



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 01:38 AM GMT

PDB ID : 2QA1
Title : Crystal structure of PgaE, an aromatic hydroxylase involved in angucycline biosynthesis
Authors : Koskiniemi, H.; Dobritzsch, D.; Metsa-Ketela, M.; Kallio, P.; Niemi, J.; Schneider, G.
Deposited on : 2007-06-14
Resolution : 1.80 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

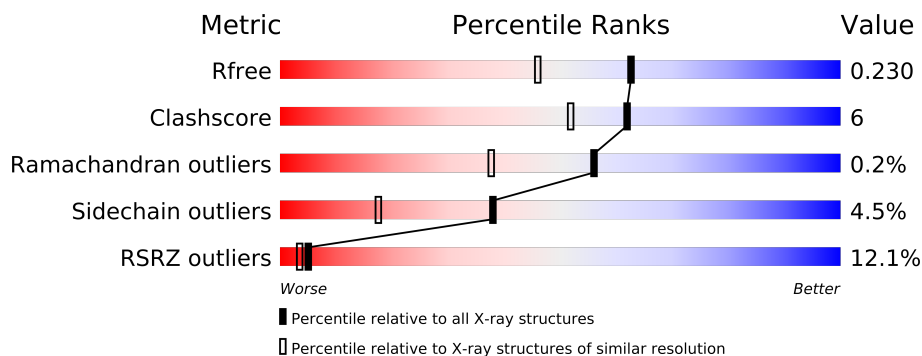
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.80 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}	66092	3513 (1.80-1.80)
Clashscore	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)
RSRZ outliers	66119	3515 (1.80-1.80)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	500	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	EDO	A	602	-	X
5	GOL	A	605	-	X
5	GOL	A	607	-	X
5	GOL	A	608	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 4166 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Polyketide oxygenase PgaE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	488	3718	2328	673	699	18	0	10	0

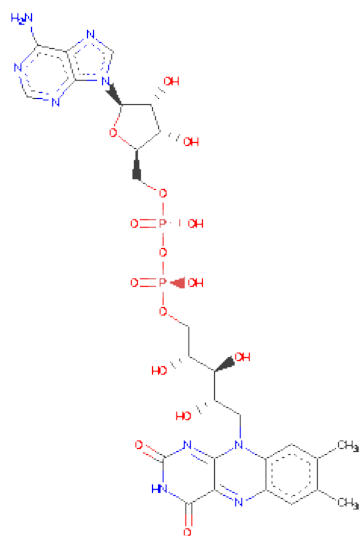
There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	ALA	-	EXPRESSION TAG	UNP Q93LY7
A	-7	HIS	-	EXPRESSION TAG	UNP Q93LY7
A	-6	HIS	-	EXPRESSION TAG	UNP Q93LY7
A	-5	HIS	-	EXPRESSION TAG	UNP Q93LY7
A	-4	HIS	-	EXPRESSION TAG	UNP Q93LY7
A	-3	HIS	-	EXPRESSION TAG	UNP Q93LY7
A	-2	HIS	-	EXPRESSION TAG	UNP Q93LY7
A	-1	HIS	-	EXPRESSION TAG	UNP Q93LY7
A	0	ARG	-	EXPRESSION TAG	UNP Q93LY7
A	1	SER	-	EXPRESSION TAG	UNP Q93LY7

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

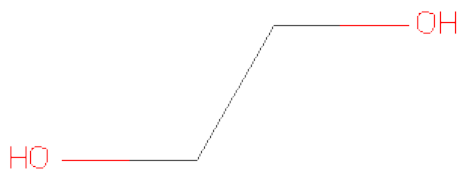
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: C₂₇H₃₃N₉O₁₅P₂).



