Full wwPDB X-ray Structure Validation Report ⓘ

Sep 26, 2023 – 06:14 AM EDT

PDB ID	:	6BIG
Title	:	Crystal structure of cobalt-substituted Synechocystis ACO
Authors	:	Sui, X.; Shi, W.; Kiser, P.D.
Deposited on	:	2017-11-02
Resolution	:	2.21 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

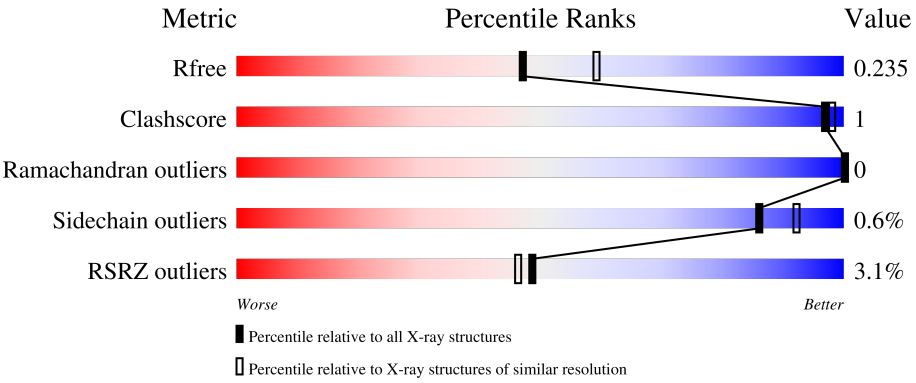
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.21 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	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

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	490	2% 94%
1	B	490	3% 94%
1	C	490	3% 95%
1	D	490	4% 94%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 15788 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 Apocarotenoid-15,15'-oxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	479	Total	C	N	O	S	0	0	0
			3764	2415	649	690	10			
1	B	479	Total	C	N	O	S	0	3	0
			3788	2430	654	694	10			
1	C	479	Total	C	N	O	S	0	0	0
			3768	2417	650	691	10			
1	D	479	Total	C	N	O	S	0	2	0
			3780	2425	651	694	10			

- Molecule 2 is COBALT (II) ION (three-letter code: CO) (formula: Co).

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

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 4 is water.

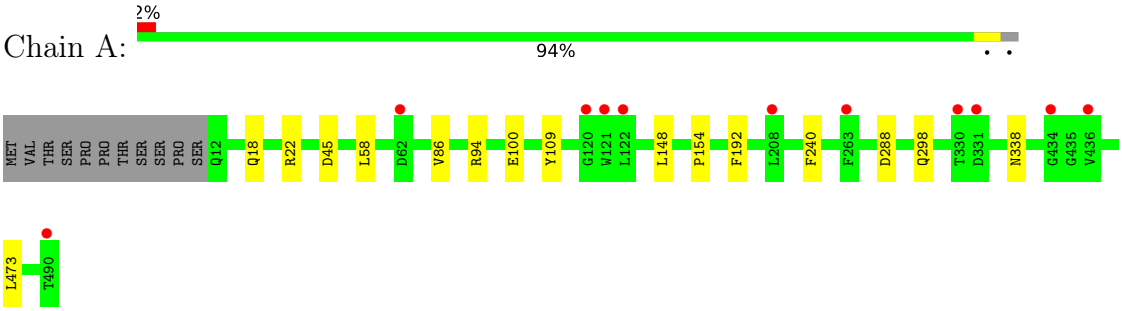
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	192	Total O 192 192	0	0
4	B	170	Total O 170 170	0	0
4	C	173	Total O 173 173	0	0
4	D	146	Total O 146 146	0	0

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.

