HD13	1.90	0.53
1:A:435:VAL:HG13	1:A:443:ASP:HB3	1.91	0.53
1:A:951:HIS:H	1:A:955:GLN:NE2	2.06	0.52
1:A:150:ARG:HG2	1:A:152:TRP:CH2	2.44	0.52
1:A:10:ARG:HD2	1:A:13:GLN:HE21	1.74	0.52
1:A:109:PRO:HB2	1:A:537:ASN:HD21	1.74	0.51
1:A:1056:LEU:HD11	1:A:1101:MET:HE2	1.92	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:139:VAL:HG13	1:A:152:TRP:HB2	1.91	0.51
1:A:711:PHE:HB2	1:A:714:ARG:HG3	1.93	0.51
1:A:522:ALA:HB3	1:A:619:ILE:HD12	1.92	0.51
1:A:1069:ILE:HG12	1:A:1070:PRO:HD3	1.92	0.51
1:A:854:ASP:HB3	1:A:917:TYR:OH	2.10	0.51
1:A:59:ASN:HD22	1:A:59:ASN:C	2.15	0.50
1:A:305:TYR:O	1:A:333:HIS:HE1	1.95	0.50
1:A:212:ILE:HD13	1:A:333:HIS:HB2	1.94	0.50
1:A:112:LYS:HG2	1:A:115:ASP:OD2	2.11	0.50
1:A:1302:GLN:OE1	1:A:1307:ILE:HD11	2.11	0.50
1:A:468:ARG:HD2	1:A:494:ARG:NH2	2.26	0.49
1:A:1179:ASP:OD2	1:A:1260[B]:ARG:NH2	2.45	0.49
1:A:950:TYR:H	1:A:955:GLN:NE2	2.04	0.49
1:A:1016:ASN:HD22	1:A:1018:HIS:HB2	1.78	0.49
1:A:143:ASP:HB3	1:A:149:ILE:CD1	2.44	0.48
1:A:102:GLY:HA3	1:A:189:VAL:HG22	1.94	0.48
1:A:954:MET:HE1	1:A:967:TYR:HB2	1.95	0.48
1:A:110:VAL:HG23	1:A:537:ASN:HD21	1.78	0.48
1:A:127:GLU:HB3	1:A:129:LYS:HG2	1.96	0.47
1:A:842:ASN:HB3	1:A:881:GLN:HE22	1.80	0.47
1:A:951:HIS:CD2	1:A:954:MET:HB2	2.49	0.47
1:A:886:GLN:HE22	1:A:896:THR:HA	1.80	0.46
1:A:1172:PHE:HB3	1:A:1294:ILE:HG22	1.97	0.46
1:A:1198:SER:HB3	1:A:1259:GLU:O	2.16	0.46
1:A:102:GLY:HA3	1:A:189:VAL:CG2	2.45	0.46
1:A:1241:GLU:HG2	1:A:1250:VAL:HB	1.97	0.46
1:A:144:PRO:HD3	1:A:177:TRP:CD1	2.50	0.46
1:A:19:ILE:O	1:A:59:ASN:HB2	2.16	0.45
1:A:1034:ILE:HD13	1:A:1040:GLU:O	2.16	0.45
1:A:950:TYR:HB2	1:A:1210:MET:HB2	1.97	0.45
1:A:953:GLU:HG3	1:A:1047:ASN:HD21	1.81	0.45
1:A:492:GLU:HG2	1:A:493:LEU:HG	1.99	0.45
1:A:1218:PRO:HG2	1:A:1299:VAL:HG21	1.98	0.44
1:A:547:ILE:HD12	1:A:653:ILE:CD1	2.47	0.44
1:A:819:ASN:O	1:A:822:THR:HG23	2.17	0.44
1:A:716:ASN:HD22	1:A:742:ARG:CZ	2.30	0.44
1:A:350:VAL:O	1:A:350:VAL:CG1	2.65	0.44
1:A:1045:THR:OG1	1:A:1051:THR:CG2	2.65	0.44
1:A:1214[A]:GLU:HG3	1:A:1250:VAL:HG22	1.99	0.44
1:A:680:GLU:HG2	1:A:1276:LYS:HA	1.98	0.44
1:A:481:VAL:HG23	1:A:578:ILE:CG2	2.47	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:724:SER:O	1:A:725:VAL:HG23	2.18	0.44
1:A:727:VAL:HG13	1:A:733:VAL:HG13	1.99	0.44
1:A:464:LYS:HG3	1:A:464:LYS:H	1.66	0.44
1:A:397:GLU:H	1:A:401:ASP:HB2	1.83	0.43
1:A:1127:ILE:HG23	1:A:1132:ILE:CD1	2.43	0.43
1:A:854:ASP:CB	1:A:917:TYR:OH	2.66	0.43
1:A:1014:MET:HB3	1:A:1061:MET:HG2	2.01	0.43
1:A:1129:SER:HG	1:A:1132:ILE:HG23	1.84	0.42
1:A:1076:VAL:O	1:A:1079:ARG:HD3	2.19	0.42
1:A:886:GLN:NE2	1:A:897:THR:H	2.17	0.42
1:A:15:TYR:HB3	1:A:64:PHE:HB2	2.02	0.42
1:A:171:THR:O	1:A:880:ASN:ND2	2.52	0.42
1:A:506:ILE:HG23	1:A:631:VAL:HG21	2.01	0.42
1:A:1303:ASN:N	1:A:1303:ASN:ND2	2.32	0.42
1:A:210:SER:HB3	1:A:223:THR:OG1	2.20	0.42
1:A:1056:LEU:O	1:A:1060:VAL:HG23	2.20	0.42
1:A:479:ARG:HG3	5:A:1366:HOH:O	2.19	0.42
1:A:212:ILE:HA	1:A:213:PRO:HD3	1.89	0.42
1:A:1282:VAL:HG12	1:A:1284:TYR:HD1	1.84	0.42
1:A:485:PHE:HB3	1:A:564:ILE:HD11	2.02	0.41
1:A:852:VAL:O	1:A:856:LEU:HG	2.20	0.41
1:A:1106:SER:HA	1:A:1107:PRO:HD3	1.94	0.41
1:A:409:LYS:HA	1:A:409:LYS:HD3	1.91	0.41
1:A:460:ASN:ND2	5:A:1364:HOH:O	2.46	0.41
1:A:267:PHE:HB3	1:A:270:PHE:HB2	2.02	0.41
1:A:1004:TYR:HA	1:A:1051:THR:HB	2.03	0.41
1:A:214:LEU:HA	1:A:303:TYR:O	2.20	0.41
1:A:975:ASN:HD22	1:A:978:ALA:H	1.68	0.40
1:A:479:ARG:O	1:A:578:ILE:HA	2.22	0.40
1:A:901:VAL:HG22	1:A:964:LEU:HD21	2.04	0.40
1:A:1079:ARG:NH2	5:A:1399:HOH:O	2.54	0.40
1:A:782:THR:OG1	1:A:1076:VAL:HG13	2.21	0.40
1:A:1110:ASN:HD22	1:A:1132:ILE:HD11	1.87	0.40
1:A:104:ILE:HD11	1:A:181:VAL:HG23	2.03	0.40
1:A:903:LEU:HD12	1:A:903:LEU:HA	1.89	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	1282/1325 (97%)	1214 (95%)	61 (5%)	7 (0%)	34 63