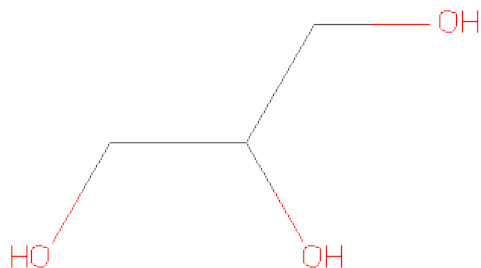
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	358	Total	O	0	0
			358	358		

4 Data and refinement statistics

Property	Value	Source
Space group	F 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	66.23Å 171.53Å 212.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.96 – 1.80 36.92 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.3 (36.96-1.80) 99.3 (36.92-1.80)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.16 (at 1.81Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, R_{free}	0.195 , 0.224 0.200 , 0.230	Depositor DCC
R_{free} test set	2801 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	27.4	Xtriage
Anisotropy	0.454	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 75.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 55590 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4166	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.96% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ZN, EDO, FAD

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	0.50	0/3814	0.61	0/5180

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	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	0	ARG	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3718	0	3714	41	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	1	0	0	0	0
3	A	53	0	31	2	0
4	A	12	0	18	6	0
5	A	24	0	32	4	0
6	A	358	0	0	6	0
All	All	4166	0	3795	43	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (43) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:275:ASP:OD2	5:A:608:GOL:H32	1.64	0.98
1:A:51[B]:MET:SD	1:A:89:GLN:O	2.34	0.86
1:A:490:THR:HA	6:A:960:HOH:O	1.81	0.80
1:A:200:MET:HB2	1:A:212:ILE:HB	1.72	0.70
1:A:126:ASP:O	1:A:128:ALA:N	2.30	0.65
1:A:401:ARG:NH1	4:A:603:EDO:H22	2.12	0.63
1:A:491:GLY:HA3	6:A:955:HOH:O	2.00	0.62
1:A:174:GLU:HB2	1:A:220:PRO:HA	1.81	0.61
1:A:343:GLY:HA3	1:A:345[A]:GLU:OE1	2.01	0.60
1:A:37[C]:ARG:NH2	1:A:104:GLU:OE1	2.34	0.59
1:A:140[A]:LYS:H	1:A:140[A]:LYS:HD2	1.66	0.59
1:A:401:ARG:HH11	4:A:603:EDO:H22	1.71	0.56
1:A:123:LEU:HD12	1:A:123:LEU:C	2.26	0.55
1:A:41:SER:HA	1:A:96:GLN:HB3	1.87	0.55
1:A:401:ARG:HH11	4:A:603:EDO:C2	2.21	0.54
1:A:140[A]:LYS:CD	1:A:140[A]:LYS:H	2.21	0.54
1:A:37[B]:ARG:NH2	1:A:104:GLU:OE1	2.40	0.53
1:A:401:ARG:NH1	4:A:603:EDO:C2	2.74	0.51
1:A:426:ARG:HD3	1:A:440:THR:OG1	2.12	0.49
1:A:173[B]:MET:HE3	1:A:175:MET:HG3	1.97	0.47
1:A:334:GLN:NE2	6:A:904:HOH:O	2.48	0.47
1:A:210:ARG:NH1	6:A:910:HOH:O	2.47	0.47
1:A:246:HIS:CD2	1:A:246:HIS:H	2.33	0.47
1:A:345[A]:GLU:CD	1:A:345[A]:GLU:H	2.19	0.46
1:A:432:TRP:CE3	1:A:481:LEU:HD13	2.50	0.46
5:A:607:GOL:H31	6:A:911:HOH:O	2.15	0.45
1:A:126:ASP:C	1:A:128:ALA:H	2.18	0.45
1:A:37[B]:ARG:NH1	1:A:100:GLU:OE1	2.48	0.44
1:A:42:ARG:H	1:A:42:ARG:HG2	1.60	0.44
1:A:413:ARG:HG2	5:A:607:GOL:H2	1.99	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:261:GLN:HB2	1:A:330:LEU:HD11	2.01	0.43
1:A:73:HIS:HE2	4:A:602:EDO:H22	1.84	0.43
1:A:42:ARG:HB2	3:A:500:FAD:C7	2.49	0.42
1:A:226:PRO:HA	1:A:227:PRO:HD3	1.82	0.42
1:A:389:LYS:HE2	6:A:653:HOH:O	2.19	0.41
3:A:500:FAD:H1'1	5:A:608:GOL:H2	2.01	0.41
1:A:125:ASP:OD1	1:A:125:ASP:C	2.58	0.41
1:A:135:ARG:HD3	1:A:140[B]:LYS:HE3	2.02	0.41
1:A:124:THR:O	1:A:130:VAL:HG13	2.21	0.41
1:A:489:LEU:HA	1:A:489:LEU:HD23	1.94	0.41
1:A:73:HIS:NE2	4:A:602:EDO:H22	2.35	0.41
1:A:180:ILE:HG22	1:A:249:PRO:HA	2.04	0.41
1:A:201:VAL:HG12	1:A:211:ILE:HG22	2.03	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	494/500 (99%)	484 (98%)	9 (2%)	1 (0%)	56	38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1	SER

