



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 4, 2024 – 12:11 PM JST

PDB ID : 8HU4
Title : Limosilactobacillus reuteri N1 GtfB
Authors : Dong, J.J.; Bai, Y.X.
Deposited on : 2022-12-22
Resolution : 2.76 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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 : 2.36
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) : 2.36

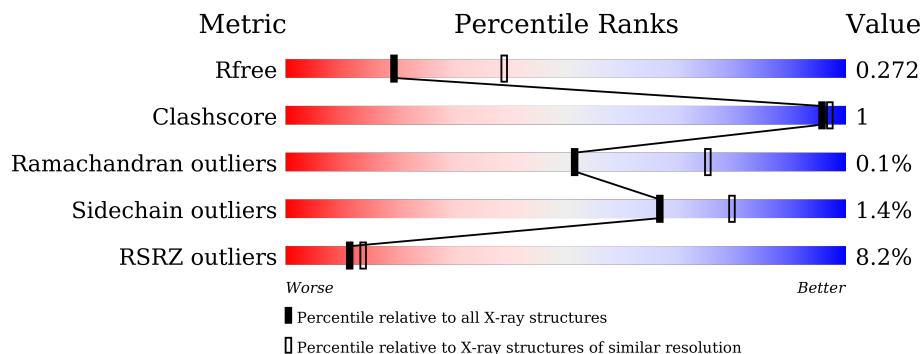
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.76 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	833	 6% 96% 5%
1	B	833	 10% 94% 5%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13042 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 dextransucrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	827	6457	4025	1095	1315	22	0	0	0
1	B	827	6457	4025	1095	1315	22	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	MET	-	initiating methionine	UNP A0A848PDI7
B	25	MET	-	initiating methionine	UNP A0A848PDI7

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

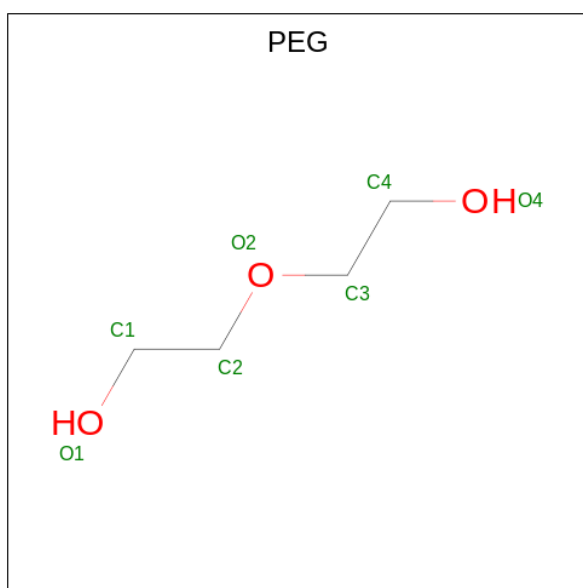
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Na	0	0
			1	1		
2	B	1	Total	Na	0	0
			1	1		

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: C₆H₈O₇).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	6	7		
3	B	1	Total	C	O	0	0
			13	6	7		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



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

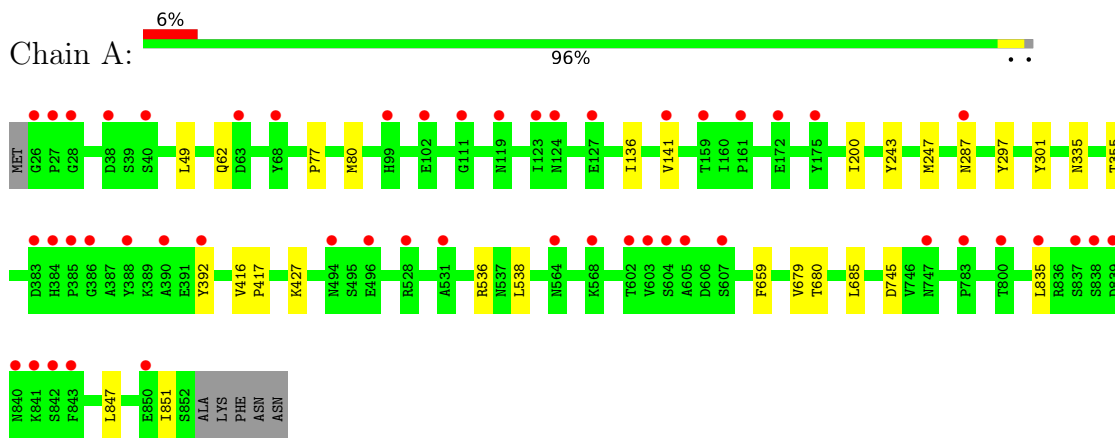
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	44	Total 44	O 44	0	0
5	B	42	Total 42	O 42	0	0

