Full wwPDB X-ray Structure Validation Report (i)

Jan 7, 2024 – 07:09 am GMT

PDB ID	:	5NA8
Title	:	Structure of DPP III from Bacteroides thetaiotaomicron in closed form
Authors	:	Sabljic, I.; Luic, M.
Deposited on	:	2017-02-27
Resolution	:	3.29 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:

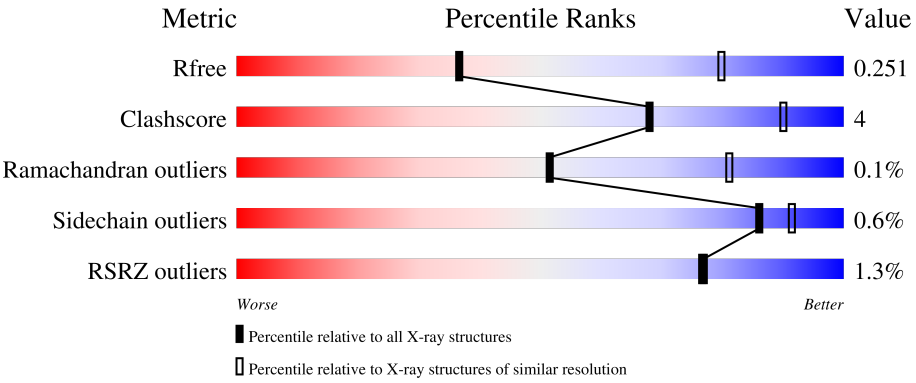
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.29 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	683	2% 85% 10% 5%
1	B	683	85% 10% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10220 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 Putative dipeptidyl-peptidase III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	651	Total	C	N	O	Se	0	0	0
			5100	3257	842	991	10			
1	B	649	Total	C	N	O	Se	0	0	0
			5098	3262	844	982	10			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	11	SER	CYS	engineered mutation	UNP Q8A6N1
A	158	SER	CYS	engineered mutation	UNP Q8A6N1
A	189	SER	CYS	engineered mutation	UNP Q8A6N1
A	425	SER	CYS	engineered mutation	UNP Q8A6N1
A	450	SER	CYS	engineered mutation	UNP Q8A6N1
A	676	LEU	-	expression tag	UNP Q8A6N1
A	677	GLU	-	expression tag	UNP Q8A6N1
A	678	HIS	-	expression tag	UNP Q8A6N1
A	679	HIS	-	expression tag	UNP Q8A6N1
A	680	HIS	-	expression tag	UNP Q8A6N1
A	681	HIS	-	expression tag	UNP Q8A6N1
A	682	HIS	-	expression tag	UNP Q8A6N1
A	683	HIS	-	expression tag	UNP Q8A6N1
B	11	SER	CYS	engineered mutation	UNP Q8A6N1
B	158	SER	CYS	engineered mutation	UNP Q8A6N1
B	189	SER	CYS	engineered mutation	UNP Q8A6N1
B	425	SER	CYS	engineered mutation	UNP Q8A6N1
B	450	SER	CYS	engineered mutation	UNP Q8A6N1
B	676	LEU	-	expression tag	UNP Q8A6N1
B	677	GLU	-	expression tag	UNP Q8A6N1
B	678	HIS	-	expression tag	UNP Q8A6N1
B	679	HIS	-	expression tag	UNP Q8A6N1
B	680	HIS	-	expression tag	UNP Q8A6N1
B	681	HIS	-	expression tag	UNP Q8A6N1
B	682	HIS	-	expression tag	UNP Q8A6N1

Continued on next page...

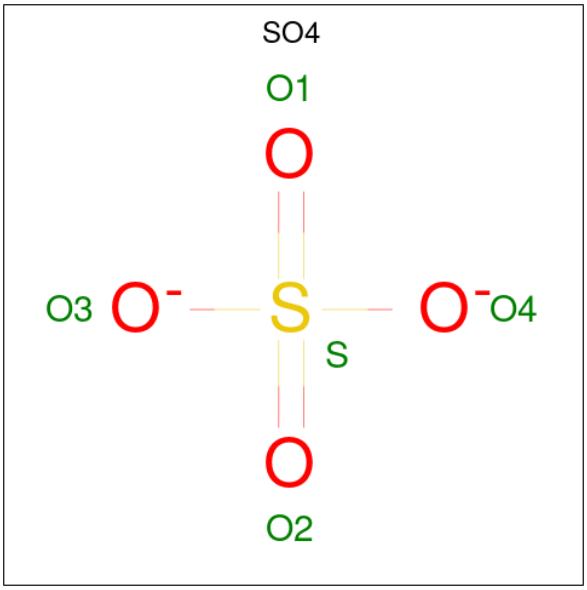
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	683	HIS	-	expression tag	UNP Q8A6N1