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	385/386 (100%)	366 (95%)	19 (5%)	35 15

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	SER
1	A	2	ASP
1	A	42	ARG
1	A	122	SER
1	A	138	GLU
1	A	140[A]	LYS
1	A	140[B]	LYS
1	A	177	LEU
1	A	183	VAL
1	A	189	MET
1	A	198	MET
1	A	199	VAL
1	A	201	VAL
1	A	345[A]	GLU
1	A	345[B]	GLU
1	A	353	LEU
1	A	390	ARG
1	A	401	ARG
1	A	482	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	56	GLN
1	A	246	HIS

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	FAD	A	500	-	58,58,58	1.25	5 (8%)	85,89,89	1.99	17 (20%)
4	EDO	A	602	-	3,3,3	0.51	0	2,2,2	0.30	0
4	EDO	A	603	-	3,3,3	0.45	0	2,2,2	0.68	0
4	EDO	A	604	-	3,3,3	0.55	0	2,2,2	0.25	0
5	GOL	A	605	-	5,5,5	0.25	0	5,5,5	0.53	0
5	GOL	A	606	-	5,5,5	0.28	0	5,5,5	0.42	0
5	GOL	A	607	-	5,5,5	0.34	0	5,5,5	0.41	0
5	GOL	A	608	-	5,5,5	0.25	0	5,5,5	0.46	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
3	FAD	A	500	-	-	0/34/50/50	0/1/6/6
4	EDO	A	602	-	-	0/1/1/1	0/0/0/0
4	EDO	A	603	-	-	0/1/1/1	0/0/0/0
4	EDO	A	604	-	-	0/1/1/1	0/0/0/0
5	GOL	A	605	-	-	0/4/4/4	0/0/0/0
5	GOL	A	606	-	-	0/4/4/4	0/0/0/0
5	GOL	A	607	-	-	0/4/4/4	0/0/0/0
5	GOL	A	608	-	-	0/4/4/4	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	500	FAD	C2A-N3A	4.39	1.40	1.32
3	A	500	FAD	C2A-N1A	3.35	1.40	1.33
3	A	500	FAD	C1'-C2'	3.19	1.54	1.51
3	A	500	FAD	C4A-N3A	2.83	1.39	1.35
3	A	500	FAD	C5X-N5	2.30	1.38	1.35

