



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 09:43 PM UTC

PDB ID : 2GUG / pdb_00002gug
Title : NAD-dependent formate dehydrogenase from Pseudomonas sp.101 in complex with formate
Authors : Filippova, E.V.; Polyakov, K.M.; Tikhonova, T.V.; Boiko, K.M.; Tishkov, V.I.; Popov, V.O.
Deposited on : 2006-04-30
Resolution : 2.28 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 ⓘ 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

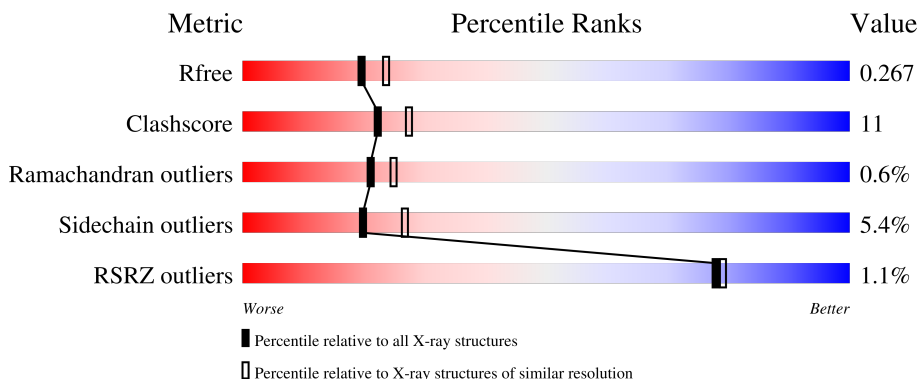
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.28 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	9078 (2.30-2.26)
Clashscore	190562	9802 (2.30-2.26)
Ramachandran outliers	187476	9690 (2.30-2.26)
Sidechain outliers	187428	9691 (2.30-2.26)
RSRZ outliers	180081	9085 (2.30-2.26)

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	401	 70% 20% • 8%
1	B	401	 70% 20% • 8%
1	C	401	 64% 24% • 8%
1	D	401	 64% 23% • 10%

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
3	FMT	A	479	-	-	X	-
4	PEG	A	476	-	-	X	-
4	PEG	A	477	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11994 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 Formate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	368	2884	1824	510	536	14	0	2	0
1	B	369	2906	1838	510	544	14	0	2	0
1	C	369	2901	1836	507	544	14	0	1	0
1	D	362	2837	1797	496	530	14	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

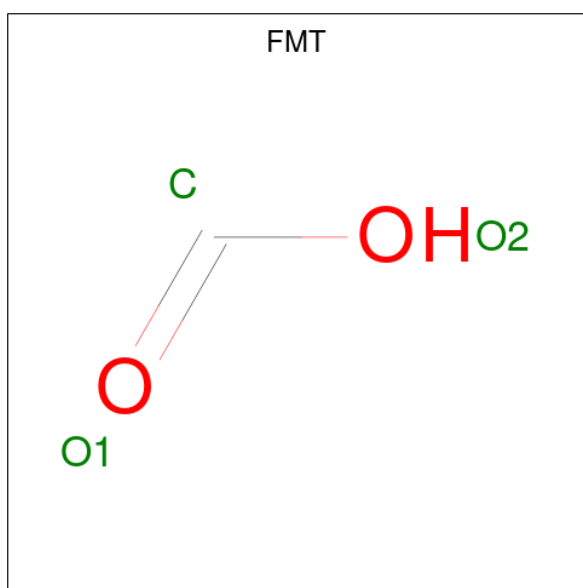
Chain	Residue	Modelled	Actual	Comment	Reference
A	354	OCS	CYS	modified residue	UNP P33160
B	354	OCS	CYS	modified residue	UNP P33160
C	354	OCS	CYS	modified residue	UNP P33160
D	354	OCS	CYS	modified residue	UNP P33160

- Molecule 2 is TETRAETHYLENE GLYCOL (CCD ID: PG4) (formula: C₈H₁₈O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	8	5		
2	D	1	Total	C	O	0	0
			13	8	5		

- Molecule 3 is FORMIC ACID (CCD ID: FMT) (formula: CH₂O₂).



