Full wwPDB X-ray Structure Validation Report (i)

Jun 13, 2024 – 04:23 pm BST

|		
PDB ID	:	8QQE
Title	:	Crystal structure of the complex between DMC1 and the PhePP domain of BRCA2
Authors	:	Miron, S.; Legrand, P.; Zinn-Justin, S.
Deposited on	:	2023-10-04
Resolution	:	3.46 Å(reported)

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 3.46 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.

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
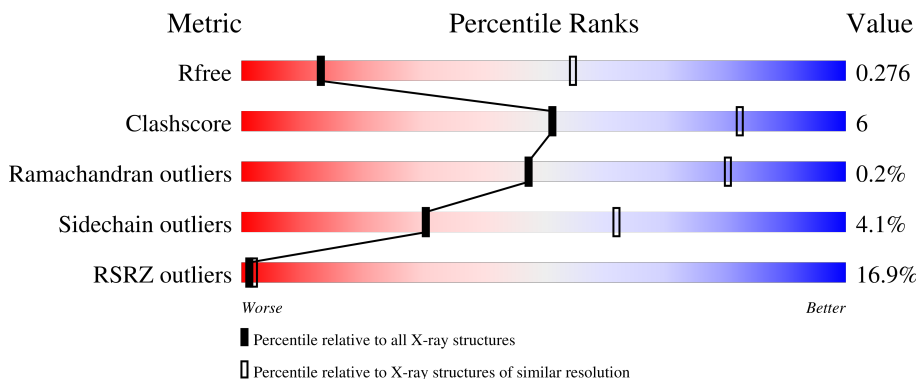
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1291 (3.52-3.40)
Clashscore	141614	1372 (3.52-3.40)
Ramachandran outliers	138981	1337 (3.52-3.40)
Sidechain outliers	138945	1338 (3.52-3.40)
RSRZ outliers	127900	1205 (3.52-3.40)

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	340	21% / 74% / 12% / 13%
1	B	340	5% / 61% / 11% / 27%
2	C	20	20% / 50% / 15% / 35%
2	D	20	10% / 20% / 5% / 75%

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
4	CL	A	502	-	-	X	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4377 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 Meiotic recombination protein DMC1/LIM15 homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	S	0	0	0
			2304	1457	405	431	11			
1	B	247	Total	C	N	O	S	0	0	0
			1929	1223	334	364	8			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP Q14565
B	1	GLY	-	expression tag	UNP Q14565

- Molecule 2 is a protein called Breast cancer type 2 susceptibility protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	13	Total	C	N	O	0	0	0
			102	70	16	16			
2	D	5	Total	C	N	O	0	5	0
			39	29	5	5			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

