Full wwPDB X-ray Structure Validation Report (i)

Oct 9, 2023 – 12:19 PM EDT

PDB ID	:	8DI0
Title	:	Bfo2290: Tannerella forsythia chondroitin sulfate A sulfatase
Authors	:	Suits, M.D.L.
Deposited on	:	2022-06-28
Resolution	:	2.85 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:

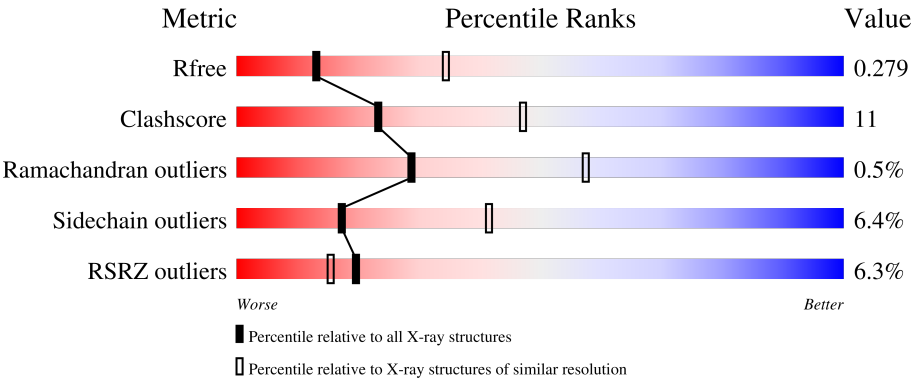
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.85 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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	516	6% 61% 22% · 14%
1	B	516	7% 54% 25% · 17%
1	C	516	3% 63% 25% · 10%

2 Entry composition

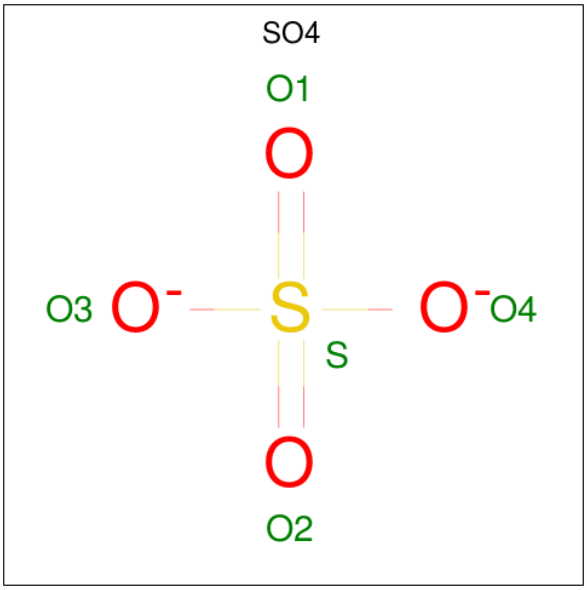
There are 3 unique types of molecules in this entry. The entry contains 10773 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Arylsulfatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	442	Total	C	N	O	S	0	0	0
			3563	2259	618	667	19			
1	B	426	Total	C	N	O	S	0	0	0
			3441	2180	600	642	19			
1	C	463	Total	C	N	O	S	0	0	0
			3745	2370	650	706	19			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 3 is water.

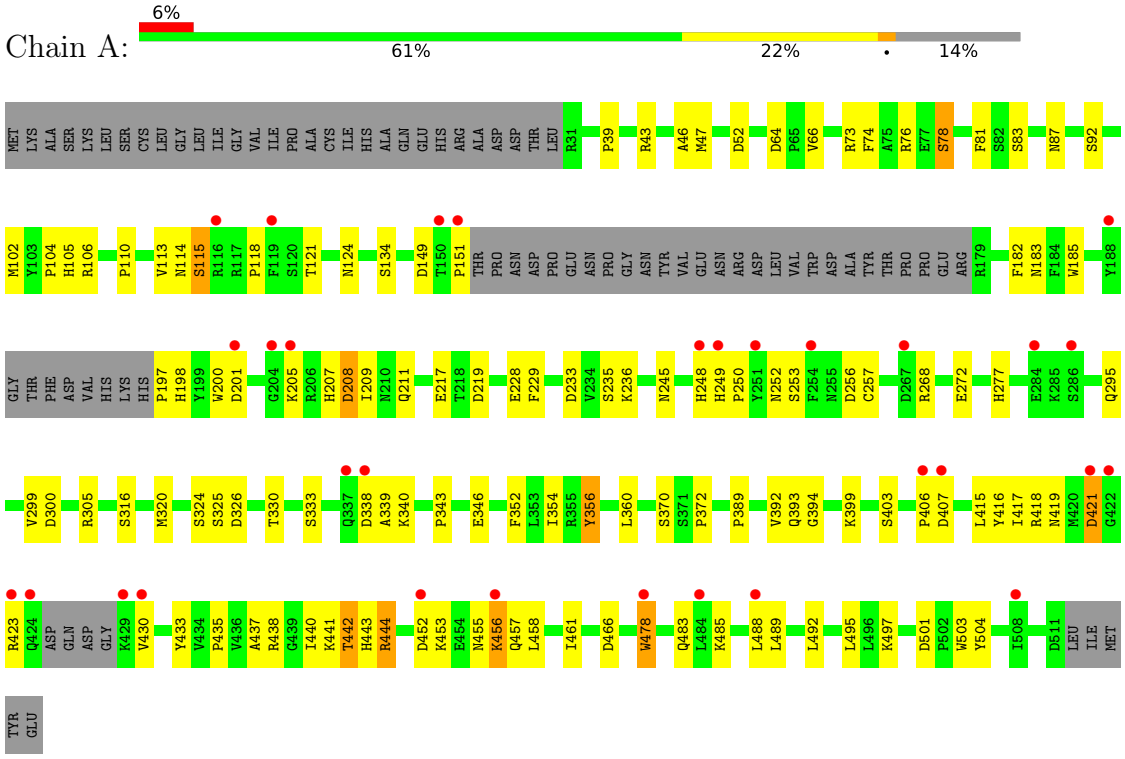
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	5	Total O 5 5	0	0
3	B	10	Total O 10 10	0	0
3	C	4	Total O 4 4	0	0

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.

