wwPDB X-ray Structure Validation Summary Report (i)

Sep 12, 2023 – 08:18 AM EDT

PDB ID	:	4NBQ
Title	:	Structure of the polynucleotide phosphorylase (CBU_0852) from Coxiella burnetii
Authors	:	Rudolph, M.J.; Cheung, J.; Franklin, M.C.; Cassidy, M.; Gary, E.; Burshteyn, F.; Love, J.
Deposited on	:	2013-10-23
Resolution	:	2.91 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:

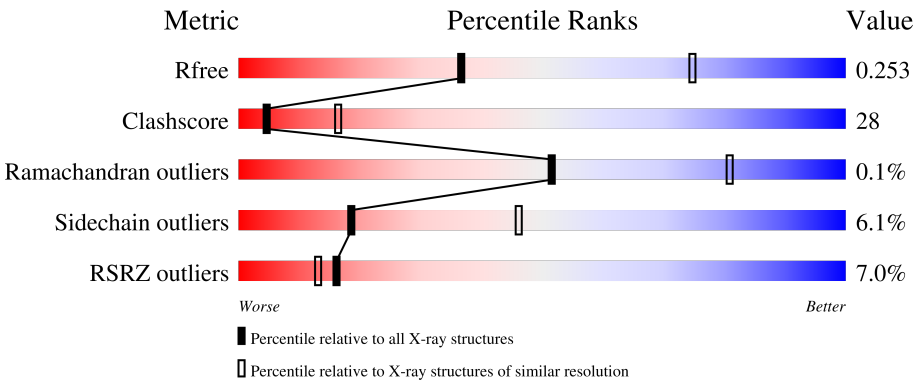
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.91 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	2307 (2.94-2.90)
Clashscore	141614	2531 (2.94-2.90)
Ramachandran outliers	138981	2462 (2.94-2.90)
Sidechain outliers	138945	2464 (2.94-2.90)
RSRZ outliers	127900	2248 (2.94-2.90)

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	715	5% 54% 35% • 8%
1	B	715	5% 53% 37% • 7%
1	C	715	9% 44% 38% • 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	A	701	-	-	X	-
2	SO4	B	701	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 14928 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 Polyribonucleotide nucleotidyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	659	5072	3189	891	973	3	16	0	1	0
1	B	668	5121	3220	892	990	3	16	0	0	0
1	C	607	4658	2930	815	895	3	15	0	0	0

There are 57 discrepancies between the modelled and reference sequences:

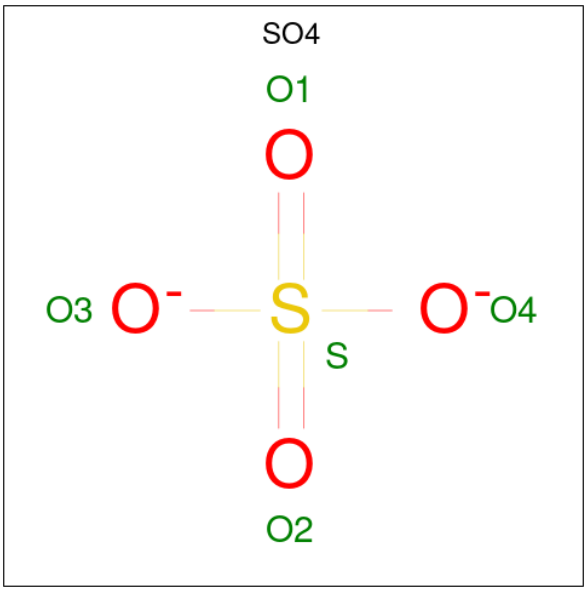
Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	MSE	-	expression tag	UNP Q83D87
A	-17	HIS	-	expression tag	UNP Q83D87
A	-16	HIS	-	expression tag	UNP Q83D87
A	-15	HIS	-	expression tag	UNP Q83D87
A	-14	HIS	-	expression tag	UNP Q83D87
A	-13	HIS	-	expression tag	UNP Q83D87
A	-12	HIS	-	expression tag	UNP Q83D87
A	-11	HIS	-	expression tag	UNP Q83D87
A	-10	HIS	-	expression tag	UNP Q83D87
A	-9	HIS	-	expression tag	UNP Q83D87
A	-8	HIS	-	expression tag	UNP Q83D87
A	-7	GLU	-	expression tag	UNP Q83D87
A	-6	ASN	-	expression tag	UNP Q83D87
A	-5	LEU	-	expression tag	UNP Q83D87
A	-4	TYR	-	expression tag	UNP Q83D87
A	-3	PHE	-	expression tag	UNP Q83D87
A	-2	GLN	-	expression tag	UNP Q83D87
A	-1	SER	-	expression tag	UNP Q83D87
A	0	ALA	-	expression tag	UNP Q83D87
B	-18	MSE	-	expression tag	UNP Q83D87
B	-17	HIS	-	expression tag	UNP Q83D87
B	-16	HIS	-	expression tag	UNP Q83D87
B	-15	HIS	-	expression tag	UNP Q83D87

