wwPDB X-ray Structure Validation Summary Report (i)

Oct 17, 2021 – 06:45 AM EDT

PDB ID	:	1KPL
Title	:	Crystal Structure of the ClC Chloride Channel from S. typhimurium
Authors	:	Dutzler, R.; Campbell, E.B.; Cadene, M.; Chait, B.T.; MacKinnon, R.
Deposited on	:	2001-12-31
Resolution	:	3.00 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 (i) symbol.

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

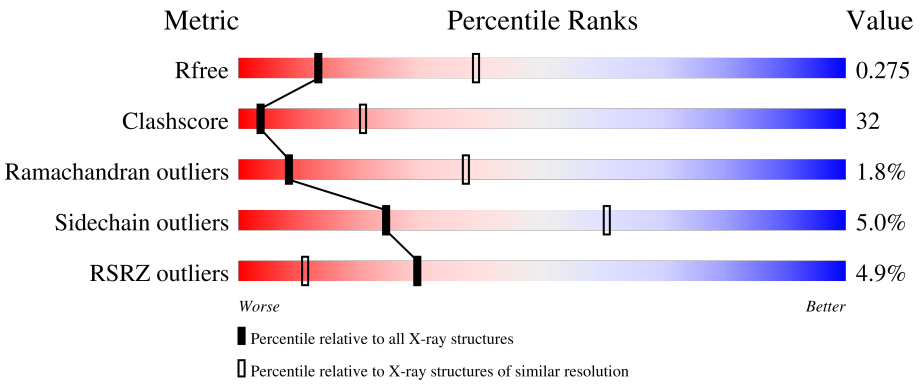
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	473	5%, 44%, 44%, • 9%
1	B	473	5%, 52%, 39%, • 5%
1	C	473	3%, 45%, 42%, • 9%
1	D	473	5%, 51%, 40%, • 5%

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
2	CL	C	603	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 13216 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 putative ClC family, chlorine transport protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	430	3193	2105	528	540	20	0	0	0
1	B	451	3379	2217	574	568	20	0	0	0
1	C	430	3193	2105	528	540	20	0	0	0
1	D	451	3379	2217	574	568	20	0	0	0

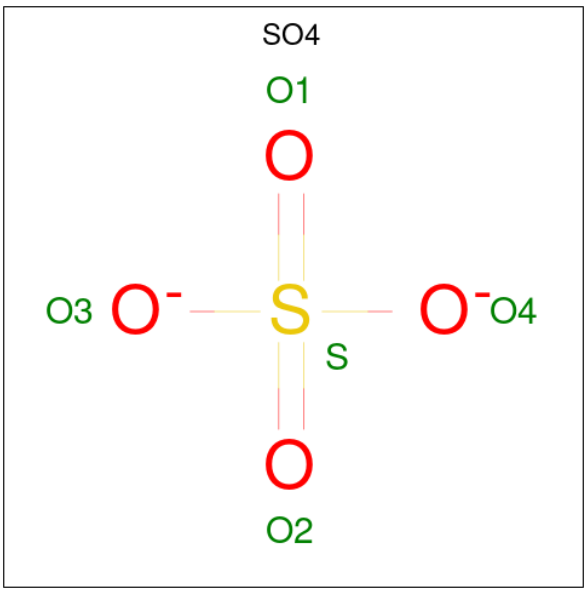
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	LEU	MET	engineered mutation	UNP Q8ZRP8
A	264	VAL	CYS	engineered mutation	UNP Q8ZRP8
B	26	LEU	MET	engineered mutation	UNP Q8ZRP8
B	264	VAL	CYS	engineered mutation	UNP Q8ZRP8
C	26	LEU	MET	engineered mutation	UNP Q8ZRP8
C	264	VAL	CYS	engineered mutation	UNP Q8ZRP8
D	26	LEU	MET	engineered mutation	UNP Q8ZRP8
D	264	VAL	CYS	engineered mutation	UNP Q8ZRP8

