



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 07:43 PM GMT

PDB ID : 1GXS
Title : CRYSTAL STRUCTURE OF HYDROXYNITRILE LYASE FROM SORGHUM BICOLOR IN COMPLEX WITH INHIBITOR BENZOIC ACID: A NOVEL CYANOGENIC ENZYME
Authors : Lauble, H.; Miehlich, B.; Foerster, S.; Wajant, H.; Effenberger, F.
Deposited on : 2002-04-11
Resolution : 2.30 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
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

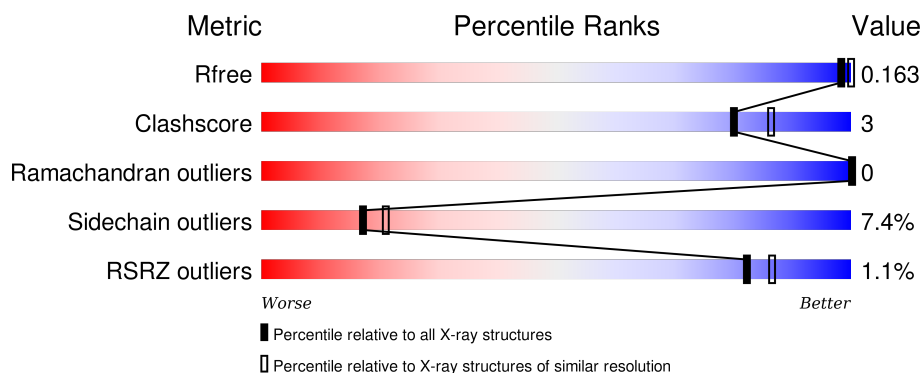
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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 ≥ 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	270	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>17%</div> <div>..</div> </div> </div>
1	C	270	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>16%</div> <div>...</div> </div> </div>
2	B	158	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>15%</div> <div>..</div> </div> </div>
2	D	158	<div> <div>%</div> <div> <div></div> <div>75%</div> <div>19%</div> <div>5%</div> <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
3	BEZ	C	602	-	-	-	X
4	DKA	A	701	-	-	-	X
4	DKA	C	702	-	-	-	X

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 7225 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 P-(S)-HYDROXYMANDELONITRILE LYASE CHAIN A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	267	Total	C	N	O	S	0	0	0
			2120	1347	356	406	11			
1	C	267	Total	C	N	O	S	0	0	0
			2120	1347	356	406	11			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	11	LEU	PRO	VARIANT	UNP Q8W4X3
A	79	ALA	PRO	VARIANT	UNP Q8W4X3
A	112	GLY	VAL	VARIANT	UNP Q8W4X3
A	203	SER	LEU	VARIANT	UNP Q8W4X3
C	11	LEU	PRO	VARIANT	UNP Q8W4X3
C	79	ALA	PRO	VARIANT	UNP Q8W4X3
C	112	GLY	VAL	VARIANT	UNP Q8W4X3
C	203	SER	LEU	VARIANT	UNP Q8W4X3

- Molecule 2 is a protein called P-(S)-HYDROXYMANDELONITRILE LYASE CHAIN B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	158	Total	C	N	O	S	0	0	0
			1254	815	206	229	4			
2	D	158	Total	C	N	O	S	0	0	0
			1254	815	206	229	4			

There are 6 discrepancies between the modelled and reference sequences:

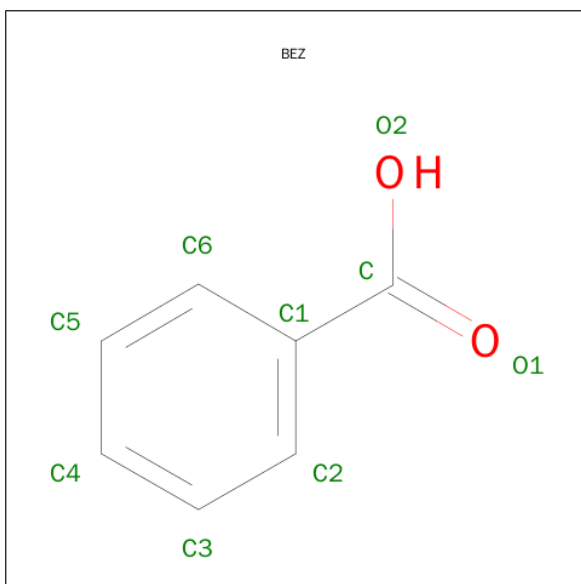
Chain	Residue	Modelled	Actual	Comment	Reference
B	408	THR	SER	VARIANT	UNP Q8W4X3
B	409	VAL	PRO	VARIANT	UNP Q8W4X3
B	410	ARG	SER	VARIANT	UNP Q8W4X3
D	408	THR	SER	VARIANT	UNP Q8W4X3

Continued on next page...

