W O R L D W I D E
wwPDB
P R O T E I N D A T A B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 28, 2021 – 03:02 pm BST

PDB ID	:	6Z1S
Title	:	Structure of Polyphenol Oxidase (mutant G292N) from Thermothelomyces thermophila
Authors	:	Dimarogona, M.; Nikolaivits, E.; Valmas, A.; Topakas, E.
Deposited on	:	2020-05-14
Resolution	:	1.53 Å(reported)

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

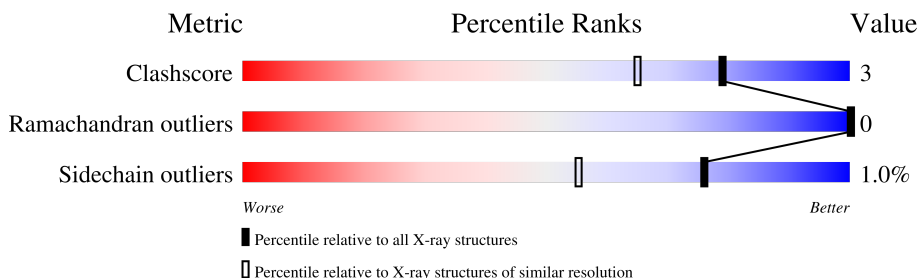|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.53 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	Percentile Ranks	Value
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		1.0%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	2634 (1.56-1.52)
Ramachandran outliers	138981	2580 (1.56-1.52)
Sidechain outliers	138945	2577 (1.56-1.52)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	425	84% 5% 10%
2	B	2	100%

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 3441 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 Tyrosinase-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	381	2980	1896	513	562	9	0	3	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLU	-	expression tag	UNP G2QLD3
A	2	PHE	-	expression tag	UNP G2QLD3
A	292	ASN	GLY	engineered mutation	UNP G2QLD3
A	405	GLU	-	expression tag	UNP G2QLD3
A	406	GLN	-	expression tag	UNP G2QLD3
A	407	LYS	-	expression tag	UNP G2QLD3
A	408	LEU	-	expression tag	UNP G2QLD3
A	409	ILE	-	expression tag	UNP G2QLD3
A	410	SER	-	expression tag	UNP G2QLD3
A	411	GLU	-	expression tag	UNP G2QLD3
A	412	GLU	-	expression tag	UNP G2QLD3
A	413	ASP	-	expression tag	UNP G2QLD3
A	414	LEU	-	expression tag	UNP G2QLD3
A	415	ASN	-	expression tag	UNP G2QLD3
A	416	SER	-	expression tag	UNP G2QLD3
A	417	ALA	-	expression tag	UNP G2QLD3
A	418	VAL	-	expression tag	UNP G2QLD3
A	419	ASP	-	expression tag	UNP G2QLD3
A	420	HIS	-	expression tag	UNP G2QLD3
A	421	HIS	-	expression tag	UNP G2QLD3
A	422	HIS	-	expression tag	UNP G2QLD3
A	423	HIS	-	expression tag	UNP G2QLD3
A	424	HIS	-	expression tag	UNP G2QLD3
A	425	HIS	-	expression tag	UNP G2QLD3

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
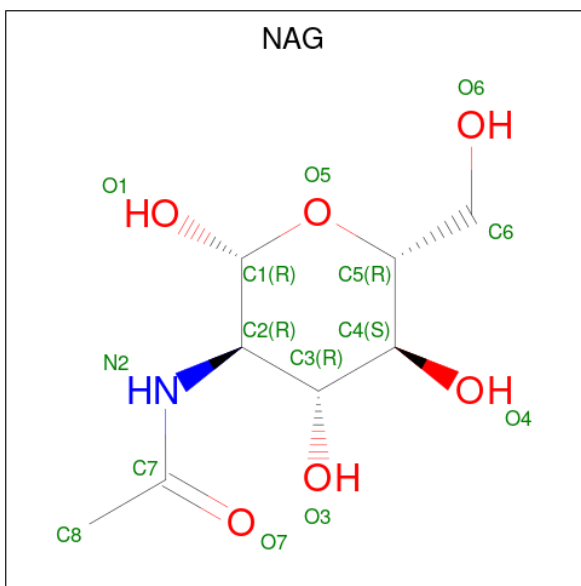
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 14	C 8	N 1	O 5	0	0
3	A	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).