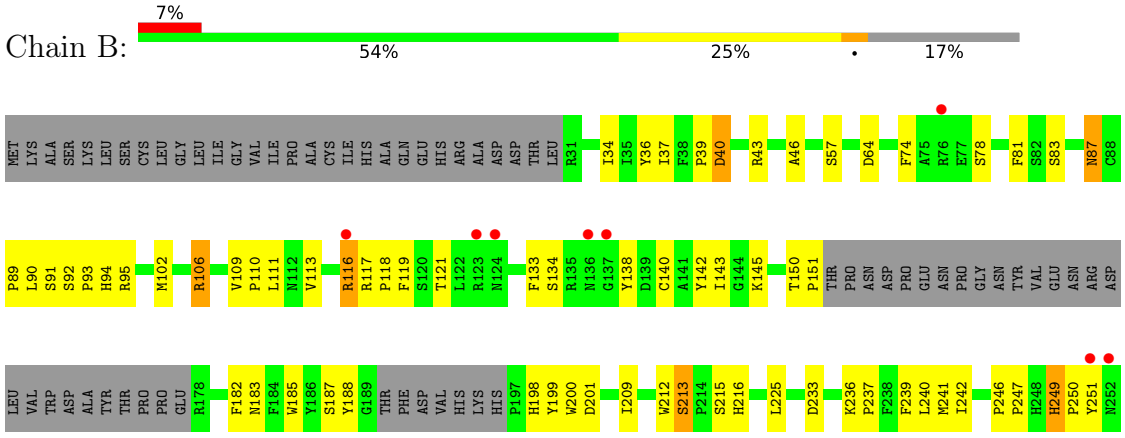
• Molecule 1: Arylsulfatase

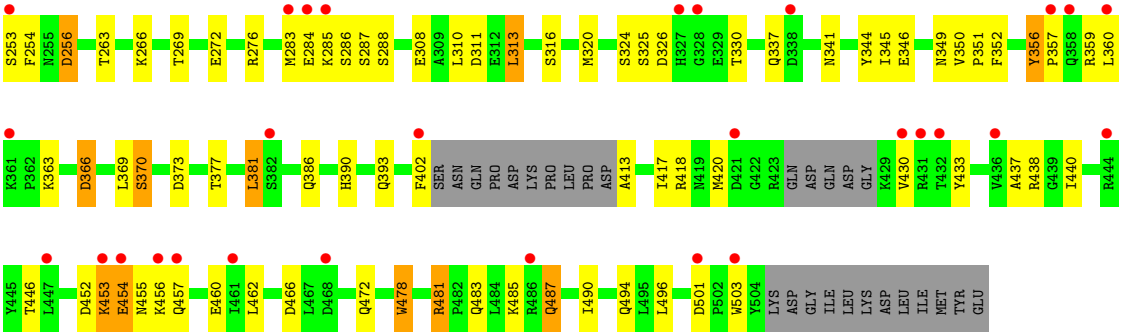
Chain A:



• Molecule 1: Arylsulfatase

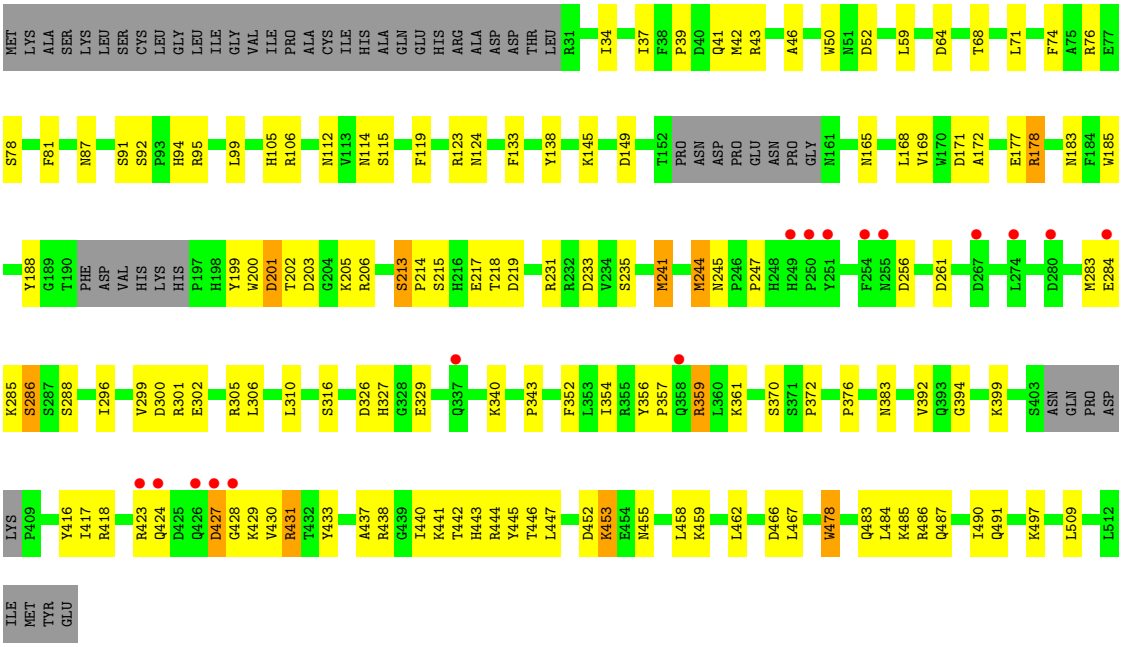
Chain B:





- Molecule 1: Arylsulfatase

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	113.18Å 121.11Å 144.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.32 – 2.85 48.32 – 2.85	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.32-2.85) 100.0 (48.32-2.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.60 (at 2.86Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.222 , 0.281 0.221 , 0.279	Depositor DCC
R_{free} test set	2355 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	69.0	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 52.9	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	0.92	EDS
Total number of atoms	10773	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.26% 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: SO4

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	0.27	0/3662	0.53	0/4971
1	B	0.27	0/3535	0.57	0/4794
1	C	0.26	0/3848	0.51	0/5225
All	All	0.27	0/11045	0.54	0/14990

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	B	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	116	ARG	Sidechain
1	B	454	GLU	Peptide

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	3563	0	3409	68	0
1	B	3441	0	3293	89	0
1	C	3745	0	3579	80	0
2	B	5	0	0	1	0
3	A	5	0	0	0	0
3	B	10	0	0	1	0
3	C	4	0	0	2	0
All	All	10773	0	10281	237	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 11.

All (237) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:456:LYS:H	1:A:456:LYS:HD3	1.47	0.79
1:C:183:ASN:ND2	3:C:601:HOH:O	2.14	0.78
1:B:452:ASP:HB3	1:B:455:ASN:HB3	1.64	0.77
1:A:418:ARG:HB3	1:A:437:ALA:HB3	1.66	0.77
1:B:418:ARG:HB3	1:B:437:ALA:HB3	1.67	0.75
1:C:37:ILE:HD11	1:C:306:LEU:HD12	1.69	0.74
1:A:421:ASP:OD1	1:A:421:ASP:N	2.21	0.73
1:C:452:ASP:OD2	1:C:455:ASN:ND2	2.22	0.72
1:B:81:PHE:HB2	1:B:352:PHE:HB3	1.72	0.70
1:B:283:MET:HG3	1:B:285:LYS:H	1.57	0.70
1:B:118:PRO:HA	1:B:150:THR:HG21	1.74	0.69
1:C:124:ASN:OD1	1:C:178:ARG:NH1	2.24	0.69
1:B:325:SER:HB3	1:B:351:PRO:HD2	1.75	0.69
1:C:105:HIS:ND1	3:C:602:HOH:O	2.26	0.68
1:A:102:MET:HG2	1:A:393:GLN:HG3	1.75	0.68
1:A:440:ILE:HG23	1:A:492:LEU:HD13	1.75	0.67
1:C:214:PRO:HB3	1:C:244:MET:HE3	1.75	0.67
1:C:41:GLN:HG3	1:C:327:HIS:CE1	2.30	0.67
1:C:203:ASP:OD2	1:C:231:ARG:NH2	2.28	0.67
1:B:283:MET:SD	1:B:284:GLU:N	2.68	0.67
1:A:52:ASP:OD2	1:A:76:ARG:NH2	2.28	0.66
1:A:219:ASP:OD1	1:A:305:ARG:NH2	2.25	0.66
1:B:142:TYR:OH	1:B:145:LYS:O	2.13	0.66
1:A:121:THR:HG23	1:A:149:ASP:HA	1.77	0.66
1:B:481:ARG:HB2	1:B:483:GLN:HE22	1.61	0.66
1:A:81:PHE:HB2	1:A:352:PHE:HB3	1.79	0.65
1:C:418:ARG:HB3	1:C:437:ALA:HB3	1.77	0.65

