WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Aug 19, 2023 – 10:46 PM EDT

PDB ID	:	2HMN
Title	:	Crystal Structure of the Naphthalene 1,2-Dioxygenase F352V Mutant Bound to Anthracene.
Authors	:	Ferraro, D.J.; Okerlund, A.L.; Mowers, J.C.; Ramaswamy, S.
Deposited on	:	2006-07-11
Resolution	:	1.70 Å(reported)

The following versions of software and data (see references ⓘ) 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
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 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 1.70 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.
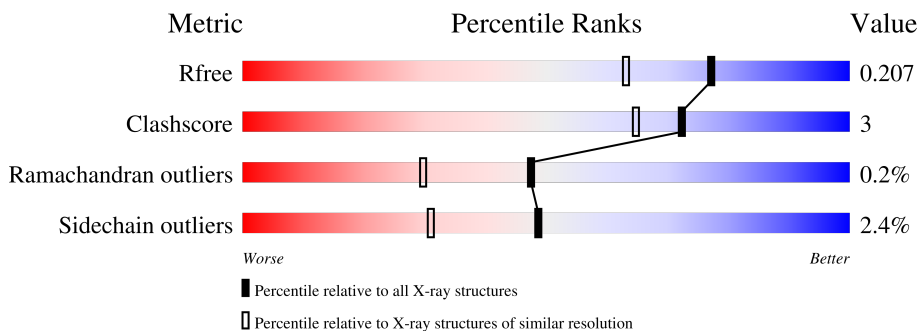
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)

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%

Mol	Chain	Length	Quality of chain
1	A	449	93% 6% ·
2	B	194	89% 9% ··

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 5776 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 Naphthalene 1,2-dioxygenase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	446	Total	C	N	O	S	0	2	0
			3474	2198	596	663	17			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	352	VAL	PHE	engineered mutation	UNP P0A111

- Molecule 2 is a protein called Naphthalene 1,2-dioxygenase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	192	Total	C	N	O	S	0	1	0
			1605	1005	302	293	5			

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Fe	0	0
			1	1		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O

O3 O⁻ — S — O⁻ O4

S

O

O2

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

- Molecule 5 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2).

FES

S1 FE2

S — Fe

Fe — S

FE1 S2
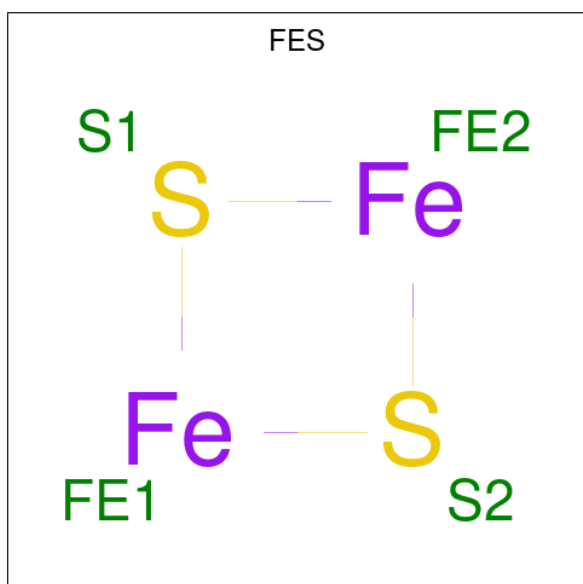
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
5	A	1	4	2	2	0	0

- Molecule 6 is ANTHRACENE (three-letter code: AN3) (formula: $C_{14}H_{10}$).
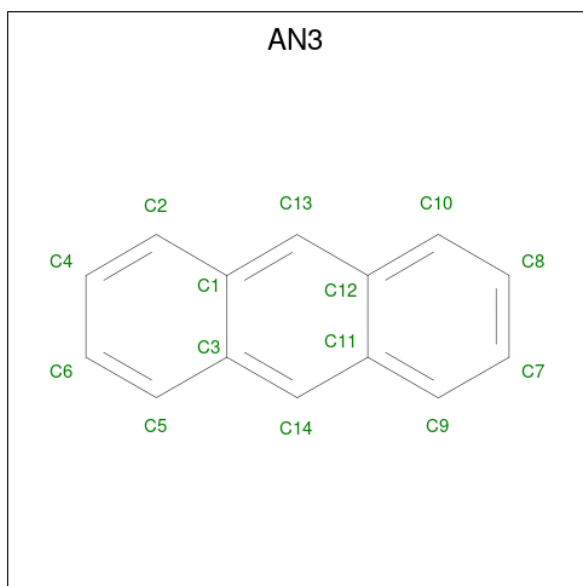
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	C		
6	A	1	14	14	0	0

