wwPDB X-ray Structure Validation Summary Report (i)

Sep 10, 2023 – 04:32 AM EDT

PDB ID	:	4I5O
Title	:	Crystal Structure of W-W-R ClpX Hexamer
Authors	:	Glynn, S.E.; Nager, A.R.; Stinson, B.S.; Schmitz, K.R.; Baker, T.A.; Sauer, R.T.
Deposited on	:	2012-11-28
Resolution	:	4.48 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

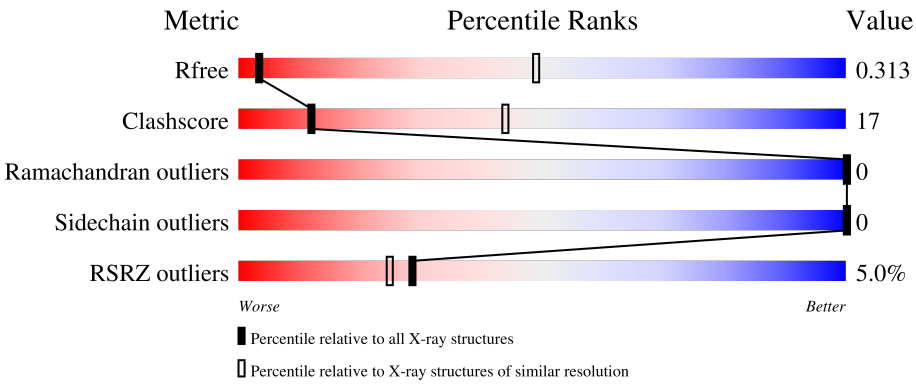
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 4.48 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	1051 (5.12-3.80)
Clashscore	141614	1119 (5.12-3.80)
Ramachandran outliers	138981	1065 (5.12-3.80)
Sidechain outliers	138945	1047 (5.12-3.80)
RSRZ outliers	127900	1099 (5.20-3.70)

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	363	5% 58% 27% 15%
1	B	363	3% 53% 31% 15%
1	C	363	7% 58% 29% 13%
1	D	363	4% 57% 27% 16%
1	E	363	4% 59% 29% 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	363	4% 56% 29% 15%

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	SO4	B	501	-	-	-	X

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 13980 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.

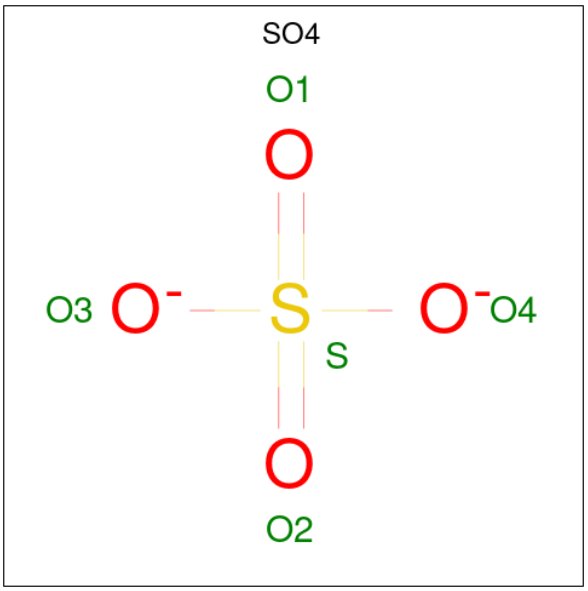
- Molecule 1 is a protein called ATP-dependent Clp protease ATP-binding subunit ClpX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	309	2317	1475	384	452	6	0	0	0
1	B	308	2312	1472	383	451	6	0	0	0
1	C	317	2342	1488	388	462	4	0	0	0
1	D	304	2294	1462	377	449	6	0	0	0
1	E	318	2365	1505	393	461	6	0	0	0
1	F	310	2320	1479	383	452	6	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	408	GLU	LYS	conflict	UNP P0A6H1
B	408	GLU	LYS	conflict	UNP P0A6H1
C	408	GLU	LYS	conflict	UNP P0A6H1
D	408	GLU	LYS	conflict	UNP P0A6H1
E	408	GLU	LYS	conflict	UNP P0A6H1
F	408	GLU	LYS	conflict	UNP P0A6H1