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

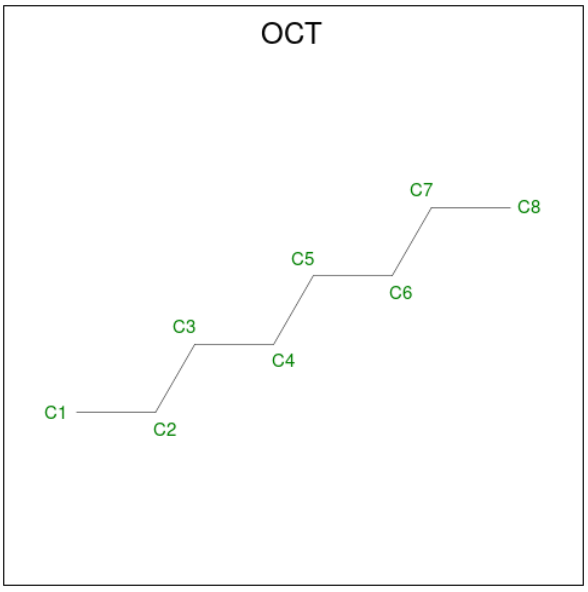
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
2	A	1	1	1	0	0
2	B	1	1	1	0	0
2	C	1	1	1	0	0
2	D	1	1	1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



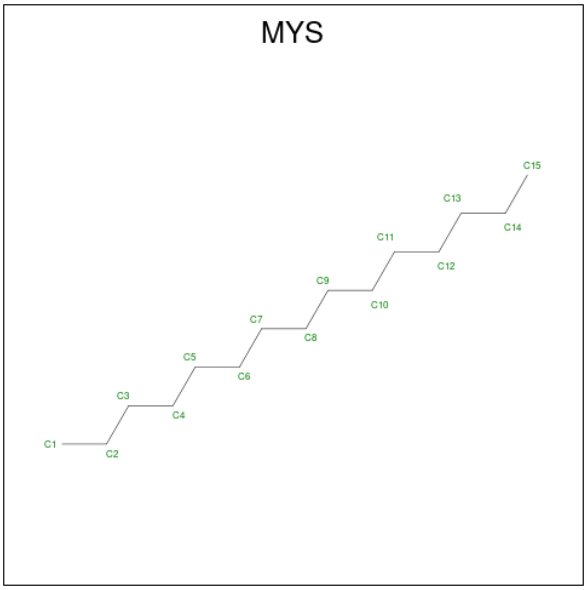
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0

- Molecule 4 is N-OCTANE (three-letter code: OCT) (formula: C_8H_{18}).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C 8 8	0	0
4	C	1	Total C 8 8	0	0

- Molecule 5 is PENTADECANE (three-letter code: MYS) (formula: $C_{15}H_{32}$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C 15 15	0	0
5	D	1	Total C 15 15	0	0

- Molecule 6 is water.

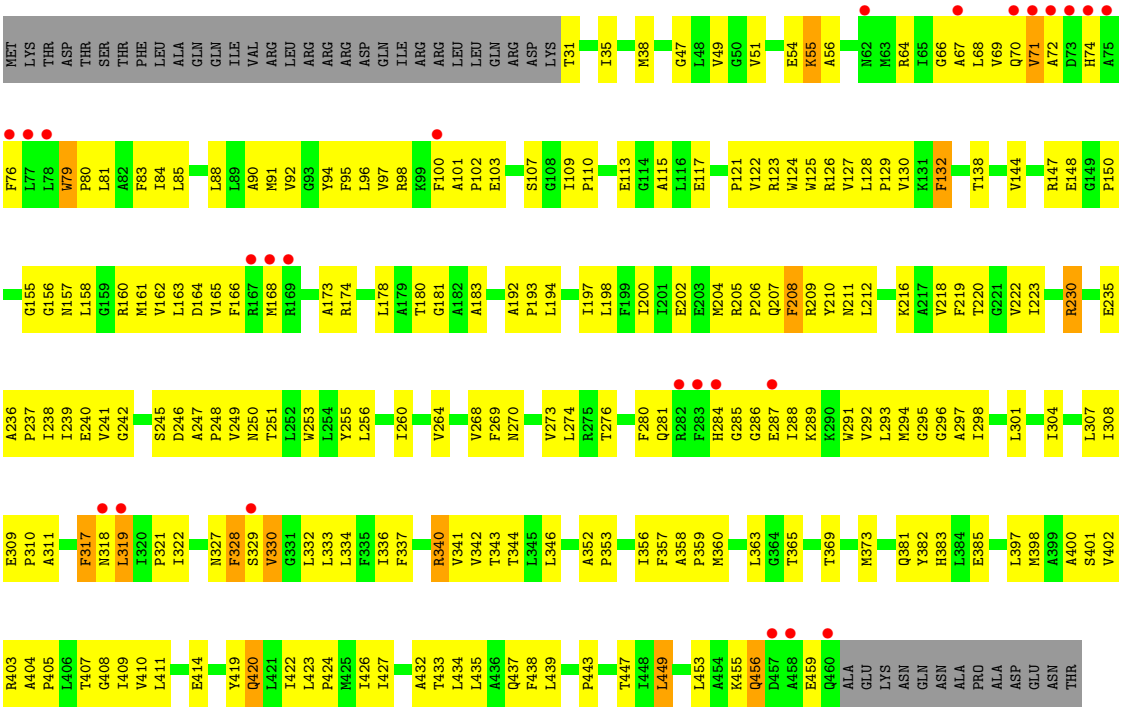
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total O 1 1	0	0
6	D	1	Total O 1 1	0	0