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).
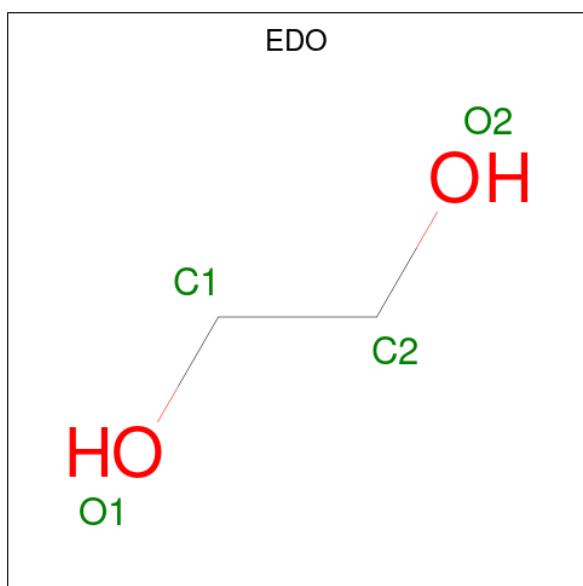
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	B	1	Total 4	C 2	O 2	0	0
7	B	1	Total 4	C 2	O 2	0	0
7	B	1	Total 4	C 2	O 2	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	405	Total 405	O 405	0	0
8	B	176	Total 176	O 176	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.

• Molecule 1: Naphthalene 1,2-dioxygenase alpha subunit

Chain A:
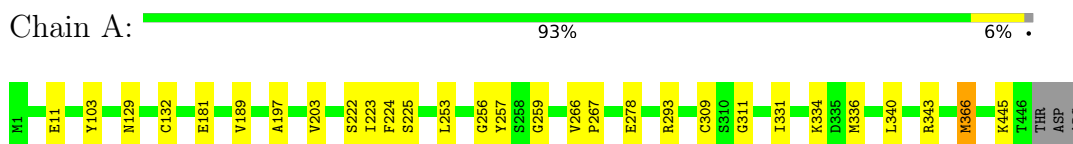
• Molecule 2: Naphthalene 1,2-dioxygenase beta subunit

Chain B:

W O R L D W I D E
wwPDB
PROTEIN DATA BANK
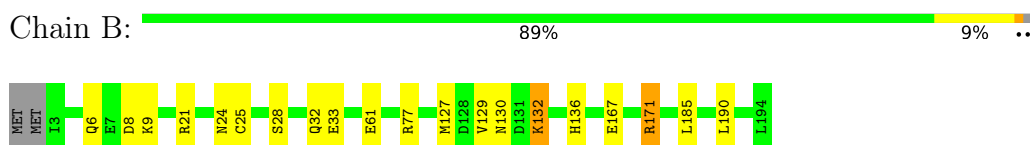
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	139.46Å 139.46Å 208.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	16.91 – 1.70 16.91 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.5 (16.91-1.70) 99.7 (16.91-1.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.69 (at 1.70Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.159 , 0.194 0.171 , 0.207	Depositor DCC
R_{free} test set	4248 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å²)	19.2	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.40 , 29.2	EDS
L-test for twinning [2]	$<|L|> = 0.51$, $<L^2> = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5776	wwPDB-VP
Average B, all atoms (Å²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.66% 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: EDO, FES, FE, AN3, 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	0.63	0/3575	0.67	1/4844 (0.0%)
2	B	0.69	0/1639	0.76	2/2211 (0.1%)
All	All	0.65	0/5214	0.70	3/7055 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	24	ASN	N-CA-C	6.50	128.54	111.00
1	A	343	ARG	NE-CZ-NH2	-5.55	117.53	120.30
2	B	25	CYS	N-CA-C	-5.36	96.53	111.00

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	3474	0	3315	28	0
2	B	1605	0	1578	11	0
3	A	1	0	0	0	0
4	A	15	0	0	0	0
4	B	10	0	0	0	0
5	A	4	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	14	0	10	3	0
7	A	60	0	90	0	0
7	B	12	0	18	0	0
8	A	405	0	0	1	0
8	B	176	0	0	2	0
All	All	5776	0	5011	35	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.