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Zn 1 1	0	0
2	B	1	Total Zn 1 1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



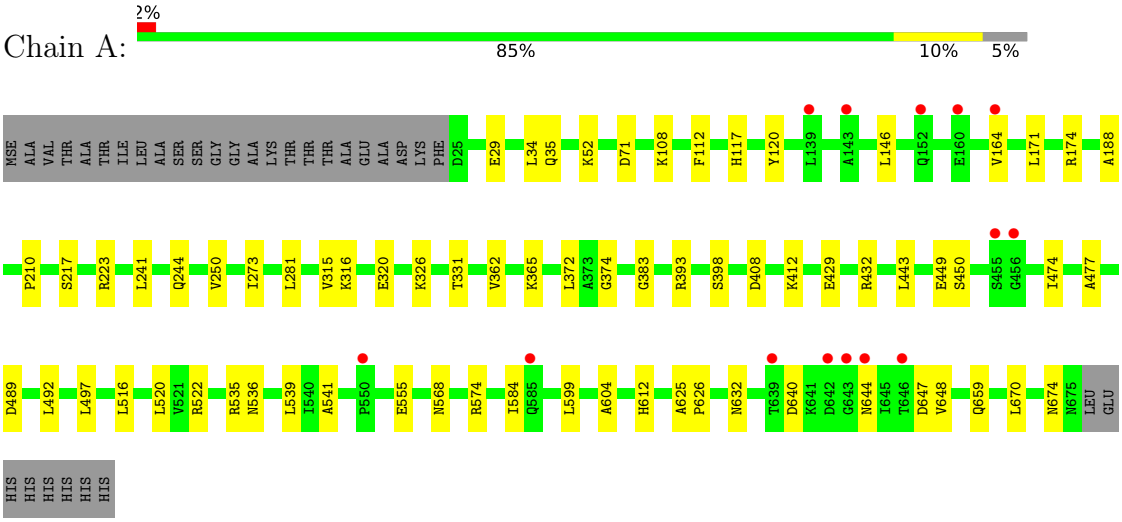
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0

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.

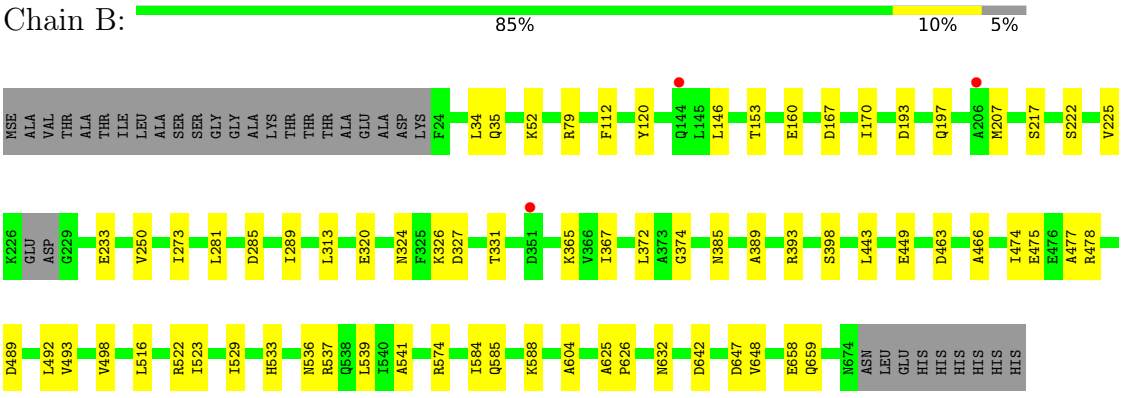
• Molecule 1: Putative dipeptidyl-peptidase III

Chain A:



• Molecule 1: Putative dipeptidyl-peptidase III

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	123.63Å 176.31Å 74.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.49 – 3.29 47.49 – 3.27	Depositor EDS
% Data completeness (in resolution range)	99.0 (47.49-3.29) 97.4 (47.49-3.27)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.06 (at 3.25Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.207 , 0.251 0.207 , 0.251	Depositor DCC
R_{free} test set	1245 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	47.0	Xtriage
Anisotropy	0.445	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 31.6	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	0.90	EDS
Total number of atoms	10220	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.91% 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, ZN

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	0.25	1/5204 (0.0%)	0.39	0/7055
1	B	0.26	1/5202 (0.0%)	0.40	0/7050
All	All	0.26	2/10406 (0.0%)	0.39	0/14105

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	146	LEU	C-N	8.55	1.50	1.34
1	A	146	LEU	C-N	5.79	1.45	1.34

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	5100	0	4875	38	0
1	B	5098	0	4900	38	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	5	0	0	0	0
3	B	15	0	0	1	0
All	All	10220	0	9775	76	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 4.