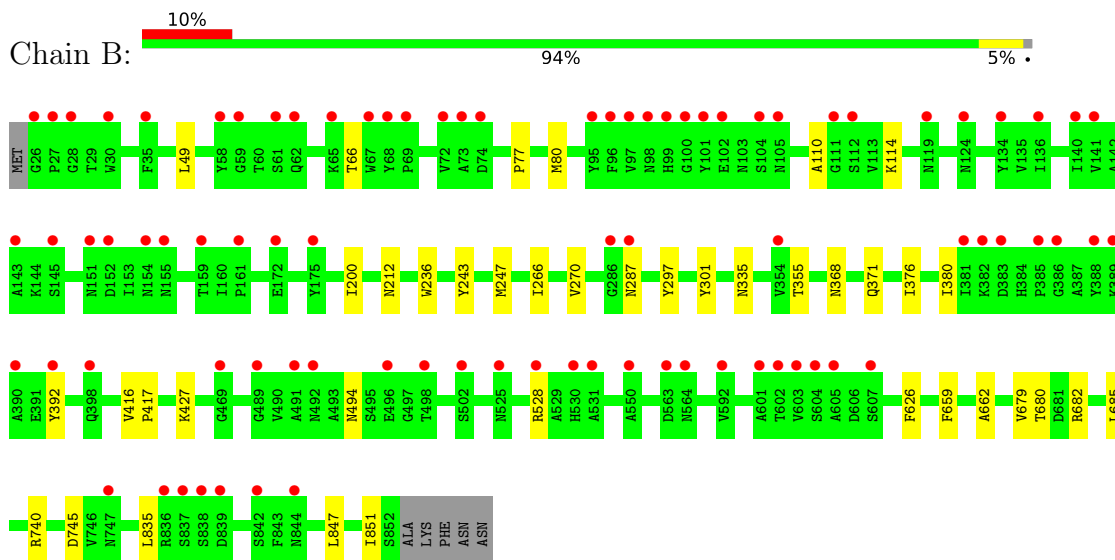
3 Residue-property plots [i](#)

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 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: dextranucrase



- Molecule 1: dextranucrase



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	204.42Å 100.73Å 106.89Å 90.00° 108.14° 90.00°	Depositor
Resolution (Å)	47.21 – 2.76 47.16 – 2.76	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.21-2.76) 99.6 (47.16-2.76)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.40 (at 2.77Å)	Xtrriage
Refinement program	REFMAC 5	Depositor
R, R_{free}	0.248 , 0.266 0.252 , 0.272	Depositor DCC
R_{free} test set	2666 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	29.6	Xtrriage
Anisotropy	0.010	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 37.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	13042	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.93% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, PEG, CIT

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	0.66	0/6597	0.71	0/8984
1	B	0.66	0/6597	0.70	0/8984
All	All	0.66	0/13194	0.71	0/17968

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 [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	6457	0	6094	11	0
1	B	6457	0	6094	16	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	13	0	5	0	0
3	B	13	0	5	0	0
4	A	7	0	10	0	0
4	B	7	0	10	0	0
5	A	44	0	0	0	0
5	B	42	0	0	0	0
All	All	13042	0	12218	27	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 1.

All (27) 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:835:LEU:HD11	1:A:847:LEU:HD21	1.69	0.73
1:B:835:LEU:HD11	1:B:847:LEU:HD21	1.70	0.72
1:B:212:ASN:O	1:B:740:ARG:NH2	2.35	0.59
1:B:243:TYR:CE1	1:B:685:LEU:HD22	2.44	0.53
1:A:243:TYR:CE1	1:A:685:LEU:HD22	2.45	0.52
1:B:851:ILE:HG22	1:B:851:ILE:O	2.10	0.51
1:B:49:LEU:HD11	1:B:835:LEU:HD13	1.93	0.49
1:A:49:LEU:HD11	1:A:835:LEU:HD13	1.94	0.49
1:A:355:THR:O	1:A:427:LYS:NZ	2.47	0.47
1:B:355:THR:O	1:B:427:LYS:NZ	2.45	0.47
1:A:77:PRO:HG2	1:A:80:MET:HG3	1.98	0.45
1:B:416:VAL:HB	1:B:417:PRO:HD3	1.98	0.45
1:B:110:ALA:O	1:B:114:LYS:HB2	2.17	0.44
1:A:416:VAL:HB	1:A:417:PRO:HD3	1.99	0.43
1:A:80:MET:SD	1:A:136:ILE:HG21	2.59	0.43
1:A:851:ILE:O	1:A:851:ILE:HG22	2.19	0.43
1:B:368:ASN:HB2	1:B:371:GLN:HB2	2.02	0.41
1:A:62:GLN:HE22	1:A:141:VAL:HG21	1.85	0.41
1:A:536:ARG:NH2	1:A:538:LEU:HD12	2.35	0.41
1:B:243:TYR:CE1	1:B:247:MET:HG3	2.54	0.41
1:B:494:ASN:HD22	1:B:528:ARG:HD2	1.86	0.41
1:B:626:PHE:CG	1:B:662:ALA:HB2	2.56	0.41
1:B:266:ILE:HD12	1:B:270:VAL:HG21	2.03	0.41
1:B:77:PRO:HG2	1:B:80:MET:HG3	2.01	0.41
1:A:243:TYR:CE1	1:A:247:MET:HG3	2.56	0.40
1:B:236:TRP:CZ2	1:B:682:ARG:HG2	2.56	0.40
1:B:376:ILE:O	1:B:380:ILE:HG12	2.21	0.40