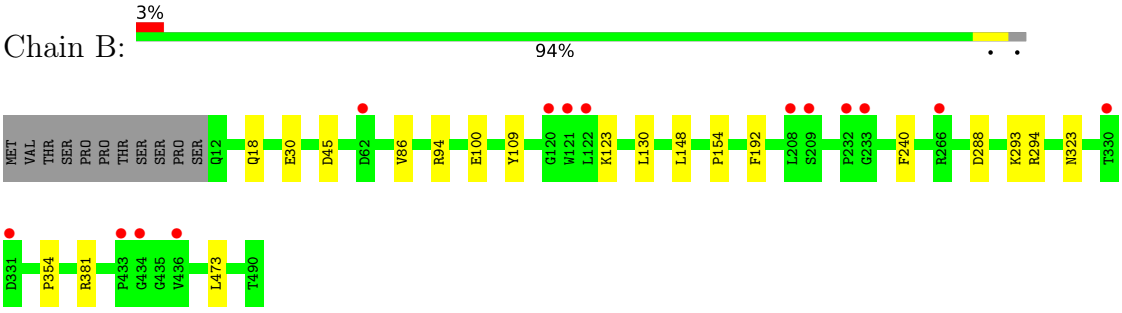
• Molecule 1: Apocarotenoid-15,15'-oxygenase

Chain A:



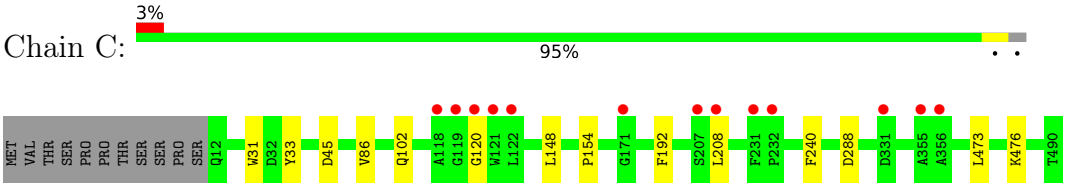
• Molecule 1: Apocarotenoid-15,15'-oxygenase

Chain B:



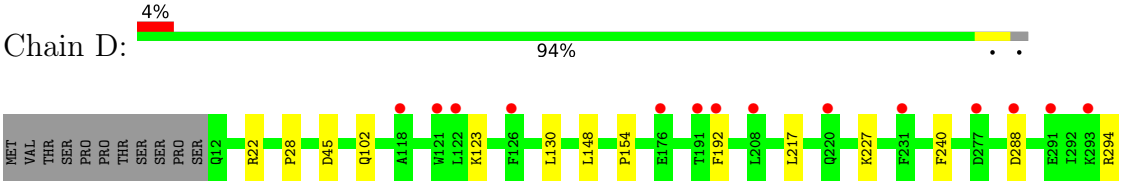
• Molecule 1: Apocarotenoid-15,15'-oxygenase

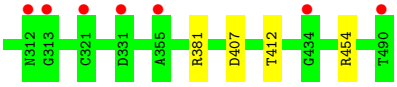
Chain C:



• Molecule 1: Apocarotenoid-15,15'-oxygenase

Chain D:




N312
G313
C321
D331
A355
R381
D407
T412
G434
R454
T490

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	118.07Å 124.53Å 202.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.39 – 2.21 48.39 – 2.21	Depositor EDS
% Data completeness (in resolution range)	94.9 (48.39-2.21) 94.9 (48.39-2.21)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.14 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0085	Depositor
R, R_{free}	0.208 , 0.229 0.214 , 0.235	Depositor DCC
R_{free} test set	6801 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	43.7	Xtriage
Anisotropy	0.252	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 41.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	15788	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.64% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, CO

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	0.71	0/3877	0.79	2/5280 (0.0%)
1	B	0.70	1/3904 (0.0%)	0.81	4/5316 (0.1%)
1	C	0.61	0/3881	0.78	1/5285 (0.0%)
1	D	0.61	0/3896	0.78	3/5306 (0.1%)
All	All	0.66	1/15558 (0.0%)	0.79	10/21187 (0.0%)

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	C	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	323	ASN	CB-CG	-5.47	1.38	1.51

