



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:53 PM GMT

PDB ID : 4MY0  
Title : Crystal Structure of GCN5-related N-acetyltransferase from *Kribbella flavida*  
Authors : Kim, Y.; Mack, J.; Endres, M.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2013-09-26  
Resolution : 2.10 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

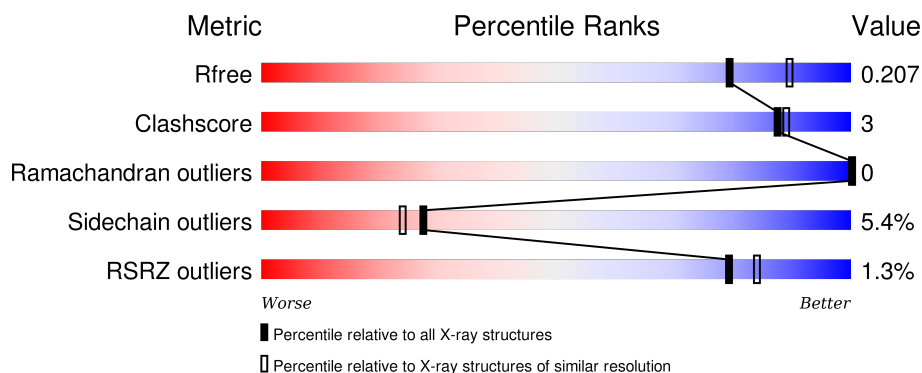
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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	392	<div> <div>90%</div> <div>8%</div> <div>..</div> </div>
1	B	392	<div> <div>91%</div> <div>8%</div> <div>.</div> </div>
1	C	392	<div> <div>91%</div> <div>8%</div> <div>.</div> </div>
1	D	392	<div> <div>2%</div> <div>86%</div> <div>12%</div> <div>..</div> </div>
1	E	392	<div> <div>3%</div> <div>88%</div> <div>10%</div> <div>..</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	392	<div> <div></div> <div>88%</div> <div>10% ..</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
4	GOL	A	402	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 18874 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 GCN5-related N-acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	386	Total	C	N	O	Se	0	1	0
			2953	1864	535	549	5			
1	B	386	Total	C	N	O	Se	0	1	0
			2953	1864	535	549	5			
1	C	392	Total	C	N	O	Se	0	3	0
			3018	1901	550	562	5			
1	D	386	Total	C	N	O	Se	0	0	0
			2947	1861	534	547	5			
1	E	386	Total	C	N	O	Se	0	0	0
			2947	1861	534	547	5			
1	F	387	Total	C	N	O	Se	0	4	0
			2992	1885	546	556	5			

There are 24 discrepancies between the modelled and reference sequences:

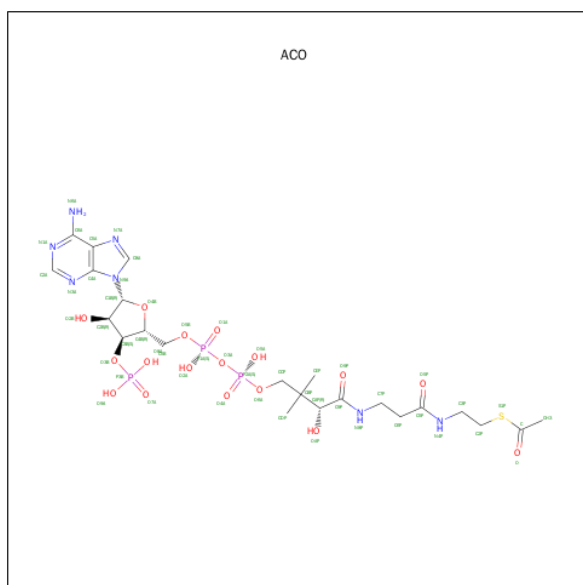
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP D2PVF8
A	-1	ASN	-	EXPRESSION TAG	UNP D2PVF8
A	0	ALA	-	EXPRESSION TAG	UNP D2PVF8
A	1	VAL	MET	SEE REMARK 999	UNP D2PVF8
B	-2	SER	-	EXPRESSION TAG	UNP D2PVF8
B	-1	ASN	-	EXPRESSION TAG	UNP D2PVF8
B	0	ALA	-	EXPRESSION TAG	UNP D2PVF8
B	1	VAL	MET	SEE REMARK 999	UNP D2PVF8
C	-2	SER	-	EXPRESSION TAG	UNP D2PVF8
C	-1	ASN	-	EXPRESSION TAG	UNP D2PVF8
C	0	ALA	-	EXPRESSION TAG	UNP D2PVF8
C	1	VAL	MET	SEE REMARK 999	UNP D2PVF8
D	-2	SER	-	EXPRESSION TAG	UNP D2PVF8
D	-1	ASN	-	EXPRESSION TAG	UNP D2PVF8
D	0	ALA	-	EXPRESSION TAG	UNP D2PVF8
D	1	VAL	MET	SEE REMARK 999	UNP D2PVF8
E	-2	SER	-	EXPRESSION TAG	UNP D2PVF8

