Full wwPDB X-ray Structure Validation Report (i)

Nov 5, 2023 – 08:43 PM EST

PDB ID	:	4WHS
Title	:	4-fluorocatechol bound to Protocatechuate 3,4-dioxygenase (pseudomonas putida) at pH 8.5
Authors	:	Knoot, C.J.; Purpero, V.M.; Lipscomb, J.D.
Deposited on	:	2014-09-23
Resolution	:	1.35 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

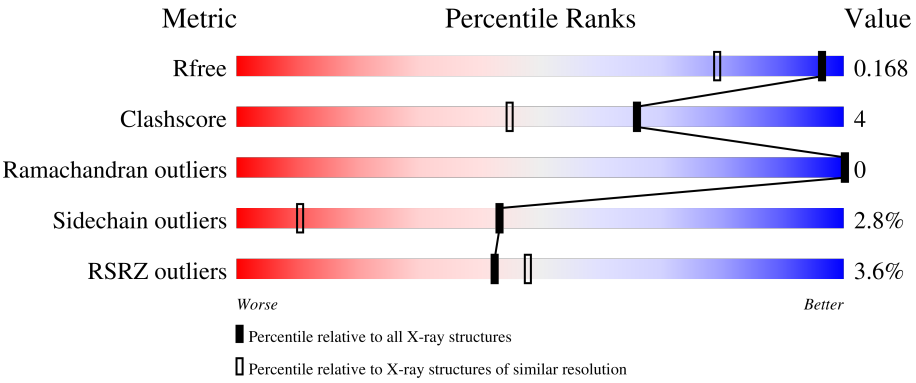
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.35 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.



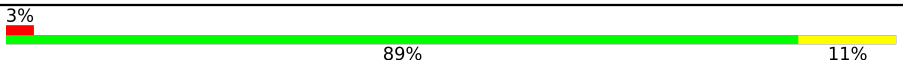
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1509 (1.38-1.34)
Clashscore	141614	1551 (1.38-1.34)
Ramachandran outliers	138981	1530 (1.38-1.34)
Sidechain outliers	138945	1530 (1.38-1.34)
RSRZ outliers	127900	1487 (1.38-1.34)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=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	200	4% 96% .
1	C	200	2% 92% 8%
1	E	200	6% 92% 7% .
2	B	238	3% 90% 9%
2	F	238	4% 90% 9% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	238	

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 12126 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 Protocatechuate 3,4-dioxygenase alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	5	0
			1614	1022	284	305	3			
1	E	200	Total	C	N	O	S	0	5	0
			1618	1023	287	305	3			
1	C	200	Total	C	N	O	S	0	5	0
			1612	1022	282	305	3			

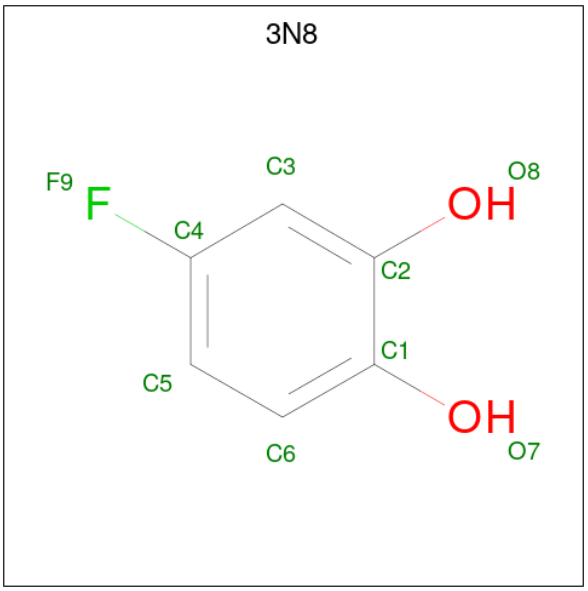
- Molecule 2 is a protein called Protocatechuate 3,4-dioxygenase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	238	Total	C	N	O	S	0	6	0
			1928	1223	350	346	9			
2	B	238	Total	C	N	O	S	0	8	0
			1944	1234	353	348	9			

- Molecule 3 is a protein called Protocatechuate 3,4-dioxygenase beta chain.