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	94	ARG	NE-CZ-NH1	6.84	123.72	120.30
1	C	208	LEU	CA-CB-CG	6.75	130.84	115.30
1	A	58	LEU	CA-CB-CG	-6.49	100.38	115.30
1	B	94	ARG	NE-CZ-NH2	-5.99	117.31	120.30
1	D	22	ARG	NE-CZ-NH1	5.90	123.25	120.30
1	D	294	ARG	NE-CZ-NH1	5.35	122.97	120.30
1	B	381	ARG	NE-CZ-NH1	5.34	122.97	120.30
1	D	381	ARG	NE-CZ-NH1	5.25	122.93	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	294	ARG	NE-CZ-NH1	5.22	122.91	120.30
1	A	22	ARG	NE-CZ-NH1	5.16	122.88	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	120	GLY	Peptide

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	3764	0	3653	8	0
1	B	3788	0	3679	9	0
1	C	3768	0	3659	6	0
1	D	3780	0	3671	8	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
3	A	2	0	0	0	0
3	B	1	0	0	0	0
4	A	192	0	0	1	1
4	B	170	0	0	3	0
4	C	173	0	0	0	1
4	D	146	0	0	3	0
All	All	15788	0	14662	27	1

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.

All (27) 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:B:293:LYS:HE2	1:B:354:PRO:O	1.89	0.72

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:86:VAL:HG11	1:B:473:LEU:HD13	1.78	0.65
1:A:86:VAL:HG11	1:A:473:LEU:HD13	1.78	0.63
1:C:86:VAL:HG11	1:C:473:LEU:HD13	1.79	0.62
1:B:100[A]:GLU:OE1	1:B:109:TYR:OH	2.15	0.61
1:A:100:GLU:OE1	1:A:109:TYR:OH	2.15	0.59
1:A:298:GLN:HG3	1:B:30:GLU:O	2.07	0.55
4:B:718:HOH:O	1:D:28:PRO:HG2	2.07	0.54
1:D:123:LYS:HG3	4:D:691:HOH:O	2.08	0.53
1:A:192:PHE:CE1	1:A:288:ASP:HB3	2.45	0.51
1:C:192:PHE:CE1	1:C:288:ASP:HB3	2.47	0.50
1:D:148:LEU:HD23	1:D:154:PRO:HB3	1.94	0.50
1:A:148:LEU:HD23	1:A:154:PRO:HB3	1.93	0.49
1:C:148:LEU:HD23	1:C:154:PRO:HB3	1.94	0.49
1:D:130:LEU:HD23	4:D:733:HOH:O	2.12	0.49
1:D:192:PHE:CE1	1:D:288:ASP:HB3	2.48	0.49
1:B:123:LYS:NZ	4:B:604:HOH:O	2.44	0.49
1:B:192:PHE:CE1	1:B:288:ASP:HB3	2.48	0.49
1:B:148:LEU:HD23	1:B:154:PRO:HB3	1.94	0.48
1:B:18:GLN:HG2	1:D:102:GLN:OE1	2.14	0.47
1:D:227:LYS:HE3	4:D:742:HOH:O	2.15	0.47
1:A:18:GLN:HG2	1:C:102:GLN:OE1	2.15	0.46
1:B:130:LEU:HD23	4:B:724:HOH:O	2.16	0.46
1:C:33:TYR:CD2	1:C:476:LYS:HE2	2.53	0.43
1:A:94:ARG:NH1	4:A:602:HOH:O	2.50	0.41
1:A:338:ASN:ND2	1:C:31:TRP:CZ3	2.88	0.40
1:D:407:ASP:HB3	1:D:412:THR:HG22	2.03	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:782:HOH:O	4:C:713:HOH:O[4_446]	2.11	0.09

