wwPDB X-ray Structure Validation Summary Report (i)

Sep 28, 2021 – 04:03 am BST

PDB ID	:	6T8F
Title	:	Crystal structure of mutant xylose isomerase (V270A/A273G) from Piromyces E2 grown in yeast, in complex with xylose
Authors	:	Rozeboom, H.J.; Janssen, D.B.
Deposited on	:	2019-10-24
Resolution	:	2.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see references (i)) were used in the production of this report:

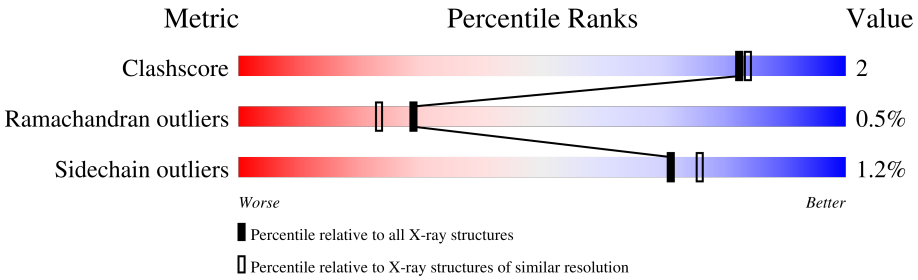
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al.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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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	437	94% 5%
1	B	437	93% 7%
1	C	437	94% 5%
1	D	437	94% 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	XLS	A	503	-	X	-	-
3	XLS	B	503	-	X	-	-
3	XLS	C	503	-	X	-	-
6	SO4	A	508	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 15847 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 Xylose isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	436	Total 3471	C 2202	N 582	O 667	S 20	0	1	0
1	B	436	Total 3471	C 2202	N 582	O 667	S 20	0	1	0
1	C	436	Total 3471	C 2202	N 582	O 667	S 20	0	1	0
1	D	436	Total 3495	C 2217	N 585	O 670	S 23	0	4	0

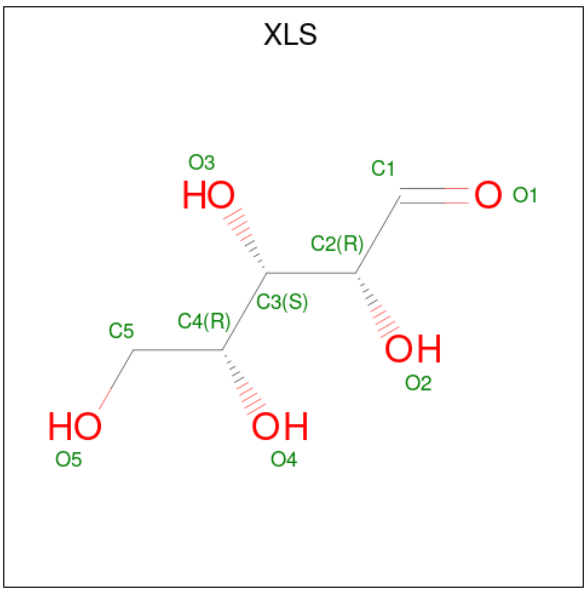
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	270	ALA	VAL	engineered mutation	UNP Q9P8C9
A	273	GLY	ALA	engineered mutation	UNP Q9P8C9
B	270	ALA	VAL	engineered mutation	UNP Q9P8C9
B	273	GLY	ALA	engineered mutation	UNP Q9P8C9
C	270	ALA	VAL	engineered mutation	UNP Q9P8C9
C	273	GLY	ALA	engineered mutation	UNP Q9P8C9
D	270	ALA	VAL	engineered mutation	UNP Q9P8C9
D	273	GLY	ALA	engineered mutation	UNP Q9P8C9