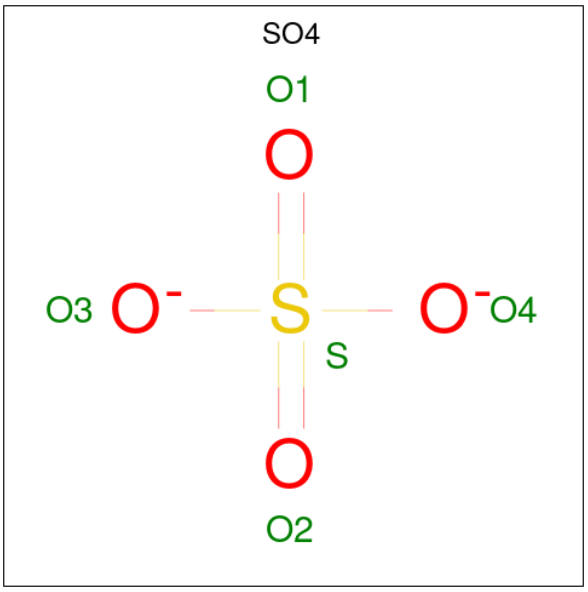
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	238	Total	C	N	O	S	0	7	0
			1941	1232	352	348	9			

- Molecule 4 is 4-fluorobenzene-1,2-diol (three-letter code: 3N8) (formula: $C_6H_5FO_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C F O 9 6 1 2	0	0
4	F	1	Total C F O 9 6 1 2	0	0
4	F	1	Total C F O 9 6 1 2	0	0
4	F	1	Total C F O 18 12 2 4	0	1
4	E	1	Total C F O 9 6 1 2	0	0
4	C	1	Total C F O 9 6 1 2	0	0
4	D	1	Total C F O 18 12 2 4	0	1
4	D	1	Total C F O 9 6 1 2	0	0
4	B	1	Total C F O 18 12 2 4	0	1
4	B	1	Total C F O 9 6 1 2	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

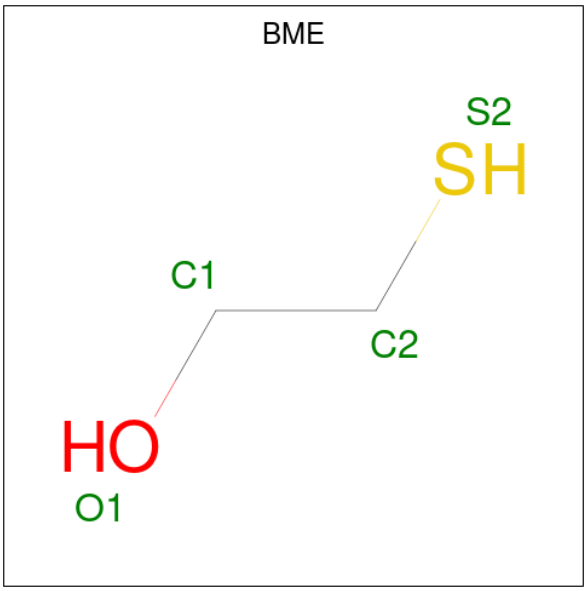


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0

- Molecule 6 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	F	1	Total Fe 1 1	0	0
6	D	1	Total Fe 1 1	0	0
6	B	1	Total Fe 1 1	0	0

- Molecule 7 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C_2H_6OS).

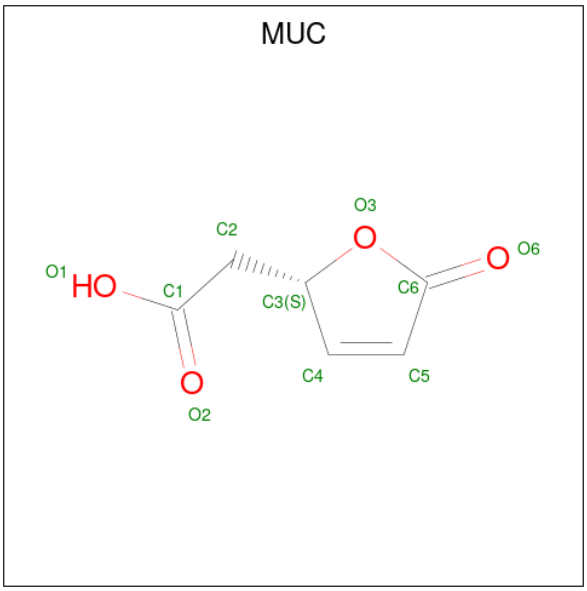


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	F	1	Total C O S 4 2 1 1	0	0
7	C	1	Total C O S 4 2 1 1	0	0
7	C	1	Total C O S 4 2 1 1	0	0
7	B	1	Total C O S 4 2 1 1	0	0

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	F	1	Total Cl 1 1	0	0
8	B	2	Total Cl 2 2	0	0

- Molecule 9 is [(2S)-5-oxo-2,5-dihydrofuran-2-yl]acetic acid (three-letter code: MUC) (formula: $C_6H_6O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	E	1	Total C O 10 6 4	0	0
9	C	1	Total C O 10 6 4	0	0

