



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 6, 2023 – 11:53 PM EST

PDB ID : 7SWS
Title : Crystal structure of the chromoprotein amilCP
Authors : Caputo, A.T.; Newman, J.; Scott, C.; Ahmed, H.
Deposited on : 2021-11-21
Resolution : 1.64 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 : 4.02b-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

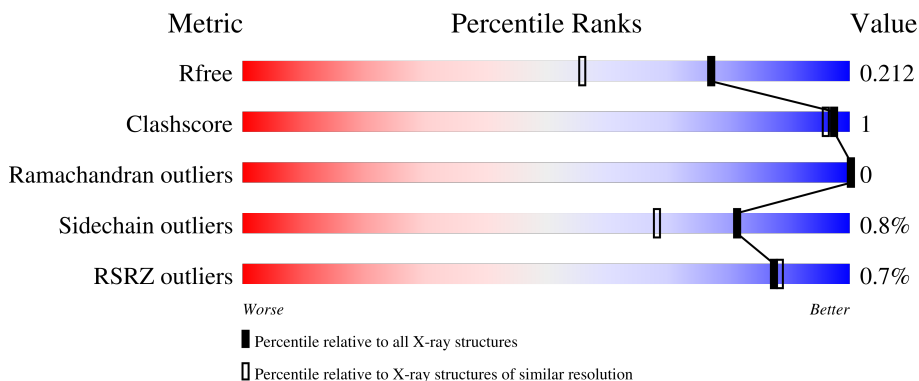
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.64 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.



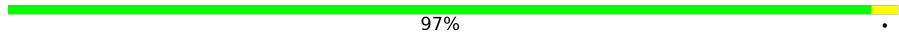
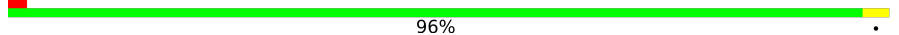
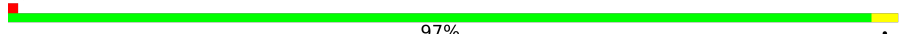
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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	219	97%
1	B	219	97%
1	C	219	95% 5%
1	D	219	97%
1	E	219	97%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	219	 97%
1	G	219	 96%
1	H	219	 97%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 28715 atoms, of which 13542 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 Chromoprotein amilCP.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	218	3451	1126	1695	289	329	12	1695	1	0
1	B	218	3454	1127	1697	289	328	13	1697	1	0
1	C	218	3437	1122	1688	288	327	12	1688	0	0
1	D	218	3451	1126	1695	289	329	12	1695	1	0
1	E	218	3454	1127	1697	289	328	13	1697	1	0
1	F	218	3437	1122	1688	288	327	12	1688	0	0
1	G	218	3448	1125	1694	289	328	12	1694	1	0
1	H	218	3437	1122	1688	288	327	12	1688	0	0

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Cl	0	0
			1	1		
2	B	1	Total	Cl	0	0
			1	1		
2	C	1	Total	Cl	0	0
			1	1		
2	D	1	Total	Cl	0	0
			1	1		
2	E	1	Total	Cl	0	0
			1	1		
2	F	1	Total	Cl	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	1	Total 1	Cl 1	0	0
2	H	2	Total 2	Cl 2	0	0

- Molecule 3 is BROMIDE ION (three-letter code: BR) (formula: Br).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total 1	Br 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	163	Total 163	O 163	0	0
4	B	150	Total 150	O 150	0	0
4	C	150	Total 150	O 150	0	0
4	D	165	Total 165	O 165	0	0
4	E	129	Total 129	O 129	0	0
4	F	142	Total 142	O 142	0	0
4	G	100	Total 100	O 100	0	0
4	H	137	Total 137	O 137	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: Chromoprotein amilCP