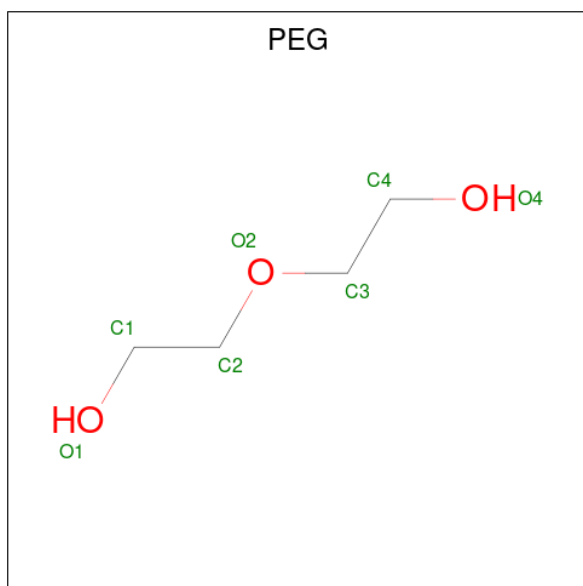
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			3	1	2		
3	B	1	Total	C	O	0	0
			3	1	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			3	1	2		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: C₄H₁₀O₃).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	112	Total	O	0	0
			112	112		
5	B	106	Total	O	0	0
			106	106		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	87	Total O 87 87	0	0
5	D	84	Total O 84 84	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	117.00Å 54.98Å 128.91Å 90.00° 95.74° 90.00°	Depositor
Resolution (Å)	100.00 – 2.28 100.00 – 2.28	Depositor EDS
% Data completeness (in resolution range)	97.5 (100.00-2.28) 97.5 (100.00-2.28)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.00 (at 2.27Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.196 , 0.269 0.197 , 0.267	Depositor DCC
R_{free} test set	3696 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	32.8	Xtrriage
Anisotropy	0.874	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 40.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11994	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.69% 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: OCS, FMT, 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	1.17	2/2956 (0.1%)	1.21	6/4025 (0.1%)
1	B	1.16	4/2980 (0.1%)	1.21	13/4062 (0.3%)
1	C	1.13	3/2970 (0.1%)	1.20	12/4049 (0.3%)
1	D	1.15	3/2897 (0.1%)	1.14	4/3947 (0.1%)
All	All	1.15	12/11803 (0.1%)	1.19	35/16083 (0.2%)

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	339	THR	N-CA	7.19	1.54	1.46
1	D	197	VAL	CA-CB	-6.99	1.45	1.53
1	D	156	MET	C-O	-6.37	1.16	1.24
1	A	211	ALA	C-O	-6.22	1.18	1.24
1	C	185	HIS	CA-C	5.60	1.60	1.53

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	255	CYS	N-CA-C	7.67	126.75	109.81
1	C	99	TRP	CA-C-N	7.54	127.81	119.90
1	C	99	TRP	C-N-CA	7.54	127.81	119.90
1	C	367	ILE	N-CA-C	-6.97	106.25	111.90
1	D	38	THR	CA-C-N	-6.88	112.56	119.92

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	2884	0	2844	68	0
1	B	2906	0	2857	42	0
1	C	2901	0	2850	72	0
1	D	2837	0	2796	65	0
2	A	13	0	18	0	0
2	D	13	0	18	5	0
3	A	3	0	1	2	0
3	B	3	0	1	0	0
3	C	3	0	1	1	0
4	A	21	0	30	24	0
4	B	7	0	10	0	0
4	C	14	0	20	0	0
5	A	112	0	0	2	0
5	B	106	0	0	3	0
5	C	87	0	0	2	0
5	D	84	0	0	3	0
All	All	11994	0	11446	245	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 11.

The worst 5 of 245 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:255:CYS:SG	1:A:256:PRO:HD3	1.81	1.20
1:A:202:ILE:HD11	4:A:476:PEG:H31	1.25	1.12
1:C:271:LYS:NZ	1:C:272:LEU:HG	1.66	1.08
1:D:122:ILE:HD11	1:D:144:TYR:H	1.23	1.01
1:A:150:VAL:HG21	4:A:476:PEG:H21	1.42	0.99

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	365/401 (91%)	345 (94%)	16 (4%)	4 (1%)	11	11
1	B	368/401 (92%)	347 (94%)	20 (5%)	1 (0%)	36	44
1	C	367/401 (92%)	335 (91%)	30 (8%)	2 (0%)	24	29
1	D	357/401 (89%)	335 (94%)	21 (6%)	1 (0%)	36	44
All	All	1457/1604 (91%)	1362 (94%)	87 (6%)	8 (0%)	21	29

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	255	CYS
1	C	260	GLU
1	A	317	LYS
1	A	242	GLU
1	B	198	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	307/328 (94%)	293 (95%)	14 (5%)	24	34
1	B	312/328 (95%)	297 (95%)	15 (5%)	23	32
1	C	311/328 (95%)	292 (94%)	19 (6%)	17	22
1	D	303/328 (92%)	284 (94%)	19 (6%)	16	21
All	All	1233/1312 (94%)	1166 (95%)	67 (5%)	20	27

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	223	HIS
1	D	232	GLU
1	D	359	ARG
1	B	223	HIS
1	B	144	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	254	ASN
1	D	369	GLN
1	B	263	HIS
1	B	369	GLN
1	C	66	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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
1	OCS	B	354	1	6,8,9	1.44	1 (16%)	7,11,13	2.02	2 (28%)
1	OCS	C	354	1	6,8,9	1.28	1 (16%)	7,11,13	1.48	1 (14%)
1	OCS	A	354	1	6,8,9	0.84	0	7,11,13	2.19	2 (28%)
1	OCS	D	354	1	6,8,9	1.07	0	7,11,13	1.71	3 (42%)