Continued from previous page...

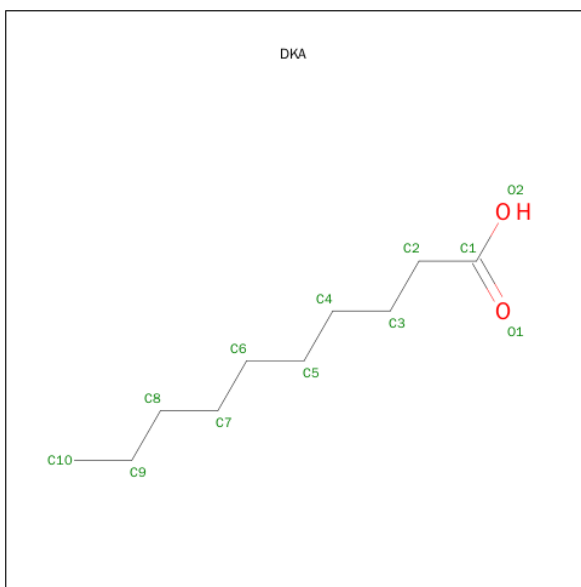
Chain	Residue	Modelled	Actual	Comment	Reference
D	409	VAL	PRO	VARIANT	UNP Q8W4X3
D	410	ARG	SER	VARIANT	UNP Q8W4X3

- Molecule 3 is BENZOIC ACID (three-letter code: BEZ) (formula: $C_7H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			9	7	2		
3	C	1	Total	C	O	0	0
			9	7	2		

- Molecule 4 is DECANOIC ACID (three-letter code: DKA) (formula: $C_{10}H_{20}O_2$).

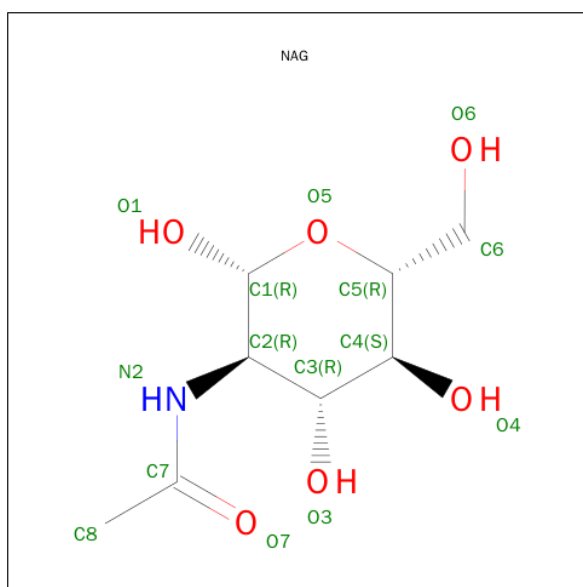


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			12	10	2		
4	C	1	Total	C	O	0	0
			12	10	2		

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	3	Total	C	N	O	0	0
			38	22	2	14		
5	C	3	Total	C	N	O	0	0
			38	22	2	14		

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	1	Total	C	N	O	0	0
			14	8	1	5		

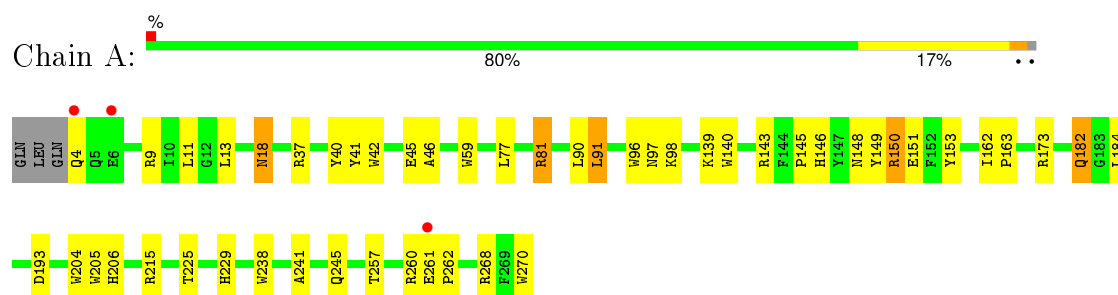
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	91	Total	O	0	0
			91	91		
7	B	73	Total	O	0	0
			73	73		
7	C	96	Total	O	0	0
			96	96		
7	D	71	Total	O	0	0
			71	71		

