



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 4, 2026 – 07:53 PM UTC

PDB ID : 5NLM / pdb_00005nlm
Title : Complex between a UDP-glucosyltransferase from *Polygonum tinctorium* capable of glucosylating indoxyl and indoxyl sulfate
Authors : Welner, D.H.; Hsu, T.; Dueber, J.; Adams, P.D.
Deposited on : 2017-04-04
Resolution : 2.14 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

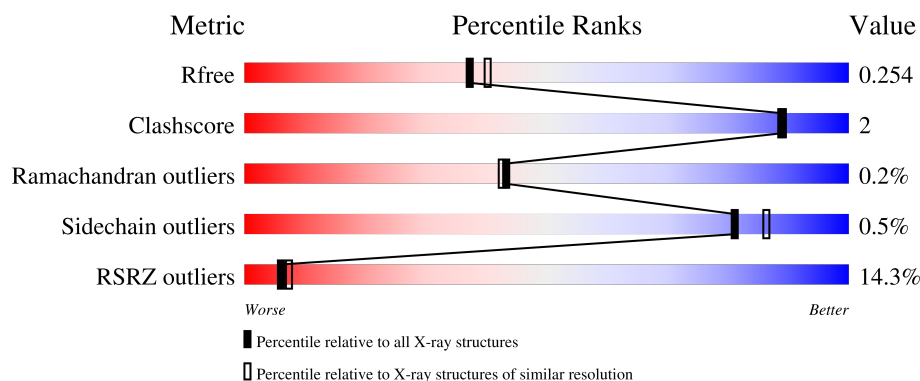
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.14 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}	180053	3689 (2.16-2.12)
Clashscore	190562	3812 (2.16-2.12)
Ramachandran outliers	187476	3773 (2.16-2.12)
Sidechain outliers	187428	3772 (2.16-2.12)
RSRZ outliers	180081	3691 (2.16-2.12)

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	478	6% 94% .
1	B	478	21% 87% 6% • 6%

2 Entry composition [i](#)

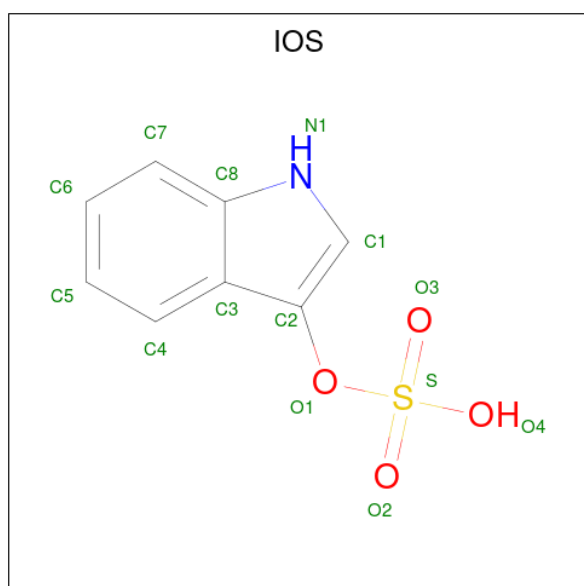
There are 4 unique types of molecules in this entry. The entry contains 14195 atoms, of which 7064 are hydrogens and 0 are deuteriums.

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 indoxyl UDP-glucosyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	459	Total	C	H	N	O	S	0	1	0
			7111	2268	3567	607	655	14			
1	B	448	Total	C	H	N	O	S	0	0	0
			6935	2210	3485	593	634	13			

- Molecule 2 is 3-SULFOOXY-1H-INDOLE (CCD ID: IOS) (formula: C₈H₇NO₄S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	S	0	0
			20	8	6	1	4	1		
2	B	1	Total	C	H	N	O	S	0	0
			20	8	6	1	4	1		

- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total 4	Mg 4	0	0
3	B	1	Total 1	Mg 1	0	0

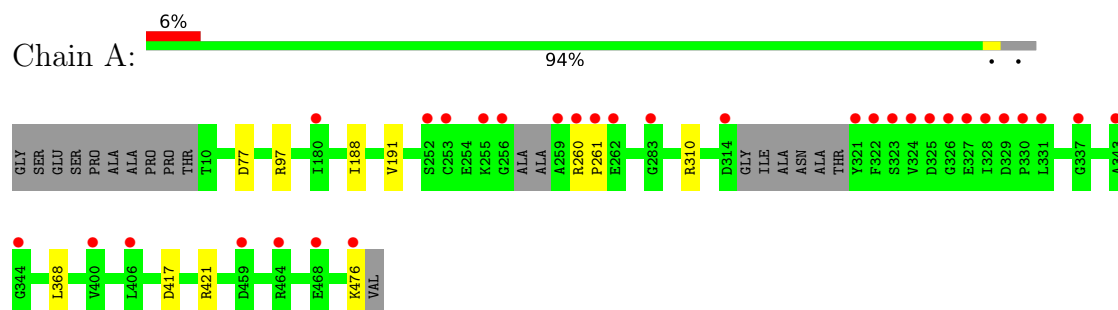
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	90	Total 90	O 90	0	0
4	B	14	Total 14	O 14	0	0