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	808	GLY
1	A	1119	LYS
1	A	200	LEU
1	A	414	ARG
1	A	1133	ASP
1	A	320	PRO
1	A	655	LYS

### 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	1147/1173 (98%)	1063 (93%)	84 (7%)	17 39

All (84) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	56	ARG
1	A	59	ASN
1	A	60	ARG
1	A	76	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	92	GLU
1	A	139	VAL
1	A	142	ARG
1	A	150	ARG
1	A	157	LEU
1	A	189	VAL
1	A	212	ILE
1	A	234	VAL
1	A	250	LYS
1	A	312	ILE
1	A	334	ASP
1	A	369	GLN
1	A	380	HIS
1	A	426	GLU
1	A	429	THR
1	A	435	VAL
1	A	436	MET
1	A	442	ILE
1	A	464	LYS
1	A	470	LYS
1	A	479	ARG
1	A	480	THR
1	A	494	ARG
1	A	499	LEU
1	A	501	ILE
1	A	540	HIS
1	A	547	ILE
1	A	559	GLU
1	A	566	HIS
1	A	570	LEU
1	A	576	ASP
1	A	625	ARG
1	A	639	TYR
1	A	653	ILE
1	A	738	LEU
1	A	782	THR
1	A	789	ASP
1	A	790	HIS
1	A	801	ILE
1	A	802	ASN
1	A	806	ASP
1	A	822	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	834	VAL
1	A	844	VAL
1	A	851	LEU
1	A	862	LYS
1	A	901	VAL
1	A	911	MET
1	A	942	ARG
1	A	945	LYS
1	A	954	MET
1	A	1034	ILE
1	A	1051	THR
1	A	1061	MET
1	A	1069	ILE
1	A	1099	THR
1	A	1106	SER
1	A	1113	THR
1	A	1128	ASN
1	A	1132	ILE
1	A	1133	ASP
1	A	1134	VAL
1	A	1138	VAL
1	A	1144	THR
1	A	1154	ILE
1	A	1167	LEU
1	A	1184	ASN
1	A	1204	THR
1	A	1206	ARG
1	A	1207	ARG
1	A	1225	ARG
1	A	1231	GLN
1	A	1234	VAL
1	A	1243	ARG
1	A	1247	THR
1	A	1260[A]	ARG
1	A	1260[B]	ARG
1	A	1294	ILE
1	A	1303	ASN
1	A	1316	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (24) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	13	GLN
1	A	59	ASN
1	A	227	ASN
1	A	333	HIS
1	A	369	GLN
1	A	458	GLN
1	A	540	HIS
1	A	716	ASN
1	A	802	ASN
1	A	880	ASN
1	A	881	GLN
1	A	886	GLN
1	A	898	ASN
1	A	951	HIS
1	A	955	GLN
1	A	975	ASN
1	A	1016	ASN
1	A	1047	ASN
1	A	1128	ASN
1	A	1226	ASN
1	A	1231	GLN
1	A	1238	GLN
1	A	1261	ASN
1	A	1303	ASN