Chain A:  97%



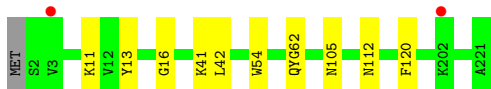
- Molecule 1: Chromoprotein amilCP

Chain B:  97%



- Molecule 1: Chromoprotein amilCP

Chain C:  95% 5%



- Molecule 1: Chromoprotein amilCP

Chain D:  97%



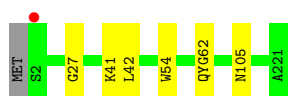
- Molecule 1: Chromoprotein amilCP

Chain E:  97%



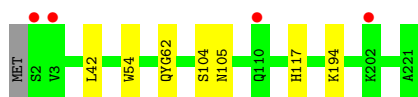
- Molecule 1: Chromoprotein amilCP

Chain F:  97%



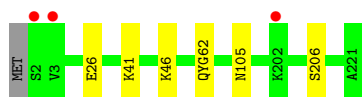
• Molecule 1: Chromoprotein amilCP

Chain G:  96%



• Molecule 1: Chromoprotein amilCP

Chain H:  97%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	71.52Å 131.74Å 93.67Å 90.00° 100.75° 90.00°	Depositor
Resolution (Å)	92.03 – 1.64 92.03 – 1.64	Depositor EDS
% Data completeness (in resolution range)	52.6 (92.03-1.64) 52.6 (92.03-1.64)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.81 (at 1.64Å)	Xtrriage
Refinement program	BUSTER 2.10.4	Depositor
R, R_{free}	0.194 , 0.218 0.187 , 0.212	Depositor DCC
R_{free} test set	5415 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å ²)	16.7	Xtrriage
Anisotropy	0.057	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 41.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	28715	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.26 % 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 7.2862e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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: CL, BR, CRQ

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	0.44	0/1780	0.64	0/2407
1	B	0.44	0/1781	0.63	0/2407
1	C	0.47	0/1773	0.63	0/2397
1	D	0.44	0/1780	0.62	0/2407
1	E	0.42	0/1781	0.62	0/2407
1	F	0.41	0/1773	0.61	0/2397
1	G	0.40	0/1781	0.62	0/2408
1	H	0.43	0/1773	0.63	0/2397
All	All	0.43	0/14222	0.62	0/19227

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	A	0	1
1	B	0	1
1	D	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	10	TYR	Sidechain
1	B	10	TYR	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	D	10	TYR	Sidechain

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	1756	1695	1694	1	0
1	B	1757	1697	1696	3	0
1	C	1749	1688	1688	7	0
1	D	1756	1695	1695	3	0
1	E	1757	1697	1696	3	0
1	F	1749	1688	1688	3	0
1	G	1754	1694	1694	2	0
1	H	1749	1688	1688	2	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	2	0	0	0	0
3	C	1	0	0	1	0
4	A	163	0	0	0	0
4	B	150	0	0	2	0
4	C	150	0	0	0	0
4	D	165	0	0	1	0
4	E	129	0	0	1	0
4	F	142	0	0	1	0
4	G	100	0	0	1	0
4	H	137	0	0	1	0
All	All	15173	13542	13539	23	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 1.

The worst 5 of 23 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:F:41:LYS:NZ	4:F:401:HOH:O	2.31	0.61
1:D:42:LEU:HD13	1:D:54:TRP:HZ2	1.68	0.59
1:H:41:LYS:NZ	4:H:401:HOH:O	2.34	0.58
1:A:42:LEU:HD13	1:A:54:TRP:HZ2	1.69	0.57
1:C:41:LYS:HE3	3:C:302:BR:BR	2.61	0.55

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	214/219 (98%)	213 (100%)	1 (0%)	0	100	100
1	B	214/219 (98%)	213 (100%)	1 (0%)	0	100	100
1	C	213/219 (97%)	211 (99%)	2 (1%)	0	100	100
1	D	214/219 (98%)	213 (100%)	1 (0%)	0	100	100
1	E	214/219 (98%)	213 (100%)	1 (0%)	0	100	100
1	F	213/219 (97%)	212 (100%)	1 (0%)	0	100	100
1	G	214/219 (98%)	213 (100%)	1 (0%)	0	100	100
1	H	213/219 (97%)	212 (100%)	1 (0%)	0	100	100
All	All	1709/1752 (98%)	1700 (100%)	9 (0%)	0	100	100

There are no Ramachandran outliers to report.

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	190/190 (100%)	189 (100%)	1 (0%)	88	80
1	B	190/190 (100%)	189 (100%)	1 (0%)	88	80
1	C	189/190 (100%)	188 (100%)	1 (0%)	88	80
1	D	190/190 (100%)	189 (100%)	1 (0%)	88	80
1	E	190/190 (100%)	189 (100%)	1 (0%)	88	80
1	F	189/190 (100%)	188 (100%)	1 (0%)	88	80
1	G	190/190 (100%)	187 (98%)	3 (2%)	62	39
1	H	189/190 (100%)	186 (98%)	3 (2%)	62	39
All	All	1517/1520 (100%)	1505 (99%)	12 (1%)	81	68

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	105	ASN
1	G	194	LYS
1	H	206	SER
1	H	46	LYS
1	D	105	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	ASN
1	D	137	GLN
1	F	137	GLN
1	G	94	ASN
1	H	137	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](#)

8 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	CRQ	C	62	1	24,25,26	1.48	2 (8%)	27,34,36	2.44	6 (22%)
1	CRQ	G	62	1	24,25,26	1.28	1 (4%)	27,34,36	2.38	6 (22%)
1	CRQ	A	62	1	24,25,26	1.46	2 (8%)	27,34,36	2.50	6 (22%)
1	CRQ	H	62	1	24,25,26	1.42	2 (8%)	27,34,36	2.30	6 (22%)
1	CRQ	F	62	1	24,25,26	1.46	2 (8%)	27,34,36	2.36	7 (25%)
1	CRQ	D	62	1	24,25,26	1.33	2 (8%)	27,34,36	2.40	7 (25%)
1	CRQ	E	62	1	24,25,26	1.37	2 (8%)	27,34,36	2.33	6 (22%)
1	CRQ	B	62	1	24,25,26	1.62	2 (8%)	27,34,36	2.38	7 (25%)