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

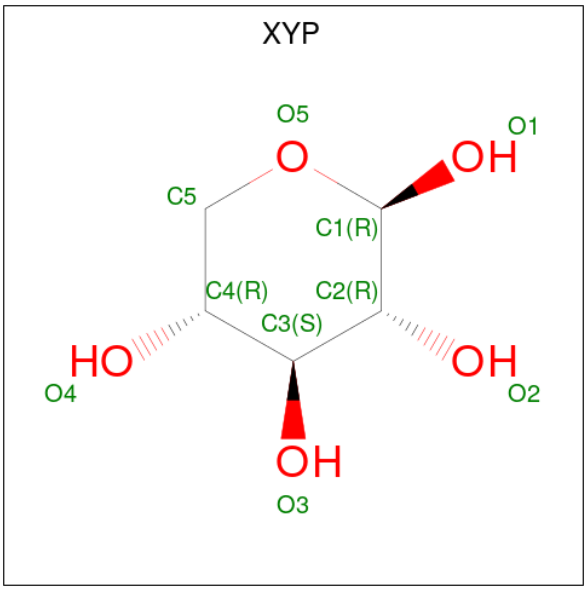
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total 2	Ca 2	0	0
2	B	2	Total 2	Ca 2	0	0
2	C	2	Total 2	Ca 2	0	0
2	D	2	Total 2	Ca 2	0	0

- Molecule 3 is D-xylose (three-letter code: XLS) (formula: $C_5H_{10}O_5$).



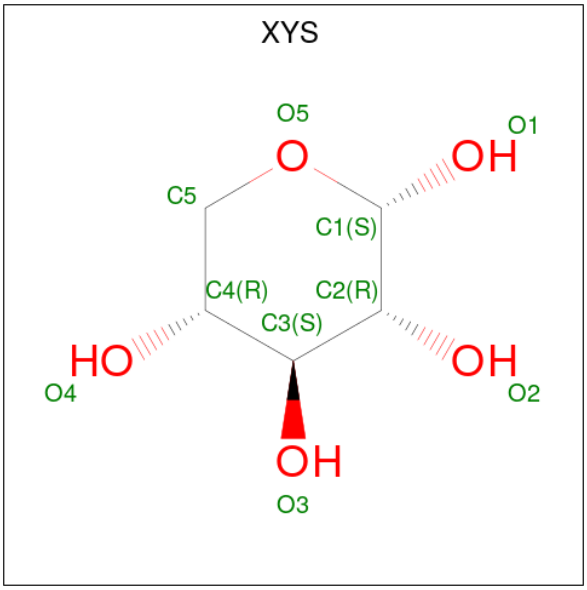
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 10 5 5	0	0
3	B	1	Total C O 10 5 5	0	0
3	C	1	Total C O 10 5 5	0	0
3	D	1	Total C O 10 5 5	0	0

- Molecule 4 is beta-D-xylopyranose (three-letter code: XYP) (formula: $C_5H_{10}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 10 5 5	0	0
4	B	1	Total C O 10 5 5	0	0
4	C	1	Total C O 10 5 5	0	0
4	C	1	Total C O 10 5 5	0	0
4	D	1	Total C O 10 5 5	0	0
4	D	1	Total C O 10 5 5	0	0

- Molecule 5 is alpha-D-xylopyranose (three-letter code: XYS) (formula: $C_5H_{10}O_5$).



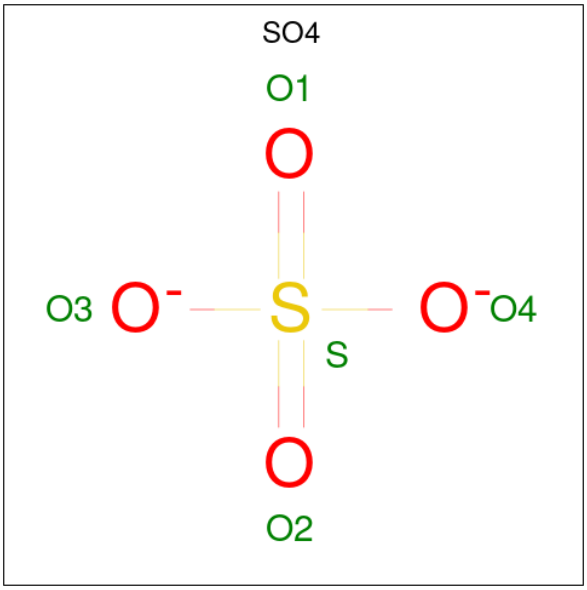
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 10 5 5	0	0
5	A	1	Total C O 10 5 5	0	0
5	B	1	Total C O 10 5 5	0	0
5	B	1	Total C O 10 5 5	0	0
5	B	1	Total C O 10 5 5	0	0
5	C	1	Total C O 10 5 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	1	Total C O 10 5 5	0	0
5	D	1	Total C O 10 5 5	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	B	1	Total O S 5 4 1	0	0
6	C	1	Total O S 5 4 1	0	0
6	D	1	Total O S 5 4 1	0	0
6	D	1	Total O S 5 4 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	415	Total O 415 415	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	433	Total O 433 433	0	0
7	C	424	Total O 424 424	0	0
7	D	449	Total O 449 449	0	0