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	HIS	-	expression tag	UNP Q83D87
B	-13	HIS	-	expression tag	UNP Q83D87
B	-12	HIS	-	expression tag	UNP Q83D87
B	-11	HIS	-	expression tag	UNP Q83D87
B	-10	HIS	-	expression tag	UNP Q83D87
B	-9	HIS	-	expression tag	UNP Q83D87
B	-8	HIS	-	expression tag	UNP Q83D87
B	-7	GLU	-	expression tag	UNP Q83D87
B	-6	ASN	-	expression tag	UNP Q83D87
B	-5	LEU	-	expression tag	UNP Q83D87
B	-4	TYR	-	expression tag	UNP Q83D87
B	-3	PHE	-	expression tag	UNP Q83D87
B	-2	GLN	-	expression tag	UNP Q83D87
B	-1	SER	-	expression tag	UNP Q83D87
B	0	ALA	-	expression tag	UNP Q83D87
C	-18	MSE	-	expression tag	UNP Q83D87
C	-17	HIS	-	expression tag	UNP Q83D87
C	-16	HIS	-	expression tag	UNP Q83D87
C	-15	HIS	-	expression tag	UNP Q83D87
C	-14	HIS	-	expression tag	UNP Q83D87
C	-13	HIS	-	expression tag	UNP Q83D87
C	-12	HIS	-	expression tag	UNP Q83D87
C	-11	HIS	-	expression tag	UNP Q83D87
C	-10	HIS	-	expression tag	UNP Q83D87
C	-9	HIS	-	expression tag	UNP Q83D87
C	-8	HIS	-	expression tag	UNP Q83D87
C	-7	GLU	-	expression tag	UNP Q83D87
C	-6	ASN	-	expression tag	UNP Q83D87
C	-5	LEU	-	expression tag	UNP Q83D87
C	-4	TYR	-	expression tag	UNP Q83D87
C	-3	PHE	-	expression tag	UNP Q83D87
C	-2	GLN	-	expression tag	UNP Q83D87
C	-1	SER	-	expression tag	UNP Q83D87
C	0	ALA	-	expression tag	UNP Q83D87

- 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	A	1	5	4	1		0	0
2	A	1	5	4	1		0	0
2	A	1	5	4	1		0	0
2	A	1	5	4	1		0	0
2	A	1	5	4	1		0	0
2	A	1	5	4	1		0	0
2	B	1	5	4	1		0	0
2	B	1	5	4	1		0	0
2	B	1	5	4	1		0	0
2	B	1	5	4	1		0	0
2	C	1	5	4	1		0	0
2	C	1	5	4	1		0	0

- Molecule 3 is water.