All (35) 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:224:PHE:CE1	1:A:253:LEU:HD23	2.09	0.87
1:A:11:GLU:HG2	8:A:1122:HOH:O	1.75	0.86
2:B:33:GLU:OE2	2:B:171:ARG:HD2	1.78	0.81
1:A:224:PHE:CZ	1:A:253:LEU:HD23	2.16	0.80
1:A:129:ASN:ND2	1:A:132:CYS:SG	2.64	0.71
1:A:224:PHE:CE1	1:A:253:LEU:CD2	2.75	0.69
1:A:336[B]:MET:HE2	1:A:340:LEU:HG	1.76	0.68
1:A:224:PHE:CD2	6:A:701:AN3:H4	2.29	0.67
1:A:336[B]:MET:CE	1:A:340:LEU:HG	2.29	0.63
2:B:127:MET:HE2	2:B:132:LYS:O	2.01	0.60
1:A:189[B]:VAL:HG12	2:B:185:LEU:HD12	1.85	0.56
2:B:28:SER:O	2:B:32:GLN:OE1	2.22	0.56
1:A:336[B]:MET:HE2	1:A:340:LEU:CG	2.37	0.55
1:A:336[B]:MET:HE2	1:A:340:LEU:CD2	2.38	0.54
1:A:224:PHE:HZ	1:A:256:GLY:HA3	1.73	0.53
1:A:224:PHE:CZ	1:A:256:GLY:HA3	2.43	0.53
1:A:189[B]:VAL:HG12	2:B:185:LEU:HB2	1.91	0.52
2:B:61:GLU:HG3	8:B:883:HOH:O	2.09	0.51
1:A:223:ILE:HD12	1:A:256:GLY:HA2	1.96	0.48
2:B:129:VAL:HG12	2:B:130:ASN:OD1	2.14	0.47
1:A:266:VAL:HB	1:A:267:PRO:HD3	1.98	0.46
1:A:189[B]:VAL:CG1	2:B:185:LEU:HD12	2.46	0.46
1:A:224:PHE:CE2	6:A:701:AN3:H4	2.52	0.44
1:A:224:PHE:CD2	6:A:701:AN3:C4	2.99	0.44
2:B:32:GLN:NE2	8:B:826:HOH:O	2.50	0.44
1:A:311:GLY:HA2	1:A:331:ILE:HG13	1.99	0.43
1:A:336[B]:MET:CE	1:A:340:LEU:CD2	2.96	0.43
1:A:336[B]:MET:HE2	1:A:340:LEU:HD23	2.00	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:21:ARG:HD3	2:B:136:HIS:CG	2.54	0.41
1:A:223:ILE:HD11	1:A:259:GLY:H	1.84	0.41
1:A:222:SER:O	1:A:225:SER:HB2	2.21	0.41
1:A:189[A]:VAL:HG13	2:B:185:LEU:HB2	2.02	0.41
1:A:181:GLU:HG3	1:A:334:LYS:HG2	2.04	0.40
1:A:197:ALA:HB1	1:A:366:MET:HE2	2.04	0.40
1:A:257:TYR:O	1:A:309:CYS:HB3	2.21	0.40

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	446/449 (99%)	434 (97%)	11 (2%)	1 (0%)	47	30
2	B	191/194 (98%)	187 (98%)	4 (2%)	0	100	100
All	All	637/643 (99%)	621 (98%)	15 (2%)	1 (0%)	47	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	445	LYS

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	368/369 (100%)	363 (99%)	5 (1%)	67	53
2	B	172/173 (99%)	164 (95%)	8 (5%)	26	10
All	All	540/542 (100%)	527 (98%)	13 (2%)	49	31

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	103	TYR
1	A	203	VAL
1	A	278	GLU
1	A	293	ARG
1	A	366	MET
2	B	6	GLN
2	B	8	ASP
2	B	9	LYS
2	B	77	ARG
2	B	132	LYS
2	B	167	GLU
2	B	171	ARG
2	B	190	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	115	GLN
1	A	129	ASN
2	B	26	HIS

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 (i)

