wwPDB X-ray Structure Validation Summary Report (i)

Sep 28, 2024 – 04:53 PM EDT

PDB ID	:	1YW0
Title	:	Crystal structure of the tryptophan 2,3-dioxygenase from Xanthomonas campestris. Northeast Structural Genomics Target XcR13.
Authors	:	Vorobiev, S.M.; Abashidze, M.; Forouhar, F.; Kuzin, A.; Xiao, R.; Ciano, M.; Aton, T.B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2005-02-16
Resolution	:	2.70 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) 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 (i)) were used in the production of this report:

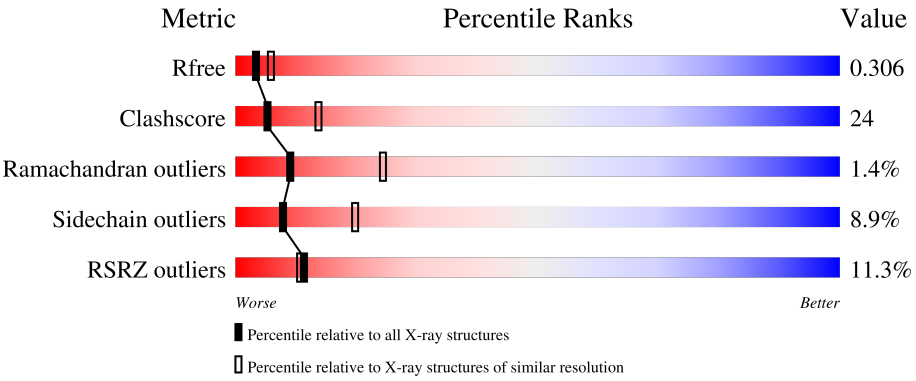
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	164625	3333 (2.70-2.70)
Clashscore	180529	3684 (2.70-2.70)
Ramachandran outliers	177936	3633 (2.70-2.70)
Sidechain outliers	177891	3633 (2.70-2.70)
RSRZ outliers	164620	3333 (2.70-2.70)

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% 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	276	7% 51% 32% • • 12%
1	B	276	12% 51% 32% • 12%
1	C	276	10% 52% 30% 5% • 12%
1	D	276	9% 49% 33% • • 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8274 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 tryptophan 2,3-dioxygenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	243	Total	C	N	O	S	Se	0	0	0
			2043	1316	358	362	2	5			
1	B	242	Total	C	N	O	S	Se	0	0	0
			2035	1312	356	360	2	5			
1	C	243	Total	C	N	O	S	Se	0	0	0
			2043	1316	358	362	2	5			
1	D	243	Total	C	N	O	S	Se	0	0	0
			2040	1315	357	361	2	5			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	MSE	MET	modified residue	UNP Q8PDA8
A	114	MSE	MET	modified residue	UNP Q8PDA8
A	145	MSE	MET	modified residue	UNP Q8PDA8
A	241	MSE	MET	modified residue	UNP Q8PDA8
A	245	MSE	MET	modified residue	UNP Q8PDA8
B	49	MSE	MET	modified residue	UNP Q8PDA8
B	114	MSE	MET	modified residue	UNP Q8PDA8
B	145	MSE	MET	modified residue	UNP Q8PDA8
B	241	MSE	MET	modified residue	UNP Q8PDA8
B	245	MSE	MET	modified residue	UNP Q8PDA8
C	49	MSE	MET	modified residue	UNP Q8PDA8
C	114	MSE	MET	modified residue	UNP Q8PDA8
C	145	MSE	MET	modified residue	UNP Q8PDA8
C	241	MSE	MET	modified residue	UNP Q8PDA8
C	245	MSE	MET	modified residue	UNP Q8PDA8
D	49	MSE	MET	modified residue	UNP Q8PDA8
D	114	MSE	MET	modified residue	UNP Q8PDA8
D	145	MSE	MET	modified residue	UNP Q8PDA8
D	241	MSE	MET	modified residue	UNP Q8PDA8
D	245	MSE	MET	modified residue	UNP Q8PDA8

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Mg 2 2	0	0
2	B	3	Total Mg 3 3	0	0
2	C	2	Total Mg 2 2	0	0
2	D	1	Total Mg 1 1	0	0

- Molecule 3 is water.