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	CRQ	C	62	1	-	3/10/32/33	0/2/2/2
1	CRQ	G	62	1	-	3/10/32/33	0/2/2/2
1	CRQ	A	62	1	-	3/10/32/33	0/2/2/2
1	CRQ	H	62	1	-	3/10/32/33	0/2/2/2
1	CRQ	F	62	1	-	3/10/32/33	0/2/2/2
1	CRQ	D	62	1	-	3/10/32/33	0/2/2/2
1	CRQ	E	62	1	-	3/10/32/33	0/2/2/2
1	CRQ	B	62	1	-	3/10/32/33	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	62	CRQ	CA2-C2	5.82	1.54	1.48
1	C	62	CRQ	CA2-C2	5.61	1.54	1.48
1	A	62	CRQ	CA2-C2	5.27	1.53	1.48
1	H	62	CRQ	CA2-C2	5.25	1.53	1.48
1	F	62	CRQ	CA2-C2	5.11	1.53	1.48

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	62	CRQ	CB2-CA2-C2	8.70	132.67	122.28
1	C	62	CRQ	CB2-CA2-C2	8.56	132.49	122.28
1	D	62	CRQ	CB2-CA2-C2	8.44	132.35	122.28
1	B	62	CRQ	CB2-CA2-C2	8.26	132.14	122.28
1	F	62	CRQ	CB2-CA2-C2	8.25	132.13	122.28

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	62	CRQ	C2-CA2-CB2-CG2
1	B	62	CRQ	C2-CA2-CB2-CG2
1	C	62	CRQ	C2-CA2-CB2-CG2
1	D	62	CRQ	C2-CA2-CB2-CG2
1	E	62	CRQ	C2-CA2-CB2-CG2

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 10 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

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	217/219 (99%)	-0.45	0 100 100	11, 19, 32, 43	0
1	B	217/219 (99%)	-0.36	2 (0%) 84 85	12, 23, 38, 53	0
1	C	217/219 (99%)	-0.47	2 (0%) 84 85	9, 19, 34, 43	0
1	D	217/219 (99%)	-0.51	0 100 100	10, 19, 30, 38	0
1	E	217/219 (99%)	-0.39	1 (0%) 91 91	12, 24, 40, 48	0
1	F	217/219 (99%)	-0.42	1 (0%) 91 91	14, 22, 36, 45	0
1	G	217/219 (99%)	-0.16	4 (1%) 68 69	12, 34, 57, 74	0
1	H	217/219 (99%)	-0.36	3 (1%) 75 76	12, 24, 39, 54	0
All	All	1736/1752 (99%)	-0.39	13 (0%) 87 88	9, 22, 41, 74	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2	SER	3.6
1	B	3	VAL	3.4
1	H	2	SER	3.1
1	H	3	VAL	2.8
1	G	202	LYS	2.8

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	CRQ	G	62	24/25	0.92	0.08	24,31,39,40	15

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
1	CRQ	F	62	24/25	0.95	0.07	14,20,23,25	15
1	CRQ	A	62	24/25	0.96	0.07	11,15,20,21	15
1	CRQ	C	62	24/25	0.96	0.07	10,15,19,21	15
1	CRQ	H	62	24/25	0.96	0.07	15,21,24,25	15
1	CRQ	B	62	24/25	0.97	0.07	14,19,25,28	15
1	CRQ	D	62	24/25	0.97	0.07	9,13,17,17	15
1	CRQ	E	62	24/25	0.97	0.06	16,20,26,30	15

6.3 Carbohydrates [i](#)

There are no monosaccharides 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(\AA^2)	Q<0.9
3	BR	C	302	1/1	0.97	0.05	59,59,59,59	0
2	CL	B	301	1/1	0.98	0.07	36,36,36,36	0
2	CL	C	301	1/1	0.99	0.05	26,26,26,26	0
2	CL	H	301	1/1	0.99	0.11	27,27,27,27	0
2	CL	A	301	1/1	0.99	0.04	38,38,38,38	0
2	CL	F	301	1/1	1.00	0.15	15,15,15,15	0
2	CL	G	301	1/1	1.00	0.09	33,33,33,33	0
2	CL	D	301	1/1	1.00	0.05	26,26,26,26	0
2	CL	H	302	1/1	1.00	0.08	16,16,16,16	0
2	CL	E	301	1/1	1.00	0.10	15,15,15,15	0

6.5 Other polymers [i](#)

There are no such residues in this entry.