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	ASN	-	EXPRESSION TAG	UNP D2PVF8
E	0	ALA	-	EXPRESSION TAG	UNP D2PVF8
E	1	VAL	MET	SEE REMARK 999	UNP D2PVF8
F	-2	SER	-	EXPRESSION TAG	UNP D2PVF8
F	-1	ASN	-	EXPRESSION TAG	UNP D2PVF8
F	0	ALA	-	EXPRESSION TAG	UNP D2PVF8
F	1	VAL	MET	SEE REMARK 999	UNP D2PVF8

- Molecule 2 is ACETYL COENZYME \*A (three-letter code: ACO) (formula:  $C_{23}H_{38}N_7O_{17}P_3S$ ).



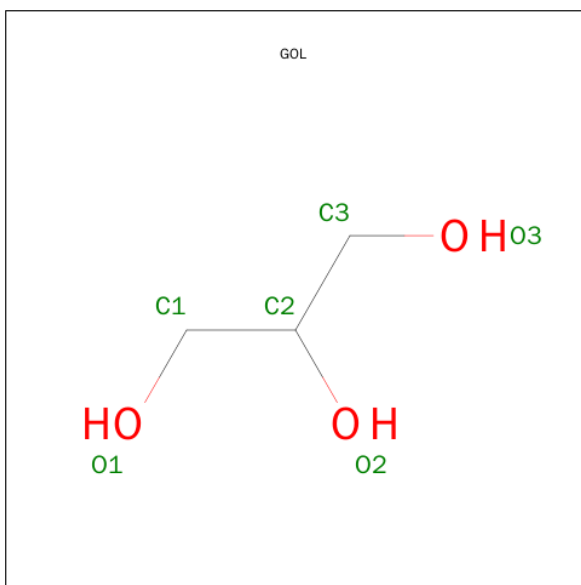
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		
2	B	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		
2	D	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		
2	E	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		
2	F	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		

- Molecule 3 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula:  $C_4H_{12}NO_3$ ).



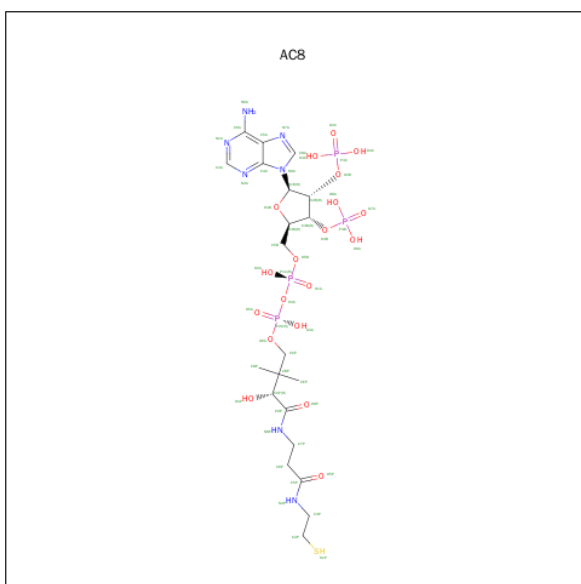
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			8	4	1	3		
3	B	1	Total	C	N	O	0	0
			8	4	1	3		
3	C	1	Total	C	N	O	0	0
			8	4	1	3		
3	D	1	Total	C	N	O	0	0
			8	4	1	3		
3	E	1	Total	C	N	O	0	0
			8	4	1	3		
3	F	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 4 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
4	A	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is [(2R,3R,4R,5R)-5-(6-AMINO-9H-PURIN-9-YL)-3,4-BIS(PHOSPHONOOXY)TETRAHYDROFURAN-2-YL]METHYL (3R)-3-HYDROXY-2,2-DIMETHYL-4-OXO-4-({3-OXO-3-[(2-SULFANYLETHYL)AMINO]PROPYL}AMINO)BUTYL DIHYDROGEN DIPHOSPHATE (three-letter code: AC8) (formula: C<sub>21</sub>H<sub>37</sub>N<sub>7</sub>O<sub>19</sub>P<sub>4</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	C	1	Total	C	N	O	P	S	0	0
			52	21	7	19	4	1		

- Molecule 6 is water.

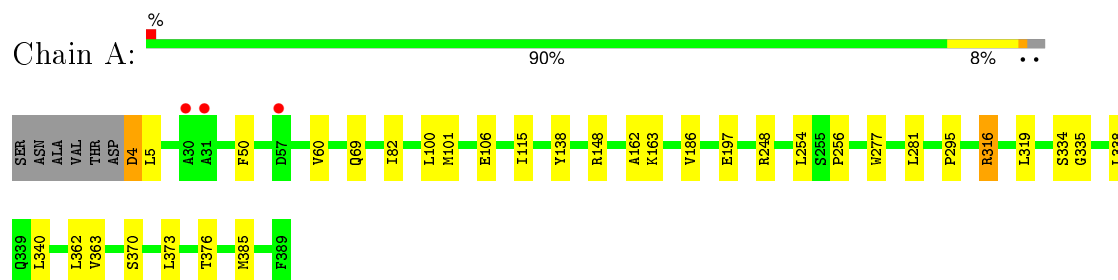
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	156	Total	O	0	0
			156	156		
6	B	142	Total	O	0	0
			142	142		
6	C	116	Total	O	0	0
			116	116		
6	D	77	Total	O	0	0
			77	77		
6	E	91	Total	O	0	0
			91	91		
6	F	130	Total	O	0	0
			130	130		