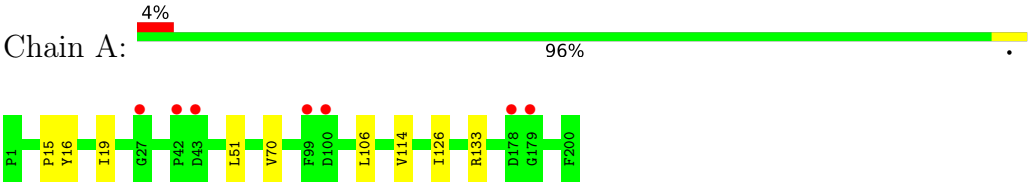
- Molecule 10 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	184	Total O 186 186	0	2
10	F	243	Total O 246 246	0	3
10	E	153	Total O 155 155	0	2
10	C	198	Total O 200 200	0	2
10	D	255	Total O 255 255	0	0
10	B	261	Total O 263 263	0	2

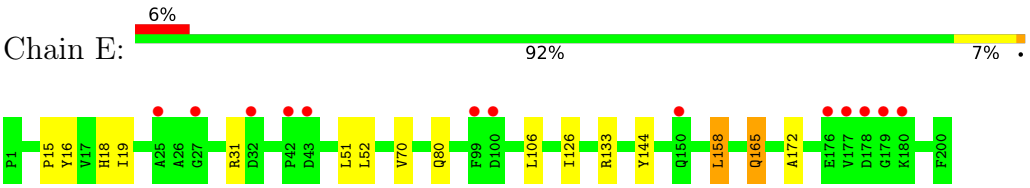
3 Residue-property plots (i)

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

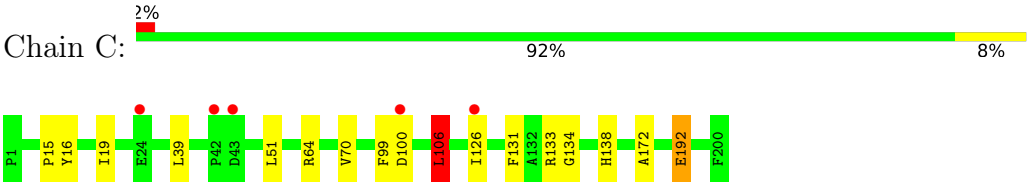
• Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



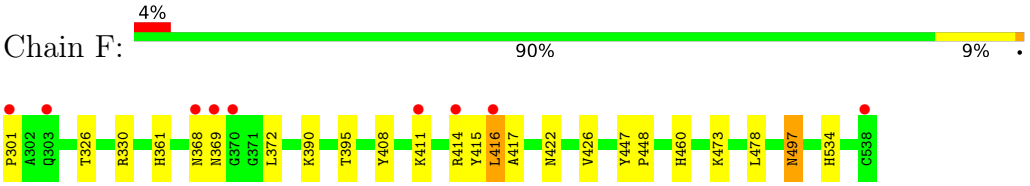
• Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



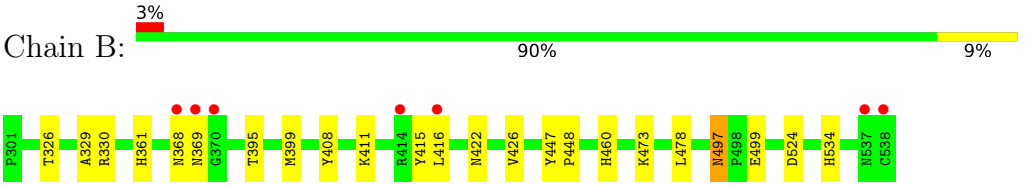
• Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain

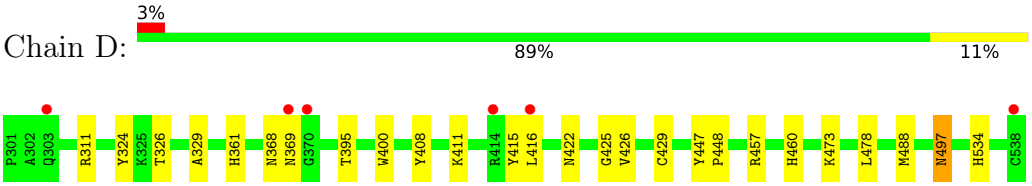


• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain



• Molecule 3: Protocatechuate 3,4-dioxygenase beta chain

Chain D:


3%
89%
11%
P301
A302
Q303
R311
Y324
K325
T326
A329
H361
N368
N369
G370
T395
W400
Y408
K411
R414
Y415
L416
N422
G425
V426
C429
Y447
P448
R457
H460
K473
L478
M488
N497
H534
C538

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	128.01Å 140.66Å 168.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.55 – 1.35 24.13 – 1.35	Depositor EDS
% Data completeness (in resolution range)	94.5 (29.55-1.35) 94.6 (24.13-1.35)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.99 (at 1.35Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.136 , 0.167 0.137 , 0.168	Depositor DCC
R_{free} test set	15583 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	12.8	Xtriage
Anisotropy	0.591	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.46 , 61.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	12126	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.88% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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: MHO, 3N8, FE, MUC, CSO, BME, SO4, CL