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: Meiotic recombination protein DMC1/LIM15 homolog
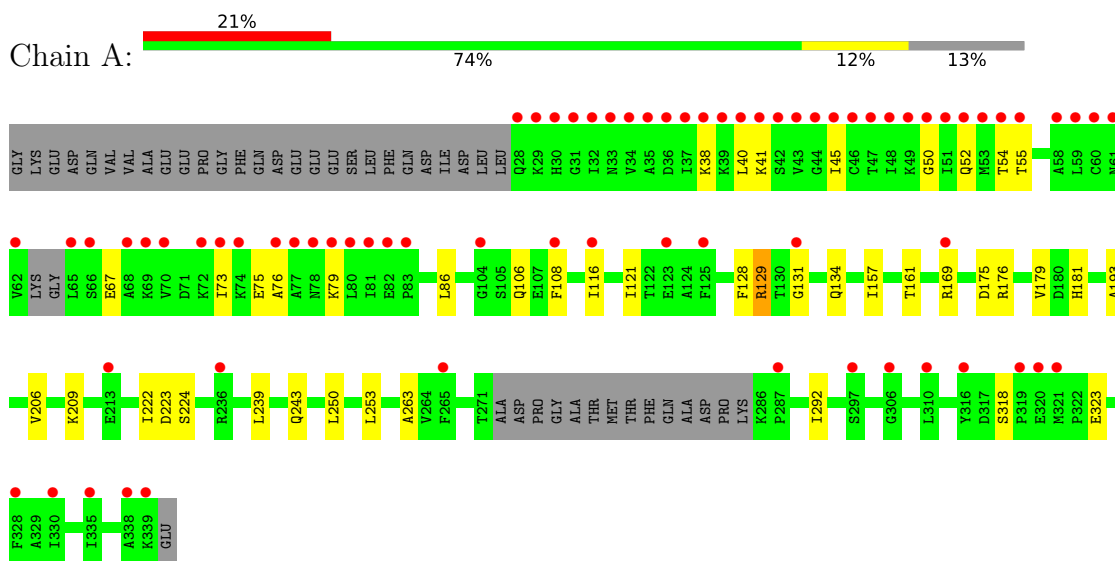
• Molecule 1: Meiotic recombination protein DMC1/LIM15 homolog
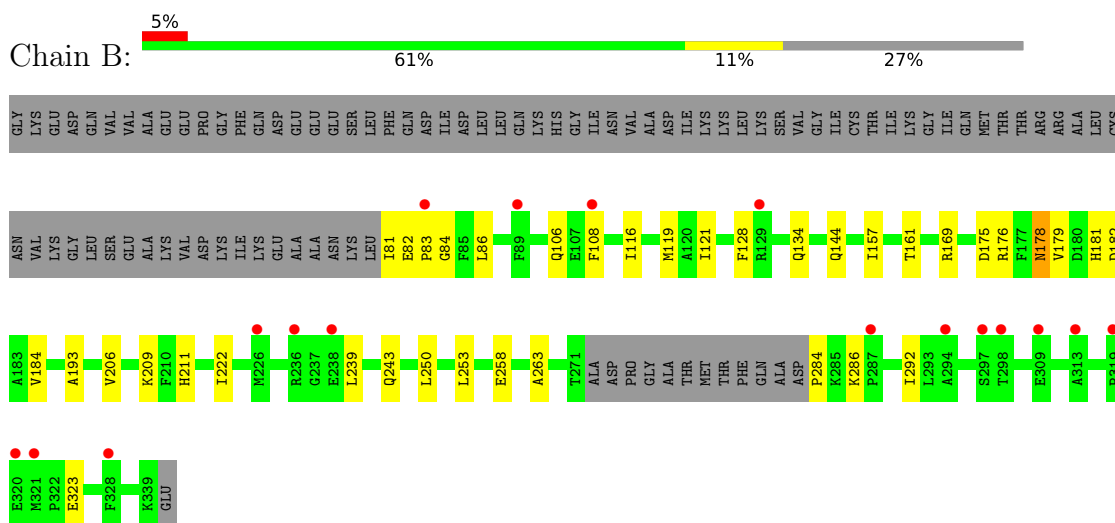
• Molecule 2: Breast cancer type 2 susceptibility protein

• Molecule 2: Breast cancer type 2 susceptibility protein

Chain D:
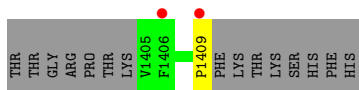
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 4 3 2	Depositor
Cell constants a, b, c, α, β, γ	234.38Å 234.38Å 234.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.84 – 3.46 47.84 – 3.46	Depositor EDS
% Data completeness (in resolution range)	95.4 (47.84-3.46) 95.4 (47.84-3.46)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.89 (at 3.48Å)	Xtriage
Refinement program	BUSTER 2.10.4	Depositor
R, R_{free}	0.248 , 0.261 0.249 , 0.276	Depositor DCC
R_{free} test set	708 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	151.3	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 124.3	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4377	wwPDB-VP
Average B, all atoms (Å^2)	169.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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: CL, 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	0.29	0/2338	0.47	0/3139
1	B	0.30	0/1964	0.50	0/2643
2	C	0.39	0/105	0.46	0/141
2	D	0.35	0/41	0.42	0/57
All	All	0.30	0/4448	0.48	0/5980

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	2304	0	2340	26	0
1	B	1929	0	1917	27	0
2	C	102	0	109	4	0
2	D	39	0	40	1	0
3	A	1	0	0	0	0
4	A	1	0	0	2	0
4	B	1	0	0	1	0
All	All	4377	0	4406	49	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 6.