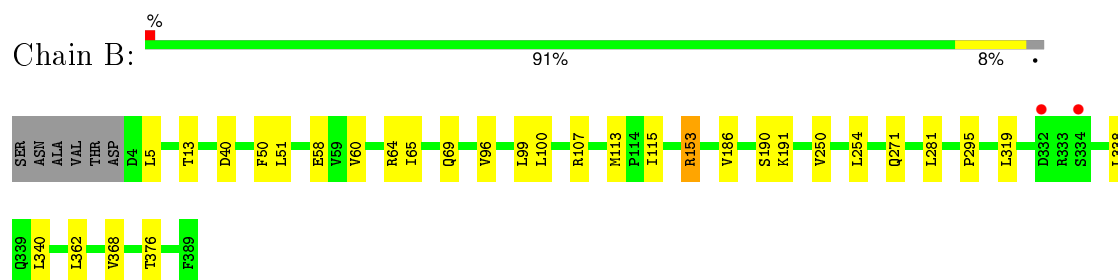
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

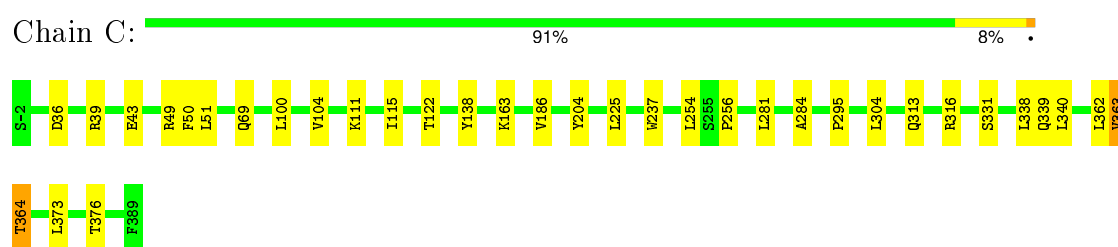
- Molecule 1: GCN5-related N-acetyltransferase



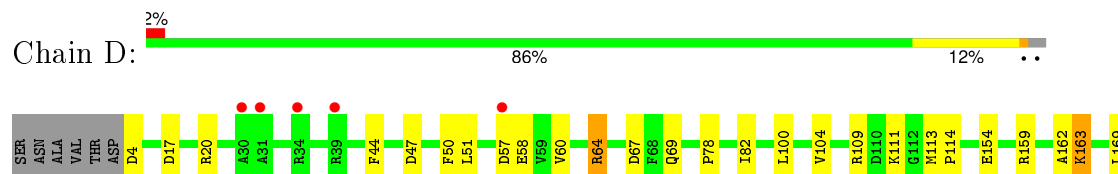
- Molecule 1: GCN5-related N-acetyltransferase

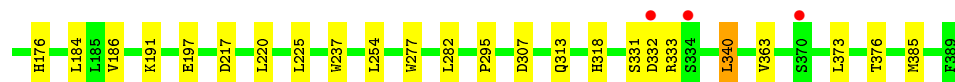


- Molecule 1: GCN5-related N-acetyltransferase

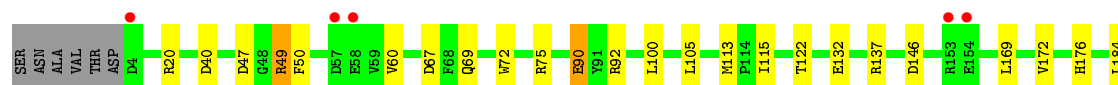
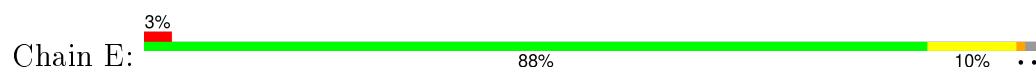


- Molecule 1: GCN5-related N-acetyltransferase

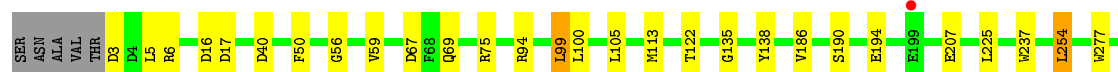
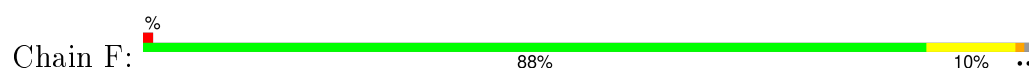