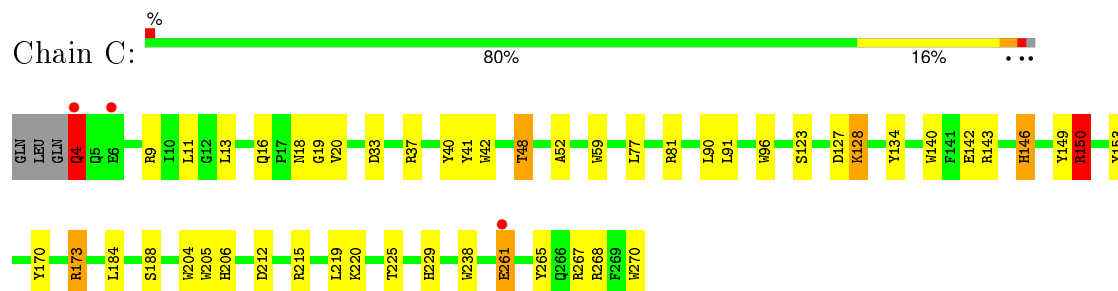
3 Residue-property plots [i](#)

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 ($\text{RSRZ} > 2$). 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.

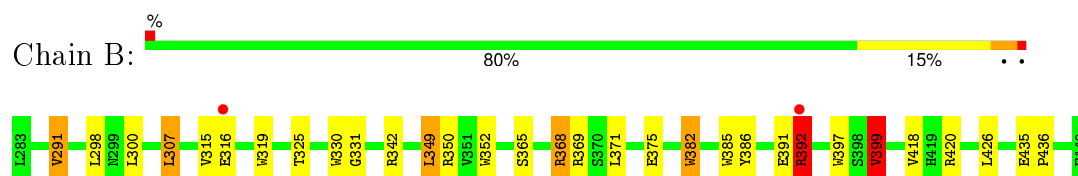
• Molecule 1: P-(S)-HYDROXYMANDELONITRILE LYASE CHAIN A



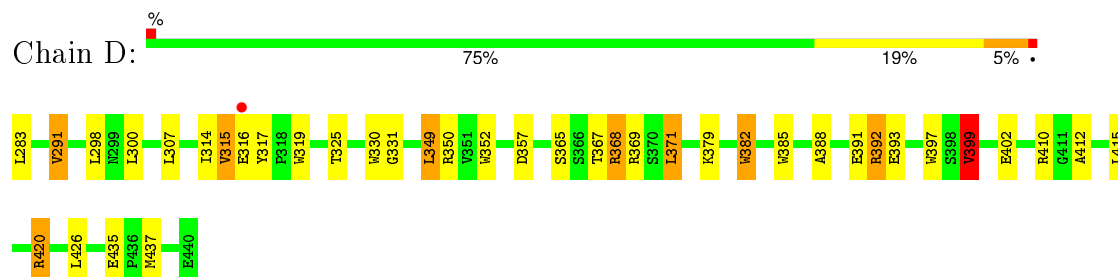
• Molecule 1: P-(S)-HYDROXYMANDELONITRILE LYASE CHAIN A



• Molecule 2: P-(S)-HYDROXYMANDELONITRILE LYASE CHAIN B



• Molecule 2: P-(S)-HYDROXYMANDELONITRILE LYASE CHAIN B



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	150.70 Å 103.70 Å 90.60 Å 90.00° 101.30° 90.00°	Depositor
Resolution (Å)	8.00 – 2.30 17.12 – 2.32	Depositor EDS
% Data completeness (in resolution range)	85.2 (8.00-2.30) 88.2 (17.12-2.32)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	8.13 (at 2.32 Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.165 , 0.222 0.164 , 0.163	Depositor DCC
R_{free} test set	2569 reflections (5.08%)	DCC
Wilson B-factor (Å ²)	21.1	Xtriage
Anisotropy	0.488	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 74.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 51746 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7225	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.93% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: BEZ, FUL, NAG, DKA

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	0.89	0/2187	1.59	43/2978 (1.4%)
1	C	0.93	0/2187	1.63	42/2978 (1.4%)
2	B	0.91	1/1296 (0.1%)	1.74	31/1781 (1.7%)
2	D	0.92	1/1296 (0.1%)	1.74	31/1781 (1.7%)
All	All	0.91	2/6966 (0.0%)	1.66	147/9518 (1.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	C	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	315	VAL	CA-CB	5.71	1.66	1.54
2	B	418	VAL	CA-CB	5.32	1.66	1.54