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	0.41	0/1655	0.67	0/2254
1	C	0.40	1/1653 (0.1%)	0.71	2/2251 (0.1%)
1	E	0.38	0/1659	0.67	2/2259 (0.1%)
2	B	0.44	0/2001	0.69	0/2723
2	F	0.40	0/1984	0.71	0/2701
3	D	0.41	0/1979	0.68	2/2691 (0.1%)
All	All	0.41	1/10931 (0.0%)	0.69	6/14879 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	192	GLU	CD-OE1	-5.12	1.20	1.25

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	311	ARG	NE-CZ-NH2	-6.02	117.29	120.30
1	E	31	ARG	NE-CZ-NH2	-5.89	117.36	120.30
3	D	457	ARG	NE-CZ-NH1	5.68	123.14	120.30
1	E	31	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	C	106[A]	LEU	CA-CB-CG	5.18	127.21	115.30
1	C	106[B]	LEU	CA-CB-CG	5.18	127.21	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	1614	0	1548	8	0
1	C	1612	0	1552	14	0
1	E	1618	0	1552	14	0
2	B	1944	0	1897	19	0
2	F	1928	0	1879	20	0
3	D	1941	0	1895	20	0
4	A	9	0	4	0	0
4	B	27	0	9	3	0
4	C	9	0	5	0	0
4	D	27	0	9	4	0
4	E	9	0	3	0	0
4	F	36	0	13	3	0
5	A	5	0	0	0	0
6	B	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
7	B	4	0	6	1	0
7	C	8	0	12	0	0
7	F	4	0	6	0	0
8	B	2	0	0	0	0
8	F	1	0	0	0	0
9	C	10	0	5	1	0
9	E	10	0	5	1	0
10	A	186	0	0	0	0
10	B	263	0	0	5	1
10	C	200	0	0	0	0
10	D	255	0	0	4	2
10	E	155	0	0	2	0
10	F	246	0	0	3	0
All	All	12126	0	10400	86	3

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 4.