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



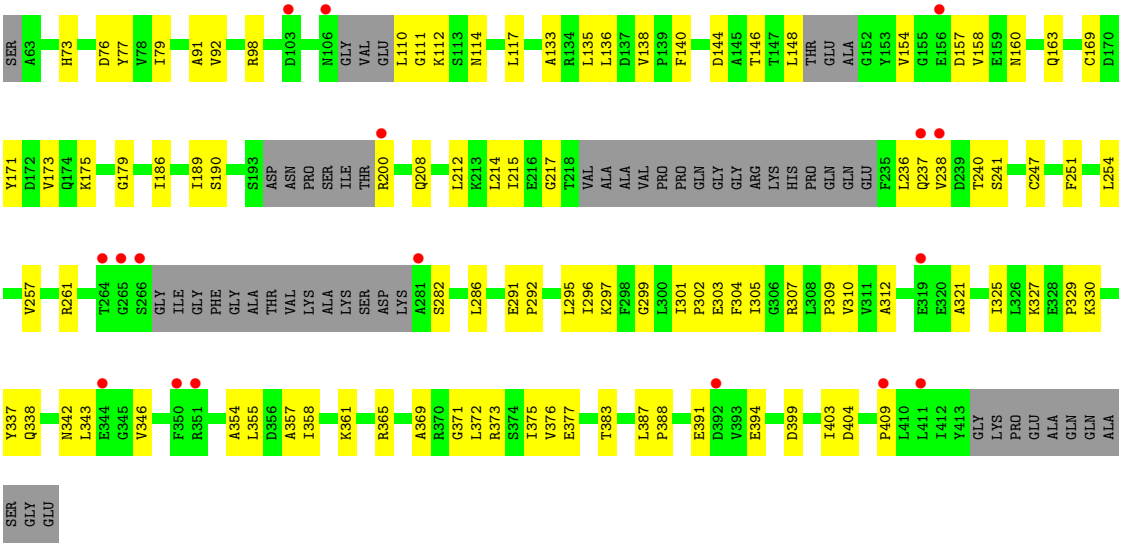
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	B	1	5	4	1	0	0
2	C	1	5	4	1	0	0
2	D	1	5	4	1	0	0
2	E	1	5	4	1	0	0
2	F	1	5	4	1	0	0

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.

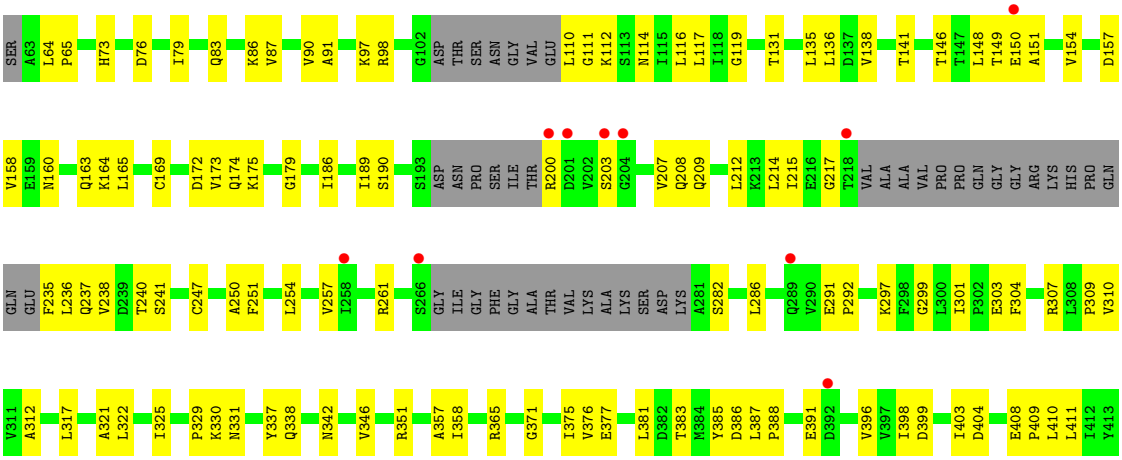
• Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