All (147) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	350	ARG	NE-CZ-NH2	-13.68	113.46	120.30
2	D	368	ARG	NE-CZ-NH2	-13.31	113.65	120.30
2	B	350	ARG	NE-CZ-NH1	12.59	126.59	120.30
1	A	81	ARG	NE-CZ-NH1	12.38	126.49	120.30
1	C	215	ARG	NE-CZ-NH2	-11.85	114.37	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	420	ARG	NE-CZ-NH1	11.59	126.10	120.30
2	B	420	ARG	NE-CZ-NH1	11.44	126.02	120.30
1	C	267	ARG	NE-CZ-NH1	11.44	126.02	120.30
1	C	215	ARG	NE-CZ-NH1	11.34	125.97	120.30
2	B	368	ARG	NE-CZ-NH2	-11.33	114.63	120.30
1	A	150	ARG	NE-CZ-NH2	-11.31	114.64	120.30
2	B	350	ARG	NE-CZ-NH2	-10.98	114.81	120.30
2	B	368	ARG	NE-CZ-NH1	10.77	125.68	120.30
1	C	150	ARG	NE-CZ-NH2	-10.55	115.03	120.30
2	D	368	ARG	NE-CZ-NH1	10.32	125.46	120.30
1	C	268	ARG	NE-CZ-NH1	10.28	125.44	120.30
1	A	215	ARG	NE-CZ-NH1	9.95	125.28	120.30
2	D	350	ARG	NE-CZ-NH1	9.90	125.25	120.30
1	C	268	ARG	NE-CZ-NH2	-9.81	115.39	120.30
2	B	399	VAL	CB-CA-C	-9.77	92.83	111.40
1	C	150	ARG	NE-CZ-NH1	9.28	124.94	120.30
2	D	399	VAL	CB-CA-C	-9.24	93.84	111.40
1	A	9	ARG	NE-CZ-NH1	8.99	124.80	120.30
2	B	397	TRP	CD1-CG-CD2	8.97	113.47	106.30
1	A	150	ARG	NE-CZ-NH1	8.96	124.78	120.30
2	D	291	VAL	CB-CA-C	-8.74	94.80	111.40
1	C	270	TRP	CD1-CG-CD2	8.51	113.11	106.30
1	C	267	ARG	NE-CZ-NH2	-8.32	116.14	120.30
1	A	270	TRP	CD1-CG-CD2	8.23	112.88	106.30
2	B	385	TRP	CD1-CG-CD2	8.21	112.87	106.30
1	C	238	TRP	CD1-CG-CD2	8.09	112.78	106.30
2	D	385	TRP	CD1-CG-CD2	8.09	112.77	106.30
1	A	140	TRP	CD1-CG-CD2	8.09	112.77	106.30
1	A	268	ARG	NE-CZ-NH1	8.07	124.33	120.30
2	D	437	MET	CG-SD-CE	7.99	112.99	100.20
2	B	319	TRP	CD1-CG-CD2	7.98	112.69	106.30
2	B	385	TRP	CE2-CD2-CG	-7.97	100.93	107.30
2	B	420	ARG	NE-CZ-NH2	-7.86	116.37	120.30
1	A	270	TRP	CE2-CD2-CG	-7.82	101.05	107.30
2	B	397	TRP	CE2-CD2-CG	-7.81	101.05	107.30
1	A	205	TRP	CD1-CG-CD2	7.72	112.48	106.30
1	C	270	TRP	CE2-CD2-CG	-7.67	101.16	107.30
1	A	238	TRP	CD1-CG-CD2	7.66	112.43	106.30
2	B	352	TRP	CD1-CG-CD2	7.61	112.39	106.30
2	D	385	TRP	CE2-CD2-CG	-7.56	101.25	107.30
1	A	140	TRP	CE2-CD2-CG	-7.50	101.30	107.30
2	D	352	TRP	CD1-CG-CD2	7.49	112.29	106.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	420	ARG	NE-CZ-NH2	-7.45	116.58	120.30
2	B	319	TRP	CE2-CD2-CG	-7.39	101.39	107.30
1	C	238	TRP	CE2-CD2-CG	-7.33	101.44	107.30
2	B	291	VAL	CB-CA-C	-7.29	97.55	111.40
2	D	330	TRP	CD1-CG-CD2	7.23	112.08	106.30
1	A	59	TRP	CE2-CD2-CG	-7.19	101.55	107.30
1	C	37	ARG	NE-CZ-NH1	7.17	123.89	120.30
1	C	134	TYR	CB-CG-CD2	-7.17	116.70	121.00
1	A	238	TRP	CE2-CD2-CG	-7.14	101.58	107.30
2	D	352	TRP	CE2-CD2-CG	-7.07	101.64	107.30
1	C	204	TRP	CD1-CG-CD2	7.06	111.95	106.30
1	C	140	TRP	CE2-CD2-CG	-7.02	101.68	107.30
1	C	204	TRP	CE2-CD2-CG	-7.02	101.68	107.30
1	A	96	TRP	CD1-CG-CD2	7.00	111.90	106.30
1	A	96	TRP	CE2-CD2-CG	-6.96	101.73	107.30
1	A	205	TRP	CE2-CD2-CG	-6.92	101.76	107.30
2	D	397	TRP	CD1-CG-CD2	6.91	111.83	106.30
1	C	59	TRP	CE2-CD2-CG	-6.89	101.78	107.30
1	A	91	LEU	CA-CB-CG	6.83	131.00	115.30
2	D	330	TRP	CE2-CD2-CG	-6.81	101.85	107.30
2	B	352	TRP	CE2-CD2-CG	-6.79	101.87	107.30
2	B	397	TRP	CB-CG-CD1	-6.79	118.18	127.00
2	D	382	TRP	CD1-CG-CD2	6.77	111.72	106.30
2	D	319	TRP	CE2-CD2-CG	-6.77	101.89	107.30
1	C	205	TRP	CE2-CD2-CG	-6.71	101.94	107.30
2	D	397	TRP	CE2-CD2-CG	-6.67	101.96	107.30
2	B	385	TRP	CG-CD2-CE3	6.67	139.90	133.90
1	A	59	TRP	CD1-CG-CD2	6.57	111.56	106.30
1	A	204	TRP	CE2-CD2-CG	-6.57	102.04	107.30
1	C	140	TRP	CD1-CG-CD2	6.54	111.53	106.30
2	B	330	TRP	CD1-CG-CD2	6.49	111.49	106.30
2	B	397	TRP	CG-CD2-CE3	6.44	139.70	133.90
1	C	205	TRP	CB-CG-CD1	-6.40	118.68	127.00
1	C	205	TRP	CD1-CG-CD2	6.35	111.38	106.30
1	A	204	TRP	CD1-CG-CD2	6.32	111.35	106.30
1	A	42	TRP	CD1-CG-CD2	6.29	111.34	106.30
2	B	385	TRP	CB-CG-CD1	-6.28	118.83	127.00
2	D	382	TRP	CE2-CD2-CG	-6.28	102.28	107.30
1	C	42	TRP	CD1-CG-CD2	6.25	111.30	106.30
1	C	42	TRP	CE2-CD2-CG	-6.25	102.30	107.30
2	D	319	TRP	CD1-CG-CD2	6.24	111.29	106.30
1	C	205	TRP	CG-CD2-CE3	6.18	139.46	133.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	342	ARG	NE-CZ-NH1	6.15	123.38	120.30
1	C	48	THR	N-CA-CB	-6.12	98.67	110.30
2	D	369	ARG	NE-CZ-NH1	6.01	123.30	120.30
2	B	330	TRP	CE2-CD2-CG	-5.99	102.51	107.30
1	A	260	ARG	NE-CZ-NH2	-5.93	117.33	120.30
2	B	382	TRP	CE2-CD2-CG	-5.93	102.56	107.30
1	A	42	TRP	CE2-CD2-CG	-5.90	102.58	107.30
2	B	397	TRP	CG-CD1-NE1	-5.88	104.22	110.10
2	B	399	VAL	N-CA-CB	5.88	124.43	111.50
2	B	382	TRP	CD1-CG-CD2	5.86	110.99	106.30
2	B	385	TRP	CG-CD1-NE1	-5.86	104.25	110.10
1	C	4	GLN	CA-C-N	5.83	130.02	117.20
1	C	96	TRP	CE2-CD2-CG	-5.78	102.67	107.30
1	C	59	TRP	CD1-CG-CD2	5.76	110.91	106.30
1	C	212	ASP	CB-CG-OD1	5.76	123.48	118.30
2	B	349	LEU	CA-CB-CG	5.71	128.42	115.30
1	A	173	ARG	NE-CZ-NH1	5.68	123.14	120.30
1	A	225	THR	CA-CB-CG2	5.68	120.36	112.40
1	A	193	ASP	CB-CG-OD1	5.67	123.41	118.30
2	D	399	VAL	N-CA-CB	5.65	123.93	111.50
1	A	96	TRP	CG-CD2-CE3	5.63	138.97	133.90
1	A	139	LYS	CA-CB-CG	5.63	125.78	113.40
1	A	260	ARG	NE-CZ-NH1	5.63	123.11	120.30
1	A	270	TRP	CG-CD2-CE3	5.59	138.93	133.90
2	D	349	LEU	CA-CB-CG	5.57	128.12	115.30
1	A	18	ASN	N-CA-C	5.54	125.97	111.00
1	C	96	TRP	CD1-CG-CD2	5.52	110.72	106.30
1	A	215	ARG	NE-CZ-NH2	-5.51	117.55	120.30
1	A	81	ARG	NE-CZ-NH2	-5.48	117.56	120.30
1	A	225	THR	CA-CB-OG1	-5.41	97.64	109.00
2	D	330	TRP	CG-CD2-CE3	5.40	138.76	133.90
1	A	98	LYS	CA-CB-CG	-5.39	101.54	113.40
1	C	204	TRP	CG-CD2-CE3	5.37	138.73	133.90
1	A	204	TRP	CG-CD2-CE3	5.36	138.72	133.90
1	A	45	GLU	CA-CB-CG	-5.35	101.64	113.40
1	A	149	TYR	CB-CG-CD2	-5.32	117.81	121.00
1	C	19	GLY	CA-C-N	5.30	128.87	117.20
2	D	317	TYR	CB-CG-CD2	-5.28	117.83	121.00
2	B	392	ARG	NE-CZ-NH1	5.23	122.91	120.30
1	A	37	ARG	NE-CZ-NH1	5.21	122.91	120.30
1	C	149	TYR	CB-CG-CD1	-5.21	117.88	121.00
2	D	357	ASP	CB-CG-OD1	5.18	122.97	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	205	TRP	CB-CG-CD1	-5.15	120.30	127.00
1	C	153	TYR	CB-CG-CD2	-5.15	117.91	121.00
1	C	140	TRP	CG-CD2-CE3	5.14	138.53	133.90
2	D	385	TRP	CA-CB-CG	5.11	123.41	113.70
2	D	291	VAL	N-CA-CB	5.10	122.72	111.50
1	A	268	ARG	NE-CZ-NH2	-5.08	117.76	120.30
2	B	386	TYR	CB-CG-CD1	-5.07	117.96	121.00
1	A	143	ARG	NE-CZ-NH1	5.06	122.83	120.30
1	C	59	TRP	CG-CD2-CE3	5.06	138.45	133.90
1	C	270	TRP	CG-CD1-NE1	-5.05	105.05	110.10
1	C	261	GLU	N-CA-C	5.04	124.62	111.00
1	C	265	TYR	CB-CG-CD2	-5.02	117.99	121.00
2	D	415	LEU	CA-CB-CG	5.01	126.82	115.30
1	C	41	TYR	CB-CG-CD1	-5.01	118.00	121.00
1	C	238	TRP	CG-CD1-NE1	-5.01	105.09	110.10
2	D	402	GLU	CA-CB-CG	5.00	124.40	113.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	40	TYR	Sidechain
1	A	41	TYR	Sidechain
1	C	40	TYR	Sidechain