All (86) 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)
2:F:447[A]:TYR:CE1	10:F:903:HOH:O	2.26	0.88
2:B:447[B]:TYR:CE1	10:B:907:HOH:O	2.29	0.84
3:D:447[A]:TYR:CE1	10:D:930:HOH:O	2.34	0.79
1:A:19:ILE:HG22	2:B:426[A]:VAL:HG22	1.67	0.74
1:C:64:ARG:NE	1:C:99:PHE:O	2.26	0.69
1:E:165:GLN:H	1:E:165:GLN:HE21	1.40	0.68
1:C:70:VAL:HG11	1:C:106[A]:LEU:HD21	1.75	0.67
2:F:426[A]:VAL:HG22	1:E:19:ILE:HG22	1.78	0.66
3:D:429:CSO:OD	10:D:934:HOH:O	2.02	0.65
1:A:133:ARG:HG2	2:B:326[A]:THR:HG21	1.79	0.65
1:C:51:LEU:HD11	1:C:126[B]:ILE:CD1	2.30	0.62
1:C:19:ILE:HG22	3:D:426[A]:VAL:HG22	1.82	0.61
3:D:447[B]:TYR:HB2	3:D:448:PRO:HD2	1.83	0.60
2:F:326[A]:THR:HG21	1:E:133:ARG:HG2	1.83	0.60
2:F:447[B]:TYR:CE2	2:F:460:HIS:HE1	2.21	0.58
3:D:361:HIS:CD2	3:D:361:HIS:H	2.23	0.57
1:C:172:ALA:HA	9:C:304:MUC:H5	1.88	0.55
3:D:447[B]:TYR:CE2	3:D:460:HIS:HE1	2.25	0.55
2:B:361:HIS:H	2:B:361:HIS:CD2	2.25	0.54
1:E:172:ALA:HA	9:E:302:MUC:H5	1.90	0.54
2:F:447[B]:TYR:CE2	2:F:460:HIS:CE1	2.96	0.53
2:B:329:ALA:HB3	10:B:913:HOH:O	2.08	0.53
3:D:408:TYR:HE2	3:D:447[B]:TYR:CZ	2.26	0.53
1:E:18[A]:HIS:CE1	10:E:404:HOH:O	2.61	0.53
1:E:70:VAL:HG11	1:E:106:LEU:HD21	1.91	0.53
1:A:19:ILE:CG2	2:B:426[A]:VAL:HG22	2.38	0.52
2:B:408:TYR:HE2	2:B:447[A]:TYR:CZ	2.27	0.52
3:D:408:TYR:HE2	3:D:447[B]:TYR:CE2	2.27	0.52
1:C:39:LEU:HD13	1:C:106[B]:LEU:HD11	1.91	0.52
2:B:447[A]:TYR:HB2	2:B:448:PRO:HD2	1.92	0.52
3:D:329:ALA:HB3	10:D:928:HOH:O	2.08	0.51
1:E:18[A]:HIS:HE1	10:E:404:HOH:O	1.92	0.51
2:F:408:TYR:HE2	2:F:447[B]:TYR:CZ	2.28	0.50
3:D:478:LEU:C	3:D:478:LEU:HD23	2.32	0.50
1:C:134:GLY:HA2	3:D:326[B]:THR:HG23	1.92	0.50
2:B:447[A]:TYR:CE2	2:B:460:HIS:HE1	2.29	0.50
2:F:361:HIS:H	2:F:361:HIS:CD2	2.28	0.50
2:B:415:TYR:HB2	10:B:907:HOH:O	2.13	0.49
4:F:605[B]:3N8:H3	1:E:16:TYR:CD2	2.47	0.49
1:C:16:TYR:CD2	4:D:602[B]:3N8:H3	2.47	0.49
1:C:15:PRO:HD2	4:D:602[B]:3N8:C6	2.42	0.49
1:A:51:LEU:HD11	1:A:126[B]:ILE:CD1	2.43	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:369:ASN:H	2:F:422:ASN:HD22	1.60	0.48
2:B:408:TYR:HE2	2:B:447[A]:TYR:CE2	2.31	0.48
7:B:602:BME:H22	10:B:713:HOH:O	2.14	0.48
2:B:326[A]:THR:HG22	2:B:330:ARG:HD2	1.96	0.48
1:E:51:LEU:HD11	1:E:126:ILE:CD1	2.44	0.48
1:A:70:VAL:HG11	1:A:106:LEU:HD11	1.96	0.47
2:F:414:ARG:HA	10:F:868:HOH:O	2.14	0.47
3:D:408:TYR:CE2	3:D:447[B]:TYR:CE2	3.02	0.47
1:A:15:PRO:HD2	4:B:603[B]:3N8:C6	2.44	0.47
2:F:301:PRO:HD2	1:E:80:GLN:HE22	1.79	0.47
2:F:447[B]:TYR:HB2	2:F:448:PRO:HD2	1.95	0.47
2:F:408:TYR:CE2	2:F:447[B]:TYR:CE2	3.03	0.47
2:F:408:TYR:HE2	2:F:447[B]:TYR:CE2	2.33	0.47
4:F:605[B]:3N8:C6	1:E:15:PRO:HD2	2.45	0.46
2:F:478:LEU:C	2:F:478:LEU:HD23	2.36	0.46
2:F:301:PRO:HD2	1:E:80:GLN:NE2	2.30	0.46
2:F:415:TYR:HB2	10:F:903:HOH:O	2.16	0.45
2:B:408:TYR:CE2	2:B:447[A]:TYR:CE2	3.04	0.45
2:B:447[A]:TYR:CE2	2:B:460:HIS:CE1	3.04	0.45
3:D:497:ASN:HD22	3:D:497:ASN:C	2.21	0.45
3:D:447[B]:TYR:CE2	3:D:460:HIS:CE1	3.05	0.44
3:D:369:ASN:H	3:D:422:ASN:HD22	1.64	0.44
2:B:369:ASN:H	2:B:422:ASN:HD22	1.63	0.44
1:A:16:TYR:CD2	4:B:603[B]:3N8:H3	2.53	0.44
3:D:324:TYR:CE1	3:D:326[B]:THR:HG22	2.53	0.44
1:C:39:LEU:HD13	1:C:106[B]:LEU:CD1	2.48	0.43
1:C:16:TYR:CD2	4:D:602[A]:3N8:H1	2.53	0.43
2:F:416:LEU:HD23	2:F:417:ALA:N	2.34	0.43
4:F:605[A]:3N8:H1	1:E:16:TYR:CD2	2.54	0.43
3:D:447[B]:TYR:HB2	3:D:448:PRO:CD	2.48	0.43
3:D:415:TYR:HB2	10:D:930:HOH:O	2.18	0.43
2:F:497:ASN:C	2:F:497:ASN:HD22	2.22	0.43
2:B:478:LEU:HD23	2:B:478:LEU:C	2.39	0.42
2:B:497:ASN:HD22	2:B:499:GLU:H	1.66	0.42
2:F:326[A]:THR:HG22	2:F:330:ARG:HD2	2.02	0.41
1:C:133:ARG:HG2	3:D:326[A]:THR:HG21	2.02	0.41
2:B:497:ASN:HD22	2:B:497:ASN:C	2.24	0.41
2:B:524:ASP:HB3	10:B:942:HOH:O	2.19	0.41
1:C:131:PHE:CE2	1:C:138:HIS:HB3	2.56	0.41
2:F:447[B]:TYR:CZ	2:F:460:HIS:HE1	2.39	0.40
1:A:16:TYR:CD2	4:B:603[A]:3N8:H1	2.55	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:15:PRO:HD2	4:D:602[A]:3N8:C3	2.50	0.40
3:D:400:TRP:HA	3:D:425:GLY:O	2.21	0.40
1:E:144:TYR:CE1	1:E:158[A]:LEU:HD13	2.56	0.40