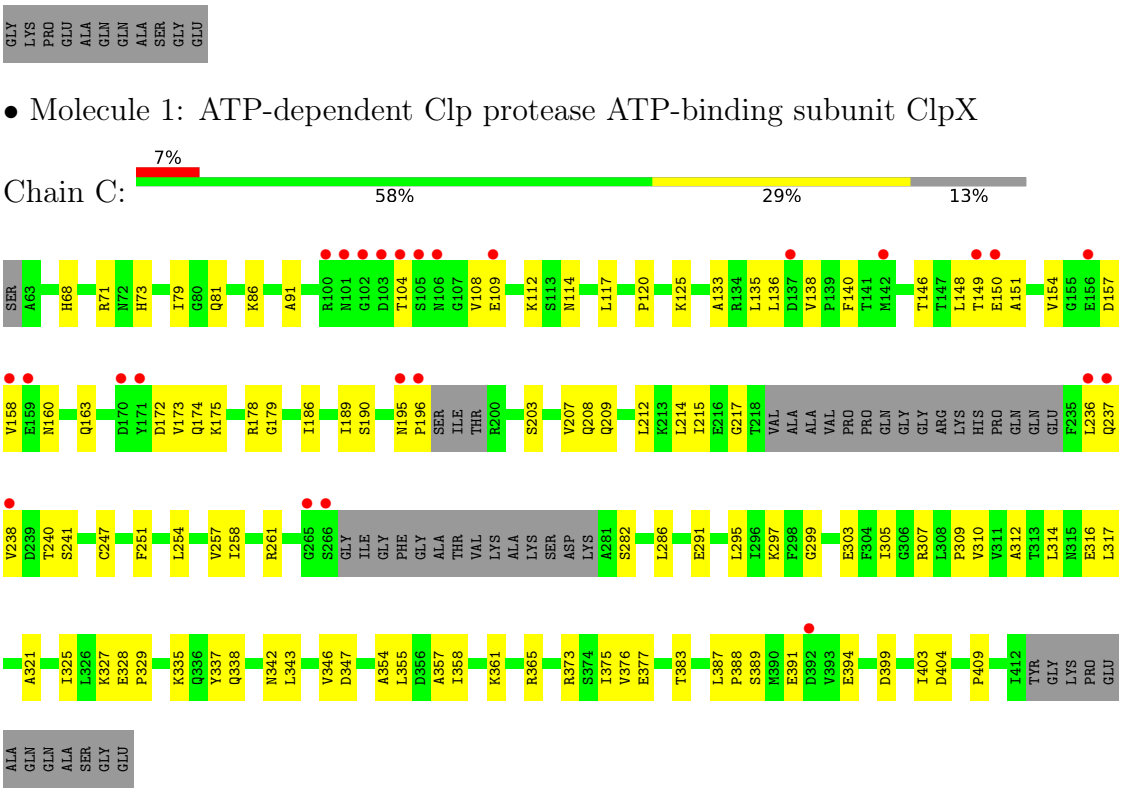
Chain A:

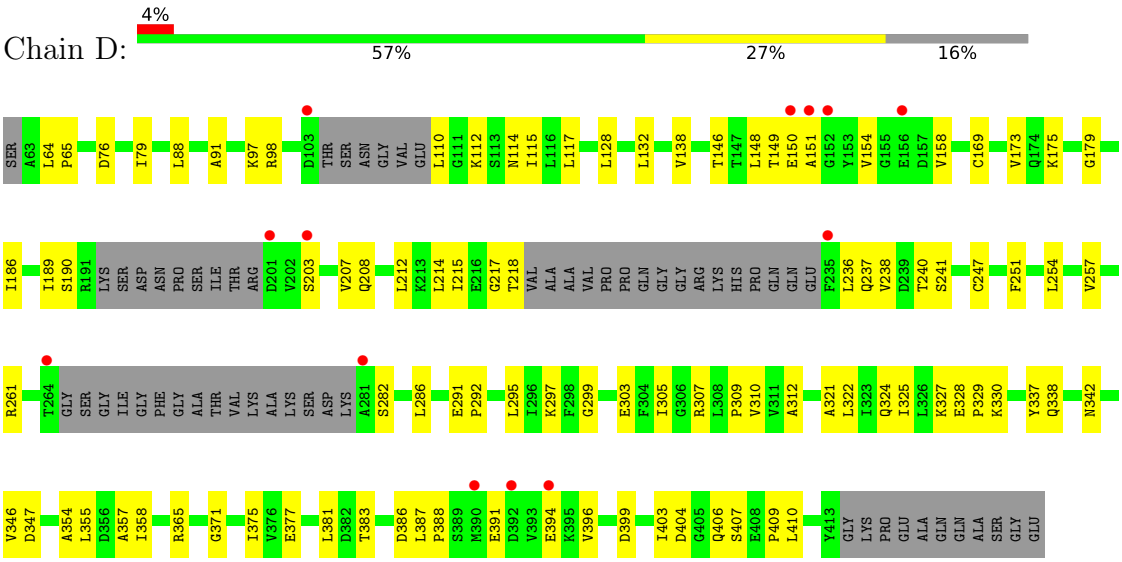


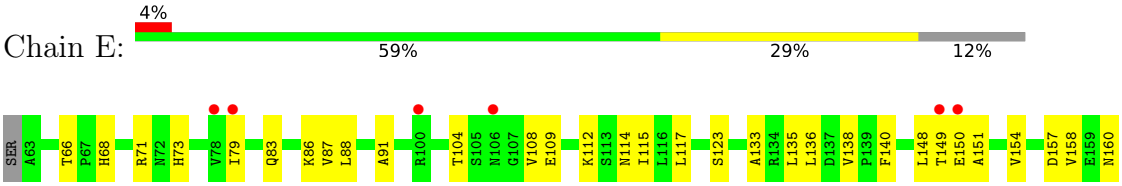
• Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

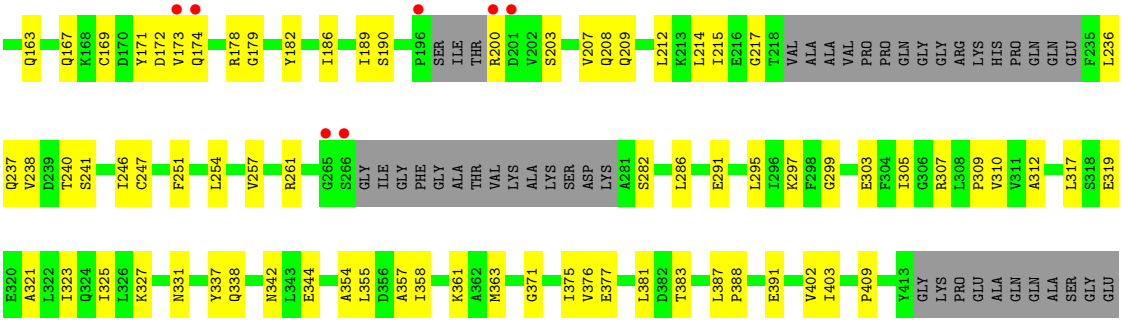
Chain B:




GLY LYS PRO GLU ALA GLN GLN ALA SER GLY GLU
• Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX
Chain C:
7%
58%
29%
13%
SER A63 H68 R71 N72 H73 I79 G80 Q81 K86 A91 R100 N101 G102 D103 T104 S105 N106 G107 V108 E109 K112 S113 N114 L117 P120 K125 A133 R134 L135 L136 D137 V138 P139 F140 T141 M142 T146 T147 L148 T149 E150 A151 V154 G155 E156 D157
V158 E159 N160 Q163 D170 Y171 D172 V173 Q174 K175 R178 G179 I186 I189 S190 N195 P196 SER ILE THR R200 S203 V207 Q208 Q209 L212 K213 L214 I215 E216 G217 T218 VAL ALA ALA VAL PRO PRO GLN GLY GLY ARG LYS HIS PRO GLN GLN GLU F235 L236 Q237
V238 D239 T240 S241 C247 F251 L254 V257 I258 R261 G265 S266 GLY ILE GLY PHE GLY ALA THR VAL LYS ALA LYS SER ASP LYS A281 S282 L286 E291 L295 I296 K297 F298 G299 E303 F304 I305 G306 R307 L308 P309 V310 V311 A312 T313 L314 N315 E316 L317
A321 I325 L326 K327 E328 P329 K335 Q336 Y337 Q338 N342 L343 V346 D347 A354 L355 D356 A357 I358 K361 R365 R373 S374 I375 V376 E377 T383 L387 P388 S389 M390 E391 D392 V393 E394 D399 I403 D404 P409 I412 TYR GLY LYS PRO GLU
ALA GLN GLN ALA SER GLY GLU

