WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 10, 2024 – 06:41 PM EST

PDB ID	:	2IDH
Title	:	Crystal Structure of human FE65 WW domain
Authors	:	Meiyappan, M.; Birrane, G.; Ladias, J.A.A.
Deposited on	:	2006-09-15
Resolution	:	2.28 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

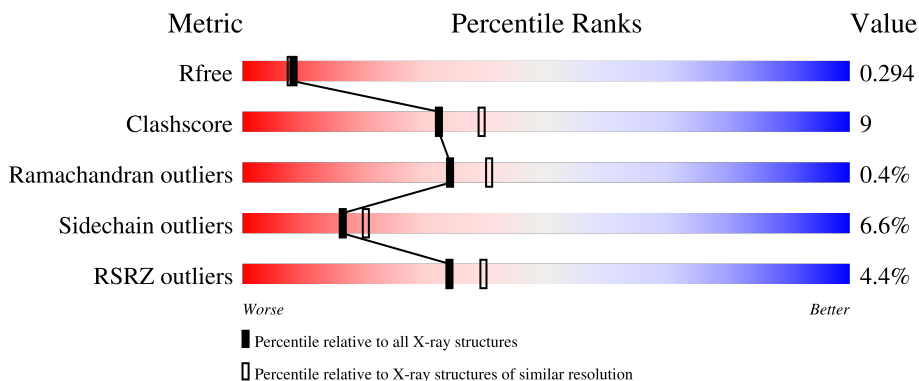|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.28 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	Percentile Ranks	Value
Rfree		0.294
Clashscore		9
Ramachandran outliers		0.4%
Sidechain outliers		6.6%
RSRZ outliers		4.4%

Worse   *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (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 <=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	38	3% 50% 26% • 21%
1	B	38	8% 61% 16% 5% 18%
1	C	38	3% 63% 13% • 21%
1	D	38	5% 74% 13% • 11%
1	E	38	71% 5% • 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	38	50% 21% 8% 21%
1	G	38	5% 71% 8% • 18%
1	H	38	5% 74% 11% • 13%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2217 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 Amyloid beta A4 protein-binding family B member 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	30	Total	C	N	O	S	0	0	0
			247	160	40	46	1			
1	B	31	Total	C	N	O	S	0	0	0
			251	162	41	47	1			
1	C	30	Total	C	N	O	S	0	1	0
			255	165	43	46	1			
1	D	34	Total	C	N	O	S	0	0	0
			272	173	47	51	1			
1	E	30	Total	C	N	O	S	0	0	0
			247	160	40	46	1			
1	F	30	Total	C	N	O	S	0	0	0
			247	160	40	46	1			
1	G	31	Total	C	N	O	S	0	1	0
			256	166	41	47	2			
1	H	33	Total	C	N	O	S	0	0	0
			261	167	43	50	1			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	252	GLY	-	expression tag	UNP O00213
B	252	GLY	-	expression tag	UNP O00213
C	252	GLY	-	expression tag	UNP O00213
D	252	GLY	-	expression tag	UNP O00213
E	252	GLY	-	expression tag	UNP O00213
F	252	GLY	-	expression tag	UNP O00213
G	252	GLY	-	expression tag	UNP O00213
H	252	GLY	-	expression tag	UNP O00213

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).

PG4

C8 OH O5

O4 C7

C5 C6

C4 O O3

O2 C3

C1

C2

HO

O1
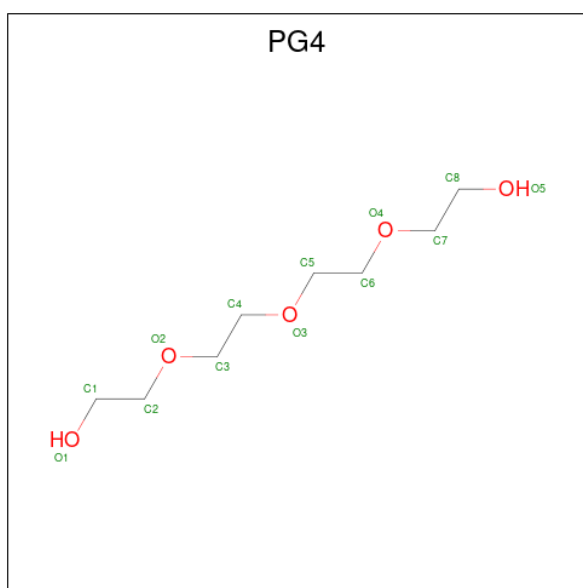
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			7	4	3		
3	C	1	Total	C	O	0	0
			7	4	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total 10	C 6	O 4	0	0
3	E	1	Total 7	C 4	O 3	0	0
3	F	1	Total 7	C 4	O 3	0	0
3	H	1	Total 7	C 4	O 3	0	0
3	H	1	Total 7	C 4	O 3	0	0

- Molecule 4 is water.