All (3) 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	Interatomic distance (Å)	Clash overlap (Å)
10:D:702:HOH:O	10:D:726:HOH:O[4_555]	1.64	0.56
10:D:726:HOH:O	10:D:726:HOH:O[4_555]	1.72	0.48
10:B:741:HOH:O	10:B:741:HOH:O[2_555]	2.10	0.10

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	203/200 (102%)	201 (99%)	2 (1%)	0	100	100
1	C	202/200 (101%)	199 (98%)	3 (2%)	0	100	100
1	E	203/200 (102%)	199 (98%)	4 (2%)	0	100	100
2	B	243/238 (102%)	238 (98%)	5 (2%)	0	100	100
2	F	241/238 (101%)	237 (98%)	4 (2%)	0	100	100
3	D	240/238 (101%)	234 (98%)	6 (2%)	0	100	100
All	All	1332/1314 (101%)	1308 (98%)	24 (2%)	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	168/163 (103%)	166 (99%)	2 (1%)	71	42
1	C	168/163 (103%)	164 (98%)	4 (2%)	49	15
1	E	168/163 (103%)	164 (98%)	4 (2%)	49	15
2	B	210/202 (104%)	202 (96%)	8 (4%)	33	4
2	F	208/202 (103%)	199 (96%)	9 (4%)	29	3
3	D	207/200 (104%)	200 (97%)	7 (3%)	37	7
All	All	1129/1093 (103%)	1095 (97%)	34 (3%)	43	9

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	114[A]	VAL
1	A	114[B]	VAL
2	F	368	ASN
2	F	372	LEU
2	F	390	LYS
2	F	395	THR
2	F	411	LYS
2	F	416	LEU
2	F	473	LYS
2	F	497	ASN
2	F	534	HIS
1	E	52	LEU
1	E	158[A]	LEU
1	E	158[B]	LEU
1	E	165	GLN
1	C	100	ASP
1	C	106[A]	LEU
1	C	106[B]	LEU
1	C	192	GLU
3	D	368	ASN
3	D	395	THR
3	D	411	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	D	416	LEU
3	D	473	LYS
3	D	497	ASN
3	D	534	HIS
2	B	368	ASN
2	B	395	THR
2	B	399	MET
2	B	411	LYS
2	B	416	LEU
2	B	473	LYS
2	B	497	ASN
2	B	534	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (23) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	165	GLN
2	F	361	HIS
2	F	368	ASN
2	F	412	ASN
2	F	422	ASN
2	F	497	ASN
2	F	503	GLN
1	E	159	ASN
1	E	163	GLN
1	E	165	GLN
3	D	361	HIS
3	D	368	ASN
3	D	412	ASN
3	D	422	ASN
3	D	497	ASN
3	D	503	GLN
2	B	361	HIS
2	B	368	ASN
2	B	369	ASN
2	B	412	ASN
2	B	422	ASN
2	B	497	ASN
2	B	503	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	$\#\|Z\| > 2$
3	CSO	D	429	3	3,6,7	0.79	0	0,6,8	-	-
3	MHO	D	488	3	7,8,9	1.04	0	4,9,11	1.53	1 (25%)

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	CSO	D	429	3	-	0/1/5/7	-
3	MHO	D	488	3	-	2/6/7/9	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	488	MHO	CE-SD-CG	-2.80	91.34	97.71

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	488	MHO	CB-CG-SD-OD1
3	D	488	MHO	CB-CG-SD-CE

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	429	CSO	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 26 ligands modelled in this entry, 6 are monoatomic - leaving 20 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	$\#\|Z\| > 2$
9	MUC	C	304	-	10,10,10	1.93	1 (10%)	13,13,13	2.52	6 (46%)
7	BME	C	302	-	3,3,3	0.32	0	1,2,2	0.16	0
4	3N8	D	603	-	9,9,9	0.82	0	12,12,12	1.14	1 (8%)
7	BME	B	602	-	3,3,3	0.30	0	1,2,2	0.03	0
4	3N8	F	604	-	9,9,9	0.78	0	12,12,12	0.97	0
7	BME	C	303	-	3,3,3	0.35	0	1,2,2	0.73	0
4	3N8	C	301	-	9,9,9	0.86	0	12,12,12	1.08	1 (8%)
5	SO4	A	302	-	4,4,4	0.32	0	6,6,6	0.10	0
4	3N8	B	603[B]	-	9,9,9	0.65	0	12,12,12	1.05	1 (8%)
7	BME	F	603	-	3,3,3	0.21	0	1,2,2	0.13	0
4	3N8	B	603[A]	-	9,9,9	0.74	0	12,12,12	1.14	1 (8%)
4	3N8	F	602	-	9,9,9	0.66	0	12,12,12	1.06	1 (8%)
4	3N8	B	604	-	9,9,9	0.71	0	12,12,12	1.06	0
9	MUC	E	302	-	10,10,10	1.81	1 (10%)	13,13,13	2.66	7 (53%)
4	3N8	F	605[B]	-	9,9,9	0.59	0	12,12,12	1.14	1 (8%)
4	3N8	F	605[A]	-	9,9,9	0.64	0	12,12,12	1.13	1 (8%)
4	3N8	D	602[B]	-	9,9,9	0.70	0	12,12,12	1.08	1 (8%)
4	3N8	A	301	-	9,9,9	0.77	0	12,12,12	1.07	1 (8%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	3N8	D	602[A]	-	9,9,9	0.76	0	12,12,12	1.16	1 (8%)
4	3N8	E	301	-	9,9,9	0.77	0	12,12,12	0.97	1 (8%)