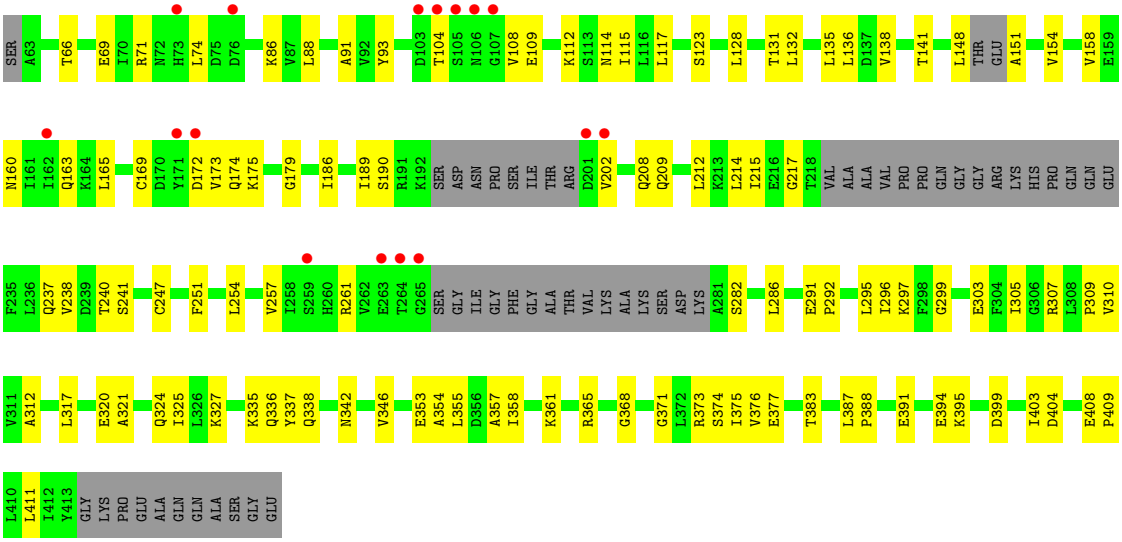

• Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX
Chain D:
4%
57%
27%
16%
SER A63 L64 P65 D76 I79 L88 A91 K97 R98 D103 THR SER ASN GLY VAL GLU L110 G111 K112 S113 N114 I115 L116 L117 L128 L132 V138 T146 T147 L148 T149 E150 A151 G152 Y153 V154 G155 E156 D157 V158 C169 V173 Q174 K175 G179
I186 I189 S190 R191 LYS SER ASP ASN PRO SER ILE THR ARG D201 V202 S203 V207 Q208 L212 K213 L214 I215 E216 G217 T218 VAL ALA ALA VAL PRO PRO GLN GLY GLY ARG LYS HIS PRO GLN GLN GLU F235 L236 Q237 V238 D239 T240 S241 C247 F251 L254 V257
R261 T264 GLY SER GLY ILE GLY PHE GLY ALA THR VAL LYS ALA LYS SER ASP LYS A281 S282 L286 E291 P292 L295 I296 K297 F298 G299 E303 F304 I305 G306 R307 L308 P309 V310 V311 A312 A321 L322 I323 Q324 I325 L326 K327 E328 P329 K330 Y337 Q338 N342
V346 D347 A354 L355 D356 A357 I358 R365 G371 I375 V376 E377 L381 D382 T383 D386 L387 P388 S389 M390 E391 D392 V393 E394 K395 V396 D399 I403 D404 G405 Q406 S407 E408 P409 L410 Y413 GLY LYS PRO GLU ALA GLN GLN ALA SER GLY GLU


• Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX
Chain E:
4%
59%
29%
12%
SER A63 T66 P67 H68 R71 N72 H73 V78 I79 Q83 K86 V87 L88 A91 R100 T104 S105 N106 G107 V108 E109 K112 S113 N114 I115 L116 L117 S123 A133 R134 L135 L136 D137 V138 P139 F140 L148 T149 E150 A151 V154 D157 V158 E159 N160



- Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.20Å 199.88Å 222.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.32 – 4.48 48.32 – 4.48	Depositor EDS
% Data completeness (in resolution range)	91.5 (48.32-4.48) 91.5 (48.32-4.48)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.26 (at 4.45Å)	Xtriage
Refinement program	PHENIX 1.5_2	Depositor
R, R_{free}	0.286 , 0.323 0.286 , 0.313	Depositor DCC
R_{free} test set	714 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	163.5	Xtriage
Anisotropy	0.461	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 186.3	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	0.87	EDS
Total number of atoms	13980	wwPDB-VP
Average B, all atoms (Å^2)	222.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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