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	825/833 (99%)	783 (95%)	41 (5%)	1 (0%)	51	75
1	B	825/833 (99%)	783 (95%)	41 (5%)	1 (0%)	51	75
All	All	1650/1666 (99%)	1566 (95%)	82 (5%)	2 (0%)	51	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	200	ILE
1	B	200	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	695/700 (99%)	686 (99%)	9 (1%)	69	81
1	B	695/700 (99%)	685 (99%)	10 (1%)	67	79
All	All	1390/1400 (99%)	1371 (99%)	19 (1%)	67	79

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

Mol	Chain	Res	Type
1	A	287	ASN
1	A	297	TYR
1	A	301	TYR
1	A	335	ASN
1	A	392	TYR
1	A	659	PHE
1	A	679	VAL
1	A	680	THR
1	A	745	ASP
1	B	66	THR
1	B	287	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	297	TYR
1	B	301	TYR
1	B	335	ASN
1	B	392	TYR
1	B	659	PHE
1	B	679	VAL
1	B	680	THR
1	B	745	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	62	GLN
1	A	98	ASN
1	A	334	ASN
1	A	335	ASN
1	A	419	GLN
1	A	638	ASN
1	A	747	ASN
1	A	765	ASN
1	A	767	GLN
1	B	98	ASN
1	B	334	ASN
1	B	335	ASN
1	B	494	ASN
1	B	638	ASN
1	B	765	ASN
1	B	767	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

Of 6 ligands modelled in this entry, 2 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
3	CIT	B	902	-	12,12,12	1.18	1 (8%)	17,17,17	1.33	2 (11%)
4	PEG	B	903	-	6,6,6	0.13	0	5,5,5	0.08	0
3	CIT	A	902	-	12,12,12	1.15	1 (8%)	17,17,17	1.29	2 (11%)
4	PEG	A	903	-	6,6,6	0.14	0	5,5,5	0.08	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	CIT	B	902	-	-	3/16/16/16	-
4	PEG	B	903	-	-	2/4/4/4	-
3	CIT	A	902	-	-	0/16/16/16	-
4	PEG	A	903	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	902	CIT	C3-C6	2.39	1.55	1.53
3	A	902	CIT	C3-C6	2.23	1.55	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	902	CIT	O5-C6-C3	-3.49	117.31	122.25
3	A	902	CIT	O5-C6-C3	-3.46	117.36	122.25
3	B	902	CIT	O6-C6-C3	2.67	117.68	113.05
3	A	902	CIT	O6-C6-C3	2.51	117.42	113.05

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	902	CIT	C1-C2-C3-O7
3	B	902	CIT	C1-C2-C3-C4
3	B	902	CIT	C1-C2-C3-C6
4	B	903	PEG	O1-C1-C2-O2
4	A	903	PEG	C4-C3-O2-C2
4	A	903	PEG	O2-C3-C4-O4
4	B	903	PEG	C4-C3-O2-C2

There are no ring outliers.

No monomer is involved in short contacts.

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	827/833 (99%)	0.54	50 (6%) 21 26	19, 29, 63, 72	0
1	B	827/833 (99%)	0.63	85 (10%) 6 7	20, 29, 65, 76	0
All	All	1654/1666 (99%)	0.59	135 (8%) 11 14	19, 29, 65, 76	0