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
9	MUC	C	304	-	-	3/4/13/13	0/1/1/1
7	BME	C	302	-	-	0/1/1/1	-
4	3N8	D	603	-	-	-	0/1/1/1
7	BME	B	602	-	-	1/1/1/1	-
4	3N8	F	604	-	-	-	0/1/1/1
7	BME	C	303	-	-	1/1/1/1	-
4	3N8	C	301	-	-	-	0/1/1/1
4	3N8	B	603[B]	-	-	-	0/1/1/1
7	BME	F	603	-	-	1/1/1/1	-
4	3N8	B	603[A]	-	-	-	0/1/1/1
4	3N8	F	602	-	-	-	0/1/1/1
4	3N8	B	604	-	-	-	0/1/1/1
9	MUC	E	302	-	-	1/4/13/13	0/1/1/1
4	3N8	F	605[B]	-	-	-	0/1/1/1
4	3N8	F	605[A]	-	-	-	0/1/1/1
4	3N8	D	602[B]	-	-	-	0/1/1/1
4	3N8	A	301	-	-	-	0/1/1/1
4	3N8	D	602[A]	-	-	-	0/1/1/1
4	3N8	E	301	-	-	-	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	C	304	MUC	O3-C6	5.62	1.46	1.36
9	E	302	MUC	O3-C6	5.27	1.45	1.36

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	E	302	MUC	C2-C3-C4	-5.47	103.42	114.36
9	C	304	MUC	C3-O3-C6	-4.92	104.40	109.14
9	E	302	MUC	C3-O3-C6	-4.54	104.76	109.14

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	C	304	MUC	C2-C3-C4	-3.83	106.69	114.36
9	C	304	MUC	O6-C6-C5	-3.57	125.17	130.91
9	C	304	MUC	O3-C6-O6	3.32	125.01	120.95
9	E	302	MUC	O6-C6-C5	-2.91	126.24	130.91
4	D	602[A]	3N8	C5-C4-C3	-2.71	119.76	123.29
4	F	605[B]	3N8	C5-C4-C3	-2.71	119.77	123.29
4	B	603[A]	3N8	C5-C4-C3	-2.61	119.90	123.29
4	D	602[B]	3N8	C5-C4-C3	-2.57	119.96	123.29
4	B	603[B]	3N8	C5-C4-C3	-2.54	120.00	123.29
9	E	302	MUC	O3-C6-C5	2.52	109.83	108.16
4	D	603	3N8	C5-C4-C3	-2.51	120.02	123.29
9	C	304	MUC	O3-C6-C5	2.50	109.82	108.16
4	A	301	3N8	C5-C4-C3	-2.48	120.07	123.29
9	E	302	MUC	O3-C6-O6	2.42	123.91	120.95
4	F	605[A]	3N8	C5-C4-C3	-2.42	120.15	123.29
9	C	304	MUC	O3-C3-C4	2.40	106.72	103.31
4	C	301	3N8	C5-C4-C3	-2.37	120.21	123.29
4	E	301	3N8	C5-C4-C3	-2.31	120.30	123.29
9	E	302	MUC	O3-C3-C4	2.19	106.43	103.31
4	F	602	3N8	C5-C4-C3	-2.16	120.48	123.29
9	E	302	MUC	O1-C1-C2	2.10	120.79	114.07

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	F	603	BME	O1-C1-C2-S2
9	E	302	MUC	C1-C2-C3-C4
9	C	304	MUC	C1-C2-C3-C4
9	C	304	MUC	C1-C2-C3-O3
7	C	303	BME	O1-C1-C2-S2
7	B	602	BME	O1-C1-C2-S2
9	C	304	MUC	O1-C1-C2-C3

There are no ring outliers.