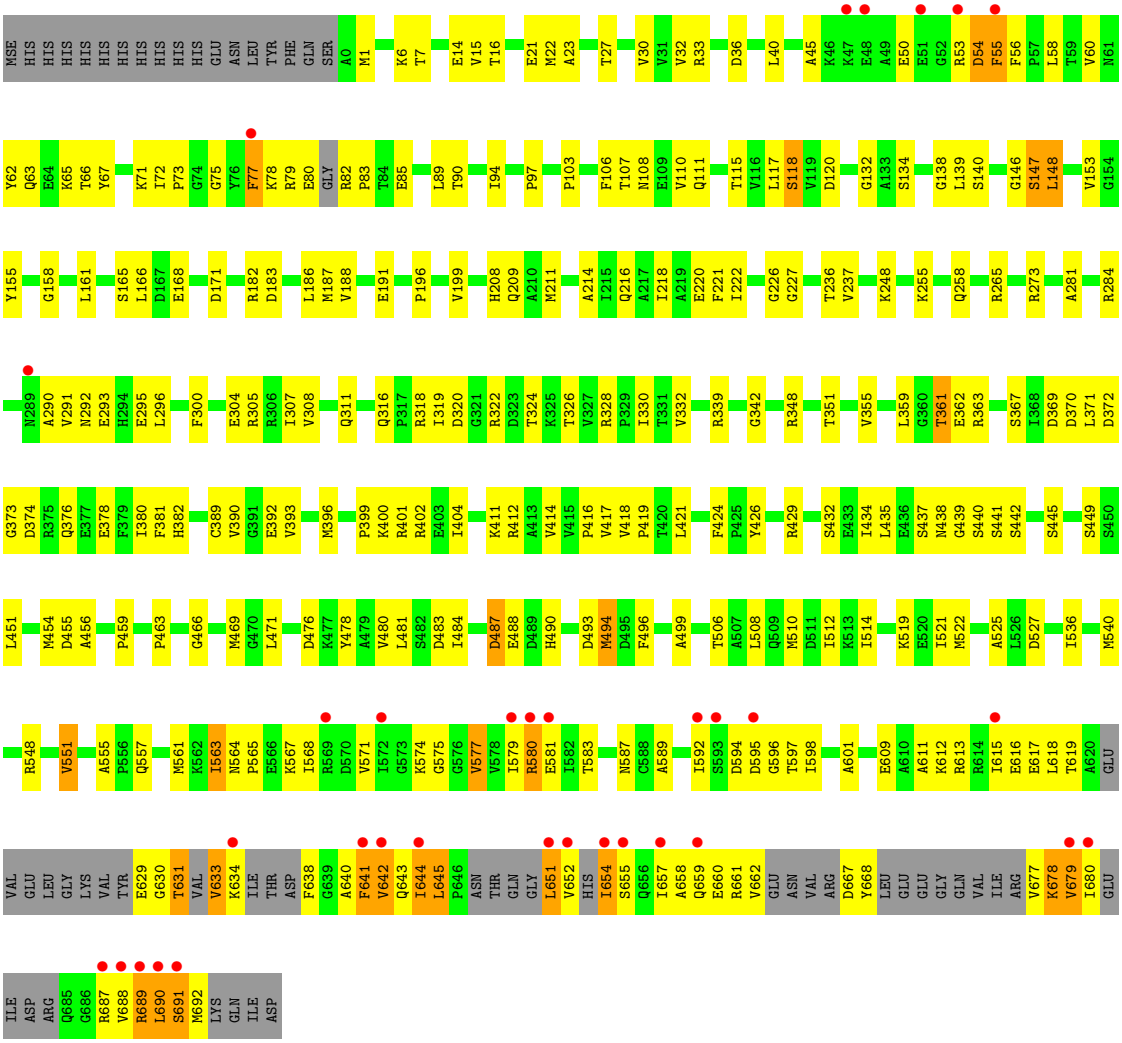
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	9	9	9	0	0
			Total	O		
3	B	3	3	3	0	0

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.