MPD

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			8	6	2		

- Molecule 5 is GLYCINE (three-letter code: GLY) (formula: $C_2H_5NO_2$).

GLY
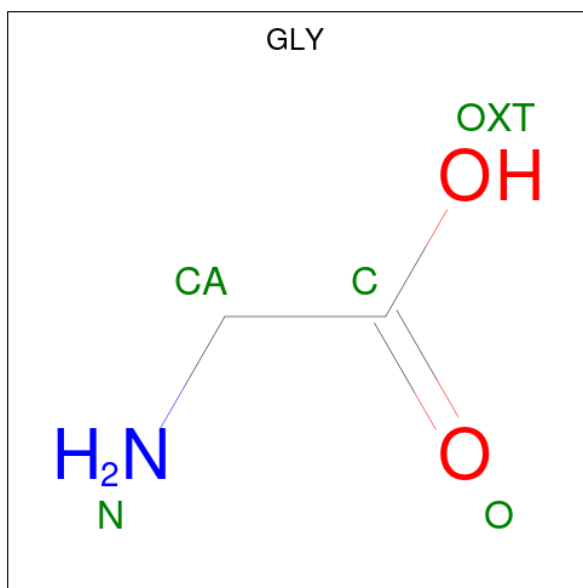
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			4	2	1	1		
5	A	1	Total	C	N	O	0	0
			5	2	1	2		
5	A	1	Total	C	N	O	0	0
			5	2	1	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 5	C 2	N 1	O 2	0	0
5	A	1	Total 5	C 2	N 1	O 2	0	0
5	A	1	Total 5	C 2	N 1	O 2	0	0

- Molecule 6 is SERINE (three-letter code: SER) (formula: $C_3H_7NO_3$).
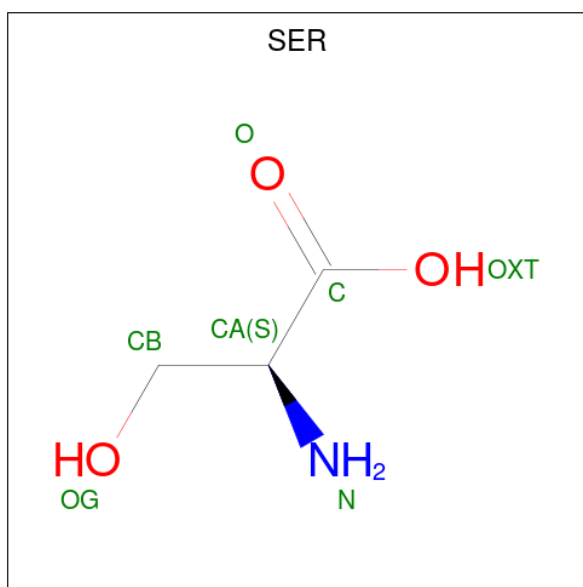
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total 7	C 3	N 1	O 3	0	0

- Molecule 7 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	2	Total 2	Cu 2	0	0

- Molecule 8 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	359	Total 359	O 359	0	0

3 Residue-property plots ⓘ

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. 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. 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.

Note EDS was not executed.

• Molecule 1: Tyrosinase-like protein

Chain A:
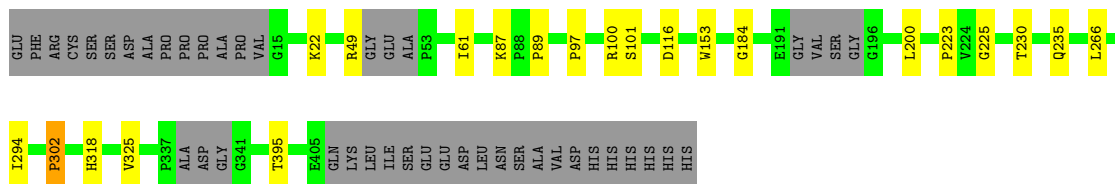
• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