- Molecule 1: GCN5-related N-acetyltransferase



- Molecule 1: GCN5-related N-acetyltransferase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.14Å 166.52Å 183.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.91 – 2.10 46.03 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.2 (36.91-2.10) 99.2 (46.03-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.78 (at 2.10Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1161)	Depositor
R, $R_{free}$	0.166 , 0.202 0.175 , 0.207	Depositor DCC
$R_{free}$ test set	8198 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.5	Xtriage
Anisotropy	0.192	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 44.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 163393 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	18874	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.33% 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: GOL, TRS, AC8, ACO

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	0.40	0/3016	0.53	0/4095
1	B	0.41	0/3016	0.53	0/4095
1	C	0.39	0/3082	0.53	0/4184
1	D	0.35	0/3010	0.51	0/4086
1	E	0.35	0/3010	0.53	0/4086
1	F	0.39	0/3055	0.55	1/4146 (0.0%)
All	All	0.38	0/18189	0.53	1/24692 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	99	LEU	CA-CB-CG	5.20	127.26	115.30

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	2953	0	2879	15	0
1	B	2953	0	2879	10	0
1	C	3018	0	2938	15	0
1	D	2947	0	2874	23	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2947	0	2874	16	0
1	F	2992	0	2914	21	0
2	A	48	0	31	0	0
2	B	48	0	31	1	0
2	D	48	0	31	0	0
2	E	48	0	31	2	0
2	F	48	0	31	0	0
3	A	8	0	12	0	0
3	B	8	0	12	0	0
3	C	8	0	12	0	0
3	D	8	0	12	0	0
3	E	8	0	12	1	0
3	F	8	0	12	0	0
4	A	6	0	8	0	0
4	E	6	0	8	0	0
5	C	52	0	31	1	0
6	A	156	0	0	0	0
6	B	142	0	0	1	0
6	C	116	0	0	0	0
6	D	77	0	0	0	0
6	E	91	0	0	0	0
6	F	130	0	0	0	0
All	All	18874	0	17632	94	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 3.