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:302:GLU:OE1	1:C:305:ARG:NH1	2.29	0.64
1:A:46:ALA:HA	1:A:64:ASP:HB2	1.79	0.64
1:A:201:ASP:HB2	1:A:205:LYS:HB2	1.78	0.64
1:C:81:PHE:HB2	1:C:352:PHE:HB3	1.80	0.64
1:C:427:ASP:OD1	1:C:427:ASP:N	2.30	0.64
1:A:441:LYS:NZ	1:A:466:ASP:OD2	2.31	0.63
1:C:283:MET:O	1:C:286:SER:OG	2.16	0.63
1:C:74:PHE:O	1:C:78:SER:OG	2.12	0.63
1:B:373:ASP:O	1:B:377:THR:OG1	2.16	0.63
1:C:376:PRO:HG3	1:C:392:VAL:HG11	1.80	0.62
1:C:442:THR:HG23	1:C:445:TYR:H	1.65	0.62
1:C:428:GLY:O	1:C:431:ARG:NH1	2.33	0.61
1:B:102:MET:HB3	1:B:393:GLN:HG3	1.83	0.61
1:A:330:THR:HA	1:A:346:GLU:HB3	1.83	0.61
1:B:359:ARG:HD3	1:B:402:PHE:HE1	1.66	0.60
1:A:443:HIS:HB2	1:A:444:ARG:HD3	1.82	0.60
1:A:415:LEU:HG	1:A:495:LEU:HD23	1.83	0.60
1:A:455:ASN:O	1:A:457:GLN:N	2.35	0.60
1:A:483:GLN:OE1	1:A:483:GLN:N	2.32	0.59
1:B:345:ILE:HG13	1:B:349:ASN:HB2	1.84	0.59
1:C:441:LYS:NZ	1:C:466:ASP:OD2	2.27	0.59
1:B:356:TYR:HB3	1:B:360:LEU:HB2	1.83	0.59
1:A:249:HIS:ND1	1:A:250:PRO:HD3	2.17	0.59
1:A:87:ASN:HD21	1:A:370:SER:HB3	1.67	0.58
1:B:151:PRO:O	3:B:701:HOH:O	2.17	0.58
1:A:209:ILE:HG13	1:A:211:GLN:H	1.68	0.58
1:B:199:TYR:HH	1:B:213:SER:HG	1.51	0.58
1:B:478:TRP:HB2	1:B:485:LYS:HB2	1.85	0.58
1:A:104:PRO:HB3	1:A:110:PRO:HA	1.86	0.58
1:A:421:ASP:HB3	1:A:430:VAL:HG11	1.85	0.58
1:A:74:PHE:O	1:A:78:SER:OG	2.22	0.57
1:A:423:ARG:HA	1:A:430:VAL:HG22	1.85	0.57
1:A:201:ASP:OD2	1:A:207:HIS:NE2	2.38	0.56
1:A:356:TYR:HB3	1:A:360:LEU:HB2	1.87	0.56
1:B:74:PHE:O	1:B:78:SER:OG	2.18	0.56
1:C:92:SER:HB2	1:C:112:ASN:HA	1.88	0.56
1:C:247:PRO:HG2	1:C:256:ASP:HB3	1.88	0.56
1:B:247:PRO:HB2	1:B:256:ASP:HB3	1.88	0.56
1:B:487:GLN:O	1:B:490:ILE:HG13	2.06	0.56
1:B:34:ILE:HG22	1:B:320:MET:HG2	1.88	0.56
1:C:453:LYS:HE3	1:C:459:LYS:NZ	2.21	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:41:GLN:HG3	1:C:327:HIS:HE1	1.70	0.55
1:C:285:LYS:O	1:C:288:SER:OG	2.17	0.55
1:B:251:TYR:CZ	1:B:285:LYS:HA	2.43	0.54
1:C:52:ASP:OD2	1:C:76:ARG:NH2	2.37	0.54
1:C:188:TYR:HB2	1:C:199:TYR:CE2	2.42	0.54
1:B:43:ARG:NH1	1:B:330:THR:O	2.40	0.54
1:C:105:HIS:CD2	1:C:106:ARG:HG2	2.43	0.54
1:B:249:HIS:HB2	1:B:250:PRO:HD2	1.90	0.53
1:C:39:PRO:HG3	1:C:299:VAL:HG11	1.89	0.53
1:B:81:PHE:CE2	1:B:369:LEU:HD12	2.43	0.53
1:B:496:LEU:HD12	1:B:501:ASP:HB3	1.91	0.53
1:A:115:SER:HB3	1:A:151:PRO:HG2	1.88	0.53
1:C:217:GLU:OE2	1:C:245:ASN:ND2	2.42	0.52
1:B:455:ASN:OD1	1:B:457:GLN:HG3	2.08	0.52
1:C:114:ASN:OD1	1:C:115:SER:N	2.42	0.52
1:B:453:LYS:HG2	1:B:454:GLU:HG3	1.91	0.52
1:A:217:GLU:OE2	1:A:245:ASN:ND2	2.41	0.52
1:C:446:THR:OG1	1:C:466:ASP:OD2	2.27	0.51
1:A:197:PRO:HB2	1:A:198:HIS:CE1	2.44	0.51
1:C:359:ARG:NH2	1:C:383:ASN:OD1	2.43	0.51
1:A:453:LYS:N	1:A:453:LYS:HD3	2.25	0.51
1:B:390:HIS:H	1:B:390:HIS:CD2	2.28	0.51
1:B:446:THR:OG1	1:B:466:ASP:OD2	2.29	0.50
1:A:399:LYS:O	1:A:403:SER:OG	2.28	0.50
1:B:133:PHE:O	1:B:138:TYR:HB2	2.11	0.50
1:C:424:GLN:HG2	1:C:430:VAL:HG23	1.94	0.50
1:C:106:ARG:O	1:C:123:ARG:NH1	2.45	0.50
1:C:486:ARG:O	1:C:490:ILE:HG12	2.12	0.49
1:C:261:ASP:CG	1:C:301:ARG:HH21	2.16	0.49
1:A:92:SER:HB3	1:A:113:VAL:HG23	1.94	0.49
1:B:455:ASN:OD1	1:B:456:LYS:N	2.41	0.49
1:C:218:THR:HG23	1:C:306:LEU:HG	1.93	0.49
1:B:438:ARG:NH2	1:B:501:ASP:OD1	2.45	0.48
1:C:68:THR:HG22	1:C:68:THR:O	2.12	0.48
1:C:359:ARG:HH22	1:C:383:ASN:HD21	1.60	0.48
1:B:46:ALA:HA	1:B:64:ASP:HB2	1.95	0.48
1:B:249:HIS:HB2	1:B:250:PRO:CD	2.43	0.48
1:C:343:PRO:HG3	1:C:416:TYR:CE2	2.49	0.48
1:B:377:THR:O	1:B:381:LEU:HD23	2.13	0.48
1:C:145:LYS:H	1:C:245:ASN:HD21	1.60	0.48
1:B:102:MET:HE3	1:B:106:ARG:HB3	1.95	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:43:ARG:NH2	1:A:64:ASP:OD1	2.39	0.48
1:A:442:THR:OG1	1:A:443:HIS:N	2.45	0.48
1:B:143:ILE:HD12	1:B:240:LEU:HD11	1.95	0.48
1:B:34:ILE:HD11	1:B:239:PHE:HD1	1.78	0.48
1:B:316:SER:O	1:B:357:PRO:HG2	2.14	0.48
1:C:171:ASP:CG	1:C:172:ALA:H	2.16	0.48
1:C:165:ASN:HA	1:C:168:LEU:HD12	1.96	0.47
1:C:185:TRP:HB3	1:C:202:THR:HA	1.96	0.47
1:A:417:ILE:HG23	1:A:438:ARG:CZ	2.44	0.47
1:A:452:ASP:HB3	1:A:455:ASN:O	2.14	0.47
1:C:119:PHE:CZ	1:C:433:TYR:HB3	2.49	0.47
1:B:344:TYR:HD1	1:B:472:GLN:HE21	1.62	0.47
1:C:442:THR:OG1	1:C:443:HIS:N	2.47	0.47
1:C:453:LYS:HE3	1:C:459:LYS:HZ1	1.80	0.47
1:B:183:ASN:OD1	1:B:183:ASN:N	2.45	0.47
1:A:118:PRO:HD2	1:A:430:VAL:O	2.15	0.47
1:B:87:ASN:ND2	1:B:370:SER:OG	2.41	0.47
1:A:352:PHE:HE2	1:A:354:ILE:HD11	1.81	0.46
1:B:283:MET:CG	1:B:285:LYS:H	2.26	0.46
1:A:182:PHE:CD2	1:A:185:TRP:HZ3	2.34	0.46
1:A:458:LEU:HD21	1:A:461:ILE:HD11	1.98	0.46
1:C:458:LEU:HD13	1:C:509:LEU:HD21	1.97	0.46
1:C:42:MET:HG3	1:C:296:ILE:HD13	1.97	0.46
1:B:43:ARG:HH21	1:B:64:ASP:CG	2.19	0.46
1:C:94:HIS:CD2	1:C:326:ASP:HA	2.50	0.46
1:C:352:PHE:HE1	1:C:354:ILE:HD11	1.81	0.46