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	500	FAD	N3A-C2A-N1A	-12.05	118.63	128.71
3	A	500	FAD	C2-N1-C10	5.94	120.96	114.98
3	A	500	FAD	C1B-N9A-C4A	-3.57	120.46	126.64
3	A	500	FAD	C4X-C10-N1	-3.36	119.37	122.73
3	A	500	FAD	C4X-N5-C5X	3.28	120.38	116.69
3	A	500	FAD	C5X-C9A-N10	3.08	119.84	116.80
3	A	500	FAD	C9A-N10-C10	-2.90	118.92	121.77
3	A	500	FAD	N3A-C4A-N9A	2.88	130.63	125.43
3	A	500	FAD	C8A-N9A-C4A	2.86	109.08	106.90
3	A	500	FAD	C4-N3-C2	-2.61	120.04	125.39
3	A	500	FAD	C1'-N10-C9A	2.44	121.25	118.87
3	A	500	FAD	N7A-C8A-N9A	-2.36	107.68	114.36
3	A	500	FAD	O2P-P-O3P	2.18	115.47	105.14
3	A	500	FAD	N6A-C6A-N1A	2.14	123.57	119.36
3	A	500	FAD	C8A-N9A-C1B	2.03	130.39	126.38
3	A	500	FAD	C2A-N3A-C4A	2.02	119.75	114.01
3	A	500	FAD	C4A-C5A-N7A	-2.01	107.80	109.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 95th 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	488/500 (97%)	0.52	60 (12%) 5 3	35, 42, 53, 75	0

All (60) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	205	PRO	6.6
1	A	1	SER	6.4
1	A	183	VAL	6.1
1	A	232	VAL	5.9
1	A	0	ARG	5.7
1	A	229	TRP	5.2
1	A	227	PRO	5.2
1	A	241	GLY	4.9
1	A	220	PRO	4.8
1	A	212	ILE	4.7
1	A	208	ILE	4.7
1	A	235	ALA	4.5
1	A	39	GLY	4.3
1	A	226	PRO	4.3
1	A	230	HIS	4.2
1	A	207	GLY	4.1
1	A	206	GLY	4.0
1	A	128	ALA	3.9
1	A	193	THR	3.9
1	A	239	LEU	3.9
1	A	181	LYS	3.8
1	A	184	GLU	3.8
1	A	217	GLY	3.7
1	A	186	GLN	3.7
1	A	248	GLU	3.6
1	A	203	PRO	3.6
1	A	-1	HIS	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	194	LEU	3.4
1	A	190	ILE	3.3
1	A	249	PRO	3.2
1	A	179	ASP	3.1
1	A	182	GLY	3.1
1	A	180	ILE	3.0
1	A	173[A]	MET	3.0
1	A	2	ASP	3.0
1	A	176	TYR	3.0
1	A	171	ALA	2.9
1	A	250	VAL	2.9
1	A	240	THR	2.8
1	A	3	ALA	2.8
1	A	196	GLY	2.7
1	A	236	TRP	2.7
1	A	233	ALA	2.6
1	A	246	HIS	2.6
1	A	214	CYS	2.6
1	A	199	VAL	2.5
1	A	247	ALA	2.5
1	A	204	LEU	2.5
1	A	38	THR	2.5
1	A	185	LEU	2.4
1	A	124	THR	2.4
1	A	219	PRO	2.4
1	A	400	THR	2.4
1	A	195	PRO	2.3
1	A	165[A]	ASP	2.3
1	A	126	ASP	2.2
1	A	237	LYS	2.2
1	A	160	LYS	2.2
1	A	41	SER	2.1
1	A	216	ARG	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 ⓘ

There are no carbohydrates in this entry.

6.4 Ligands

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
5	GOL	A	608	6/6	0.23	7.65	50,53,53,55	0
4	EDO	A	602	4/4	0.23	5.41	46,47,49,49	0
5	GOL	A	605	6/6	0.29	5.22	69,70,71,71	0
5	GOL	A	607	6/6	0.19	3.42	54,55,55,55	0
3	FAD	A	500	53/53	0.18	1.85	34,41,71,73	0
4	EDO	A	604	4/4	0.14	1.52	39,42,45,45	0
4	EDO	A	603	4/4	0.23	1.31	59,60,60,60	0
5	GOL	A	606	6/6	0.17	-0.04	82,83,83,83	0
2	ZN	A	601	1/1	0.10	-0.75	62,62,62,62	0

6.5 Other polymers

There are no such residues in this entry.