5.2 Too-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 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	2120	0	1951	13	0
1	C	2120	0	1951	12	0
2	B	1254	0	1213	10	0
2	D	1254	0	1213	13	0
3	A	9	0	5	0	0
3	C	9	0	5	0	0
4	A	12	0	19	0	0
4	C	12	0	19	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	38	0	34	1	0
5	C	38	0	34	2	0
6	B	14	0	13	1	0
6	D	14	0	13	0	0
7	A	91	0	0	4	0
7	B	73	0	0	1	0
7	C	96	0	0	5	0
7	D	71	0	0	0	0
All	All	7225	0	6470	46	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 (46) 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:C:4:GLN:HG3	1:C:143:ARG:HG3	1.63	0.80
2:D:392:ARG:HE	2:D:393:GLU:H	1.41	0.69
2:B:392:ARG:H	2:B:392:ARG:HE	1.42	0.67
5:A:1118:NAG:H2	5:A:1119:FUL:H5	1.79	0.64
1:A:229:HIS:HE1	2:B:331:GLY:O	1.80	0.63
1:C:229:HIS:HE1	2:D:331:GLY:O	1.84	0.60
1:A:206:HIS:HD2	2:D:365:SER:OG	1.88	0.57
1:C:142:GLU:HG3	7:C:2048:HOH:O	2.05	0.56
1:C:206:HIS:HE1	7:C:2065:HOH:O	1.87	0.56
2:D:388:ALA:HB3	2:D:391:GLU:HG2	1.87	0.55
1:A:261:GLU:HG3	1:A:262:PRO:HD3	1.88	0.55
5:C:1118:NAG:H2	5:C:1119:FUL:H5	1.89	0.54
1:A:13:LEU:HD21	2:B:307:LEU:HD13	1.90	0.54
2:D:382:TRP:CE3	2:D:399:VAL:HG22	2.43	0.53
1:C:146:HIS:HB3	7:C:2049:HOH:O	2.07	0.53
2:D:412:ALA:HB2	2:D:420:ARG:HG3	1.92	0.51
2:B:391:GLU:HA	2:B:392:ARG:HH21	1.76	0.51
1:A:145:PRO:HA	1:A:148:ASN:ND2	2.26	0.51
1:A:153:TYR:CE1	1:A:182:GLN:HB3	2.47	0.50
7:A:2062:HOH:O	2:D:368:ARG:HD2	2.12	0.49
2:D:392:ARG:HE	2:D:393:GLU:N	2.11	0.48
2:B:325:THR:HG23	7:B:2023:HOH:O	2.13	0.48
2:B:368:ARG:HD2	7:C:2068:HOH:O	2.13	0.48
1:A:151:GLU:HG2	1:A:182:GLN:HG2	1.95	0.47
1:C:170:TYR:O	1:C:173:ARG:HG3	2.15	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:382:TRP:HE3	2:D:399:VAL:HG22	1.80	0.46
1:C:13:LEU:H	1:C:16:GLN:NE2	2.13	0.46
1:A:229:HIS:HD2	7:A:2071:HOH:O	1.99	0.46
1:A:162:ILE:HB	1:A:163:PRO:HD3	1.98	0.45
2:B:315:VAL:HG22	6:B:1310:NAG:HN2	1.83	0.44
1:C:225:THR:HG22	7:C:2073:HOH:O	2.18	0.44
1:C:52:ALA:HA	1:C:150:ARG:NH2	2.33	0.44
1:A:241:ALA:O	1:A:245:GLN:HG3	2.18	0.43
2:B:382:TRP:CE3	2:B:399:VAL:HG22	2.53	0.43
2:D:410:ARG:O	2:D:420:ARG:HD2	2.18	0.43
1:C:127:ASP:HB2	1:C:128:LYS:HZ2	1.83	0.43
2:B:435:GLU:HA	2:B:436:PRO:HD3	1.85	0.43
5:C:1118:NAG:H83	2:D:314:ILE:HD11	2.01	0.43
1:A:11:LEU:HD11	7:A:2002:HOH:O	2.19	0.42
1:A:81:ARG:HD3	7:A:2023:HOH:O	2.20	0.41
2:D:283:LEU:HA	2:D:283:LEU:HD12	1.89	0.41
1:C:9:ARG:HH21	1:C:11:LEU:HA	1.86	0.41
1:C:220:LYS:HE2	1:C:220:LYS:HB3	1.85	0.41
2:B:365:SER:O	2:B:369:ARG:HG3	2.20	0.41
1:A:46:ALA:HA	1:A:97:ASN:O	2.21	0.41
2:D:367:THR:HG22	2:D:371:LEU:HD22	2.02	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	265/270 (98%)	251 (95%)	14 (5%)	0	100	100
1	C	265/270 (98%)	251 (95%)	14 (5%)	0	100	100
2	B	156/158 (99%)	155 (99%)	1 (1%)	0	100	100
2	D	156/158 (99%)	156 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	842/856 (98%)	813 (97%)	29 (3%)	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	223/226 (99%)	213 (96%)	10 (4%)	34	46
1	C	223/226 (99%)	205 (92%)	18 (8%)	15	18
2	B	135/135 (100%)	124 (92%)	11 (8%)	15	18
2	D	135/135 (100%)	121 (90%)	14 (10%)	9	10
All	All	716/722 (99%)	663 (93%)	53 (7%)	17	21