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
1	OCS	B	354	1	-	3/4/7/9	-
1	OCS	C	354	1	-	0/4/7/9	-
1	OCS	A	354	1	-	0/4/7/9	-
1	OCS	D	354	1	-	0/4/7/9	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	354	OCS	CB-CA	-2.79	1.51	1.53
1	C	354	OCS	OD1-SG	-2.15	1.39	1.45

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	354	OCS	OD1-SG-CB	4.81	113.94	106.76
1	B	354	OCS	CA-CB-SG	-3.48	108.43	113.61
1	D	354	OCS	OD2-SG-CB	3.10	111.96	105.97
1	C	354	OCS	OD3-SG-CB	2.57	110.59	106.76
1	D	354	OCS	OD3-SG-CB	2.19	110.03	106.76

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	354	OCS	CA-CB-SG-OD2
1	B	354	OCS	CA-CB-SG-OD1
1	B	354	OCS	CA-CB-SG-OD3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	354	OCS	1	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry

11 ligands are modelled in this entry.

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
4	PEG	B	481	-	6,6,6	0.80	0	5,5,5	0.61	0
4	PEG	A	480	-	6,6,6	0.57	0	5,5,5	1.02	0
4	PEG	A	477	-	6,6,6	1.14	0	5,5,5	1.78	2 (40%)
3	FMT	C	401	-	2,2,2	0.46	0	1,1,1	0.33	0
3	FMT	B	401	-	2,2,2	0.65	0	1,1,1	0.08	0
3	FMT	A	479	-	2,2,2	0.41	0	1,1,1	0.00	0
4	PEG	C	483	-	6,6,6	0.44	0	5,5,5	0.47	0
4	PEG	A	476	-	6,6,6	0.78	0	5,5,5	0.90	0
2	PG4	D	480	-	12,12,12	0.75	0	11,11,11	0.91	0
4	PEG	C	482	-	6,6,6	0.66	0	5,5,5	0.47	0
2	PG4	A	478	-	12,12,12	0.68	0	11,11,11	0.51	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
4	PEG	B	481	-	-	4/4/4/4	-
4	PEG	A	480	-	-	4/4/4/4	-
4	PEG	A	477	-	-	2/4/4/4	-
4	PEG	C	483	-	-	1/4/4/4	-
4	PEG	A	476	-	-	3/4/4/4	-
2	PG4	D	480	-	-	6/10/10/10	-
4	PEG	C	482	-	-	4/4/4/4	-
2	PG4	A	478	-	-	6/10/10/10	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	477	PEG	C3-O2-C2	2.73	125.19	113.26
4	A	477	PEG	O4-C4-C3	2.05	123.88	111.82

There are no chirality outliers.

5 of 30 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	477	PEG	O2-C3-C4-O4
4	A	480	PEG	O2-C3-C4-O4
2	D	480	PG4	O1-C1-C2-O2
4	B	481	PEG	O2-C3-C4-O4
4	A	476	PEG	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 30 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	480	PEG	1	0
4	A	477	PEG	13	0
3	C	401	FMT	1	0
3	A	479	FMT	2	0
4	A	476	PEG	10	0
2	D	480	PG4	5	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	367/401 (91%)	-0.37	4 (1%) 78 79	17, 32, 48, 70	2 (0%)
1	B	368/401 (91%)	-0.29	5 (1%) 73 75	17, 34, 51, 73	2 (0%)
1	C	368/401 (91%)	-0.24	4 (1%) 78 79	19, 36, 53, 73	1 (0%)
1	D	361/401 (90%)	-0.19	3 (0%) 82 84	21, 38, 57, 72	0
All	All	1464/1604 (91%)	-0.27	16 (1%) 78 79	17, 35, 54, 73	5 (0%)

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	260	GLU	4.1
1	A	263	HIS	3.6
1	A	256	PRO	3.0
1	C	259	PRO	2.9
1	D	256	PRO	2.9

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	OCS	D	354	9/10	0.96	0.07	44,46,48,49	0
1	OCS	B	354	9/10	0.97	0.07	31,36,39,39	0
1	OCS	A	354	9/10	0.97	0.06	33,35,38,40	0
1	OCS	C	354	9/10	0.98	0.06	36,40,42,43	0

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(Å ²)	Q<0.9
4	PEG	C	482	7/7	0.64	0.14	60,62,65,66	0
4	PEG	B	481	7/7	0.77	0.12	55,57,58,60	0
4	PEG	A	476	7/7	0.78	0.21	40,47,51,57	0
3	FMT	B	401	3/3	0.83	0.10	52,52,52,53	0
2	PG4	A	478	13/13	0.85	0.11	45,54,58,58	0
3	FMT	C	401	3/3	0.87	0.12	39,39,42,43	0
4	PEG	A	477	7/7	0.87	0.18	27,31,38,40	0
2	PG4	D	480	13/13	0.91	0.08	31,35,45,45	0
4	PEG	A	480	7/7	0.91	0.08	35,37,42,46	0
4	PEG	C	483	7/7	0.91	0.10	42,46,50,52	0
3	FMT	A	479	3/3	0.93	0.08	31,31,34,36	0

6.5 Other polymers [i](#)

There are no such residues in this entry.