1:A:389:PRO:HB2	1:A:392:VAL:HG23	1.98	0.46
1:B:212:TRP:NE1	1:B:256:ASP:O	2.47	0.46
1:C:201:ASP:HB2	1:C:205:LYS:H	1.81	0.46
1:A:497:LYS:HE3	1:A:504:TYR:CE2	2.51	0.45
1:B:337:GLN:OE1	1:B:337:GLN:N	2.38	0.45
1:A:478:TRP:CE2	1:A:485:LYS:HE3	2.51	0.45
1:B:89:PRO:O	1:B:90:LEU:HD23	2.16	0.45
1:C:177:GLU:H	1:C:177:GLU:CD	2.17	0.45
1:C:233:ASP:OD1	1:C:235:SER:OG	2.22	0.45
1:C:50:TRP:HB3	1:C:59:LEU:HD22	1.98	0.45
1:C:133:PHE:O	1:C:138:TYR:HB2	2.16	0.45
1:C:447:LEU:HA	1:C:462:LEU:O	2.16	0.45
1:C:149:ASP:HB2	1:C:169:VAL:HG11	1.98	0.45
1:B:36:TYR:HB3	1:B:241:MET:HG3	1.98	0.45
1:B:233:ASP:HB3	1:B:236:LYS:HG3	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:68:THR:CG2	1:C:71:LEU:HB2	2.46	0.45
1:B:37:ILE:HG23	1:B:242:ILE:HB	1.99	0.45
1:C:87:ASN:OD1	1:C:370:SER:HB3	2.17	0.44
1:A:489:LEU:HD23	1:A:489:LEU:HA	1.85	0.44
1:B:93:PRO:HA	1:B:109:VAL:HG13	1.99	0.44
1:B:417:ILE:HG23	1:B:438:ARG:CZ	2.47	0.44
1:C:444:ARG:CZ	1:C:467:LEU:HD23	2.47	0.44
1:A:233:ASP:OD1	1:A:235:SER:OG	2.28	0.44
1:B:89:PRO:HD2	1:B:341:ASN:OD1	2.17	0.44
1:A:47:MET:HB2	1:A:47:MET:HE2	1.94	0.44
1:A:497:LYS:HE3	1:A:497:LYS:HB2	1.73	0.44
1:C:483:GLN:OE1	1:C:483:GLN:N	2.46	0.44
1:B:116:ARG:C	1:B:118:PRO:HD3	2.39	0.44
1:C:219:ASP:OD1	1:C:305:ARG:NH2	2.41	0.44
1:C:326:ASP:OD1	1:C:327:HIS:N	2.50	0.44
1:A:228:GLU:HB3	1:A:229:PHE:CD1	2.53	0.43
1:C:306:LEU:HD23	1:C:306:LEU:HA	1.78	0.43
1:C:478:TRP:CD1	1:C:485:LYS:HB2	2.53	0.43
1:A:435:PRO:HB2	1:A:503:TRP:CD1	2.54	0.43
1:B:413:ALA:HB1	1:B:440:ILE:HD11	2.00	0.43
1:B:209:ILE:HD12	1:B:216:HIS:NE2	2.34	0.43
1:B:94:HIS:CE1	1:B:326:ASP:HA	2.54	0.43
1:B:138:TYR:CE2	1:B:237:PRO:HB2	2.54	0.43
1:C:200:TRP:CZ3	1:C:206:ARG:HB2	2.53	0.43
1:C:372:PRO:O	1:C:394:GLY:HA3	2.19	0.43
1:B:454:GLU:HG3	1:B:454:GLU:H	1.67	0.43
1:B:40:ASP:OD1	1:B:40:ASP:N	2.50	0.43
1:B:269:THR:HG22	1:B:272:GLU:HG3	2.01	0.43
1:C:199:TYR:HH	1:C:213:SER:HG	1.62	0.43
1:B:276:ARG:NE	1:B:346:GLU:OE2	2.41	0.43
1:C:316:SER:O	1:C:357:PRO:HG2	2.18	0.43
1:B:110:PRO:HG3	1:B:433:TYR:CD2	2.53	0.42
1:B:254:PHE:HE2	1:B:287:SER:HB2	1.84	0.42
1:C:34:ILE:HD12	1:C:133:PHE:CZ	2.54	0.42
1:A:39:PRO:CG	1:A:299:VAL:HG11	2.49	0.42
1:B:209:ILE:HD12	1:B:216:HIS:CE1	2.54	0.42
1:C:487:GLN:O	1:C:491:GLN:HG3	2.19	0.42
1:A:200:TRP:HA	1:A:205:LYS:O	2.19	0.42
1:A:233:ASP:HB3	1:A:236:LYS:HD3	2.01	0.42
1:B:109:VAL:HG13	1:B:109:VAL:O	2.20	0.42
1:C:99:LEU:HD21	1:C:241:MET:HE1	2.00	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:34:ILE:HD11	1:B:239:PHE:CD1	2.54	0.42
1:B:185:TRP:NE1	1:B:187:SER:HB3	2.35	0.42
1:A:105:HIS:CD2	1:A:106:ARG:HG2	2.55	0.42
1:B:249:HIS:HE1	1:B:256:ASP:CG	2.23	0.42
1:B:313:LEU:HD12	1:B:313:LEU:HA	1.80	0.42
1:A:43:ARG:HH21	1:A:64:ASP:CG	2.21	0.42
1:A:114:ASN:OD1	1:A:115:SER:N	2.53	0.42
1:B:92:SER:HB2	1:B:93:PRO:HD3	2.02	0.42
1:B:109:VAL:HG22	1:B:111:LEU:O	2.20	0.42
1:A:268:ARG:HE	1:A:272:GLU:HB3	1.85	0.42
1:B:359:ARG:HD3	1:B:402:PHE:CE1	2.49	0.42
1:B:39:PRO:HD2	1:B:324:SER:O	2.19	0.42
1:B:366:ASP:OD1	1:B:366:ASP:N	2.52	0.42
1:C:417:ILE:HG23	1:C:438:ARG:CZ	2.50	0.42
1:A:248:HIS:CE1	1:A:295:GLN:NE2	2.88	0.41
1:A:326:ASP:OD1	1:A:326:ASP:N	2.50	0.41
1:A:343:PRO:HG3	1:A:416:TYR:CE2	2.54	0.41
1:B:438:ARG:HB2	1:B:503:TRP:CZ3	2.55	0.41
1:A:372:PRO:O	1:A:394:GLY:HA3	2.20	0.41
1:B:95:ARG:NH2	2:B:601:SO4:O4	2.53	0.41
1:B:460:GLU:OE1	1:B:462:LEU:HD11	2.20	0.41
1:A:324:SER:OG	1:A:325:SER:N	2.53	0.41
1:A:438:ARG:NE	1:A:501:ASP:OD2	2.48	0.41
1:B:249:HIS:HE1	1:B:256:ASP:OD2	2.04	0.41
1:B:344:TYR:HD1	1:B:472:GLN:NE2	2.19	0.41
1:C:359:ARG:HH22	1:C:383:ASN:ND2	2.18	0.41
1:C:440:ILE:HD13	1:C:440:ILE:HA	1.84	0.41
1:A:421:ASP:CB	1:A:430:VAL:HG11	2.49	0.41
1:B:140:CYS:HB3	1:B:182:PHE:CD1	2.56	0.41
1:A:208:ASP:OD1	1:A:208:ASP:N	2.54	0.41
1:A:338:ASP:O	1:A:340:LYS:N	2.54	0.41
1:A:440:ILE:HD11	1:A:488:LEU:HD22	2.03	0.41
1:B:225:LEU:HD13	1:B:310:LEU:HD13	2.03	0.41
1:B:246:PRO:HA	1:B:247:PRO:HD3	1.97	0.41
1:B:490:ILE:O	1:B:494:GLN:HG3	2.20	0.41
1:C:43:ARG:HD3	1:C:329:GLU:O	2.21	0.41
1:A:419:ASN:HB3	1:A:433:TYR:HD1	1.85	0.41
1:B:182:PHE:CD2	1:B:185:TRP:HZ3	2.38	0.41
1:B:119:PHE:HE2	1:B:433:TYR:HD2	1.68	0.40
1:C:46:ALA:HA	1:C:64:ASP:HB2	2.03	0.40
1:A:330:THR:HB	1:A:333: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:91:SER:HB3	1:C:95:ARG:HH22	1.86	0.40
1:C:483:GLN:HG2	1:C:484:LEU:N	2.37	0.40
1:B:263:THR:OG1	1:B:266:LYS:HE2	2.21	0.40
1:C:310:LEU:HD12	1:C:310:LEU:HA	1.89	0.40
1:B:92:SER:HB3	1:B:113:VAL:HG23	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	434/516 (84%)	412 (95%)	18 (4%)	4 (1%)	17	43
1	B	416/516 (81%)	397 (95%)	17 (4%)	2 (0%)	29	57
1	C	455/516 (88%)	439 (96%)	15 (3%)	1 (0%)	47	75
All	All	1305/1548 (84%)	1248 (96%)	50 (4%)	7 (0%)	29	57