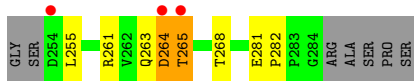| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	12	Total 12	O 12	0	0
4	B	6	Total 6	O 6	0	0
4	C	7	Total 7	O 7	0	0
4	D	15	Total 15	O 15	0	0
4	E	16	Total 16	O 16	0	0
4	F	22	Total 22	O 22	0	0
4	G	17	Total 17	O 17	0	0
4	H	24	Total 24	O 24	0	0

3 Residue-property plots ⓘ

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.
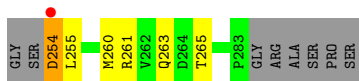
• Molecule 1: Amyloid beta A4 protein-binding family B member 1

Chain A:

• Molecule 1: Amyloid beta A4 protein-binding family B member 1

Chain B:

• Molecule 1: Amyloid beta A4 protein-binding family B member 1

Chain C:

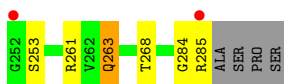• Molecule 1: Amyloid beta A4 protein-binding family B member 1
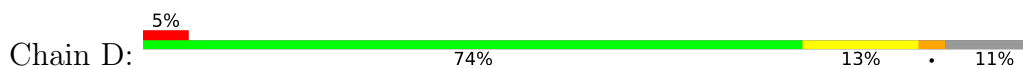
Chain D:

• Molecule 1: Amyloid beta A4 protein-binding family B member 1
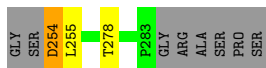
Chain E:

• Molecule 1: Amyloid beta A4 protein-binding family B member 1

Chain F:

50%   21%   8%   21%

• Molecule 1: Amyloid beta A4 protein-binding family B member 1

Chain G:

5%
71%   8%   ·   18%

• Molecule 1: Amyloid beta A4 protein-binding family B member 1

Chain H:

5%
74%   11%   ·   13%

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	75.61Å 75.61Å 226.49Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.29 – 2.28 37.29 – 2.28	Depositor EDS
% Data completeness (in resolution range)	99.6 (37.29-2.28) 99.6 (37.29-2.28)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	6.89 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.217 , 0.282 0.230 , 0.294	Depositor DCC
R_{free} test set	924 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	39.9	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 52.1	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	0.93	EDS
Total number of atoms	2217	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.49% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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: SO4, PG4

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	0.98	0/259	0.86	0/359
1	B	0.97	0/263	0.97	1/364 (0.3%)
1	C	0.91	0/270	0.99	0/373
1	D	0.98	0/284	0.83	0/391
1	E	0.93	0/259	0.87	0/359
1	F	1.09	2/259 (0.8%)	0.97	1/359 (0.3%)
1	G	0.98	0/271	0.89	1/374 (0.3%)
1	H	1.11	0/273	1.18	2/377 (0.5%)
All	All	0.99	2/2138 (0.1%)	0.95	5/2956 (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
1	C	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	271	TRP	CB-CG	-5.68	1.40	1.50
1	F	276	GLY	N-CA	5.40	1.54	1.46

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	261	ARG	NE-CZ-NH1	9.05	124.83	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	261	ARG	NE-CZ-NH2	-8.72	115.94	120.30
1	B	265	THR	N-CA-C	7.18	130.39	111.00
1	F	275	THR	C-N-CA	-6.27	109.12	122.30
1	G	255	LEU	CB-CG-CD1	5.48	120.32	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	264	ASP	Peptide
1	C	254	ASP	Peptide

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	247	0	219	6	0
1	B	251	0	222	4	0
1	C	255	0	232	9	0
1	D	272	0	243	10	0
1	E	247	0	219	4	0
1	F	247	0	219	5	0
1	G	256	0	231	2	0
1	H	261	0	230	4	0
2	A	5	0	0	1	0
2	D	5	0	0	0	0
3	A	7	0	9	0	0
3	C	7	0	9	0	0
3	D	10	0	13	1	0
3	E	7	0	9	0	0
3	F	7	0	9	0	0
3	H	14	0	18	1	0
4	A	12	0	0	0	0
4	B	6	0	0	1	0
4	C	7	0	0	0	0
4	D	15	0	0	1	0
4	E	16	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	22	0	0	0	0
4	G	17	0	0	0	0
4	H	24	0	0	0	0
All	All	2217	0	1882	35	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 9.

The worst 5 of 35 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:H:252:GLY:N	1:H:253:SER:HA	1.71	1.01
1:C:261[A]:ARG:NH2	1:D:263:GLN:OE1	2.05	0.89
1:E:254:ASP:CG	1:E:254:ASP:O	2.15	0.86
1:H:252:GLY:N	1:H:253:SER:CA	2.41	0.82
1:D:284:GLY:O	1:D:285:ARG:HB2	1.86	0.74

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	28/38 (74%)	27 (96%)	1 (4%)	0	100	100
1	B	29/38 (76%)	27 (93%)	1 (3%)	1 (3%)	3	2
1	C	29/38 (76%)	25 (86%)	4 (14%)	0	100	100
1	D	32/38 (84%)	32 (100%)	0	0	100	100
1	E	28/38 (74%)	28 (100%)	0	0	100	100
1	F	28/38 (74%)	27 (96%)	1 (4%)	0	100	100
1	G	30/38 (79%)	30 (100%)	0	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	H	31/38 (82%)	30 (97%)	1 (3%)	0	100	100
All	All	235/304 (77%)	226 (96%)	8 (3%)	1 (0%)	34	40