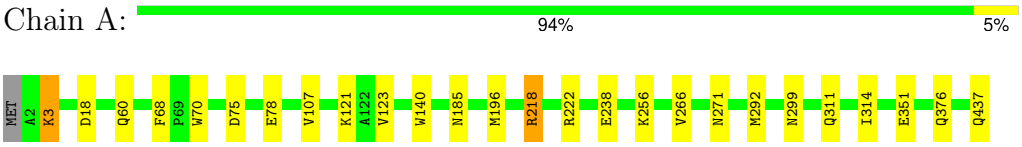
3 Residue-property plots (i)

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. 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.

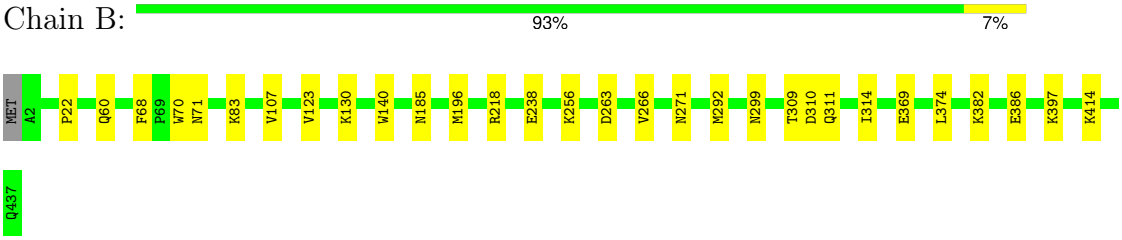
• Molecule 1: Xylose isomerase

Chain A:



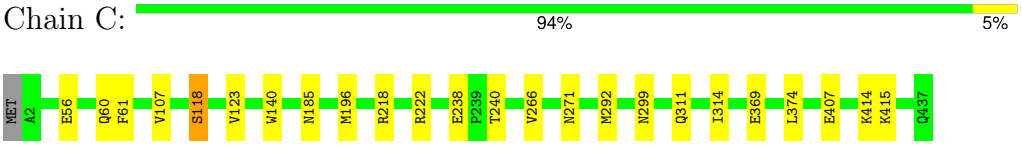
• Molecule 1: Xylose isomerase

Chain B:



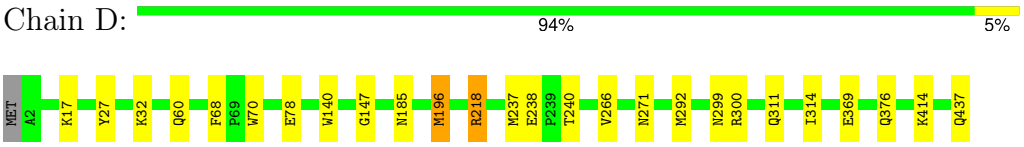
• Molecule 1: Xylose isomerase

Chain C:



• Molecule 1: Xylose isomerase

Chain D:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	78.60Å 79.37Å 91.98Å 115.46° 89.98° 117.13°	Depositor
Resolution (Å)	46.60 – 2.00	Depositor
% Data completeness (in resolution range)	95.3 (46.60-2.00)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.149 , 0.185	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	15847	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: XLS, CA, XYP, XYS, 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	0.73	0/3552	0.81	2/4780 (0.0%)
1	B	0.73	1/3552 (0.0%)	0.79	0/4780
1	C	0.73	3/3552 (0.1%)	0.81	3/4780 (0.1%)
1	D	0.73	0/3576	0.80	1/4810 (0.0%)
All	All	0.73	4/14232 (0.0%)	0.80	6/19150 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	369	GLU	CD-OE1	-7.23	1.17	1.25
1	C	369	GLU	CD-OE1	-5.52	1.19	1.25
1	C	407	GLU	CD-OE1	5.18	1.31	1.25
1	C	56	GLU	CD-OE2	5.07	1.31	1.25