All (94) 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:F:67:ASP:HA	1:F:113:MSE:HE1	1.63	0.79
1:A:338:LEU:HD11	1:A:363:VAL:HG13	1.71	0.71
1:D:163:LYS:H	1:D:163:LYS:HD2	1.56	0.70
1:A:162:ALA:HB2	1:A:197:GLU:HG3	1.75	0.66
1:E:69:GLN:HB2	1:E:186:VAL:HB	1.76	0.66
1:F:6[B]:ARG:HH12	1:F:56:GLY:H	1.42	0.66
1:A:148:ARG:NH2	5:C:400:AC8:O3C	2.29	0.65
1:D:162:ALA:HB2	1:D:197:GLU:HG3	1.81	0.63
1:E:67:ASP:HA	1:E:113:MSE:HE1	1.81	0.62
1:C:304:LEU:HD23	1:C:339[A]:GLN:HG2	1.83	0.60
1:F:302:ASP:OD2	1:F:316:ARG:NH2	2.34	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:295:PRO:O	1:B:376:THR:HG21	2.04	0.58
1:B:69:GLN:HB2	1:B:186:VAL:HB	1.86	0.58
1:E:47:ASP:OD2	1:E:49:ARG:NH1	2.37	0.57
1:F:6[B]:ARG:NH2	1:F:56:GLY:O	2.38	0.57
1:A:316:ARG:NH1	1:A:335:GLY:O	2.39	0.56
1:B:115:ILE:HG22	1:B:281:LEU:HD22	1.88	0.56
1:D:69:GLN:HB2	1:D:186:VAL:HB	1.87	0.56
1:F:316:ARG:NH1	1:F:335:GLY:O	2.38	0.56
1:C:69:GLN:HB2	1:C:186:VAL:HB	1.87	0.56
1:C:36:ASP:OD2	1:C:39:ARG:NH2	2.36	0.56
1:D:67:ASP:HA	1:D:113:MSE:HE1	1.88	0.55
1:A:115:ILE:HG22	1:A:281:LEU:HD22	1.87	0.54
1:F:75:ARG:NH2	1:F:289:ALA:O	2.41	0.53
1:C:204:TYR:CE1	1:E:90:GLU:HG3	2.43	0.53
1:D:44:PHE:CD1	1:D:64:ARG:HG2	2.44	0.53
1:F:6[B]:ARG:HH12	1:F:56:GLY:N	2.08	0.52
1:A:376:THR:HG21	1:B:295:PRO:O	2.10	0.52
1:F:17:ASP:HB3	1:F:59:VAL:HB	1.92	0.51
1:D:340:LEU:HB3	1:D:363:VAL:HG22	1.93	0.51
1:B:58:GLU:OE2	1:D:159:ARG:NH2	2.44	0.51
1:E:376:THR:HG21	1:F:295:PRO:O	2.11	0.50
1:E:176:HIS:CE1	1:E:184:LEU:HD13	2.46	0.50
1:E:295:PRO:O	1:F:376:THR:HG21	2.11	0.50
1:D:4:ASP:N	1:D:4:ASP:OD2	2.43	0.50
1:C:295:PRO:O	1:D:376:THR:HG21	2.12	0.49
1:B:64:ARG:NH1	6:B:614:HOH:O	2.44	0.49
1:D:176:HIS:CE1	1:D:184:LEU:HD13	2.47	0.49
1:D:109:ARG:HG2	1:D:282:LEU:HB3	1.95	0.49
1:F:303:LEU:HD22	1:F:340:LEU:HD11	1.94	0.49
1:A:69:GLN:HB2	1:A:186:VAL:HB	1.94	0.49
2:E:400:ACO:S1P	3:E:401:TRS:O2	2.66	0.48
1:F:69:GLN:HB2	1:F:186:VAL:HB	1.93	0.48
1:C:338:LEU:HD11	1:C:363:VAL:HG13	1.96	0.48
1:E:137:ARG:HD3	1:E:275:GLN:OE1	2.13	0.47
1:F:225:LEU:HD13	1:F:237:TRP:CE2	2.50	0.47
1:F:6[B]:ARG:NH1	1:F:56:GLY:H	2.09	0.47
1:B:65:ILE:HG21	1:B:113:MSE:HE1	1.96	0.47
1:E:67:ASP:HA	1:E:113:MSE:CE	2.45	0.47
1:C:316:ARG:HG2	1:C:331:SER:HB3	1.96	0.47
1:A:370:SER:HA	1:A:373:LEU:HD12	1.96	0.47
1:E:132:GLU:HB2	1:E:343:ARG:HB3	1.95	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:277:TRP:CE2	1:A:385:MSE:HE3	2.51	0.46
1:E:115:ILE:HG22	1:E:281:LEU:HD22	1.98	0.45
1:E:75:ARG:NH2	1:E:289:ALA:O	2.47	0.45
1:B:51:LEU:HD11	1:B:107:ARG:HG2	1.98	0.45
1:C:284:ALA:N	1:C:313:GLN:OE1	2.49	0.45
1:F:3:ASP:HB3	1:F:5:LEU:H	1.81	0.45
1:A:82:ILE:HD13	1:A:101:MSE:HE1	1.98	0.45
1:C:376:THR:HG21	1:D:295:PRO:O	2.16	0.45
1:D:78:PRO:O	1:D:114:PRO:HD2	2.16	0.45
1:F:135:GLY:HA3	1:F:277:TRP:CZ2	2.52	0.45
1:F:105:LEU:HD22	1:F:282:LEU:HD21	1.98	0.45
1:B:153:ARG:HG2	1:B:153:ARG:H	1.52	0.45
1:D:277:TRP:CD2	1:D:385:MSE:HE3	2.52	0.45
1:D:307:ASP:OD1	1:D:313:GLN:NE2	2.50	0.44
1:D:17:ASP:OD1	1:D:20:ARG:NH2	2.51	0.44
1:F:301:VAL:HG21	1:F:374:LEU:HD21	2.00	0.44
1:E:146:ASP:CG	1:E:264:VAL:HG22	2.39	0.43
1:A:82:ILE:CD1	1:A:101:MSE:HE1	2.48	0.43
1:C:138:TYR:CD2	1:C:256:PRO:HD3	2.52	0.43
1:E:172:VAL:HG13	1:E:176:HIS:CD2	2.53	0.43
1:A:338:LEU:HD11	1:A:363:VAL:CG1	2.46	0.43
1:C:49:ARG:HG3	1:C:111:LYS:HE3	2.00	0.43
1:C:339[A]:GLN:HB2	1:C:364:THR:HG22	2.00	0.43
1:D:113:MSE:HA	1:D:114:PRO:HD3	1.81	0.43
1:D:47:ASP:O	1:D:111:LYS:NZ	2.32	0.43
1:C:115:ILE:HG22	1:C:281:LEU:HD22	2.01	0.43
1:C:225:LEU:HD13	1:C:237:TRP:CE2	2.54	0.43
1:D:82:ILE:HD12	1:D:104:VAL:HG11	2.00	0.43
1:E:72:TRP:HB3	1:E:291:ARG:HG3	2.01	0.43
1:F:277:TRP:CE2	1:F:385:MSE:HE3	2.55	0.42
1:B:5:LEU:HD13	1:B:96:VAL:HG22	2.00	0.42
1:D:169:LEU:HD23	1:D:169:LEU:HA	1.91	0.41
1:D:191:LYS:HA	1:D:191:LYS:HD3	1.92	0.41
1:A:138:TYR:CD2	1:A:256:PRO:HD3	2.56	0.41
1:F:138:TYR:HB2	1:F:254:LEU:O	2.20	0.41
1:A:4:ASP:OD1	1:A:4:ASP:N	2.54	0.41
1:E:92:ARG:HG2	2:E:400:ACO:OAP	2.20	0.41
1:F:190:SER:O	1:F:194:GLU:HG3	2.21	0.41
1:D:225:LEU:HD13	1:D:237:TRP:CE2	2.56	0.41
2:B:400:ACO:N4P	2:B:400:ACO:H131	2.36	0.41
1:D:318:HIS:HB2	1:D:331:SER:HB2	2.02	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:51:LEU:HD12	1:C:104:VAL:HA	2.04	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	385/392 (98%)	384 (100%)	1 (0%)	0	100	100
1	B	385/392 (98%)	383 (100%)	2 (0%)	0	100	100
1	C	393/392 (100%)	391 (100%)	2 (0%)	0	100	100
1	D	384/392 (98%)	383 (100%)	1 (0%)	0	100	100
1	E	384/392 (98%)	383 (100%)	1 (0%)	0	100	100
1	F	389/392 (99%)	388 (100%)	1 (0%)	0	100	100
All	All	2320/2352 (99%)	2312 (100%)	8 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	289/288 (100%)	275 (95%)	14 (5%)	31	29
1	B	289/288 (100%)	272 (94%)	17 (6%)	24	20