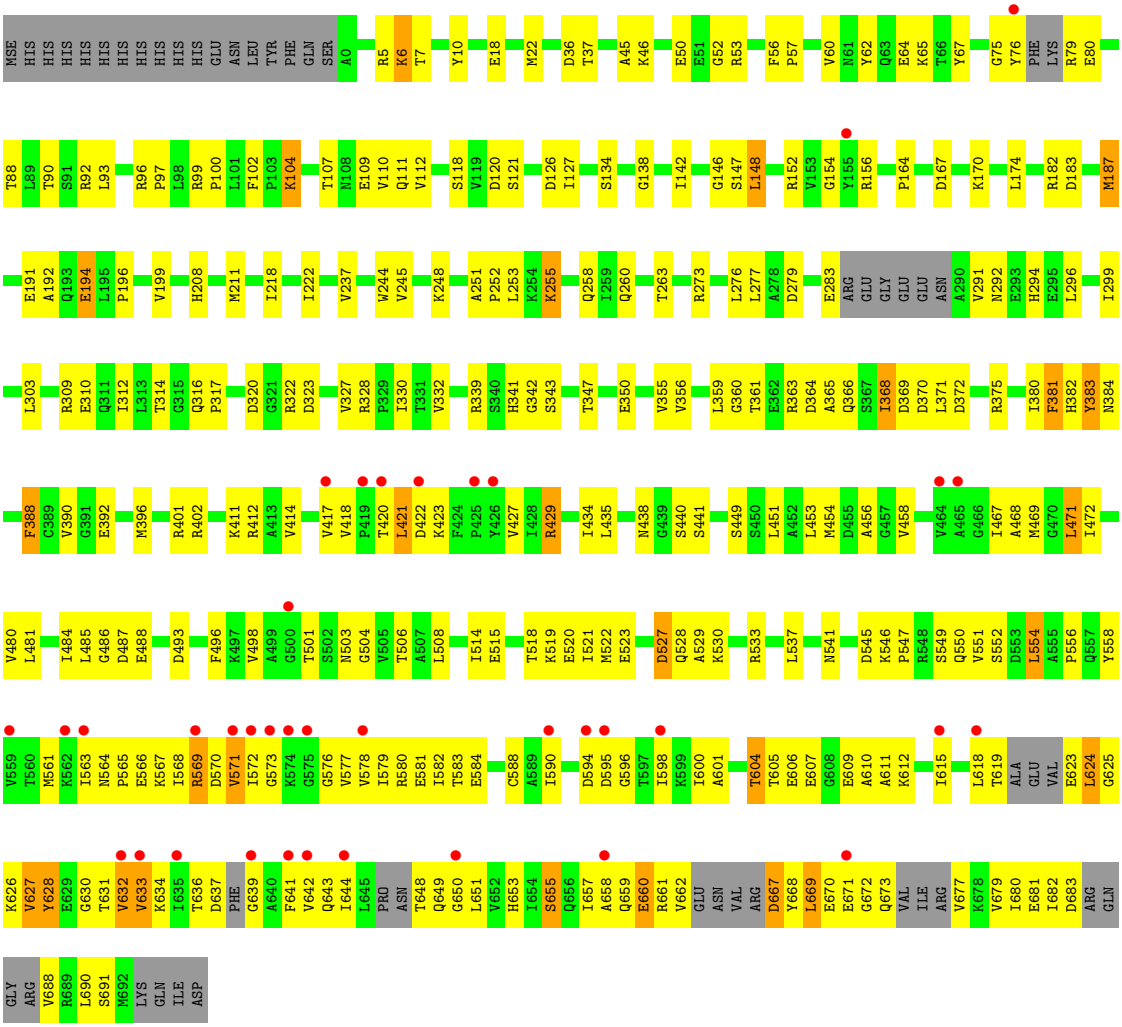
• Molecule 1: Polyribonucleotide nucleotidyltransferase