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252	ASN
1	A	339	ALA
1	A	456	LYS
1	B	253	SER
1	C	429	LYS
1	B	249	HIS
1	A	406	PRO

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	394/465 (85%)	372 (94%)	22 (6%)	21 47
1	B	380/465 (82%)	346 (91%)	34 (9%)	9 26
1	C	414/465 (89%)	394 (95%)	20 (5%)	25 55
All	All	1188/1395 (85%)	1112 (94%)	76 (6%)	17 41

All (76) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66	VAL
1	A	73	ARG
1	A	78	SER
1	A	83	SER
1	A	115	SER
1	A	124	ASN
1	A	134	SER
1	A	183	ASN
1	A	208	ASP
1	A	253	SER
1	A	256	ASP
1	A	257	CYS
1	A	277	HIS
1	A	300	ASP
1	A	316	SER
1	A	320	MET
1	A	356	TYR
1	A	407	ASP
1	A	421	ASP
1	A	442	THR
1	A	444	ARG
1	A	478	TRP
1	B	40	ASP
1	B	57	SER
1	B	83	SER
1	B	87	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	91	SER
1	B	106	ARG
1	B	117	ARG
1	B	121	THR
1	B	134	SER
1	B	188	TYR
1	B	198	HIS
1	B	200	TRP
1	B	201	ASP
1	B	213	SER
1	B	215	SER
1	B	256	ASP
1	B	286	SER
1	B	288	SER
1	B	308	GLU
1	B	311	ASP
1	B	313	LEU
1	B	350	VAL
1	B	356	TYR
1	B	363	LYS
1	B	366	ASP
1	B	370	SER
1	B	381	LEU
1	B	386	GLN
1	B	420	MET
1	B	430	VAL
1	B	453	LYS
1	B	478	TRP
1	B	481	ARG
1	B	487	GLN
1	C	178	ARG
1	C	201	ASP
1	C	213	SER
1	C	215	SER
1	C	241	MET
1	C	244	MET
1	C	284	GLU
1	C	286	SER
1	C	300	ASP
1	C	340	LYS
1	C	356	TYR
1	C	359	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	361	LYS
1	C	399	LYS
1	C	423	ARG
1	C	427	ASP
1	C	431	ARG
1	C	453	LYS
1	C	478	TRP
1	C	497	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	41	GLN
1	A	248	HIS
1	A	295	GLN
1	A	327	HIS
1	A	487	GLN
1	A	494	GLN
1	B	207	HIS
1	B	390	HIS
1	B	443	HIS
1	B	472	GLN
1	C	245	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