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	296/288 (103%)	285 (96%)	11 (4%)	41	41
1	D	288/288 (100%)	272 (94%)	16 (6%)	26	22
1	E	288/288 (100%)	267 (93%)	21 (7%)	17	13
1	F	293/288 (102%)	279 (95%)	14 (5%)	31	29
All	All	1743/1728 (101%)	1650 (95%)	93 (5%)	27	25

All (93) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	ASP
1	A	5	LEU
1	A	50	PHE
1	A	60	VAL
1	A	100	LEU
1	A	106	GLU
1	A	163	LYS
1	A	248	ARG
1	A	254	LEU
1	A	316	ARG
1	A	319	LEU
1	A	334	SER
1	A	340	LEU
1	A	362	LEU
1	B	13	THR
1	B	40	ASP
1	B	50	PHE
1	B	60	VAL
1	B	99	LEU
1	B	100	LEU
1	B	153	ARG
1	B	190	SER
1	B	191	LYS
1	B	250	VAL
1	B	254	LEU
1	B	271	GLN
1	B	319	LEU
1	B	338	LEU
1	B	340	LEU
1	B	362	LEU
1	B	368	VAL
1	C	43	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	50	PHE
1	C	100	LEU
1	C	122	THR
1	C	163	LYS
1	C	254	LEU
1	C	340	LEU
1	C	362	LEU
1	C	363	VAL
1	C	364	THR
1	C	373	LEU
1	D	50	PHE
1	D	51	LEU
1	D	57	ASP
1	D	58	GLU
1	D	60	VAL
1	D	64	ARG
1	D	100	LEU
1	D	154	GLU
1	D	163	LYS
1	D	217	ASP
1	D	220	LEU
1	D	254	LEU
1	D	332	ASP
1	D	333	ARG
1	D	340	LEU
1	D	373	LEU
1	E	20	ARG
1	E	40	ASP
1	E	49	ARG
1	E	50	PHE
1	E	60	VAL
1	E	90	GLU
1	E	100	LEU
1	E	105	LEU
1	E	122	THR
1	E	169	LEU
1	E	220	LEU
1	E	254	LEU
1	E	255	SER
1	E	264	VAL
1	E	319	LEU
1	E	325	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	336	ASP
1	E	340	LEU
1	E	343	ARG
1	E	368	VAL
1	E	373	LEU
1	F	16	ASP
1	F	40	ASP
1	F	50	PHE
1	F	94	ARG
1	F	99	LEU
1	F	100	LEU
1	F	122	THR
1	F	207	GLU
1	F	254	LEU
1	F	333	ARG
1	F	340	LEU
1	F	362	LEU
1	F	363	VAL
1	F	373	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	266	HIS
1	A	271	GLN
1	A	273	HIS
1	B	271	GLN
1	B	339	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

14 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$
2	ACO	A	400	-	40,50,53	0.91	2 (5%)	50,75,79	1.56	5 (10%)
3	TRS	A	401	-	7,7,7	0.46	0	9,9,9	6.86	6 (66%)
4	GOL	A	402	-	5,5,5	0.26	0	5,5,5	0.48	0
2	ACO	B	400	-	40,50,53	0.87	2 (5%)	50,75,79	2.03	11 (22%)
3	TRS	B	401	-	7,7,7	0.63	0	9,9,9	0.69	0
5	AC8	C	400	-	44,54,54	0.68	1 (2%)	56,82,82	1.93	7 (12%)
3	TRS	C	401	-	7,7,7	0.46	0	9,9,9	6.83	6 (66%)
2	ACO	D	400	-	40,50,53	0.81	1 (2%)	50,75,79	1.48	2 (4%)
3	TRS	D	401	-	7,7,7	0.42	0	9,9,9	7.34	7 (77%)
2	ACO	E	400	-	40,50,53	0.85	1 (2%)	50,75,79	1.51	5 (10%)
3	TRS	E	401	-	7,7,7	0.39	0	9,9,9	7.12	7 (77%)
4	GOL	E	402	-	5,5,5	0.33	0	5,5,5	0.27	0
2	ACO	F	400	-	40,50,53	0.82	2 (5%)	50,75,79	1.63	4 (8%)
3	TRS	F	401	-	7,7,7	0.48	0	9,9,9	1.06	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	ACO	A	400	-	-	0/44/64/67	0/3/3/3
3	TRS	A	401	-	-	0/9/9/9	0/0/0/0
4	GOL	A	402	-	-	0/4/4/4	0/0/0/0
2	ACO	B	400	-	-	0/44/64/67	0/3/3/3
3	TRS	B	401	-	-	0/9/9/9	0/0/0/0
5	AC8	C	400	-	-	0/49/69/69	0/3/3/3
3	TRS	C	401	-	-	0/9/9/9	0/0/0/0
2	ACO	D	400	-	-	0/44/64/67	0/3/3/3