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	222	ARG	NE-CZ-NH1	6.28	123.44	120.30
1	C	222	ARG	NE-CZ-NH1	6.13	123.36	120.30
1	C	415	LYS	CB-CA-C	-5.41	99.58	110.40
1	C	222	ARG	NE-CZ-NH2	-5.18	117.71	120.30
1	A	222	ARG	NE-CZ-NH2	-5.18	117.71	120.30

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	3471	0	3354	18	0
1	B	3471	0	3354	18	0
1	C	3471	0	3354	9	0
1	D	3495	0	3378	17	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	10	0	8	2	0
3	B	10	0	8	2	0
3	C	10	0	8	1	0
3	D	10	0	8	1	0
4	A	10	0	0	0	0
4	B	10	0	0	0	0
4	C	20	0	0	0	0
4	D	20	0	0	0	0
5	A	20	0	20	0	0
5	B	30	0	30	1	0
5	C	10	0	10	0	0
5	D	20	0	20	0	0
6	A	10	0	0	3	0
6	B	5	0	0	0	0
6	C	5	0	0	0	0
6	D	10	0	0	1	0
7	A	415	0	0	13	0
7	B	433	0	0	5	0
7	C	424	0	0	2	0
7	D	449	0	0	5	0
All	All	15847	0	13552	68	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 2.

The worst 5 of 68 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:437:GLN:OE1	7:A:601:HOH:O	1.71	1.06
1:B:196:MET:CE	7:B:944:HOH:O	2.14	0.96
1:A:196:MET:CE	7:A:921:HOH:O	2.20	0.89
1:B:271:ASN:HD21	1:B:309:THR:H	1.26	0.80
1:D:32:LYS:HE2	1:D:369:GLU:OE2	1.84	0.78

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	435/437 (100%)	429 (99%)	4 (1%)	2 (0%)	29	23
1	B	435/437 (100%)	428 (98%)	5 (1%)	2 (0%)	29	23
1	C	435/437 (100%)	425 (98%)	8 (2%)	2 (0%)	29	23
1	D	438/437 (100%)	429 (98%)	7 (2%)	2 (0%)	29	23
All	All	1743/1748 (100%)	1711 (98%)	24 (1%)	8 (0%)	29	23

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	238	GLU
1	B	238	GLU
1	C	238	GLU
1	D	238	GLU
1	A	314	ILE

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	357/357 (100%)	353 (99%)	4 (1%)	73	78
1	B	357/357 (100%)	353 (99%)	4 (1%)	73	78
1	C	357/357 (100%)	353 (99%)	4 (1%)	73	78
1	D	360/357 (101%)	354 (98%)	6 (2%)	60	65
All	All	1431/1428 (100%)	1413 (99%)	18 (1%)	71	74

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	196[B]	MET
1	D	414	LYS
1	D	271	ASN
1	C	118	SER
1	D	196[A]	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	437	GLN
1	B	71	ASN
1	B	271	ASN
1	D	437	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)