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	477/490 (97%)	468 (98%)	9 (2%)	0	100	100
1	B	480/490 (98%)	471 (98%)	9 (2%)	0	100	100
1	C	477/490 (97%)	469 (98%)	8 (2%)	0	100	100
1	D	479/490 (98%)	470 (98%)	9 (2%)	0	100	100
All	All	1913/1960 (98%)	1878 (98%)	35 (2%)	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	398/410 (97%)	396 (100%)	2 (0%)	88	94
1	B	401/410 (98%)	399 (100%)	2 (0%)	88	94
1	C	399/410 (97%)	397 (100%)	2 (0%)	88	94
1	D	401/410 (98%)	397 (99%)	4 (1%)	76	85
All	All	1599/1640 (98%)	1589 (99%)	10 (1%)	86	92

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	ASP
1	A	240	PHE
1	B	45	ASP
1	B	240	PHE
1	C	45	ASP
1	C	240	PHE
1	D	45	ASP
1	D	217	LEU
1	D	240	PHE
1	D	454	ARG

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

Of 7 ligands modelled in this entry, 7 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

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	479/490 (97%)	-0.03	11 (2%)	60	58	31, 57, 86, 116	0
1	B	479/490 (97%)	-0.09	14 (2%)	51	49	31, 56, 83, 121	0
1	C	479/490 (97%)	0.01	13 (2%)	54	52	37, 58, 89, 105	0
1	D	479/490 (97%)	0.18	21 (4%)	34	32	38, 63, 98, 125	0
All	All	1916/1960 (97%)	0.02	59 (3%)	49	46	31, 58, 89, 125	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	121	TRP	4.9
1	D	355	ALA	4.9
1	B	121	TRP	4.7
1	D	121	TRP	4.7
1	A	122	LEU	4.6
1	D	192	PHE	4.3
1	C	121	TRP	4.0
1	C	231	PHE	4.0
1	C	208	LEU	4.0
1	B	434	GLY	4.0
1	B	209	SER	3.9
1	B	331	ASP	3.7
1	A	436	VAL	3.6
1	A	330	THR	3.6
1	C	118	ALA	3.4
1	D	220	GLN	3.4
1	D	122	LEU	3.1
1	A	434	GLY	3.1
1	B	330	THR	3.0
1	D	288	ASP	3.0
1	D	313	GLY	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	232	PRO	2.9
1	C	356	ALA	2.8
1	D	331	ASP	2.8
1	C	331	ASP	2.8
1	D	126	PHE	2.8
1	C	120	GLY	2.7
1	D	176	GLU	2.7
1	C	355	ALA	2.7
1	D	434	GLY	2.7
1	D	118	ALA	2.6
1	D	312	ASN	2.6
1	C	119	GLY	2.6
1	A	331	ASP	2.5
1	B	122	LEU	2.5
1	D	321	CYS	2.4
1	A	120	GLY	2.4
1	B	436	VAL	2.4
1	D	291	GLU	2.4
1	D	208	LEU	2.4
1	B	233	GLY	2.4
1	C	232	PRO	2.4
1	D	231	PHE	2.4
1	C	171	GLY	2.3
1	D	191	THR	2.3
1	C	122	LEU	2.3
1	C	207	SER	2.3
1	D	293	LYS	2.2
1	D	490	THR	2.1
1	B	62	ASP	2.1
1	A	490	THR	2.1
1	A	62	ASP	2.1
1	A	208	LEU	2.1
1	B	120	GLY	2.1
1	B	433	PRO	2.1
1	A	263	PHE	2.1
1	D	277[A]	ASP	2.0
1	B	266	ARG	2.0
1	B	208	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

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, 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
3	CL	A	503	1/1	0.94	0.12	66,66,66,66	0
3	CL	B	502	1/1	0.96	0.08	58,58,58,58	0
3	CL	A	502	1/1	0.98	0.11	61,61,61,61	0
2	CO	B	501	1/1	0.99	0.14	34,34,34,34	0
2	CO	A	501	1/1	0.99	0.13	35,35,35,35	0
2	CO	D	501	1/1	1.00	0.18	40,40,40,40	0
2	CO	C	501	1/1	1.00	0.15	39,39,39,39	0

6.5 Other polymers (i)

There are no such residues in this entry.