All (49) 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:131:GLY:HA2	4:A:502:CL:CL	1.88	1.11
1:A:179:VAL:HG13	2:D:1409[A]:PRO:HB2	1.72	0.72
1:A:38:LYS:HG2	1:A:41:LYS:HZ3	1.63	0.63
1:A:129:ARG:HA	1:A:129:ARG:HE	1.69	0.57
1:A:38:LYS:HG2	1:A:41:LYS:NZ	2.21	0.56
1:A:318:SER:HA	1:B:128:PHE:HZ	1.71	0.55
1:A:55:THR:HG23	1:B:211:HIS:CD2	2.41	0.55
1:B:81:ILE:HG22	1:B:83:PRO:HD3	1.90	0.54
1:A:55:THR:CG2	1:B:211:HIS:CD2	2.91	0.53
1:B:284:PRO:HB2	1:B:286:LYS:NZ	2.25	0.52
1:B:239:LEU:O	1:B:243:GLN:HG2	2.10	0.52
1:A:38:LYS:HA	1:A:41:LYS:HD2	1.92	0.52
1:A:239:LEU:O	1:A:243:GLN:HG2	2.10	0.51
1:A:40:LEU:HD11	1:A:73:ILE:HD13	1.92	0.50
1:B:144:GLN:OE1	2:C:2409:PRO:HG3	2.12	0.50
1:B:175:ASP:OD1	1:B:181:HIS:NE2	2.35	0.50
1:A:134:GLN:HG2	4:A:502:CL:CL	2.48	0.50
1:B:284:PRO:CG	1:B:286:LYS:HZ3	2.25	0.50
1:A:116:ILE:HG22	1:A:263:ALA:HB1	1.94	0.49
1:B:116:ILE:HG22	1:B:263:ALA:HB1	1.94	0.49
1:A:40:LEU:HD22	1:A:45:ILE:HD12	1.94	0.48
1:B:284:PRO:HB2	1:B:286:LYS:HZ3	1.78	0.48
1:A:134:GLN:HE22	1:A:169:ARG:NE	2.12	0.47
1:B:182:ASP:HB3	2:C:2413:LYS:HE2	1.95	0.47
1:B:134:GLN:HE22	1:B:169:ARG:NE	2.13	0.47
1:A:175:ASP:OD1	1:A:181:HIS:NE2	2.34	0.46
1:A:223:ASP:HA	1:A:224:SER:HA	1.77	0.46
1:B:222:ILE:HD12	1:B:253:LEU:HD11	1.98	0.46
1:A:222:ILE:HD12	1:A:253:LEU:HD11	1.98	0.46
1:B:284:PRO:CB	1:B:286:LYS:NZ	2.80	0.45
1:A:76:ALA:HA	1:A:79:LYS:HD2	1.98	0.45
1:B:179:VAL:HG11	2:C:2409:PRO:HG2	1.99	0.45
1:A:75:GLU:HG2	1:A:79:LYS:HE3	1.98	0.45
1:B:178:ASN:O	1:B:178:ASN:CG	2.55	0.44
1:A:250:LEU:HD12	1:A:292:ILE:HG22	2.00	0.43
1:B:284:PRO:CB	1:B:286:LYS:HZ3	2.31	0.43
1:B:179:VAL:HG13	2:C:2409:PRO:HB2	2.01	0.43
1:A:121:ILE:HD11	1:A:253:LEU:HB3	2.01	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:157:ILE:HG21	1:B:206:VAL:HG22	2.01	0.43
1:A:50:GLY:O	1:A:54:THR:OG1	2.29	0.42
1:B:250:LEU:HD12	1:B:292:ILE:HG22	2.00	0.42
1:A:157:ILE:HG21	1:A:206:VAL:HG22	2.01	0.42
1:B:128:PHE:O	4:B:401:CL:CL	2.75	0.42
1:B:119:MET:HG2	1:B:258:GLU:HA	2.02	0.42
1:B:121:ILE:HD11	1:B:253:LEU:HB3	2.02	0.41
1:B:179:VAL:HG12	1:B:184:VAL:HG23	2.02	0.41
1:A:161:THR:HG22	1:A:193:ALA:HB3	2.01	0.41
1:B:161:THR:HG22	1:B:193:ALA:HB3	2.03	0.41
1:A:55:THR:HG21	1:B:211:HIS:CD2	2.57	0.40

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	290/340 (85%)	283 (98%)	7 (2%)	0	100	100
1	B	243/340 (72%)	239 (98%)	3 (1%)	1 (0%)	34	70
2	C	11/20 (55%)	9 (82%)	2 (18%)	0	100	100
2	D	3/20 (15%)	3 (100%)	0	0	100	100
All	All	547/720 (76%)	534 (98%)	12 (2%)	1 (0%)	47	80

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	84	GLY

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	243/279 (87%)	233 (96%)	10 (4%)	30	62
1	B	201/279 (72%)	193 (96%)	8 (4%)	31	63
2	C	12/19 (63%)	11 (92%)	1 (8%)	11	38
2	D	5/19 (26%)	5 (100%)	0	100	100
All	All	461/596 (77%)	442 (96%)	19 (4%)	30	62