Chain A: 5% 54% 35% • 8%



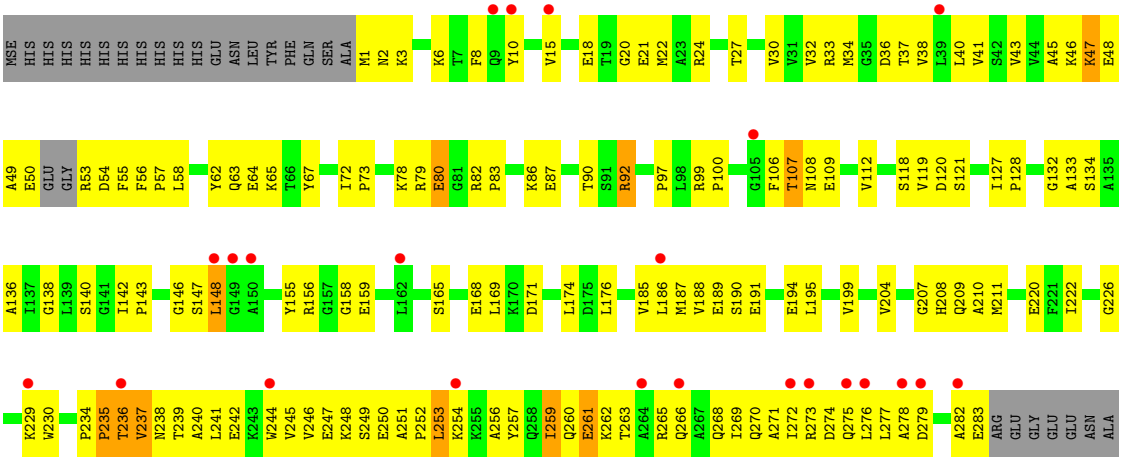
• Molecule 1: Polyribonucleotide nucleotidyltransferase

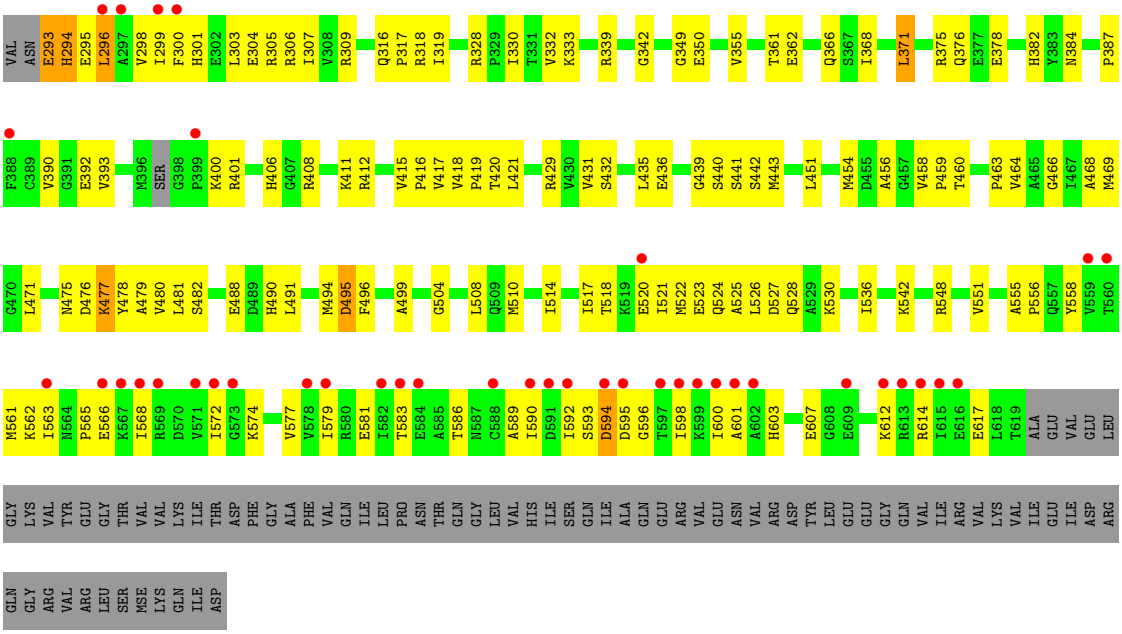
Chain B: 5% 53% 37% • 7%



- Molecule 1: Polyribonucleotide nucleotidyltransferase

Chain C:




VAL ASN E293 H294 E295 L296 A297 V298 I299 F300 H301 E302 L303 E304 R305 R306 I307 V308 R309 Q316 P317 R318 I319 R328 P329 I330 T331 V332 K333 R339 G342 G349 E350 V355 T361 E362 Q366 S367 I368 L371 R375 Q376 E377 E378 H382 Y383 N384 P387
F388 C389 V390 G391 E392 V393 M396 SER G398 P399 K400 R401 H406 G407 R408 K411 R412 V415 P416 V417 V418 P419 T420 L421 R429 V430 V431 S432 L435 E436 G439 S440 S441 S442 M443 L451 M454 D455 A456 G457 V458 P459 T460 P463 V464 A465 G466 I467 A468 M469
G470 L471 N475 D476 K477 Y478 A479 V480 L481 S482 E488 D489 H490 L491 M494 D496 F496 A499 G504 L508 Q509 M510 I514 I517 T518 K519 E520 I521 M522 E523 Q524 A525 L526 D527 Q528 A529 K530 I536 K542 R548 V551 A555 P556 Q557 Y558 V559 T560
M561 K562 I563 N564 P565 E566 K567 I568 R569 D570 V571 I572 G573 K574 V577 V578 I579 R580 E581 I582 T583 E584 A585 T586 N587 C588 A589 I590 D591 I592 S593 D594 D595 G596 T597 I598 K599 I600 A601 A602 H603 E607 G608 E609 K612 R613 R614 I615 E616 E617 L618 T619 ALA GLU VAL GLU LEU
GLY LYS VAL TYR GLU GLY THR VAL VAL LYS ILE THR ASP PHE GLY ALA PHE VAL GLN ILE LEU PRO ASN THR GLN GLY LEU VAL HIS ILE SER GLN ILE ALA GLN GLU ARG VAL GLU ASN VAL ARG ASP TYR LEU GLU GLU GLY GLN VAL ILE ARG VAL LYS VAL ILE GLU ILE ASP ARG
GLN GLY ARG VAL ARG LEU SER MSE LYS GLN ILE ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	99.71Å 111.36Å 219.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.63 – 2.91 49.63 – 2.91	Depositor EDS
% Data completeness (in resolution range)	94.3 (49.63-2.91) 94.1 (49.63-2.91)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.56 (at 2.91Å)	Xtriage
Refinement program	PHENIX 1.7.1_743	Depositor
R, R_{free}	0.208 , 0.260 0.203 , 0.253	Depositor DCC
R_{free} test set	2590 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	67.7	Xtriage
Anisotropy	0.118	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 64.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	14928	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.78% 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	0.29	0/5124	0.56	1/6887 (0.0%)
1	B	0.26	0/5173	0.51	0/6959
1	C	0.25	0/4714	0.48	0/6348
All	All	0.27	0/15011	0.52	1/20194 (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	A	0	3
1	C	0	1
All	All	0	4

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	679	VAL	CB-CA-C	-6.76	98.56	111.40

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	437	SER	Peptide
1	A	678	LYS	Peptide
1	A	689	ARG	Peptide
1	C	235	PRO	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	5072	0	5156	297	3
1	B	5121	0	5203	296	2
1	C	4658	0	4737	298	0
2	A	35	0	0	3	0
2	B	20	0	0	4	0
2	C	10	0	0	0	0
3	A	9	0	0	1	0
3	B	3	0	0	1	0
All	All	14928	0	15096	847	3

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 28.

The worst 5 of 847 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:241:LEU:HD23	1:C:244:TRP:HE1	1.20	1.03
1:A:654:ILE:HG22	1:A:657:ILE:HG22	1.42	1.02
1:B:623:GLU:O	1:B:628:TYR:OH	1.75	1.01
1:A:633:VAL:HG23	1:A:642:VAL:HA	1.39	1.00
1:B:660:GLU:OE2	1:B:668:TYR:OH	1.80	0.99

All (3) 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:A:171:ASP:OD2	1:B:661:ARG:NH1[4_445]	1.96	0.24
1:A:284:ARG:O	1:B:653:HIS:NE2[2_555]	2.10	0.10
1:A:227:GLY:O	1:A:661:ARG:NH1[2_555]	2.17	0.03