Of 26 ligands modelled in this entry, 1 is monoatomic - leaving 25 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	#$|Z| > 2$
7	EDO	A	711	-	3,3,3	0.43	0	2,2,2	0.43	0
7	EDO	A	704	-	3,3,3	0.59	0	2,2,2	0.50	0
4	SO4	A	715	-	4,4,4	0.17	0	6,6,6	0.29	0
7	EDO	A	722	-	3,3,3	0.45	0	2,2,2	0.19	0
7	EDO	B	716	-	3,3,3	0.45	0	2,2,2	0.13	0
7	EDO	A	703	-	3,3,3	0.54	0	2,2,2	0.20	0
7	EDO	B	717	-	3,3,3	0.63	0	2,2,2	0.35	0
5	FES	A	725	1	0,4,4	-	-	-		
7	EDO	A	705	-	3,3,3	0.78	0	2,2,2	0.43	0
4	SO4	B	721	-	4,4,4	0.21	0	6,6,6	0.55	0
4	SO4	A	724	-	4,4,4	0.16	0	6,6,6	0.17	0
7	EDO	A	707	-	3,3,3	0.43	0	2,2,2	0.21	0
7	EDO	A	710	-	3,3,3	0.60	0	2,2,2	0.27	0
7	EDO	A	714	-	3,3,3	0.33	0	2,2,2	0.11	0
7	EDO	A	713	-	3,3,3	0.47	0	2,2,2	0.24	0
7	EDO	A	706	-	3,3,3	0.46	0	2,2,2	0.49	0
7	EDO	A	719	-	3,3,3	0.43	0	2,2,2	0.37	0
4	SO4	A	723	-	4,4,4	0.22	0	6,6,6	0.76	0
6	AN3	A	701	-	16,16,16	1.31	1 (6%)	22,22,22	1.40	4 (18%)
4	SO4	B	720	-	4,4,4	0.27	0	6,6,6	0.17	0
7	EDO	A	712	-	3,3,3	0.56	0	2,2,2	0.29	0
7	EDO	A	709	-	3,3,3	0.56	0	2,2,2	0.23	0
7	EDO	A	708	-	3,3,3	0.50	0	2,2,2	0.29	0
7	EDO	B	718	-	3,3,3	0.35	0	2,2,2	0.39	0
7	EDO	A	702	-	3,3,3	0.81	0	2,2,2	0.45	0

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
7	EDO	A	711	-	-	1/1/1/1	-
7	EDO	A	704	-	-	1/1/1/1	-
7	EDO	A	722	-	-	0/1/1/1	-
7	EDO	B	716	-	-	0/1/1/1	-
7	EDO	A	703	-	-	0/1/1/1	-
7	EDO	B	717	-	-	1/1/1/1	-
5	FES	A	725	1	-	-	0/1/1/1
7	EDO	A	705	-	-	1/1/1/1	-
7	EDO	A	707	-	-	0/1/1/1	-
7	EDO	A	710	-	-	1/1/1/1	-
7	EDO	A	714	-	-	0/1/1/1	-
7	EDO	A	713	-	-	0/1/1/1	-
7	EDO	A	706	-	-	1/1/1/1	-
7	EDO	A	719	-	-	0/1/1/1	-
6	AN3	A	701	-	-	-	0/3/3/3
7	EDO	A	712	-	-	0/1/1/1	-
7	EDO	A	709	-	-	1/1/1/1	-
7	EDO	A	708	-	-	1/1/1/1	-
7	EDO	B	718	-	-	0/1/1/1	-
7	EDO	A	702	-	-	1/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	701	AN3	C12-C11	-2.29	1.37	1.42

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	701	AN3	C5-C3-C14	-2.67	117.63	122.00
6	A	701	AN3	C12-C13-C1	-2.55	118.23	121.92
6	A	701	AN3	C9-C11-C14	-2.44	118.00	122.00
6	A	701	AN3	C11-C14-C3	-2.29	118.61	121.92

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	710	EDO	O1-C1-C2-O2
7	A	711	EDO	O1-C1-C2-O2
7	A	704	EDO	O1-C1-C2-O2
7	A	708	EDO	O1-C1-C2-O2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
7	A	702	EDO	O1-C1-C2-O2
7	A	706	EDO	O1-C1-C2-O2
7	B	717	EDO	O1-C1-C2-O2
7	A	705	EDO	O1-C1-C2-O2
7	A	709	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	701	AN3	3	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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3   Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4   Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5   Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.