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 32 ligands modelled in this entry, 8 are monoatomic - leaving 24 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	$\#\|Z\| > 2$
6	SO4	B	508	-	4,4,4	0.30	0	6,6,6	0.08	0
5	XYS	B	506	-	10,10,10	1.38	1 (10%)	14,14,14	1.36	3 (21%)
3	XLS	D	503	2	8,9,9	0.89	1 (12%)	10,11,11	3.32	8 (80%)
6	SO4	D	509	-	4,4,4	0.39	0	6,6,6	0.10	0
4	XYP	D	507	-	10,10,10	0.49	0	14,14,14	1.95	4 (28%)
5	XYS	D	506	-	10,10,10	1.31	2 (20%)	14,14,14	1.07	2 (14%)
6	SO4	C	507	-	4,4,4	0.35	0	6,6,6	0.10	0
6	SO4	A	507	-	4,4,4	0.36	0	6,6,6	0.06	0
3	XLS	C	503	2	8,9,9	1.56	1 (12%)	10,11,11	3.65	9 (90%)
4	XYP	C	506	-	10,10,10	0.75	0	14,14,14	1.76	4 (28%)
6	SO4	D	508	-	4,4,4	0.29	0	6,6,6	0.13	0
6	SO4	A	508	-	4,4,4	0.37	0	6,6,6	0.06	0
4	XYP	C	504	-	10,10,10	0.85	0	14,14,14	1.30	2 (14%)
4	XYP	D	504	-	10,10,10	0.55	0	14,14,14	1.39	3 (21%)
5	XYS	C	505	-	10,10,10	1.28	1 (10%)	14,14,14	1.55	2 (14%)
5	XYS	B	507	-	10,10,10	0.90	0	14,14,14	2.74	5 (35%)
5	XYS	D	505	-	10,10,10	1.33	1 (10%)	14,14,14	1.50	3 (21%)
3	XLS	A	503	2	8,9,9	0.98	1 (12%)	10,11,11	2.87	7 (70%)
4	XYP	B	504	-	10,10,10	0.88	0	14,14,14	1.38	1 (7%)
5	XYS	A	506	-	10,10,10	0.47	0	14,14,14	2.23	3 (21%)
5	XYS	B	505	-	10,10,10	1.19	0	14,14,14	1.60	4 (28%)
4	XYP	A	504	-	10,10,10	0.85	0	14,14,14	1.58	4 (28%)
3	XLS	B	503	2	8,9,9	1.00	1 (12%)	10,11,11	3.76	9 (90%)
5	XYS	A	505	-	10,10,10	1.33	2 (20%)	14,14,14	1.24	2 (14%)

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	XLS	A	503	2	-	9/10/12/12	-
4	XYP	B	504	-	-	-	0/1/1/1
5	XYS	A	506	-	-	-	0/1/1/1
5	XYS	B	506	-	-	-	0/1/1/1
3	XLS	D	503	2	-	3/10/12/12	-
5	XYS	B	505	-	-	-	0/1/1/1
4	XYP	A	504	-	-	-	0/1/1/1
4	XYP	C	506	-	-	-	0/1/1/1
4	XYP	D	507	-	-	-	0/1/1/1
5	XYS	D	506	-	-	-	0/1/1/1
3	XLS	B	503	2	-	5/10/12/12	-
4	XYP	C	504	-	-	-	0/1/1/1
4	XYP	D	504	-	-	-	0/1/1/1
5	XYS	C	505	-	-	-	0/1/1/1
5	XYS	A	505	-	-	-	0/1/1/1
3	XLS	C	503	2	-	6/10/12/12	-
5	XYS	B	507	-	-	-	0/1/1/1
5	XYS	D	505	-	-	-	0/1/1/1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	503	XLS	O3-C3	3.20	1.50	1.43
5	D	505	XYS	C4-C3	3.19	1.57	1.52
5	B	506	XYS	O5-C1	2.92	1.47	1.43
3	A	503	XLS	C3-C2	-2.51	1.49	1.53
5	C	505	XYS	C4-C3	2.40	1.56	1.52

The worst 5 of 75 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	507	XYS	O5-C5-C4	-5.90	101.66	110.77
3	D	503	XLS	C5-C4-C3	-5.61	100.25	112.41
3	B	503	XLS	O4-C4-C3	5.55	122.58	109.10
3	C	503	XLS	O3-C3-C4	5.19	121.35	108.81
3	B	503	XLS	C5-C4-C3	-5.06	101.44	112.41

There are no chirality outliers.

5 of 23 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	503	XLS	C1-C2-C3-C4
3	B	503	XLS	O2-C2-C3-O3
3	C	503	XLS	O2-C2-C3-O3
3	D	503	XLS	C1-C2-C3-C4
3	C	503	XLS	C3-C4-C5-O5

There are no ring outliers.

7 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	503	XLS	1	0
6	D	509	SO4	1	0
3	C	503	XLS	1	0
6	A	508	SO4	3	0
5	B	507	XYS	1	0
3	A	503	XLS	2	0
3	B	503	XLS	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

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.