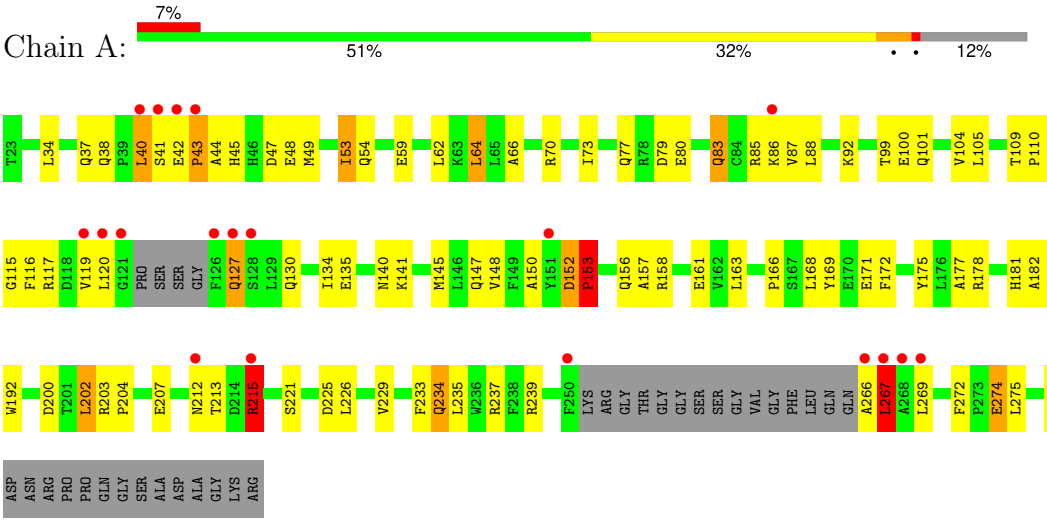
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	31	Total O 31 31	0	0
3	B	25	Total O 25 25	0	0
3	C	20	Total O 20 20	0	0
3	D	29	Total O 29 29	0	0

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.

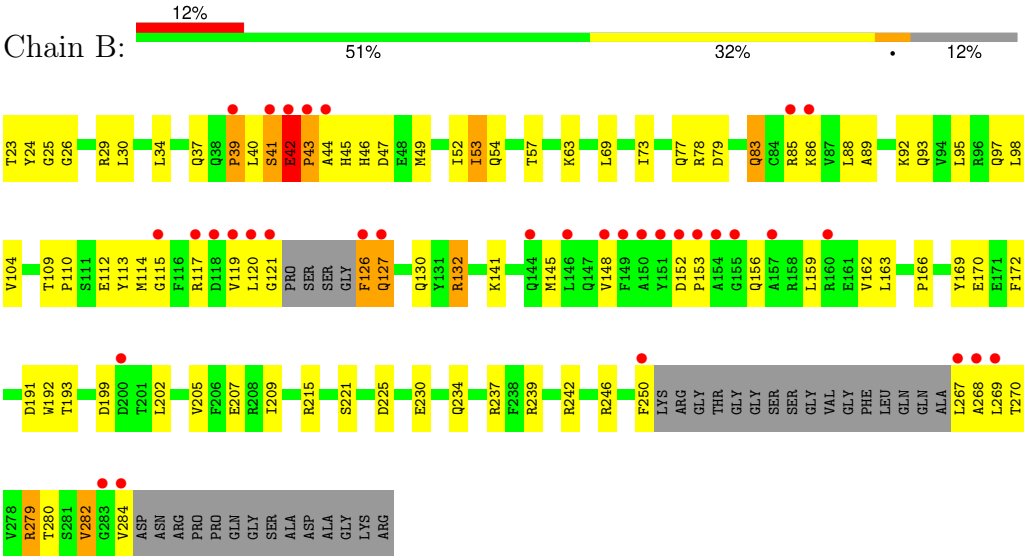
• Molecule 1: tryptophan 2,3-dioxygenase

Chain A:



