



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 29, 2021 – 10:29 AM EDT

PDB ID : 7M3K
Title : Crystal Structure of Galactonate dehydratase from Brucella melitensis biovar Abortus 2308
Authors : Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on : 2021-03-18
Resolution : 2.30 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 ⓘ symbol.

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	:	2.18
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)	:	2.18

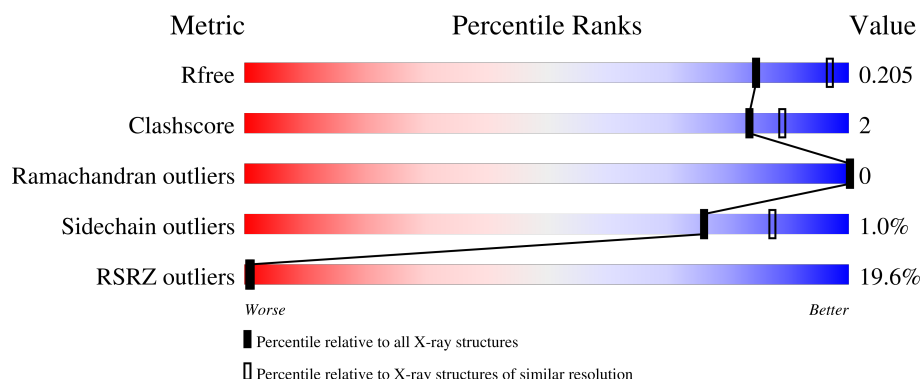
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 $\leq 5\%$. 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	609	<div> <div>19%</div> <div>91%</div> <div>6%</div> <div>.</div> </div>

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
4	PG4	A	706	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 4577 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 Galactonate dehydratase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	587	Total	C	N	O	S	0	5	0
			4270	2670	747	832	21			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP Q2YJ21
A	-6	ALA	-	expression tag	UNP Q2YJ21
A	-5	HIS	-	expression tag	UNP Q2YJ21
A	-4	HIS	-	expression tag	UNP Q2YJ21
A	-3	HIS	-	expression tag	UNP Q2YJ21
A	-2	HIS	-	expression tag	UNP Q2YJ21
A	-1	HIS	-	expression tag	UNP Q2YJ21
A	0	HIS	-	expression tag	UNP Q2YJ21
A	244	MET	THR	conflict	UNP Q2YJ21

- 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 CHLORIDE ION (three-letter code: CL) (formula: Cl).

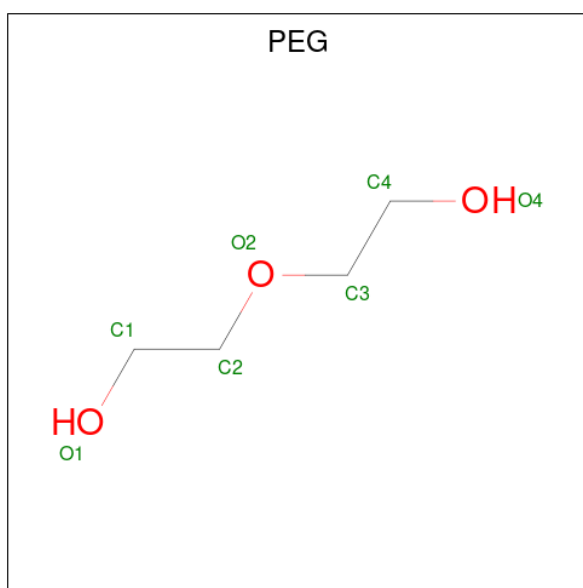
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	3	Total	Cl	0	0
			3	3		

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C₈H₁₈O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			10	6	4		
4	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			10	6	4		

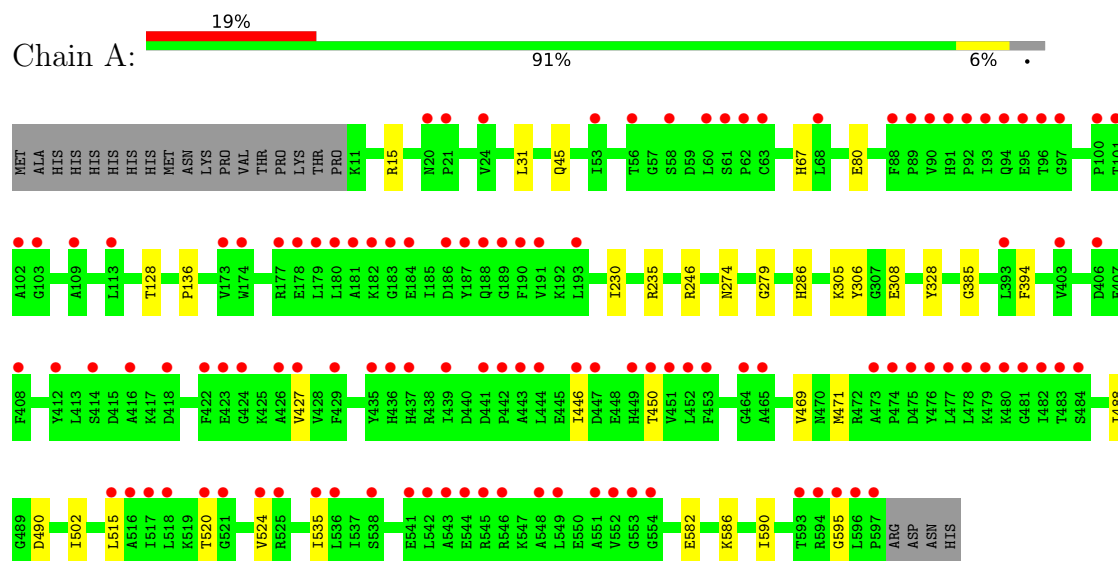
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	268	Total	O	0	1
			269	269		