All (53) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	18	ASN
1	A	77	LEU
1	A	90	LEU
1	A	91	LEU
1	A	146	HIS
1	A	150	ARG
1	A	182	GLN
1	A	184	LEU
1	A	257	THR
2	B	291	VAL
2	B	298	LEU
2	B	300	LEU
2	B	307	LEU
2	B	316	GLU
2	B	349	LEU
2	B	371	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	375	GLU
2	B	392	ARG
2	B	399	VAL
2	B	426	LEU
1	C	4	GLN
1	C	18	ASN
1	C	20	VAL
1	C	33	ASP
1	C	48	THR
1	C	77	LEU
1	C	81	ARG
1	C	90	LEU
1	C	91	LEU
1	C	123	SER
1	C	128	LYS
1	C	146	HIS
1	C	150	ARG
1	C	173	ARG
1	C	184	LEU
1	C	188	SER
1	C	219	LEU
1	C	261	GLU
2	D	291	VAL
2	D	298	LEU
2	D	300	LEU
2	D	307	LEU
2	D	315	VAL
2	D	316	GLU
2	D	325	THR
2	D	349	LEU
2	D	371	LEU
2	D	379	LYS
2	D	392	ARG
2	D	399	VAL
2	D	426	LEU
2	D	435	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	148	ASN
1	A	164	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	206	HIS
1	A	229	HIS
1	A	245	GLN
2	B	430	GLN
1	C	16	GLN
1	C	83	HIS
1	C	148	ASN
1	C	164	GLN
1	C	206	HIS
1	C	229	HIS
1	C	245	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 ⓘ