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	640/715 (90%)	606 (95%)	33 (5%)	1 (0%)	47	77
1	B	650/715 (91%)	617 (95%)	32 (5%)	1 (0%)	47	77
1	C	599/715 (84%)	547 (91%)	52 (9%)	0	100	100
All	All	1889/2145 (88%)	1770 (94%)	117 (6%)	2 (0%)	51	81

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	644	ILE
1	B	627	VAL

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	543/577 (94%)	508 (94%)	35 (6%)	17	44
1	B	550/577 (95%)	516 (94%)	34 (6%)	18	45
1	C	499/577 (86%)	471 (94%)	28 (6%)	21	50
All	All	1592/1731 (92%)	1495 (94%)	97 (6%)	18	46

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	569	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	24	ARG
1	B	594	ASP
1	B	655	SER
1	C	107	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	145	ASN
1	C	366	GLN
1	C	603	HIS
1	C	382	HIS
1	B	384	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)

13 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	A	701	-	4,4,4	0.13	0	6,6,6	0.18	0
2	SO4	A	707	-	4,4,4	0.12	0	6,6,6	0.10	0
2	SO4	B	701	-	4,4,4	0.15	0	6,6,6	0.10	0
2	SO4	B	704	-	4,4,4	0.14	0	6,6,6	0.08	0
2	SO4	A	705	-	4,4,4	0.13	0	6,6,6	0.06	0
2	SO4	C	702	-	4,4,4	0.15	0	6,6,6	0.06	0
2	SO4	A	706	-	4,4,4	0.15	0	6,6,6	0.07	0
2	SO4	B	703	-	4,4,4	0.16	0	6,6,6	0.13	0
2	SO4	A	703	-	4,4,4	0.13	0	6,6,6	0.15	0
2	SO4	B	702	-	4,4,4	0.15	0	6,6,6	0.06	0
2	SO4	C	701	-	4,4,4	0.14	0	6,6,6	0.09	0
2	SO4	A	702	-	4,4,4	0.15	0	6,6,6	0.05	0
2	SO4	A	704	-	4,4,4	0.13	0	6,6,6	0.08	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.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	SO4	2	0
2	B	701	SO4	2	0
2	A	705	SO4	1	0
2	B	703	SO4	1	0
2	B	702	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	643/715 (89%)	0.17	33 (5%)	28	24	32, 53, 133, 198	0
1	B	652/715 (91%)	0.30	37 (5%)	23	20	37, 75, 123, 156	0
1	C	592/715 (82%)	0.52	63 (10%)	6	5	43, 83, 139, 184	0
All	All	1887/2145 (87%)	0.32	133 (7%)	16	13	32, 73, 132, 198	0

The worst 5 of 133 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	597	THR	6.2
1	C	571	VAL	6.1
1	B	575	GLY	6.0
1	C	583	THR	6.0
1	C	591	ASP	5.4

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	703	5/5	0.85	0.24	95,106,109,110	0
2	SO4	A	705	5/5	0.86	0.22	140,142,143,143	0
2	SO4	A	706	5/5	0.88	0.35	149,150,153,154	0
2	SO4	C	702	5/5	0.88	0.28	136,137,137,137	0
2	SO4	B	702	5/5	0.89	0.21	139,140,140,140	0
2	SO4	A	702	5/5	0.90	0.13	112,116,117,120	0
2	SO4	B	701	5/5	0.91	0.19	119,121,123,123	0
2	SO4	C	701	5/5	0.93	0.12	103,106,110,110	0
2	SO4	A	707	5/5	0.93	0.30	122,125,127,128	0
2	SO4	A	701	5/5	0.95	0.12	82,84,86,89	0
2	SO4	A	704	5/5	0.95	0.14	95,101,102,105	0
2	SO4	A	703	5/5	0.96	0.15	71,74,79,82	0
2	SO4	B	704	5/5	0.97	0.09	96,99,103,103	0

6.5 Other polymers (i)

There are no such residues in this entry.