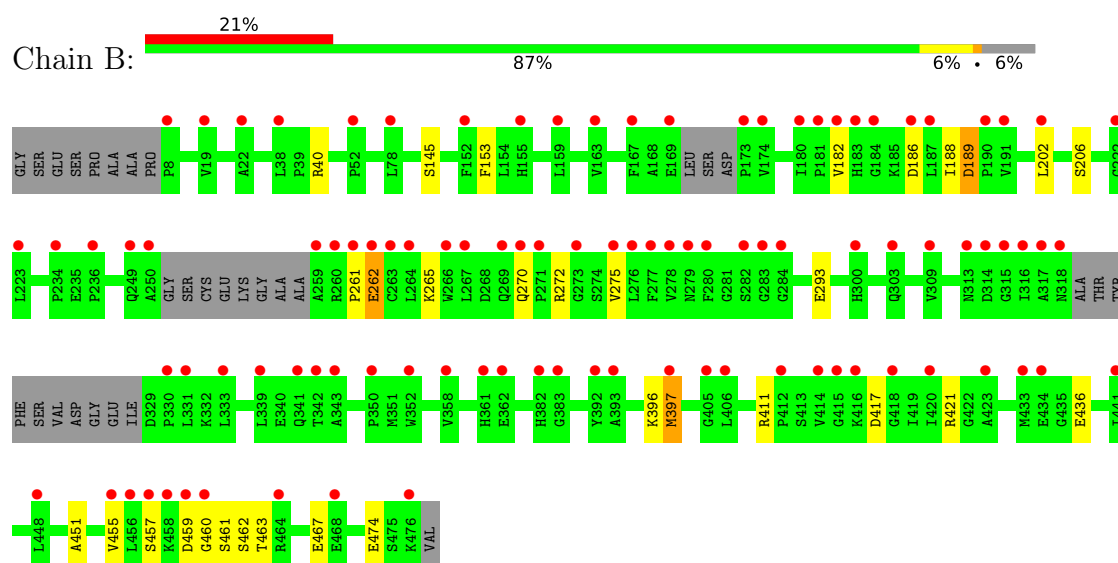
3 Residue-property plots [i](#)

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: indoxyl UDP-glucosyltransferase



- Molecule 1: indoxyl UDP-glucosyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	121.00Å 172.82Å 48.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.41 – 2.14 48.41 – 2.14	Depositor EDS
% Data completeness (in resolution range)	98.0 (48.41-2.14) 98.0 (48.41-2.14)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.31 (at 2.14Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, R_{free}	0.225 , 0.252 0.227 , 0.254	Depositor DCC
R_{free} test set	2780 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å ²)	42.5	Xtriage
Anisotropy	0.486	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 49.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	14195	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.80% 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: IOS, MG

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	0.11	0/3632	0.28	0/4933
1	B	0.11	0/3533	0.28	0/4800
All	All	0.11	0/7165	0.28	0/9733

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	3544	3567	3569	6	0
1	B	3450	3485	3485	21	0
2	A	14	6	7	0	0
2	B	14	6	7	1	0
3	A	4	0	0	0	0
3	B	1	0	0	0	0
4	A	90	0	0	1	0
4	B	14	0	0	0	0
All	All	7131	7064	7068	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 2.