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

Mol	Chain	Res	Type
1	A	52	GLN
1	A	67	GLU
1	A	86	LEU
1	A	106	GLN
1	A	108	PHE
1	A	128	PHE
1	A	129	ARG
1	A	176	ARG
1	A	209	LYS
1	A	323	GLU
1	B	82	GLU
1	B	86	LEU
1	B	106	GLN
1	B	108	PHE
1	B	176	ARG
1	B	178	ASN
1	B	209	LYS
1	B	323	GLU
2	C	2403	THR

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	78	ASN
1	B	211	HIS

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 ⓘ

Of 3 ligands modelled in this entry, 3 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 ⓘ

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 (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	296/340 (87%)	1.46	72 (24%)	0	0	134, 169, 258, 261	0
1	B	247/340 (72%)	0.75	17 (6%)	16	19	116, 149, 191, 203	0
2	C	13/20 (65%)	1.22	4 (30%)	0	0	147, 158, 173, 175	0
2	D	5/20 (25%)	1.38	2 (40%)	0	0	217, 218, 221, 221	0
All	All	561/720 (77%)	1.14	95 (16%)	1	2	116, 158, 256, 261	0

All (95) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	45	ILE	8.7
1	A	62	VAL	8.7
1	A	46	CYS	7.7
1	A	36	ASP	7.5
1	A	338	ALA	7.0
1	A	30	HIS	6.7
1	A	80	LEU	6.6
1	A	59	LEU	6.4
1	A	40	LEU	6.3
1	A	47	THR	6.1
1	A	44	GLY	5.9
1	A	52	GLN	5.8
1	A	328	PHE	5.6
1	A	35	ALA	5.5
1	A	33	ASN	5.4
1	A	29	LYS	5.3
1	A	320	GLU	5.1
1	A	51	ILE	5.0
1	A	65	LEU	5.0
1	B	129	ARG	4.9
1	A	66	SER	4.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	61	ASN	4.8
1	A	28	GLN	4.8
1	A	39	LYS	4.7
1	A	34	VAL	4.7
1	A	335	ILE	4.7
1	A	73	ILE	4.6
1	A	48	ILE	4.6
1	A	32	ILE	4.6
1	A	43	VAL	4.5
1	A	60	CYS	4.5
1	A	339	LYS	4.4
1	A	76	ALA	4.4
1	A	42	SER	4.3
1	A	54	THR	4.3
1	A	37	ILE	4.1
1	A	58	ALA	4.0
1	A	77	ALA	3.9
1	A	319	PRO	3.8
1	A	213	GLU	3.8
1	A	38	LYS	3.8
1	A	31	GLY	3.8
1	A	50	GLY	3.6
1	B	319	PRO	3.6
1	A	53	MET	3.5
1	B	238	GLU	3.5
1	A	125	PHE	3.5
1	B	321	MET	3.5
2	C	2413	LYS	3.4
1	B	287	PRO	3.2
2	C	2402	PRO	3.2
1	A	78	ASN	3.2
1	A	55	THR	3.2
1	A	79	LYS	3.1
1	B	320	GLU	3.0
1	A	83	PRO	3.0
1	A	81	ILE	2.9
1	A	49	LYS	2.8
1	B	297	SER	2.8
1	B	89	PHE	2.8
1	A	70	VAL	2.8
1	A	82	GLU	2.7
1	A	169	ARG	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	74	LYS	2.7
1	A	72	LYS	2.7
1	B	309	GLU	2.6
1	A	287	PRO	2.6
1	A	316	TYR	2.6
2	D	1406[A]	PHE	2.5
1	B	83	PRO	2.5
1	A	41	LYS	2.5
1	B	313	ALA	2.5
1	A	108	PHE	2.4
1	A	116	ILE	2.4
1	B	294	ALA	2.4
1	A	68	ALA	2.4
2	C	2414	SER	2.4
1	A	321	MET	2.4
1	A	131	GLY	2.4
1	B	328	PHE	2.4
1	B	298	THR	2.4
1	A	123	GLU	2.4
1	A	69	LYS	2.3
1	A	306	GLY	2.3
1	A	236	ARG	2.2
1	B	236	ARG	2.2
1	A	330	ILE	2.2
1	B	226	MET	2.2
1	A	297	SER	2.1
2	C	2404	LYS	2.1
1	B	108	PHE	2.1
1	A	310	LEU	2.1
1	A	104	GLY	2.1
1	A	265	PHE	2.0
2	D	1409[A]	PRO	2.0

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, 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	CL	A	502	1/1	0.82	0.21	171,171,171,171	0
4	CL	B	401	1/1	0.84	0.43	169,169,169,169	0
3	MG	A	501	1/1	0.96	0.23	175,175,175,175	0

6.5 Other polymers ⓘ

There are no such residues in this entry.