1 ligand is 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	SO4	B	601	-	4,4,4	0.13	0	6,6,6	0.07	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

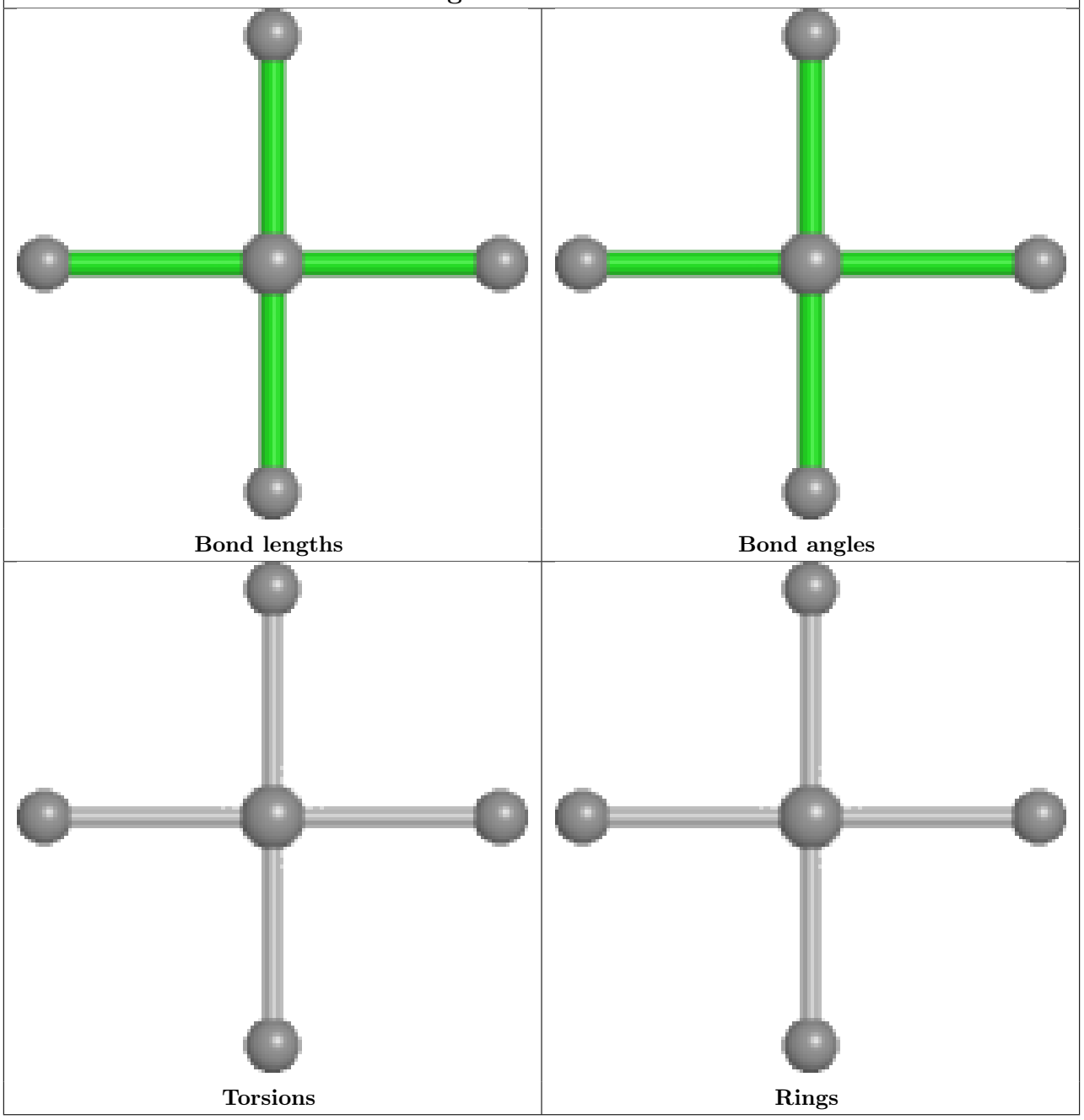
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	601	SO4	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand SO4 B 601



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	442/516 (85%)	0.37	31 (7%)	16	12	47, 67, 106, 142	0
1	B	426/516 (82%)	0.65	37 (8%)	10	7	54, 80, 110, 135	0
1	C	463/516 (89%)	0.12	16 (3%)	44	38	42, 61, 89, 144	0
All	All	1331/1548 (85%)	0.38	84 (6%)	20	15	42, 69, 107, 144	0