All (135) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	26	GLY	10.8
1	A	603	VAL	8.6
1	B	603	VAL	8.1
1	A	838	SER	7.8
1	A	837	SER	7.5
1	B	838	SER	7.1
1	B	385	PRO	6.9
1	B	28	GLY	6.9
1	A	124	ASN	6.6
1	A	385	PRO	6.4
1	A	26	GLY	5.9
1	B	141	VAL	5.6
1	A	388	TYR	5.2
1	B	390	ALA	5.1
1	A	27	PRO	4.7
1	B	386	GLY	4.6
1	B	388	TYR	4.5
1	A	604	SER	4.5
1	B	97	VAL	4.4
1	B	124	ASN	4.3
1	A	842	SER	4.3
1	B	836	ARG	4.2
1	B	27	PRO	4.1
1	B	68	TYR	4.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	95	TYR	4.1
1	B	837	SER	4.1
1	A	386	GLY	4.1
1	A	747	ASN	4.0
1	A	839	ASP	4.0
1	B	604	SER	3.9
1	A	392	TYR	3.9
1	B	605	ALA	3.9
1	B	74	ASP	3.8
1	B	382	LYS	3.8
1	B	550	ALA	3.8
1	B	58	TYR	3.8
1	A	172	GLU	3.8
1	A	607	SER	3.6
1	B	151	ASN	3.6
1	B	99	HIS	3.6
1	B	607	SER	3.5
1	B	65	LYS	3.5
1	B	389	LYS	3.5
1	A	605	ALA	3.5
1	A	102	GLU	3.4
1	B	69	PRO	3.4
1	B	172	GLU	3.4
1	A	528	ARG	3.3
1	B	491	ALA	3.3
1	B	104	SER	3.3
1	B	111	GLY	3.3
1	B	73	ALA	3.2
1	B	61	SER	3.2
1	B	747	ASN	3.2
1	A	68	TYR	3.1
1	B	496	GLU	3.1
1	B	140	ILE	3.1
1	A	496	GLU	3.1
1	B	839	ASP	3.1
1	A	28	GLY	3.0
1	A	38	ASP	3.0
1	A	602	THR	2.9
1	B	59	GLY	2.9
1	B	35	PHE	2.8
1	A	141	VAL	2.8
1	B	67	TRP	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	159	THR	2.8
1	B	469	GLY	2.8
1	A	127	GLU	2.8
1	B	392	TYR	2.7
1	B	136	ILE	2.7
1	A	123	ILE	2.7
1	B	498	THR	2.7
1	A	390	ALA	2.7
1	B	161	PRO	2.7
1	B	354	VAL	2.7
1	B	601	ALA	2.7
1	B	62	GLN	2.7
1	A	63	ASP	2.7
1	B	134	TYR	2.6
1	B	152	ASP	2.6
1	B	602	THR	2.6
1	B	287	ASN	2.6
1	A	568	LYS	2.6
1	A	843	PHE	2.6
1	B	175	TYR	2.6
1	B	842	SER	2.6
1	B	101	TYR	2.6
1	A	783	PRO	2.6
1	B	502	SER	2.6
1	A	159	THR	2.5
1	B	105	ASN	2.5
1	B	112	SER	2.5
1	A	531	ALA	2.5
1	A	119	ASN	2.5
1	B	154	ASN	2.5
1	B	528	ARG	2.5
1	A	384	HIS	2.4
1	B	96	PHE	2.4
1	B	844	ASN	2.4
1	B	592	VAL	2.4
1	B	155	ASN	2.4
1	A	383	ASP	2.4
1	B	525	ASN	2.4
1	B	530	HIS	2.4
1	B	145	SER	2.4
1	B	531	ALA	2.4
1	B	398	GLN	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	111	GLY	2.3
1	B	100	GLY	2.3
1	B	119	ASN	2.3
1	A	841	LYS	2.3
1	B	489	GLY	2.3
1	A	564	ASN	2.3
1	A	850	GLU	2.3
1	A	161	PRO	2.2
1	B	102	GLU	2.2
1	B	72	VAL	2.2
1	B	383	ASP	2.2
1	A	494	ASN	2.2
1	A	99	HIS	2.2
1	A	800	THR	2.2
1	B	492	ASN	2.1
1	A	835	LEU	2.1
1	B	564	ASN	2.1
1	A	175	TYR	2.1
1	A	287	ASN	2.1
1	A	40	SER	2.1
1	B	286	GLY	2.1
1	B	381	ILE	2.1
1	B	98	ASN	2.1
1	B	563	ASP	2.1
1	A	840	ASN	2.0
1	B	30	TRP	2.0
1	B	143	ALA	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(Å ²)	Q<0.9
3	CIT	A	902	13/13	0.80	0.28	47,48,48,48	0
3	CIT	B	902	13/13	0.86	0.23	46,46,46,46	0
4	PEG	A	903	7/7	0.89	0.25	36,36,36,36	0
4	PEG	B	903	7/7	0.89	0.23	29,29,29,29	0
2	NA	A	901	1/1	0.92	0.09	23,23,23,23	0
2	NA	B	901	1/1	0.95	0.15	22,22,22,22	0

6.5 Other polymers [i](#)

There are no such residues in this entry.