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	TRS	D	401	-	-	0/9/9/9	0/0/0/0
2	ACO	E	400	-	-	0/44/64/67	0/3/3/3
3	TRS	E	401	-	-	0/9/9/9	0/0/0/0
4	GOL	E	402	-	-	0/4/4/4	0/0/0/0
2	ACO	F	400	-	-	0/44/64/67	0/3/3/3
3	TRS	F	401	-	-	0/9/9/9	0/0/0/0

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	400	ACO	O4B-C1B	2.01	1.43	1.41
5	C	400	AC8	O4B-C1B	2.16	1.43	1.41
2	F	400	ACO	O4B-C1B	2.21	1.44	1.41
2	A	400	ACO	O4B-C1B	2.53	1.44	1.41
2	F	400	ACO	C5A-C4A	2.84	1.46	1.40
2	D	400	ACO	C5A-C4A	3.10	1.47	1.40
2	B	400	ACO	C5A-C4A	3.10	1.47	1.40
2	E	400	ACO	C5A-C4A	3.23	1.47	1.40
2	A	400	ACO	C5A-C4A	3.26	1.47	1.40

All (60) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	401	TRS	C2-C-N	-13.26	83.96	108.09
3	D	401	TRS	C3-C-N	-13.08	84.28	108.09
3	C	401	TRS	C3-C-N	-12.30	85.71	108.09
3	D	401	TRS	C2-C-N	-11.95	86.34	108.09
3	A	401	TRS	C3-C-N	-11.61	86.96	108.09
3	E	401	TRS	C3-C-N	-11.37	87.39	108.09
5	C	400	AC8	N3A-C2A-N1A	-11.05	120.43	128.89
3	A	401	TRS	C2-C-N	-10.70	88.61	108.09
3	A	401	TRS	C1-C-N	-10.68	88.65	108.09
3	D	401	TRS	C1-C-N	-10.52	88.94	108.09
3	C	401	TRS	C1-C-N	-10.37	89.21	108.09
3	C	401	TRS	C2-C-N	-10.07	89.75	108.09
3	E	401	TRS	C1-C-N	-9.41	90.95	108.09
2	F	400	ACO	N3A-C2A-N1A	-7.94	122.81	128.89
2	A	400	ACO	N3A-C2A-N1A	-7.39	123.24	128.89
2	B	400	ACO	N3A-C2A-N1A	-7.27	123.33	128.89
2	E	400	ACO	N3A-C2A-N1A	-7.24	123.35	128.89
2	D	400	ACO	N3A-C2A-N1A	-7.15	123.42	128.89
2	B	400	ACO	O3A-P2A-O6A	-4.86	90.04	102.94