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	0.21	0/2341	0.36	0/3168
1	B	0.21	0/2337	0.38	0/3164
1	C	0.21	0/2368	0.38	0/3215
1	D	0.21	0/2319	0.38	0/3140
1	E	0.21	0/2392	0.38	0/3243
1	F	0.21	0/2345	0.37	0/3175
All	All	0.21	0/14102	0.37	0/19105

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 (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	2317	0	2312	81	0
1	B	2312	0	2314	91	0
1	C	2342	0	2310	90	0
1	D	2294	0	2301	79	0
1	E	2365	0	2348	85	0
1	F	2320	0	2323	88	0
2	A	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	5	0	0	0	0
2	C	5	0	0	1	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
All	All	13980	0	13908	479	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 17.

The worst 5 of 479 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:C:150:GLU:HA	1:C:203:SER:HB2	1.42	1.00
1:B:215:ILE:HG21	1:B:307:ARG:HB3	1.45	0.97
1:D:215:ILE:HG21	1:D:307:ARG:HB3	1.49	0.94
1:A:215:ILE:HG21	1:A:307:ARG:HB3	1.49	0.94
1:F:215:ILE:HG21	1:F:307:ARG:HB3	1.50	0.94

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	297/363 (82%)	284 (96%)	13 (4%)	0	100	100
1	B	298/363 (82%)	283 (95%)	15 (5%)	0	100	100
1	C	309/363 (85%)	289 (94%)	20 (6%)	0	100	100
1	D	294/363 (81%)	279 (95%)	15 (5%)	0	100	100
1	E	310/363 (85%)	293 (94%)	17 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	300/363 (83%)	283 (94%)	17 (6%)	0	100	100
All	All	1808/2178 (83%)	1711 (95%)	97 (5%)	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	240/299 (80%)	240 (100%)	0	100	100
1	B	240/299 (80%)	240 (100%)	0	100	100
1	C	239/299 (80%)	239 (100%)	0	100	100
1	D	240/299 (80%)	240 (100%)	0	100	100
1	E	242/299 (81%)	242 (100%)	0	100	100
1	F	240/299 (80%)	240 (100%)	0	100	100
All	All	1441/1794 (80%)	1441 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	209	GLN
1	F	114	ASN
1	F	324	GLN
1	C	73	HIS
1	D	114	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)

6 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	$\#\|Z\| > 2$
2	SO4	B	501	-	4,4,4	0.13	0	6,6,6	0.06	0
2	SO4	A	501	-	4,4,4	0.14	0	6,6,6	0.10	0
2	SO4	E	501	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	F	501	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	C	501	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	D	501	-	4,4,4	0.14	0	6,6,6	0.04	0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	501	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	309/363 (85%)	0.40	17 (5%)	25	22	176, 221, 267, 325	0
1	B	308/363 (84%)	0.13	10 (3%)	47	37	175, 216, 300, 377	0
1	C	317/363 (87%)	0.38	25 (7%)	12	11	144, 208, 331, 421	0
1	D	304/363 (83%)	0.18	13 (4%)	35	29	151, 195, 302, 370	0
1	E	318/363 (87%)	0.17	13 (4%)	37	30	177, 231, 301, 357	0
1	F	310/363 (85%)	0.19	16 (5%)	27	24	168, 214, 290, 348	0
All	All	1866/2178 (85%)	0.24	94 (5%)	28	25	144, 218, 301, 421	0

The worst 5 of 94 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	105	SER	8.8
1	A	265	GLY	8.7
1	A	266	SER	7.3
1	B	200	ARG	7.0
1	A	264	THR	6.7

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	SO4	B	501	5/5	0.71	0.43	201,205,209,212	0
2	SO4	D	501	5/5	0.81	0.39	194,195,199,206	0
2	SO4	F	501	5/5	0.90	0.17	174,177,179,180	0
2	SO4	E	501	5/5	0.92	0.39	204,204,209,215	0
2	SO4	A	501	5/5	0.95	0.23	203,210,212,216	0
2	SO4	C	501	5/5	0.96	0.12	161,161,162,168	0

6.5 Other polymers (i)

There are no such residues in this entry.