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: Galactonate dehydratase



4 Data and refinement statistics

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α , β , γ	144.95Å 144.95Å 167.57Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.24 – 2.30 47.45 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.24-2.30) 100.0 (47.45-2.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.39 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.19	Depositor
R, R_{free}	0.180 , 0.208 0.178 , 0.205	Depositor DCC
R_{free} test set	2039 reflections (4.37%)	wwPDB-VP
Wilson B-factor (Å ²)	41.0	Xtriage
Anisotropy	0.437	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 61.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4577	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: CL, PGE, ZN, PG4, PEG

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	0.32	0/4371	0.55	0/5962

There are no bond length outliers.

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	4270	0	3999	17	0
2	A	1	0	0	0	0
3	A	3	0	0	1	0
4	A	17	0	22	3	0
5	A	7	0	10	0	0
6	A	10	0	14	0	0
7	A	269	0	0	1	0
All	All	4577	0	4045	20	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.

All (20) 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)
3:A:704:CL:CL	7:A:1041:HOH:O	2.36	0.81
1:A:582:GLU:HG3	4:A:705:PG4:H12	1.68	0.75
1:A:488:ILE:HG12	1:A:502:ILE:HB	1.84	0.60
1:A:31:LEU:HD11	1:A:45:GLN:NE2	2.24	0.52
1:A:67:HIS:CD2	1:A:128:THR:HB	2.47	0.50
1:A:427:VAL:HA	1:A:520:THR:HG23	1.94	0.50
1:A:305:LYS:HE2	1:A:306:TYR:CZ	2.46	0.50
1:A:67:HIS:CG	1:A:128:THR:HB	2.47	0.49
1:A:469:VAL:O	1:A:471:MET:N	2.45	0.48
1:A:136:PRO:HB3	1:A:286:HIS:HB3	1.94	0.48
1:A:446:ILE:HG23	1:A:450:THR:HG21	1.96	0.48
1:A:490:ASP:HB3	1:A:515:LEU:HD12	1.98	0.45
1:A:80:GLU:HG2	1:A:246[A]:ARG:HH11	1.85	0.42
4:A:705:PG4:H51	4:A:705:PG4:H32	1.63	0.42
1:A:230:ILE:HG12	1:A:328:TYR:CE1	2.55	0.41
4:A:705:PG4:H11	4:A:705:PG4:H31	1.79	0.41
1:A:590:ILE:HB	1:A:595:GLY:HA2	2.02	0.41
1:A:274:ASN:OD1	1:A:279:GLY:HA3	2.20	0.41
1:A:524:VAL:HG12	1:A:535:ILE:HA	2.04	0.40
1:A:308:GLU:HA	1:A:385:GLY:HA3	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	590/609 (97%)	564 (96%)	26 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	414/488 (85%)	410 (99%)	4 (1%)	76 87

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15	ARG
1	A	235	ARG
1	A	394	PHE
1	A	586	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	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 8 ligands modelled in this entry, 4 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
6	PGE	A	708	-	9,9,9	0.34	0	8,8,8	0.26	0
4	PG4	A	705	-	9,9,12	0.20	0	8,8,11	0.36	0
5	PEG	A	707	-	6,6,6	0.11	0	5,5,5	0.14	0
4	PG4	A	706	-	6,6,12	0.15	0	5,5,11	0.55	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
6	PGE	A	708	-	-	2/7/7/7	-
4	PG4	A	705	-	-	6/7/7/10	-
5	PEG	A	707	-	-	3/4/4/4	-
4	PG4	A	706	-	-	2/4/4/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	705	PG4	C3-C4-O3-C5
4	A	705	PG4	C1-C2-O2-C3
4	A	705	PG4	O2-C3-C4-O3
6	A	708	PGE	C3-C4-O3-C5
5	A	707	PEG	C1-C2-O2-C3
4	A	705	PG4	C6-C5-O3-C4
4	A	705	PG4	C4-C3-O2-C2
4	A	706	PG4	O4-C7-C8-O5
5	A	707	PEG	O1-C1-C2-O2
6	A	708	PGE	O2-C3-C4-O3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	A	707	PEG	C4-C3-O2-C2
4	A	706	PG4	O3-C5-C6-O4
4	A	705	PG4	O3-C5-C6-O4