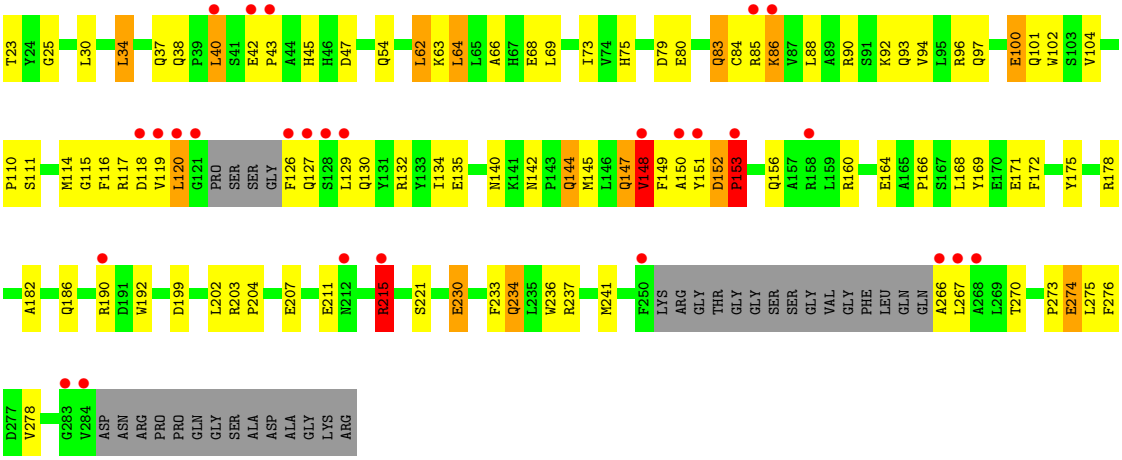
• Molecule 1: tryptophan 2,3-dioxygenase

Chain B:



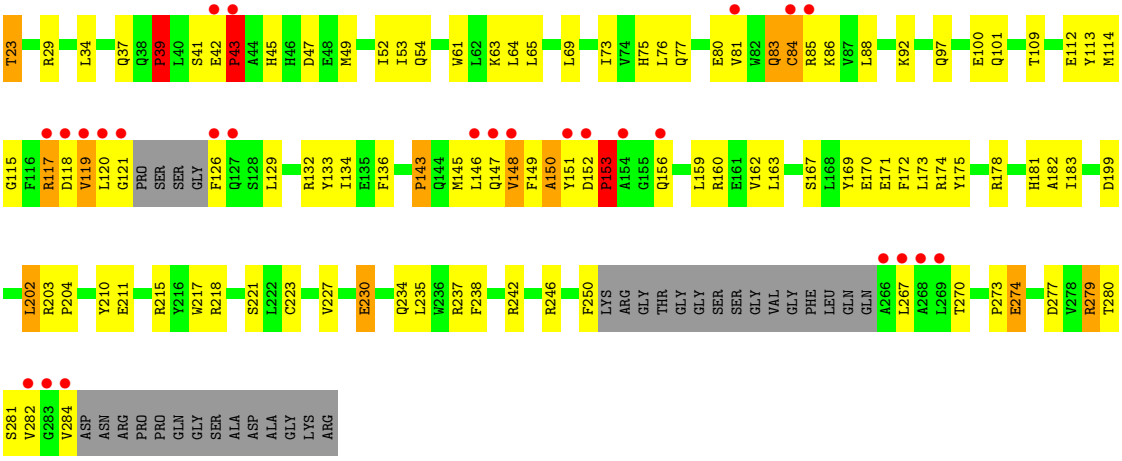
• Molecule 1: tryptophan 2,3-dioxygenase

Chain C:



• Molecule 1: tryptophan 2,3-dioxygenase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	89.70Å 110.22Å 149.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.32 – 2.70 29.32 – 2.70	Depositor EDS
% Data completeness (in resolution range)	81.8 (29.32-2.70) 84.9 (29.32-2.70)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 2.68Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.250 , 0.294 0.262 , 0.306	Depositor DCC
R_{free} test set	1510 reflections (3.85%)	wwPDB-VP
Wilson B-factor (Å^2)	21.8	Xtriage
Anisotropy	0.240	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 43.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	8274	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 48.78 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.1917e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: 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	0.54	1/2090 (0.0%)	0.78	7/2825 (0.2%)
1	B	0.49	1/2082 (0.0%)	0.80	6/2814 (0.2%)
1	C	0.55	0/2090	0.75	6/2825 (0.2%)
1	D	0.56	2/2087 (0.1%)	0.81	9/2821 (0.3%)
All	All	0.54	4/8349 (0.0%)	0.78	28/11285 (0.2%)