4 Data and refinement statistics ⓘ

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	86.40Å 97.18Å 47.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.24 – 1.53	Depositor
% Data completeness (in resolution range)	49.7 (43.24-1.53)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.157 , 0.207	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3441	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, CU, MPD

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	0.72	0/3075	0.84	1/4195 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	302	PRO	N-CA-CB	-5.36	96.70	102.60

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	2980	0	2829	13	0
2	B	28	0	25	0	0
3	A	28	0	26	0	0
4	A	8	0	14	0	0
5	A	29	0	12	5	0
6	A	7	0	4	0	0
7	A	2	0	0	0	0
8	A	359	0	0	7	0
All	All	3441	0	2910	17	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.

The worst 5 of 17 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:116:ASP:OD2	8:A:603:HOH:O	1.99	0.81
5:A:510:GLY:OXT	8:A:604:HOH:O	2.10	0.69
1:A:395:THR:HG1	5:A:509:GLY:N	1.92	0.67
5:A:512:GLY:CA	8:A:606:HOH:O	2.56	0.54
1:A:235:GLN:NE2	8:A:609:HOH:O	2.33	0.53

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	376/425 (88%)	364 (97%)	12 (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	311/343 (91%)	308 (99%)	3 (1%)	76	55

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	22	LYS
1	A	49	ARG
1	A	302	PRO

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)

2 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	NAG	B	1	2,1	14,14,15	0.95	1 (7%)	17,19,21	1.24	1 (5%)
2	NAG	B	2	2	14,14,15	0.84	0	17,19,21	3.31	8 (47%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	B	1	2,1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	B	2	2	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	NAG	C1-C2	2.32	1.55	1.52

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2	NAG	C1-O5-C5	9.68	125.30	112.19
2	B	2	NAG	O5-C5-C6	-6.09	97.66	107.20
2	B	2	NAG	C1-C2-N2	-3.88	103.86	110.49
2	B	2	NAG	O7-C7-C8	3.13	127.87	122.06
2	B	2	NAG	C8-C7-N2	-2.93	111.13	116.10