All (76) 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:B:477:ALA:HB2	1:B:536:ASN:HB2	1.72	0.70
1:A:408:ASP:OD1	1:A:412:LYS:NZ	2.31	0.63
1:A:522:ARG:HG2	1:A:626:PRO:HB2	1.81	0.63
1:A:429:GLU:OE1	1:A:432:ARG:NH1	2.31	0.62
1:A:539:LEU:HB2	1:A:604:ALA:HB2	1.82	0.62
1:A:315:VAL:HG13	1:A:316:LYS:HD2	1.81	0.61
1:A:35:GLN:HB2	1:A:625:ALA:HB3	1.82	0.61
1:B:516:LEU:HG	1:B:541:ALA:HB1	1.83	0.60
1:B:217:SER:O	1:B:281:LEU:HD13	2.02	0.60
1:B:372:LEU:O	1:B:632:ASN:ND2	2.34	0.59
1:A:632:ASN:O	1:A:659:GLN:NE2	2.38	0.57
1:A:244:GLN:HB2	1:A:670:LEU:HD23	1.87	0.56
1:A:516:LEU:HG	1:A:541:ALA:HB1	1.87	0.56
1:B:79:ARG:NH1	1:B:160:GLU:OE2	2.39	0.55
1:A:326:LYS:HE2	1:A:365:LYS:HD2	1.88	0.54
1:B:167:ASP:HB3	1:B:170:ILE:HG22	1.91	0.53
1:B:489:ASP:OD2	1:B:574:ARG:NH2	2.42	0.53
1:A:477:ALA:HB2	1:A:536:ASN:HB2	1.91	0.53
1:B:522:ARG:NH1	3:B:703:SO4:O2	2.40	0.53
1:A:372:LEU:O	1:A:632:ASN:ND2	2.38	0.52
1:A:492:LEU:HG	1:A:497:LEU:HD12	1.91	0.52
1:B:539:LEU:HB2	1:B:604:ALA:HB2	1.92	0.51
1:B:385:ASN:ND2	1:B:449:GLU:O	2.42	0.51
1:B:250:VAL:HG13	1:B:273:ILE:HG23	1.93	0.50
1:B:523:ILE:HD11	1:B:529:ILE:HD13	1.92	0.50
1:B:474:ILE:HG23	1:B:584:ILE:HG23	1.94	0.50
1:B:533:HIS:O	1:B:537:ARG:HG3	2.12	0.49
1:B:320:GLU:HA	1:B:374:GLY:HA3	1.94	0.49
1:B:443:LEU:HD21	1:B:492:LEU:HD23	1.95	0.49
1:A:217:SER:O	1:A:281:LEU:HD13	2.14	0.48
1:B:493:VAL:HG12	1:B:498:VAL:HG23	1.96	0.48
1:A:362:VAL:HG21	1:A:450:SER:O	2.14	0.48
1:A:647:ASP:OD1	1:A:648:VAL:N	2.47	0.48
1:A:331:THR:HG23	1:A:365:LYS:HA	1.96	0.47
1:A:640:ASP:OD1	1:A:644:ASN:N	2.47	0.47
1:A:52:LYS:HB3	1:A:648:VAL:HB	1.96	0.47
1:B:475:GLU:OE1	1:B:588:LYS:NZ	2.47	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:71:ASP:OD2	1:A:674:ASN:ND2	2.47	0.46
1:A:443:LEU:HD21	1:A:492:LEU:HD23	1.98	0.46
1:A:29:GLU:OE2	1:A:108:LYS:NZ	2.32	0.46
1:A:164:VAL:HG13	1:A:171:LEU:HD12	1.98	0.46
1:A:489:ASP:OD2	1:A:574:ARG:NH2	2.49	0.46
1:B:285:ASP:O	1:B:289:ILE:HG12	2.16	0.45
1:A:117:HIS:CG	1:A:174:ARG:HB2	2.51	0.45
1:B:34:LEU:HD13	1:B:112:PHE:CG	2.52	0.45
1:B:331:THR:HG23	1:B:365:LYS:HA	1.97	0.45
1:B:647:ASP:OD1	1:B:648:VAL:N	2.49	0.45
1:B:477:ALA:N	1:B:536:ASN:HD22	2.14	0.45
1:A:383:GLY:HA3	1:A:449:GLU:HG3	2.00	0.44
1:B:324:ASN:HB2	1:B:367:ILE:HB	1.99	0.44
1:B:52:LYS:HB3	1:B:648:VAL:HB	1.98	0.44
1:B:327:ASP:O	1:B:331:THR:OG1	2.29	0.44
1:B:632:ASN:O	1:B:659:GLN:NE2	2.51	0.44
1:A:535:ARG:HG2	1:A:604:ALA:O	2.17	0.43
1:A:393:ARG:HD2	1:A:398:SER:HB3	2.01	0.43
1:A:474:ILE:HG23	1:A:584:ILE:HG23	2.00	0.43
1:A:250:VAL:HG13	1:A:273:ILE:HG23	2.00	0.43
1:A:326:LYS:HE3	1:A:326:LYS:HB2	1.87	0.43
1:A:555:GLU:OE2	1:A:568:ASN:ND2	2.47	0.43
1:A:34:LEU:HD13	1:A:112:PHE:CG	2.54	0.43
1:A:320:GLU:HA	1:A:374:GLY:HA3	2.01	0.43
1:A:539:LEU:HD21	1:A:599:LEU:HG	2.00	0.43
1:B:393:ARG:HE	1:B:398:SER:HB3	1.83	0.43
1:B:536:ASN:OD1	1:B:537:ARG:N	2.51	0.42
1:A:174:ARG:HA	1:A:188:ALA:HB3	2.00	0.42
1:B:522:ARG:HG2	1:B:626:PRO:HB2	2.01	0.42
1:B:35:GLN:HB2	1:B:625:ALA:HB3	2.02	0.41
1:B:193:ASP:HB3	1:B:225:VAL:HG12	2.02	0.41
1:B:326:LYS:HE2	1:B:365:LYS:HD2	2.03	0.41
1:B:463:ASP:HB3	1:B:466:ALA:HB2	2.03	0.41
1:A:520:LEU:HB3	1:A:612:HIS:CD2	2.56	0.40
1:B:222:SER:HB2	1:B:233:GLU:HB3	2.03	0.40
1:B:478:ARG:HD2	1:B:585:GLN:OE1	2.21	0.40
1:B:197:GLN:HG3	1:B:313:LEU:HD21	2.03	0.40
1:A:223:ARG:CZ	1:A:241:LEU:HD21	2.52	0.40
1:B:389:ALA:O	1:B:393:ARG:HG3	2.21	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	649/683 (95%)	631 (97%)	17 (3%)	1 (0%)	47 77
1	B	645/683 (94%)	631 (98%)	14 (2%)	0	100 100
All	All	1294/1366 (95%)	1262 (98%)	31 (2%)	1 (0%)	51 81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	210	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	528/569 (93%)	527 (100%)	1 (0%)	93 97
1	B	529/569 (93%)	524 (99%)	5 (1%)	78 87
All	All	1057/1138 (93%)	1051 (99%)	6 (1%)	86 91