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.

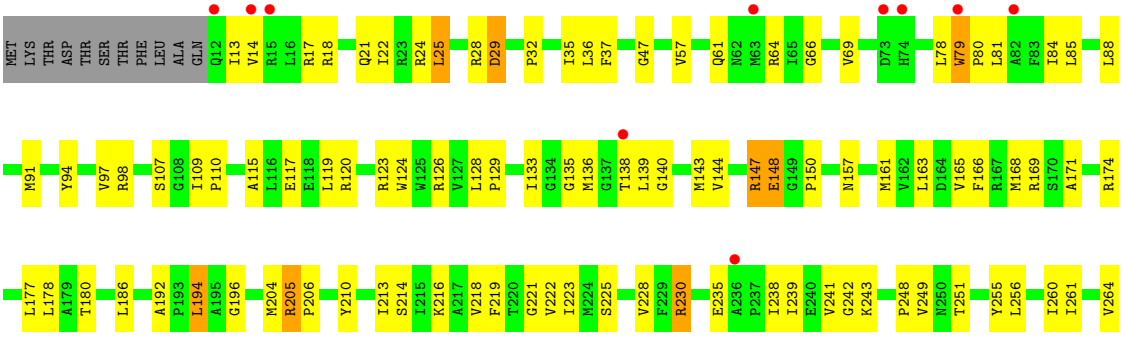
• Molecule 1: putative ClC family, chlorine transport protein

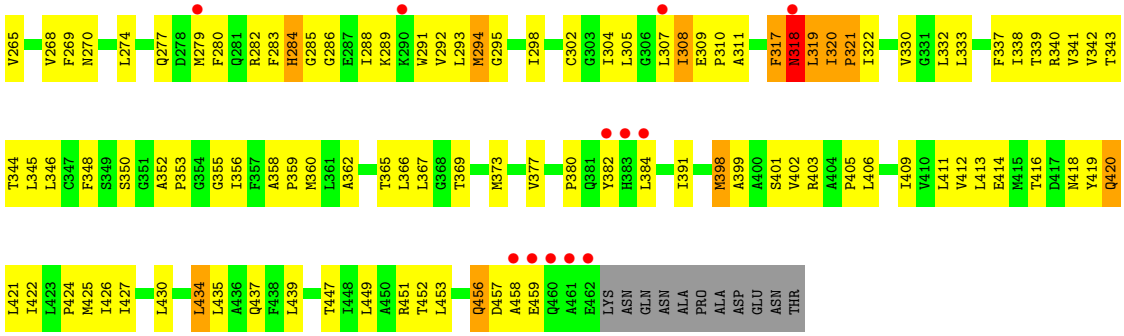
Chain A:



• Molecule 1: putative ClC family, chlorine transport protein

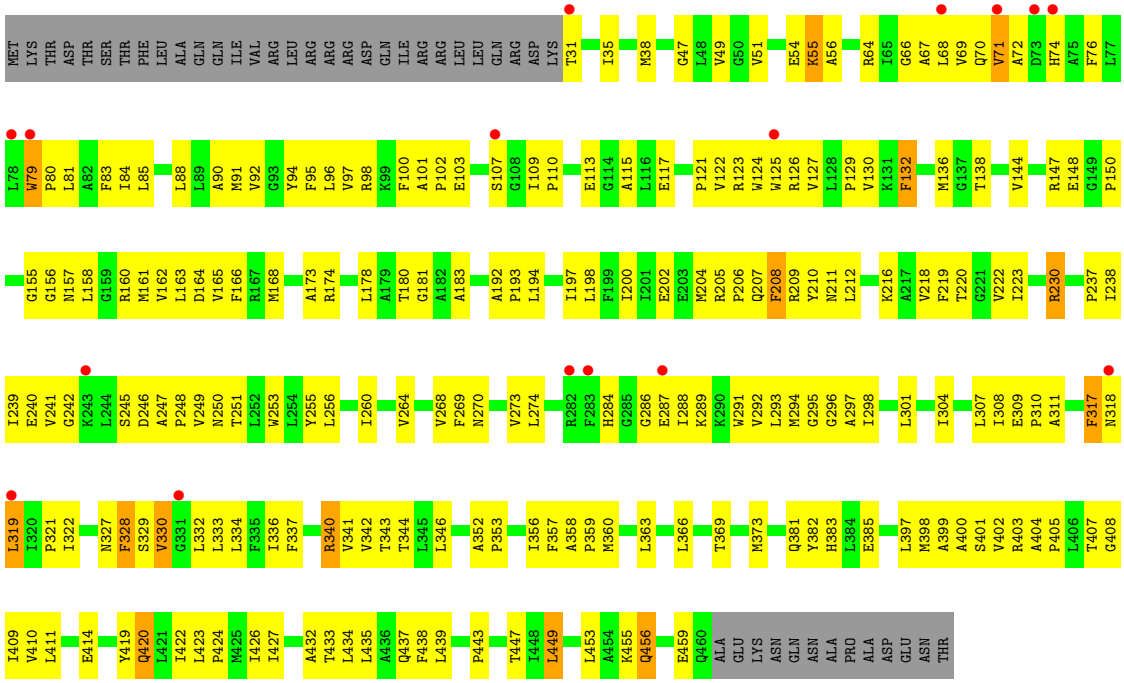
Chain B:





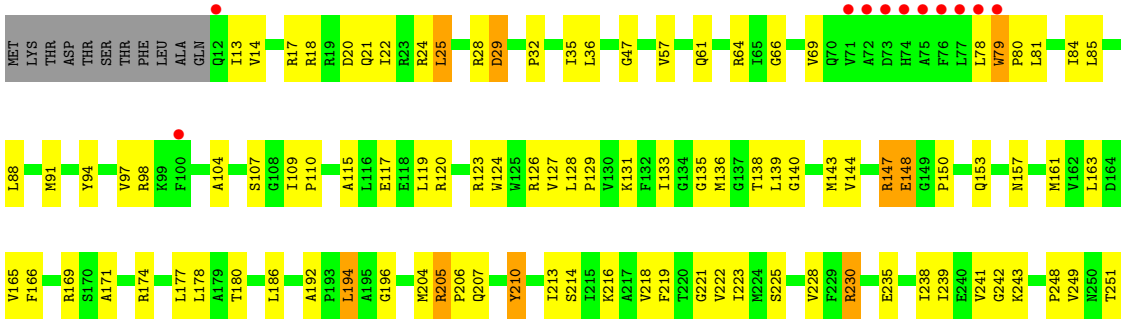
• Molecule 1: putative ClC family, chlorine transport protein

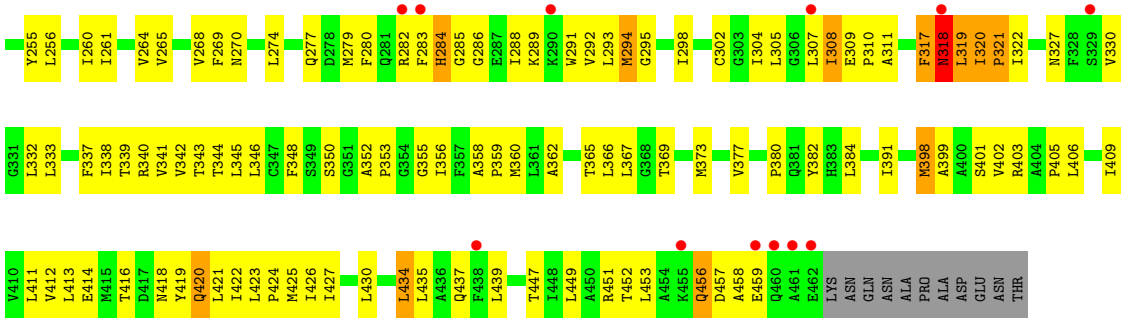
Chain C:



• Molecule 1: putative ClC family, chlorine transport protein

Chain D:




Y255 L256 I260 I261 V264 V265 V268 F269 N270 L274 Q277 D278 M279 F280 Q281 R282 F283 H284 G285 G286 E287 I288 K289 K290 W291 V292 L293 M294 G295 I298 C302 G303 I304 L305 G306 L307 I308 E309 P310 A311 F317 N318 L319 I320 P321 I322 N327 F328 S329 V330
G331 L332 L333 F337 I338 T339 R340 V341 V342 T343 T344 L345 L346 C347 F348 S349 S350 G351 A352 P353 G354 G355 I356 F357 A358 P359 M360 L361 A362 T365 L366 L367 G368 T369 M373 V377 P380 Q381 Y382 H383 L384 I391 M398 A399 A400 S401 V402 R403 A404 P405 L406 I409
V410 L411 V412 L413 E414 M415 T416 D417 N418 Y419 Q420 L421 I422 L423 P424 M425 I426 I427 L430 L434 L435 A436 Q437 F438 L439 T447 I448 L449 A450 R451 T452 L453 A454 K455 Q456 D457 A458 E459 Q460 A461 E462 LYS ASN GLN ASN ALA PRO ALA ASP GLU ASN THR

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	185.72Å 90.16Å 80.75Å 90.00° 98.95° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 34.66 – 3.00	Depositor EDS
% Data completeness (in resolution range)	94.4 (20.00-3.00) 94.5 (34.66-3.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 3.00Å)	Xtriage
Refinement program	CNS 0.3	Depositor
R, R_{free}	0.254 , 0.289 0.245 , 0.275	Depositor DCC
R_{free} test set	4991 reflections (9.97%)	wwPDB-VP
Wilson B-factor (Å^2)	85.9	Xtriage
Anisotropy	0.524	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 64.3	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	0.91	EDS
Total number of atoms	13216	wwPDB-VP
Average B, all atoms (Å^2)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% 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: SO4, CL, MYS, OCT

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	0/3262	0.69	0/4432
1	B	0.42	0/3448	0.71	3/4678 (0.1%)
1	C	0.44	0/3262	0.69	0/4432
1	D	0.42	0/3448	0.71	3/4678 (0.1%)
All	All	0.43	0/13420	0.70	6/18220 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	210	TYR	N-CA-C	-6.89	92.41	111.00
1	B	210	TYR	N-CA-C	-6.88	92.42	111.00
1	D	320	ILE	N-CA-C	6.33	128.08	111.00
1	B	320	ILE	N-CA-C	6.32	128.06	111.00
1	B	243	LYS	N-CA-C	-5.33	96.61	111.00

There are no chirality outliers.

There are no planarity outliers.

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	3193	0	3349	226	0
1	B	3379	0	3556	232	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3193	0	3349	224	2
1	D	3379	0	3556	237	2
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	5	0	0	0	0
3	B	10	0	0	2	0
3	C	5	0	0	0	0
4	A	8	0	18	0	0
4	C	8	0	18	0	0
5	B	15	0	32	1	0
5	D	15	0	32	1	0
6	B	1	0	0	1	0
6	D	1	0	0	1	0
All	All	13216	0	13910	858	2

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 32.

The worst 5 of 858 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:317:PHE:HA	1:B:321:PRO:HD2	1.33	1.11
1:A:230:ARG:HH11	1:A:230:ARG:HB3	1.18	1.07
1:D:317:PHE:HA	1:D:321:PRO:HD2	1.33	1.07
1:C:230:ARG:HB3	1:C:230:ARG:HH11	1.18	1.06
1:D:298:ILE:HD11	1:D:346:LEU:HB3	1.38	1.04

All (2) 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)
1:C:328:PHE:CA	1:D:327:ASN:ND2[2_545]	1.75	0.45
1:C:328:PHE:CD1	1:D:327:ASN:ND2[2_545]	1.87	0.33