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:B:293:GLU:OE2	1:B:421:ARG:NH1	2.20	0.74
1:B:40:ARG:NH2	1:B:474:GLU:OE1	2.24	0.69
1:B:451:ALA:O	1:B:455:VAL:HG23	1.94	0.68
1:B:460:GLY:O	1:B:462:SER:N	2.33	0.62
1:B:40:ARG:NH1	1:B:467:GLU:OE1	2.35	0.60
1:B:189:ASP:OD1	1:B:189:ASP:N	2.34	0.60
1:B:457:SER:O	1:B:463:THR:OG1	2.12	0.59
1:A:417:ASP:OD2	1:A:421:ARG:NH2	2.38	0.55
1:B:396:LYS:O	1:B:397:MET:HB2	2.08	0.54
1:B:396:LYS:O	1:B:397:MET:CB	2.57	0.52
1:B:145:SER:HB2	2:B:1000:IOS:H5	1.97	0.47
1:B:153:PHE:O	1:B:206:SER:OG	2.28	0.47
1:A:77:ASP:O	1:A:97:ARG:NH1	2.50	0.45
1:B:188:ILE:HD11	1:B:202:LEU:HD21	1.98	0.45
1:B:262:GLU:HA	1:B:262:GLU:OE1	2.16	0.45
1:A:260:ARG:N	1:A:261:PRO:CD	2.80	0.45
1:A:310:ARG:NH1	4:A:1104:HOH:O	2.50	0.44
1:B:182:VAL:HG22	1:B:186:ASP:HB2	1.99	0.44
1:B:270:GLN:OE1	1:B:275:VAL:HG22	2.18	0.44
1:B:455:VAL:O	1:B:460:GLY:O	2.34	0.43
1:B:459:ASP:O	1:B:463:THR:OG1	2.35	0.43
1:B:261:PRO:HA	1:B:265:LYS:HE2	2.00	0.43
1:B:417:ASP:OD2	1:B:421:ARG:NH2	2.49	0.43
1:B:455:VAL:HG13	1:B:461:SER:HB3	2.02	0.42
1:B:436:GLU:OE1	1:B:436:GLU:N	2.54	0.41
1:A:188:ILE:O	1:A:191:VAL:HG22	2.21	0.41
1:A:368:LEU:HD23	1:A:368:LEU:C	2.47	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	454/478 (95%)	436 (96%)	18 (4%)	0	100	100
1	B	440/478 (92%)	414 (94%)	24 (6%)	2 (0%)	24	19
All	All	894/956 (94%)	850 (95%)	42 (5%)	2 (0%)	43	42

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	397	MET
1	B	262	GLU

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	389/399 (98%)	388 (100%)	1 (0%)	86	90
1	B	378/399 (95%)	375 (99%)	3 (1%)	73	79
All	All	767/798 (96%)	763 (100%)	4 (0%)	81	86

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

Mol	Chain	Res	Type
1	A	476	LYS
1	B	189	ASP
1	B	272	ARG
1	B	411	ARG

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

Mol	Chain	Res	Type
1	A	446	GLN
1	B	446	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 5 are monoatomic - leaving 2 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	$\# Z > 2$
2	IOS	A	1000	-	15,15,15	2.20	5 (33%)	18,22,22	1.28	2 (11%)
2	IOS	B	1000	-	15,15,15	2.19	5 (33%)	18,22,22	1.14	2 (11%)

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
2	IOS	A	1000	-	-	0/1/5/5	0/2/2/2
2	IOS	B	1000	-	-	1/1/5/5	0/2/2/2

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1000	IOS	O1-S	5.04	1.66	1.58

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1000	IOS	O1-S	4.92	1.65	1.58
2	B	1000	IOS	C3-C2	-4.56	1.41	1.46
2	A	1000	IOS	C3-C2	-4.44	1.41	1.46
2	A	1000	IOS	C1-N1	-2.97	1.32	1.37
2	B	1000	IOS	C1-N1	-2.90	1.32	1.37
2	B	1000	IOS	C3-C8	-2.51	1.38	1.41
2	A	1000	IOS	C3-C8	-2.47	1.38	1.41
2	A	1000	IOS	C8-N1	-2.13	1.34	1.37
2	B	1000	IOS	C8-N1	-2.13	1.34	1.37

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1000	IOS	O3-S-O2	-2.94	100.91	112.24
2	B	1000	IOS	C7-C8-C3	-2.30	119.96	122.19
2	A	1000	IOS	C7-C8-C3	-2.27	119.99	122.19
2	B	1000	IOS	O4-S-O2	-2.10	101.21	108.56

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1000	IOS	C2-O1-S-O4

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1000	IOS	1	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	459/478 (96%)	0.43	31 (6%) 23 27	25, 57, 102, 143	1 (0%)
1	B	448/478 (93%)	1.38	99 (22%) 2 3	52, 98, 145, 192	0
All	All	907/956 (94%)	0.90	130 (14%) 6 7	25, 77, 135, 192	1 (0%)