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	400	ACO	CDP-CBP-CCP	-3.30	104.23	108.50
2	A	400	ACO	C4A-C5A-N7A	-3.30	106.45	109.48
2	B	400	ACO	C4A-C5A-N7A	-3.24	106.50	109.48
2	B	400	ACO	O3A-P1A-O5B	-3.24	94.35	102.94
2	F	400	ACO	C4A-C5A-N7A	-3.12	106.61	109.48
2	F	400	ACO	P2A-O3A-P1A	-3.10	124.03	132.73
2	E	400	ACO	C4A-C5A-N7A	-3.05	106.67	109.48
2	D	400	ACO	C4A-C5A-N7A	-3.04	106.69	109.48
3	D	401	TRS	O3-C3-C	-2.95	105.21	111.18
2	B	400	ACO	C7P-C6P-C5P	-2.76	107.77	112.31
2	E	400	ACO	P2A-O3A-P1A	-2.70	125.15	132.73
3	E	401	TRS	O3-C3-C	-2.47	106.18	111.18
5	C	400	AC8	C1B-N9A-C4A	-2.45	123.24	126.94
2	A	400	ACO	C2B-C1B-N9A	-2.19	110.95	114.29
5	C	400	AC8	C6P-C7P-N8P	-2.04	107.41	111.88
2	A	400	ACO	P2A-O3A-P1A	-2.01	127.09	132.73
2	B	400	ACO	C2A-N1A-C6A	2.02	122.37	118.77
2	E	400	ACO	CDP-CBP-CCP	2.02	111.12	108.50
2	A	400	ACO	O9A-P3B-O8A	2.03	115.11	107.38
2	B	400	ACO	O2A-P1A-O1A	2.04	123.57	112.53
2	E	400	ACO	C2A-N1A-C6A	2.05	122.42	118.77
5	C	400	AC8	C7P-C6P-C5P	2.17	115.90	112.31
2	B	400	ACO	O9A-P3B-O8A	2.19	115.72	107.38
2	F	400	ACO	CDP-CBP-CCP	2.24	111.40	108.50
3	C	401	TRS	C2-C-C1	2.52	116.24	110.78
5	C	400	AC8	C2P-C3P-N4P	2.61	117.53	112.37
3	E	401	TRS	C3-C-C1	2.71	116.64	110.78
5	C	400	AC8	C7P-N8P-C9P	2.82	128.12	122.53
2	B	400	ACO	CDP-CBP-CAP	2.91	114.65	109.34
5	C	400	AC8	C3P-N4P-C5P	3.27	129.22	122.79
3	D	401	TRS	C3-C-C2	3.59	118.55	110.78
3	A	401	TRS	C3-C-C1	3.76	118.91	110.78
3	A	401	TRS	C3-C-C2	4.16	119.80	110.78
3	D	401	TRS	C3-C-C1	4.30	120.10	110.78
3	D	401	TRS	C2-C-C1	4.38	120.27	110.78
3	E	401	TRS	C2-C-C1	4.57	120.67	110.78
3	A	401	TRS	C2-C-C1	4.70	120.96	110.78
3	C	401	TRS	C3-C-C2	4.96	121.53	110.78
3	C	401	TRS	C3-C-C1	5.15	121.93	110.78
3	E	401	TRS	C3-C-C2	5.21	122.06	110.78
2	B	400	ACO	CEP-CBP-CCP	6.10	116.41	108.50

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	400	ACO	1	0
5	C	400	AC8	1	0
2	E	400	ACO	2	0
3	E	401	TRS	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	381/392 (97%)	-0.14	3 (0%) 87 90	17, 32, 59, 82	0
1	B	381/392 (97%)	-0.18	2 (0%) 91 93	15, 32, 58, 79	0
1	C	387/392 (98%)	-0.23	0 100 100	18, 31, 54, 72	0
1	D	381/392 (97%)	0.02	8 (2%) 67 72	20, 41, 69, 92	0
1	E	381/392 (97%)	0.03	11 (2%) 55 63	18, 40, 68, 87	0
1	F	382/392 (97%)	-0.29	5 (1%) 79 84	17, 33, 59, 83	0
All	All	2293/2352 (97%)	-0.13	29 (1%) 79 84	15, 34, 62, 92	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	335	GLY	5.4
1	D	30	ALA	5.4
1	D	31	ALA	4.9
1	E	334	SER	4.0
1	B	332	ASP	3.7
1	E	57	ASP	3.6
1	E	332	ASP	3.6
1	E	153	ARG	3.4
1	E	336	ASP	3.4
1	D	57	ASP	3.4
1	D	334	SER	3.4
1	E	4	ASP	3.2
1	D	332	ASP	3.1
1	F	335	GLY	2.9
1	F	336	ASP	2.7
1	F	334	SER	2.7
1	B	334	SER	2.6
1	E	337	VAL	2.6
1	A	57	ASP	2.6

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	332	ASP	2.6
1	F	199	GLU	2.4
1	D	39	ARG	2.4
1	E	58	GLU	2.3
1	A	30	ALA	2.3
1	D	34	ARG	2.3
1	A	31	ALA	2.2
1	D	370	SER	2.1
1	E	383	PRO	2.1
1	E	154	GLU	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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	GOL	A	402	6/6	0.85	0.22	3.62	50,53,55,55	0
3	TRS	B	401	8/8	0.97	0.10	1.41	15,19,24,27	0
3	TRS	A	401	8/8	0.95	0.10	0.54	23,25,31,32	0
2	ACO	B	400	48/51	0.97	0.09	0.29	22,29,45,48	0
3	TRS	E	401	8/8	0.96	0.11	0.21	22,29,31,31	0
5	AC8	C	400	52/52	0.98	0.11	-0.07	14,23,70,72	0
2	ACO	A	400	48/51	0.98	0.09	-0.09	21,27,44,46	0
2	ACO	F	400	48/51	0.97	0.08	-0.16	15,24,50,51	0
2	ACO	D	400	48/51	0.98	0.09	-0.23	24,33,56,60	0
2	ACO	E	400	48/51	0.97	0.08	-0.69	22,34,63,66	0
3	TRS	C	401	8/8	0.94	0.09	-0.74	19,20,25,26	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
3	TRS	D	401	8/8	0.94	0.08	-0.95	28,32,34,35	0
3	TRS	F	401	8/8	0.98	0.08	-1.28	18,21,24,24	0
4	GOL	E	402	6/6	0.52	0.32	-	74,77,80,81	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.