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	428/473 (90%)	371 (87%)	48 (11%)	9 (2%)	7	33
1	B	449/473 (95%)	400 (89%)	42 (9%)	7 (2%)	9	40
1	C	428/473 (90%)	371 (87%)	48 (11%)	9 (2%)	7	33
1	D	449/473 (95%)	400 (89%)	42 (9%)	7 (2%)	9	40
All	All	1754/1892 (93%)	1542 (88%)	180 (10%)	32 (2%)	8	37

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	318	ASN
1	C	318	ASN
1	B	29	ASP
1	B	318	ASN
1	B	458	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	322/361 (89%)	309 (96%)	13 (4%)	31	68
1	B	341/361 (94%)	321 (94%)	20 (6%)	19	54
1	C	322/361 (89%)	309 (96%)	13 (4%)	31	68
1	D	341/361 (94%)	321 (94%)	20 (6%)	19	54
All	All	1326/1444 (92%)	1260 (95%)	66 (5%)	24	60

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	317	PHE
1	D	332	LEU
1	D	456	GLN
1	B	317	PHE
1	B	308	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	281	GLN
1	D	157	ASN
1	C	318	ASN
1	D	62	ASN
1	D	233	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)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 for Mogul analysis.

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	OCT	C	602	-	7,7,7	0.52	0	6,6,6	0.49	0
5	MYS	B	501	-	14,14,14	0.45	0	13,13,13	0.61	0
4	OCT	A	502	-	7,7,7	0.51	0	6,6,6	0.49	0
5	MYS	D	601	-	14,14,14	0.45	0	13,13,13	0.61	0
3	SO4	B	605	-	4,4,4	0.20	0	6,6,6	0.29	0
3	SO4	C	606	-	4,4,4	0.19	0	6,6,6	0.15	0
3	SO4	A	506	-	4,4,4	0.19	0	6,6,6	0.15	0
3	SO4	B	505	-	4,4,4	0.20	0	6,6,6	0.29	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
5	MYS	D	601	-	-	3/12/12/12	-
5	MYS	B	501	-	-	3/12/12/12	-
4	OCT	C	602	-	-	1/5/5/5	-
4	OCT	A	502	-	-	1/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	501	MYS	C2-C3-C4-C5
5	D	601	MYS	C2-C3-C4-C5
5	D	601	MYS	C1-C2-C3-C4
5	B	501	MYS	C1-C2-C3-C4
5	B	501	MYS	C7-C8-C9-C10

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	501	MYS	1	0
5	D	601	MYS	1	0
3	B	605	SO4	1	0
3	B	505	SO4	1	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	430/473 (90%)	-0.05	25 (5%)	23	7	49, 74, 114, 155	0
1	B	451/473 (95%)	-0.02	22 (4%)	29	11	49, 74, 118, 154	0
1	C	430/473 (90%)	-0.03	16 (3%)	41	17	49, 74, 114, 155	0
1	D	451/473 (95%)	-0.15	23 (5%)	28	10	49, 74, 118, 154	0
All	All	1762/1892 (93%)	-0.06	86 (4%)	29	11	49, 74, 117, 155	0

The worst 5 of 86 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	461	ALA	9.4
1	D	74	HIS	6.7
1	A	74	HIS	6.3
1	B	462	GLU	6.3
1	A	73	ASP	6.1

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, 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
2	CL	C	603	1/1	0.72	1.24	100,100,100,100	0
4	OCT	C	602	8/8	0.75	0.25	43,53,71,79	0
4	OCT	A	502	8/8	0.78	0.21	43,53,71,79	0
5	MYS	B	501	15/15	0.78	0.25	55,75,88,92	0
5	MYS	D	601	15/15	0.81	0.25	55,75,88,92	0
2	CL	A	503	1/1	0.90	1.08	100,100,100,100	0
2	CL	B	504	1/1	0.92	0.36	81,81,81,81	0
2	CL	D	604	1/1	0.92	0.47	81,81,81,81	0
3	SO4	B	605	5/5	0.94	0.14	72,87,93,99	0
3	SO4	B	505	5/5	0.96	0.18	72,87,93,99	0
3	SO4	C	606	5/5	0.97	0.14	59,62,66,70	0
3	SO4	A	506	5/5	0.98	0.12	59,62,66,70	0

6.5 Other polymers (i)

There are no such residues in this entry.