All (130) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	256	GLY	6.7
1	A	321	TYR	5.8
1	A	322	PHE	5.8
1	B	250	ALA	5.6
1	A	324	VAL	5.2
1	B	261	PRO	5.0
1	B	316	ILE	5.0
1	B	393	ALA	4.9
1	B	333	LEU	4.8
1	B	278	VAL	4.7
1	B	279	ASN	4.6
1	B	223	LEU	4.4
1	B	8	PRO	4.4
1	B	191	VAL	4.4
1	B	266	TRP	4.3
1	B	167	PHE	4.1
1	A	328	ILE	4.1
1	A	325	ASP	4.0
1	B	317	ALA	4.0
1	B	173	PRO	3.9
1	A	323	SER	3.9
1	B	182	VAL	3.9
1	B	280	PHE	3.8
1	A	476	LYS	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	259	ALA	3.7
1	B	282	SER	3.6
1	B	202	LEU	3.6
1	B	456	LEU	3.5
1	B	283	GLY	3.5
1	B	392	TYR	3.5
1	B	277	PHE	3.5
1	A	326	GLY	3.5
1	B	358	VAL	3.5
1	B	183	HIS	3.4
1	A	180	ILE	3.4
1	B	267	LEU	3.3
1	B	276	LEU	3.3
1	A	464	ARG	3.3
1	B	416	LYS	3.3
1	B	159	LEU	3.2
1	B	180	ILE	3.2
1	B	271	PRO	3.2
1	B	190	PRO	3.2
1	B	263	CYS	3.2
1	B	339	LEU	3.1
1	A	253	CYS	3.1
1	B	459	ASP	3.1
1	B	397	MET	3.0
1	B	269	GLN	2.9
1	A	259	ALA	2.9
1	B	184	GLY	2.8
1	B	476	LYS	2.8
1	B	270	GLN	2.8
1	B	22	ALA	2.8
1	B	78	LEU	2.8
1	A	252	SER	2.7
1	B	458	LYS	2.7
1	B	260	ARG	2.7
1	B	434	GLU	2.7
1	B	433	MET	2.7
1	B	174	VAL	2.7
1	B	455	VAL	2.6
1	B	448	LEU	2.6
1	A	400	VAL	2.6
1	B	406	LEU	2.6
1	A	262	GLU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	318	ASN	2.5
1	B	412	PRO	2.5
1	A	331	LEU	2.5
1	B	314	ASP	2.5
1	B	383	GLY	2.5
1	B	415	GLY	2.5
1	B	262	GLU	2.5
1	A	255	LYS	2.5
1	B	300	HIS	2.5
1	A	327	GLU	2.4
1	A	260	ARG	2.4
1	A	283	GLY	2.4
1	B	303	GLN	2.4
1	A	329	ASP	2.4
1	A	459	ASP	2.4
1	B	52	PRO	2.4
1	A	314	ASP	2.4
1	B	163	VAL	2.4
1	B	457	SER	2.4
1	B	187	LEU	2.3
1	B	331	LEU	2.3
1	B	343	ALA	2.3
1	B	423	ALA	2.3
1	B	418	GLY	2.3
1	B	460	GLY	2.3
1	B	181	PRO	2.3
1	B	362	GLU	2.3
1	B	169	GLU	2.2
1	B	361	HIS	2.2
1	B	341	GLN	2.2
1	B	352	TRP	2.2
1	B	330	PRO	2.2
1	B	309	VAL	2.2
1	B	315	GLY	2.2
1	B	236	PRO	2.1
1	A	468	GLU	2.1
1	B	275	VAL	2.1
1	A	344	GLY	2.1
1	B	350	PRO	2.1
1	A	406	LEU	2.1
1	A	337	GLY	2.1
1	B	222	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	273	GLY	2.1
1	B	420	ILE	2.1
1	B	264	LEU	2.1
1	B	468	GLU	2.1
1	B	313	ASN	2.1
1	A	330	PRO	2.1
1	B	284	GLY	2.1
1	B	19	VAL	2.1
1	A	343	ALA	2.1
1	B	186	ASP	2.0
1	B	464	ARG	2.0
1	B	155	HIS	2.0
1	B	382	HIS	2.0
1	A	261	PRO	2.0
1	B	342	THR	2.0
1	B	441	ILE	2.0
1	B	152	PHE	2.0
1	B	414	VAL	2.0
1	B	38	LEU	2.0
1	B	249	GLN	2.0
1	B	234	PRO	2.0
1	B	405	GLY	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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
3	MG	B	1001	1/1	0.55	0.19	52,52,52,52	1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	MG	A	1004	1/1	0.72	0.18	100,100,100,100	0
3	MG	A	1001	1/1	0.77	0.24	88,88,88,88	0
2	IOS	A	1000	14/14	0.81	0.22	113,124,140,141	0
2	IOS	B	1000	14/14	0.85	0.21	153,160,195,195	0
3	MG	A	1003	1/1	0.86	0.14	70,70,70,70	0
3	MG	A	1002	1/1	0.99	0.03	29,29,29,29	0

6.5 Other polymers [i](#)

There are no such residues in this entry.