9 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	C	304	MUC	1	0
7	B	602	BME	1	0
4	B	603[B]	3N8	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	603[A]	3N8	1	0
9	E	302	MUC	1	0
4	F	605[B]	3N8	2	0
4	F	605[A]	3N8	1	0
4	D	602[B]	3N8	2	0
4	D	602[A]	3N8	2	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^{th} 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	200/200 (100%)	-0.27	7 (3%)	44	49	12, 17, 32, 52	0
1	C	200/200 (100%)	-0.25	5 (2%)	57	63	12, 17, 28, 49	0
1	E	200/200 (100%)	-0.02	13 (6%)	18	21	13, 21, 38, 59	0
2	B	238/238 (100%)	-0.33	7 (2%)	51	59	13, 15, 25, 45	0
2	F	238/238 (100%)	-0.28	9 (3%)	40	45	13, 17, 29, 44	1 (0%)
3	D	236/238 (99%)	-0.33	6 (2%)	57	63	13, 15, 26, 45	0
All	All	1312/1314 (99%)	-0.25	47 (3%)	42	48	12, 17, 30, 59	1 (0%)

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	178	ASP	6.2
1	E	100	ASP	4.5
3	D	414	ARG	4.4
2	F	301	PRO	4.3
1	E	179	GLY	4.2
2	B	414	ARG	3.9
2	F	414	ARG	3.8
2	F	370	GLY	3.6
1	A	42	PRO	3.6
2	F	538[A]	CYS	3.6
3	D	538[A]	CYS	3.5
1	A	178	ASP	3.5
1	E	27	GLY	3.4
1	A	27	GLY	3.3
1	A	179	GLY	3.3
2	F	303	GLN	3.1
1	C	42	PRO	3.0
2	B	369	ASN	3.0
1	C	126[A]	ILE	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	D	370	GLY	2.9
1	E	150	GLN	2.8
1	A	99	PHE	2.8
1	E	99	PHE	2.8
1	E	176	GLU	2.8
1	C	24	GLU	2.7
1	A	100	ASP	2.7
2	B	370	GLY	2.7
3	D	416	LEU	2.7
3	D	303	GLN	2.6
2	B	538[A]	CYS	2.6
2	F	411	LYS	2.5
1	E	42	PRO	2.5
3	D	369	ASN	2.5
1	E	177	VAL	2.5
1	C	100	ASP	2.4
1	A	43	ASP	2.4
2	F	416	LEU	2.3
2	B	416	LEU	2.3
1	C	43	ASP	2.3
1	E	32	ASP	2.2
2	B	537	ASN	2.2
1	E	25	ALA	2.2
1	E	43	ASP	2.2
1	E	180	LYS	2.2
2	F	368	ASN	2.1
2	F	369	ASN	2.1
2	B	368	ASN	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
3	MHO	D	488	9/10	0.97	0.08	16,18,31,37	0
3	CSO	D	429	7/8	0.98	0.05	13,14,18,22	1

6.3 Carbohydrates (i)

There are no monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
7	BME	F	603	4/4	0.62	0.18	32,33,43,43	0
7	BME	C	303	4/4	0.67	0.21	26,35,40,56	0
9	MUC	E	302	10/10	0.82	0.15	29,35,52,52	0
9	MUC	C	304	10/10	0.85	0.17	24,33,56,57	0
7	BME	B	602	4/4	0.86	0.23	29,32,43,51	0
4	3N8	C	301	9/9	0.88	0.17	27,28,32,32	0
4	3N8	F	602	9/9	0.89	0.09	30,31,39,49	0
7	BME	C	302	4/4	0.89	0.20	36,38,43,49	0
4	3N8	F	604	9/9	0.90	0.09	24,28,35,41	0
4	3N8	D	603	9/9	0.90	0.15	32,36,42,54	0
4	3N8	E	301	9/9	0.91	0.11	29,30,37,38	0
4	3N8	A	301	9/9	0.93	0.12	23,25,30,31	0
5	SO4	A	302	5/5	0.93	0.16	26,31,35,36	5
4	3N8	B	603[A]	9/9	0.95	0.11	15,16,18,19	9
4	3N8	B	603[B]	9/9	0.95	0.11	16,18,19,23	9
4	3N8	B	604	9/9	0.95	0.07	25,27,35,42	0
4	3N8	F	605[B]	9/9	0.95	0.10	16,17,18,22	9
4	3N8	F	605[A]	9/9	0.95	0.10	15,16,18,21	9
4	3N8	D	602[B]	9/9	0.96	0.10	15,16,18,20	9
4	3N8	D	602[A]	9/9	0.96	0.10	15,17,18,18	9
8	CL	F	606	1/1	0.98	0.07	43,43,43,43	0
8	CL	B	605	1/1	0.98	0.12	49,49,49,49	0
8	CL	B	606	1/1	0.99	0.06	35,35,35,35	0
6	FE	F	601	1/1	1.00	0.03	14,14,14,14	1
6	FE	D	601	1/1	1.00	0.03	13,13,13,13	1
6	FE	B	601	1/1	1.00	0.04	14,14,14,14	1

6.5 Other polymers (i)

There are no such residues in this entry.