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	120	TYR
1	B	120	TYR
1	B	153	THR
1	B	207	MSE
1	B	642	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	658	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
3	SO4	B	703	-	4,4,4	0.14	0	6,6,6	0.04	0
3	SO4	B	704	-	4,4,4	0.15	0	6,6,6	0.05	0
3	SO4	A	702	1	4,4,4	0.17	0	6,6,6	0.07	0
3	SO4	B	702	-	4,4,4	0.15	0	6,6,6	0.04	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
3	B	703	SO4	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

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	641/683 (93%)	-0.01	14 (2%) 62 60	22, 41, 66, 83	0
1	B	639/683 (93%)	-0.09	3 (0%) 91 91	19, 37, 62, 83	0
All	All	1280/1366 (93%)	-0.05	17 (1%) 77 77	19, 39, 65, 83	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	143	ALA	3.4
1	A	644	ASN	2.9
1	A	642	ASP	2.6
1	A	160	GLU	2.5
1	A	639	THR	2.5
1	A	643	GLY	2.4
1	B	206	ALA	2.4
1	A	456	GLY	2.3
1	A	164	VAL	2.3
1	A	550	PRO	2.3
1	A	455	SER	2.2
1	A	139	LEU	2.2
1	B	144	GLN	2.1
1	B	351	ASP	2.1
1	A	646	THR	2.1
1	A	585	GLN	2.0
1	A	152	GLN	2.0

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

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
3	SO4	B	704	5/5	0.74	0.28	47,57,78,90	0
3	SO4	A	702	5/5	0.92	0.16	29,29,45,56	0
3	SO4	B	703	5/5	0.95	0.27	50,61,72,74	0
2	ZN	A	701	1/1	0.95	0.18	47,47,47,47	0
3	SO4	B	702	5/5	0.97	0.12	22,30,34,47	0
2	ZN	B	701	1/1	0.98	0.22	42,42,42,42	0

6.5 Other polymers (i)

There are no such residues in this entry.