6 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	$\# Z > 2$
5	NAG	A	1117	1,5	14,14,15	0.56	0	15,19,21	0.73	0
5	NAG	A	1118	5	14,14,15	0.34	0	15,19,21	0.79	1 (6%)
5	FUL	A	1119	5	10,10,11	0.45	0	14,14,16	0.59	0
5	NAG	C	1117	1,5	14,14,15	0.45	0	15,19,21	0.74	1 (6%)
5	NAG	C	1118	5	14,14,15	0.38	0	15,19,21	0.80	1 (6%)
5	FUL	C	1119	5	10,10,11	0.42	0	14,14,16	0.47	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	A	1117	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1118	5	-	0/6/23/26	0/1/1/1
5	FUL	A	1119	5	-	0/0/17/20	0/1/1/1
5	NAG	C	1117	1,5	-	0/6/23/26	0/1/1/1
5	NAG	C	1118	5	-	0/6/23/26	0/1/1/1
5	FUL	C	1119	5	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1118	NAG	C4-C3-C2	-2.15	107.89	111.23
5	C	1118	NAG	C4-C3-C2	-2.13	107.92	111.23
5	C	1117	NAG	C2-N2-C7	-2.01	120.45	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1118	NAG	1	0
5	A	1119	FUL	1	0
5	C	1118	NAG	2	0
5	C	1119	FUL	1	0