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	2	NAG	O5-C5-C6-O6
2	B	2	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.
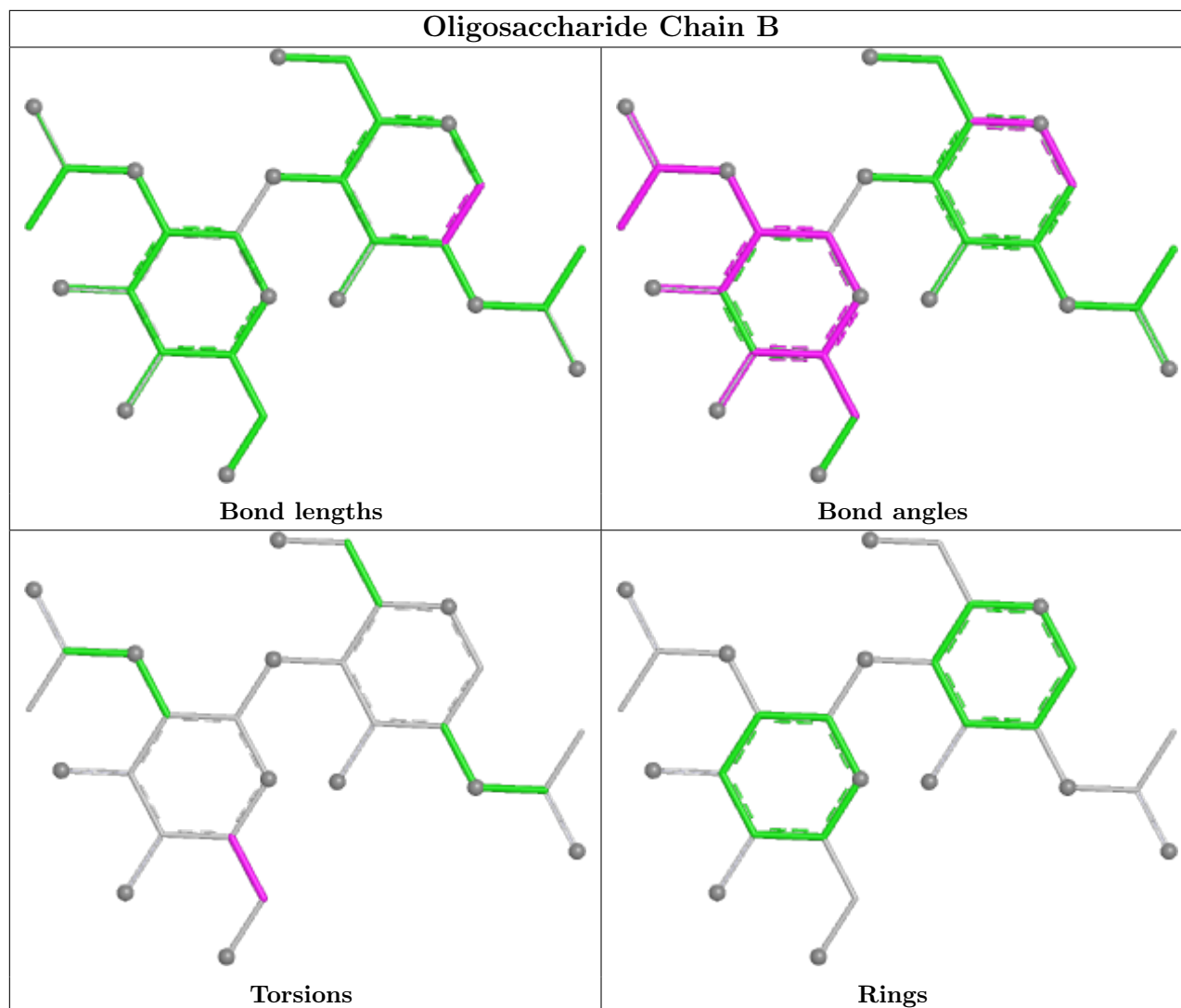
5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 2 are monoatomic - leaving 10 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	#$|Z| > 2$
5	GLY	A	506	-	3,3,4	0.69	0	0,2,4	-	-
5	GLY	A	509	-	1,4,4	0.20	0	0,4,4	-	-
5	GLY	A	512	-	1,4,4	0.11	0	0,4,4	-	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	MPD	A	504	-	7,7,7	0.18	0	9,10,10	0.77	0
3	NAG	A	501	1	14,14,15	0.59	0	17,19,21	1.16	2 (11%)
6	SER	A	507	-	3,6,6	0.79	0	1,7,7	0.76	0
5	GLY	A	510	-	1,4,4	0.16	0	0,4,4	-	-
5	GLY	A	511	-	1,4,4	0.06	0	0,4,4	-	-
3	NAG	A	505	1	14,14,15	0.46	0	17,19,21	1.60	3 (17%)
5	GLY	A	508	-	1,4,4	0.08	0	0,4,4	-	-

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	GLY	A	506	-	-	0/0/1/2	-
5	GLY	A	509	-	-	0/0/2/2	-
5	GLY	A	512	-	-	0/0/2/2	-
4	MPD	A	504	-	-	4/5/5/5	-
3	NAG	A	501	1	-	0/6/23/26	0/1/1/1
6	SER	A	507	-	-	1/2/6/6	-
5	GLY	A	510	-	-	0/0/2/2	-
5	GLY	A	511	-	-	0/0/2/2	-
3	NAG	A	505	1	-	2/6/23/26	0/1/1/1
5	GLY	A	508	-	-	0/0/2/2	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	505	NAG	O5-C5-C6	3.86	113.25	107.20
3	A	505	NAG	C4-C3-C2	-3.24	106.27	111.02
3	A	505	NAG	O5-C1-C2	-2.56	107.25	111.29
3	A	501	NAG	C3-C4-C5	-2.26	106.21	110.24
3	A	501	NAG	O5-C5-C4	-2.05	105.84	110.83

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	504	MPD	O2-C2-C3-C4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	505	NAG	O5-C5-C6-O6
3	A	505	NAG	C4-C5-C6-O6
6	A	507	SER	C-CA-CB-OG
4	A	504	MPD	C1-C2-C3-C4

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	509	GLY	1	0
5	A	512	GLY	3	0
5	A	510	GLY	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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.