All (84) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	427	ASP	6.2
1	A	338	ASP	5.8
1	B	453	LYS	5.7
1	B	116	ARG	5.3
1	B	503	TRP	4.7
1	A	484	LEU	4.4
1	A	406	PRO	4.3
1	B	284	GLU	4.3
1	C	428	GLY	4.3
1	C	426	GLN	4.2
1	B	454	GLU	4.2
1	A	407	ASP	4.0
1	A	249	HIS	4.0
1	B	421	ASP	4.0
1	B	251	TYR	3.7
1	A	429	LYS	3.7
1	A	430	VAL	3.7
1	C	337	GLN	3.7
1	B	283	MET	3.6
1	B	457	GLN	3.6
1	B	486	ARG	3.6
1	A	116	ARG	3.5
1	B	358	GLN	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	328	GLY	3.3
1	B	123	ARG	3.3
1	A	251	TYR	3.3
1	C	250	PRO	3.3
1	B	436	VAL	3.3
1	B	285	LYS	3.2
1	B	402	PHE	3.2
1	B	382	SER	3.1
1	A	151	PRO	3.1
1	A	424	GLN	3.1
1	A	248	HIS	3.0
1	A	337	GLN	3.0
1	B	361	LYS	3.0
1	B	252	ASN	3.0
1	C	255	ASN	2.9
1	B	431	ARG	2.9
1	C	280	ASP	2.9
1	C	424	GLN	2.9
1	A	452	ASP	2.7
1	B	430	VAL	2.7
1	A	508	ILE	2.7
1	B	444	ARG	2.7
1	C	358	GLN	2.7
1	A	478	TRP	2.7
1	B	456	LYS	2.6
1	C	267	ASP	2.6
1	B	357	PRO	2.6
1	A	422	GLY	2.5
1	A	423	ARG	2.5
1	C	249	HIS	2.4
1	B	124	ASN	2.4
1	A	456	LYS	2.4
1	B	432	THR	2.4
1	A	204	GLY	2.4
1	B	338	ASP	2.3
1	B	76	ARG	2.3
1	C	251	TYR	2.3
1	B	501	ASP	2.2
1	B	360	LEU	2.2
1	A	188	TYR	2.2
1	B	327	HIS	2.2
1	C	254	PHE	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	150	THR	2.1
1	A	119	PHE	2.1
1	A	254	PHE	2.1
1	A	284	GLU	2.1
1	B	253	SER	2.1
1	A	201	ASP	2.1
1	A	205	LYS	2.1
1	C	423	ARG	2.1
1	A	286	SER	2.1
1	A	267	ASP	2.1
1	B	461	ILE	2.1
1	B	447	LEU	2.1
1	B	137	GLY	2.1
1	C	274	LEU	2.0
1	C	284	GLU	2.0
1	B	136	ASN	2.0
1	A	421	ASP	2.0
1	B	468	ASP	2.0
1	A	488	LEU	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 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.

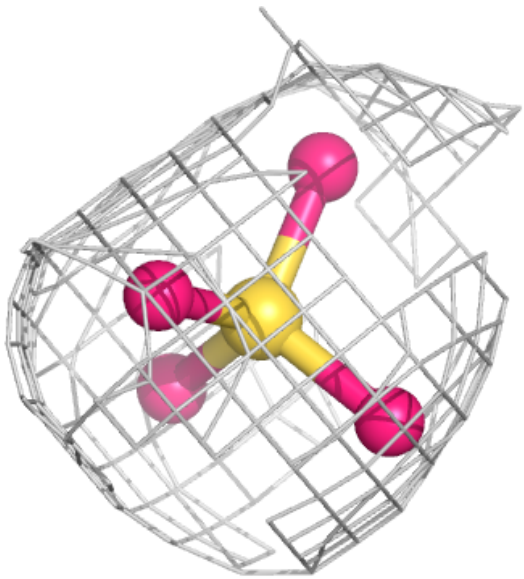
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	B	601	5/5	0.94	0.12	84,88,92,102	0

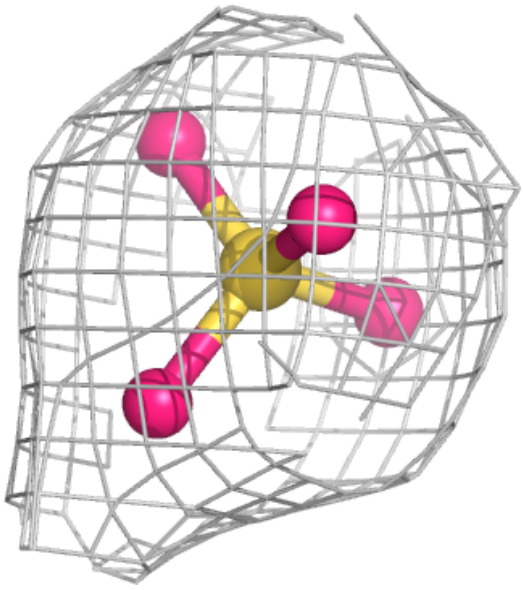
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

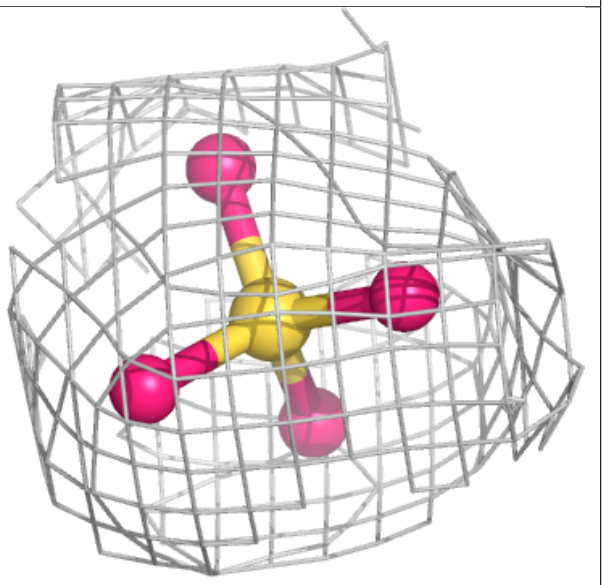
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around SO4 B 601:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.