### 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 ⓘ

2 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	$\# Z  > 2$
3	NAG	A	1329	1,3	14,14,15	0.58	0	15,19,21	0.83	0
3	NAG	A	1330	3	14,14,15	0.45	0	15,19,21	0.76	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	NAG	A	1329	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1330	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 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	$\# Z  > 2$
2	NAG	A	1326	1	14,14,15	0.45	0	15,19,21	2.04	2 (13%)
2	NAG	A	1327	1	14,14,15	0.43	0	15,19,21	1.27	3 (20%)
2	NAG	A	1328	1	14,14,15	0.56	0	15,19,21	0.73	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
2	NAG	A	1326	1	-	0/6/23/26	0/1/1/1
2	NAG	A	1327	1	-	0/6/23/26	0/1/1/1
2	NAG	A	1328	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1327	NAG	C2-N2-C7	-2.34	120.03	123.04
2	A	1327	NAG	O4-C4-C5	2.03	114.61	109.24
2	A	1327	NAG	C1-O5-C5	2.61	115.57	112.25
2	A	1326	NAG	C3-C4-C5	3.59	116.46	110.20
2	A	1326	NAG	C1-O5-C5	6.21	120.12	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1326	NAG	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 ⓘ

### 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	1283/1325 (96%)	0.48	78 (6%) 25 23	30, 58, 109, 147	0

All (78) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	808	GLY	7.3
1	A	1109	ARG	6.8
1	A	198	TYR	6.4
1	A	806	ASP	5.5
1	A	536	PHE	5.4
1	A	540	HIS	5.2
1	A	805	ALA	5.0
1	A	503	GLU	4.7
1	A	45	VAL	4.4
1	A	539	ASN	4.4
1	A	621	ARG	4.4
1	A	46	LEU	4.3
1	A	541	ASP	4.2
1	A	84	GLY	4.0
1	A	1122	ALA	3.9
1	A	491	GLN	3.9
1	A	1131	GLN	3.8
1	A	492	GLU	3.7
1	A	801	ILE	3.6
1	A	1316	GLU	3.6
1	A	1154	ILE	3.5
1	A	1119	LYS	3.4
1	A	1146	LYS	3.4
1	A	792	ASN	3.4
1	A	200	LEU	3.3
1	A	372	GLU	3.2
1	A	1317	CYS	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1126	LYS	3.2
1	A	628	ILE	3.2
1	A	85	PHE	3.1
1	A	505	GLU	3.1
1	A	133	ARG	3.1
1	A	812	ILE	3.1
1	A	560	ASN	3.0
1	A	27	SER	3.0
1	A	559	GLU	3.0
1	A	1139	ASP	3.0
1	A	388	PHE	2.9
1	A	230	PHE	2.9
1	A	158	TYR	2.9
1	A	802	ASN	2.8
1	A	199	VAL	2.8
1	A	1128	ASN	2.8
1	A	86	SER	2.6
1	A	789	ASP	2.6
1	A	51	MET	2.6
1	A	1137	PHE	2.6
1	A	862	LYS	2.5
1	A	1168	ASN	2.5
1	A	25	GLN	2.5
1	A	71	THR	2.4
1	A	1166	ASN	2.4
1	A	173[A]	MET	2.4
1	A	390	TYR	2.4
1	A	639	TYR	2.4
1	A	1184	ASN	2.4
1	A	201	SER	2.3
1	A	267	PHE	2.3
1	A	534	LEU	2.3
1	A	809	SER	2.3
1	A	1142	GLU	2.3
1	A	70	LEU	2.3
1	A	1135	GLU	2.2
1	A	11	ALA	2.2
1	A	782	THR	2.2
1	A	830[A]	HIS	2.2
1	A	1301	LYS	2.2
1	A	246	ASP	2.2
1	A	269	ASN	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	248	LYS	2.1
1	A	83	ARG	2.1
1	A	320	PRO	2.1
1	A	790	HIS	2.1
1	A	1	LEU	2.0
1	A	351	ARG	2.0
1	A	156	LYS	2.0
1	A	798	ASN	2.0
1	A	1124	TYR	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](#)

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(Å <sup>2</sup> )	Q<0.9
3	NAG	A	1329	14/15	0.90	0.29	1.16	79,83,85,88	0
3	NAG	A	1330	14/15	0.82	0.30	-	89,91,92,93	0

## 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(Å <sup>2</sup> )	Q<0.9
2	NAG	A	1326	14/15	0.83	0.41	3.25	66,72,74,75	0
2	NAG	A	1328	14/15	0.87	0.23	2.89	58,60,63,65	0
4	NA	A	1333	1/1	0.69	0.16	-0.86	53,53,53,53	0
4	NA	A	1331	1/1	1.00	0.07	-2.74	21,21,21,21	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NA	A	1332	1/1	0.96	0.36	-	43,43,43,43	0
2	NAG	A	1327	14/15	0.90	0.23	-	52,55,57,58	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.