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	705	PG4	3	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 ⓘ

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	587/609 (96%)	0.63	115 (19%) 1 1	28, 48, 124, 148	0

All (115) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	473	ALA	7.5
1	A	103	GLY	7.3
1	A	182	LYS	6.9
1	A	597	PRO	6.6
1	A	181	ALA	6.5
1	A	521	GLY	6.1
1	A	549	LEU	6.1
1	A	518	LEU	5.7
1	A	595	GLY	5.7
1	A	596	LEU	5.6
1	A	447	ASP	5.3
1	A	476	TYR	5.2
1	A	102	ALA	5.1
1	A	446	ILE	5.1
1	A	187	TYR	5.0
1	A	190	PHE	5.0
1	A	180	LEU	4.9
1	A	594	ARG	4.9
1	A	183	GLY	4.8
1	A	477	LEU	4.6
1	A	524	VAL	4.6
1	A	191	VAL	4.5
1	A	542	LEU	4.5
1	A	553	GLY	4.3
1	A	475	ASP	4.3
1	A	188	GLN	4.2
1	A	450	THR	4.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	414	SER	4.2
1	A	442	PRO	4.2
1	A	544	GLU	4.1
1	A	481	GLY	4.1
1	A	96	THR	4.0
1	A	174	TRP	4.0
1	A	427	VAL	4.0
1	A	436	HIS	3.9
1	A	426	ALA	3.9
1	A	178	GLU	3.9
1	A	520	THR	3.9
1	A	21	PRO	3.8
1	A	444	LEU	3.7
1	A	536	LEU	3.7
1	A	422	PHE	3.7
1	A	97	GLY	3.7
1	A	554	GLY	3.6
1	A	483	THR	3.5
1	A	479	LYS	3.4
1	A	90	VAL	3.3
1	A	94	GLN	3.3
1	A	541	GLU	3.3
1	A	101	THR	3.3
1	A	193	LEU	3.3
1	A	516	ALA	3.3
1	A	478	LEU	3.3
1	A	482	ILE	3.3
1	A	551	ALA	3.2
1	A	441	ASP	3.2
1	A	464	GLY	3.1
1	A	443	ALA	3.1
1	A	423	GLU	3.0
1	A	465	ALA	3.0
1	A	406	ASP	3.0
1	A	24	VAL	3.0
1	A	177	ARG	3.0
1	A	189	GLY	2.9
1	A	593	THR	2.9
1	A	412	TYR	2.8
1	A	546	ARG	2.8
1	A	184	GLU	2.8
1	A	186	ASP	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	435	TYR	2.7
1	A	93	ILE	2.7
1	A	552	VAL	2.7
1	A	68	LEU	2.7
1	A	89	PRO	2.6
1	A	437	HIS	2.6
1	A	424	GLY	2.6
1	A	56	THR	2.5
1	A	474	PRO	2.5
1	A	60	LEU	2.5
1	A	517	ILE	2.5
1	A	91	HIS	2.5
1	A	453	PHE	2.5
1	A	179	LEU	2.5
1	A	63	CYS	2.5
1	A	416	ALA	2.5
1	A	543	ALA	2.5
1	A	408	PHE	2.4
1	A	429	PHE	2.4
1	A	484	SER	2.4
1	A	515	LEU	2.4
1	A	95	GLU	2.4
1	A	173	VAL	2.4
1	A	452	LEU	2.3
1	A	53	ILE	2.3
1	A	88	PHE	2.3
1	A	525	ARG	2.3
1	A	100	PRO	2.3
1	A	545	ARG	2.2
1	A	451	VAL	2.2
1	A	109	ALA	2.2
1	A	548	ALA	2.2
1	A	449	HIS	2.1
1	A	439	ILE	2.1
1	A	58	SER	2.1
1	A	61[A]	SER	2.1
1	A	538	SER	2.1
1	A	92	PRO	2.1
1	A	113	LEU	2.1
1	A	403	VAL	2.1
1	A	393	LEU	2.1
1	A	20	ASN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	62	PRO	2.1
1	A	418	ASP	2.0
1	A	480	LYS	2.0
1	A	535	ILE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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, 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	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	PG4	A	706	7/13	0.72	0.42	58,59,65,66	0
5	PEG	A	707	7/7	0.85	0.21	46,54,57,66	0
6	PGE	A	708	10/10	0.85	0.15	52,53,55,55	10
4	PG4	A	705	10/13	0.90	0.30	56,60,64,65	0
2	ZN	A	701	1/1	0.91	0.13	55,55,55,55	1
3	CL	A	703	1/1	0.93	0.51	77,77,77,77	0
3	CL	A	702	1/1	0.96	0.07	40,40,40,40	1
3	CL	A	704	1/1	0.99	0.04	52,52,52,52	1

6.5 Other polymers [i](#)

There are no such residues in this entry.