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	B	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	42	GLU	CB-CG	-6.87	1.39	1.52
1	D	42	GLU	CA-C	-5.85	1.37	1.52
1	D	43	PRO	N-CD	-5.82	1.39	1.47
1	A	267	LEU	CA-CB	-5.01	1.42	1.53

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	43	PRO	CA-N-CD	-16.77	88.02	111.50
1	D	42	GLU	C-N-CD	-12.68	92.71	120.60
1	B	42	GLU	C-N-CD	-10.99	96.43	120.60
1	A	267	LEU	N-CA-CB	-10.36	89.69	110.40
1	B	43	PRO	N-CA-C	-9.89	86.40	112.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	42	GLU	Mainchain

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	2043	0	1998	102	0
1	B	2035	0	1989	96	0
1	C	2043	0	1998	117	0
1	D	2040	0	1994	113	0
2	A	2	0	0	0	0
2	B	3	0	0	0	0
2	C	2	0	0	0	0
2	D	1	0	0	0	0
3	A	31	0	0	1	0
3	B	25	0	0	1	0
3	C	20	0	0	0	0
3	D	29	0	0	2	0
All	All	8274	0	7979	384	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 24.

The worst 5 of 384 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:C:150:ALA:CB	1:C:156:GLN:HE22	1.32	1.41
1:C:150:ALA:HA	1:C:156:GLN:NE2	1.56	1.20
1:C:150:ALA:CB	1:C:156:GLN:NE2	2.02	1.19
1:C:215:ARG:O	1:C:215:ARG:HD2	1.40	1.18
1:C:150:ALA:CA	1:C:156:GLN:NE2	2.11	1.14

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	237/276 (86%)	220 (93%)	13 (6%)	4 (2%)	7 20
1	B	236/276 (86%)	213 (90%)	21 (9%)	2 (1%)	16 38
1	C	237/276 (86%)	219 (92%)	16 (7%)	2 (1%)	16 38
1	D	237/276 (86%)	214 (90%)	18 (8%)	5 (2%)	5 15
All	All	947/1104 (86%)	866 (91%)	68 (7%)	13 (1%)	9 24

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	148	VAL
1	D	150	ALA
1	B	282	VAL
1	D	148	VAL
1	D	182	ALA

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	216/234 (92%)	199 (92%)	17 (8%)	10 25
1	B	215/234 (92%)	198 (92%)	17 (8%)	10 25
1	C	216/234 (92%)	194 (90%)	22 (10%)	6 15
1	D	215/234 (92%)	194 (90%)	21 (10%)	6 16
All	All	862/936 (92%)	785 (91%)	77 (9%)	8 20

5 of 77 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	39	PRO
1	D	215	ARG
1	D	64	LEU
1	D	119	VAL
1	D	274	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	83	GLN
1	C	232	GLN
1	C	101	GLN
1	C	140	ASN
1	D	32	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 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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, 95^{th} 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	238/276 (86%)	0.28	20 (8%) 18 17	4, 19, 53, 78	0
1	B	237/276 (85%)	0.54	34 (14%) 7 7	3, 22, 68, 96	0
1	C	238/276 (86%)	0.35	27 (11%) 11 11	5, 18, 56, 91	0
1	D	238/276 (86%)	0.47	26 (10%) 12 11	4, 22, 65, 88	0
All	All	951/1104 (86%)	0.41	107 (11%) 11 11	3, 20, 63, 96	0

The worst 5 of 107 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	266	ALA	11.2
1	B	284	VAL	8.7
1	D	284	VAL	7.6
1	C	284	VAL	7.3
1	B	151	TYR	5.9

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 monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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
2	MG	A	406	1/1	0.91	0.08	19,19,19,19	0
2	MG	B	405	1/1	0.92	0.07	13,13,13,13	0
2	MG	C	407	1/1	0.92	0.30	20,20,20,20	0
2	MG	C	400	1/1	0.94	0.15	5,5,5,5	0
2	MG	D	402	1/1	0.94	0.09	14,14,14,14	0
2	MG	B	404	1/1	0.95	0.16	15,15,15,15	0
2	MG	B	403	1/1	0.96	0.08	15,15,15,15	0
2	MG	A	401	1/1	0.96	0.11	10,10,10,10	0

6.5 Other polymers (i)

There are no such residues in this entry.