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	265	THR

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	26/31 (84%)	22 (85%)	4 (15%)	2	2
1	B	26/31 (84%)	25 (96%)	1 (4%)	33	44
1	C	27/31 (87%)	26 (96%)	1 (4%)	34	45
1	D	28/31 (90%)	27 (96%)	1 (4%)	35	47
1	E	26/31 (84%)	25 (96%)	1 (4%)	33	44
1	F	26/31 (84%)	23 (88%)	3 (12%)	5	5
1	G	27/31 (87%)	26 (96%)	1 (4%)	34	45
1	H	27/31 (87%)	25 (93%)	2 (7%)	13	16
All	All	213/248 (86%)	199 (93%)	14 (7%)	16	20

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	254	ASP
1	F	254	ASP
1	H	262	VAL
1	G	255	LEU
1	H	260	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3　RNA ⓘ

There are no RNA molecules in this entry.

5.4　Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5　Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6　Ligand geometry ⓘ

9 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	#$|Z| > 2$
3	PG4	A	302	-	6,6,12	0.51	0	5,5,11	0.33	0
3	PG4	D	305	-	9,9,12	0.60	0	8,8,11	0.31	0
3	PG4	H	307	-	6,6,12	0.53	0	5,5,11	0.39	0
3	PG4	F	306	-	6,6,12	0.43	0	5,5,11	0.73	0
3	PG4	E	301	-	6,6,12	0.51	0	5,5,11	0.35	0
2	SO4	A	202	-	4,4,4	0.11	0	6,6,6	0.32	0
3	PG4	C	303	-	6,6,12	0.49	0	5,5,11	0.75	0
3	PG4	H	304	-	6,6,12	0.80	0	5,5,11	0.80	0
2	SO4	D	201	-	4,4,4	0.19	0	6,6,6	0.26	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	PG4	A	302	-	-	4/4/4/10	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PG4	D	305	-	-	<mark>4/7/7/10</mark>	-
3	PG4	H	307	-	-	1/4/4/10	-
3	PG4	F	306	-	-	<mark>2/4/4/10</mark>	-
3	PG4	E	301	-	-	<mark>2/4/4/10</mark>	-
3	PG4	C	303	-	-	<mark>2/4/4/10</mark>	-
3	PG4	H	304	-	-	<mark>2/4/4/10</mark>	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	F	306	PG4	O2-C3-C4-O3
3	A	302	PG4	O1-C1-C2-O2
3	H	304	PG4	O2-C3-C4-O3
3	C	303	PG4	O1-C1-C2-O2
3	E	301	PG4	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	305	PG4	1	0
2	A	202	SO4	1	0
3	H	304	PG4	1	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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(Å²)	Q<0.9
1	A	30/38 (78%)	-0.17	1 (3%)	46	52	41, 47, 61, 74	2 (6%)
1	B	31/38 (81%)	0.32	3 (9%)	7	10	43, 49, 65, 74	7 (22%)
1	C	30/38 (78%)	-0.00	1 (3%)	46	52	40, 48, 60, 64	4 (13%)
1	D	34/38 (89%)	0.16	2 (5%)	22	27	45, 49, 69, 69	3 (8%)
1	E	30/38 (78%)	0.03	0	100	100	40, 46, 56, 62	3 (10%)
1	F	30/38 (78%)	-0.13	0	100	100	36, 43, 53, 58	2 (6%)
1	G	31/38 (81%)	-0.10	2 (6%)	18	23	32, 41, 52, 60	2 (6%)
1	H	33/38 (86%)	0.31	2 (6%)	21	26	30, 38, 60, 61	4 (12%)
All	All	249/304 (81%)	0.06	11 (4%)	34	40	30, 47, 63, 74	27 (10%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	252	GLY	5.7
1	D	285	ARG	5.3
1	B	265	THR	4.2
1	H	253	SER	3.7
1	G	254	ASP	3.6

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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, 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	PG4	C	303	7/13	0.63	0.31	80,82,86,86	0
3	PG4	D	305	10/13	0.69	0.24	78,82,88,88	0
3	PG4	H	304	7/13	0.78	0.19	67,69,71,71	0
3	PG4	E	301	7/13	0.79	0.17	84,86,91,91	0
2	SO4	D	201	5/5	0.83	0.20	67,68,69,69	5
3	PG4	H	307	7/13	0.84	0.17	76,78,80,81	0
3	PG4	F	306	7/13	0.85	0.18	57,63,68,69	0
2	SO4	A	202	5/5	0.87	0.22	63,64,64,66	5
3	PG4	A	302	7/13	0.87	0.22	64,69,75,77	0

6.5 Other polymers ⓘ

There are no such residues in this entry.