5.6 Ligand geometry

6 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	BEZ	A	601	-	6,9,9	0.51	0	8,11,11	0.60	0
4	DKA	A	701	-	8,11,11	0.23	0	8,11,11	1.08	1 (12%)
6	NAG	B	1310	2	14,14,15	0.49	0	15,19,21	1.11	1 (6%)
3	BEZ	C	602	-	6,9,9	0.78	0	8,11,11	0.77	0
4	DKA	C	702	-	8,11,11	0.23	0	8,11,11	1.18	1 (12%)
6	NAG	D	1310	2	14,14,15	0.46	0	15,19,21	1.02	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	BEZ	A	601	-	-	0/0/4/4	0/1/1/1
4	DKA	A	701	-	-	0/7/9/9	0/0/0/0
6	NAG	B	1310	2	-	0/6/23/26	0/1/1/1
3	BEZ	C	602	-	-	0/0/4/4	0/1/1/1
4	DKA	C	702	-	-	0/7/9/9	0/0/0/0
6	NAG	D	1310	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	701	DKA	C5-C4-C3	-2.53	101.46	114.53
4	C	702	DKA	C5-C4-C3	-2.41	102.09	114.53
6	B	1310	NAG	O7-C7-C8	-2.16	118.09	122.06

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
6	B	1310	NAG	1	0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th 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 ²)	Q<0.9
1	A	267/270 (98%)	-0.68	3 (1%) 82 86	5, 16, 40, 85	0
1	C	267/270 (98%)	-0.71	3 (1%) 82 86	4, 15, 38, 79	0
2	B	158/158 (100%)	-0.80	2 (1%) 79 84	5, 14, 35, 52	0
2	D	158/158 (100%)	-0.76	1 (0%) 90 93	5, 14, 35, 53	0
All	All	850/856 (99%)	-0.73	9 (1%) 82 86	4, 15, 38, 85	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	261	GLU	5.1
1	C	4	GLN	3.9
1	C	261	GLU	3.6
1	A	4	GLN	3.2
1	A	6	GLU	2.8
1	C	6	GLU	2.3
2	B	316	GLU	2.2
2	D	316	GLU	2.0
2	B	392	ARG	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, 95th 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(\AA^2)	Q<0.9
5	NAG	A	1117	14/15	0.90	0.15	1.37	35,41,53,55	0
5	NAG	C	1117	14/15	0.92	0.13	0.39	37,44,53,59	0
5	NAG	C	1118	14/15	0.86	0.27	-	64,70,75,81	0
5	NAG	A	1118	14/15	0.87	0.24	-	55,61,66,67	0
5	FUL	A	1119	10/11	0.80	0.28	-	56,61,67,71	0
5	FUL	C	1119	10/11	0.87	0.23	-	57,59,65,72	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, 95th 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(\AA^2)	Q<0.9
4	DKA	A	701	12/12	0.94	0.18	5.15	17,27,36,40	0
4	DKA	C	702	12/12	0.95	0.14	3.25	19,22,27,29	0
3	BEZ	C	602	9/9	0.93	0.14	3.20	18,22,27,32	0
3	BEZ	A	601	9/9	0.95	0.11	1.41	23,29,34,37	0
6	NAG	B	1310	14/15	0.95	0.10	-0.23	7,24,31,40	0
6	NAG	D	1310	14/15	0.94	0.09	-0.30	11,22,29,34	0

6.5 Other polymers [i](#)

There are no such residues in this entry.