wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 07:51 AM GMT

PDB ID	:	3CES
Title	:	Crystal Structure of E.coli MnmG (GidA), a Highly-Conserved tRNA Modifying Enzyme
Authors	:	Shi, R.; Matte, A.; Cygler, M.; Montreal-Kingston Bacterial Structural Genomics Initiative (BSGI)
Deposited on	:	2008-02-29
Resolution	:	2.41 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

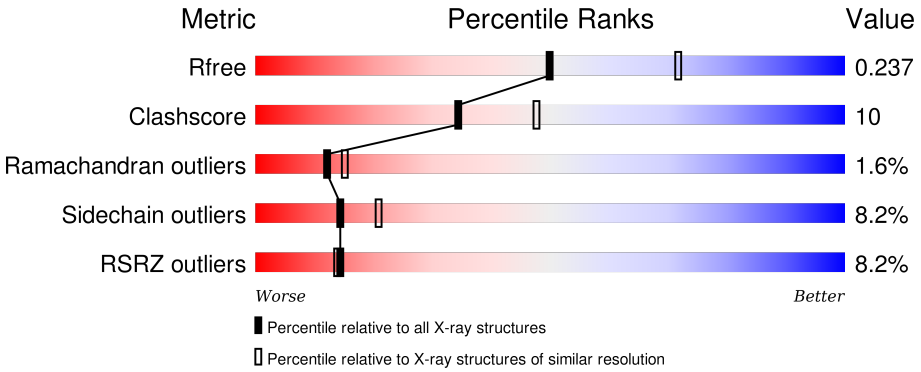
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.41 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}	91344	3386 (2.44-2.40)
Clashscore	102246	3897 (2.44-2.40)
Ramachandran outliers	100387	3837 (2.44-2.40)
Sidechain outliers	100360	3838 (2.44-2.40)
RSRZ outliers	91569	3396 (2.44-2.40)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	651	% 62% 15% • 18%
1	B	651	7% 62% 15% • 20%
1	C	651	9% 59% 17% • 21%
1	D	651	10% 59% 18% • 22%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 15980 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 tRNA uridine 5-carboxymethylaminomethyl modification enzyme gidA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	531	Total	C	N	O	S	0	1	0
			4098	2562	739	778	19			
1	B	519	Total	C	N	O	S	0	1	0
			3895	2447	696	735	17			
1	C	516	Total	C	N	O	S	0	0	0
			3884	2433	699	734	18			
1	D	507	Total	C	N	O	S	0	1	0
			3771	2370	678	705	18			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	EXPRESSION TAG	UNP P0A6U3
A	-20	GLY	-	EXPRESSION TAG	UNP P0A6U3
A	-19	SER	-	EXPRESSION TAG	UNP P0A6U3
A	-18	SER	-	EXPRESSION TAG	UNP P0A6U3
A	-17	HIS	-	EXPRESSION TAG	UNP P0A6U3
A	-16	HIS	-	EXPRESSION TAG	UNP P0A6U3
A	-15	HIS	-	EXPRESSION TAG	UNP P0A6U3
A	-14	HIS	-	EXPRESSION TAG	UNP P0A6U3
A	-13	HIS	-	EXPRESSION TAG	UNP P0A6U3
A	-12	HIS	-	EXPRESSION TAG	UNP P0A6U3
A	-11	HIS	-	EXPRESSION TAG	UNP P0A6U3
A	-10	HIS	-	EXPRESSION TAG	UNP P0A6U3
A	-9	SER	-	EXPRESSION TAG	UNP P0A6U3
A	-8	SER	-	EXPRESSION TAG	UNP P0A6U3
A	-7	GLY	-	EXPRESSION TAG	UNP P0A6U3
A	-6	LEU	-	EXPRESSION TAG	UNP P0A6U3
A	-5	VAL	-	EXPRESSION TAG	UNP P0A6U3
A	-4	PRO	-	EXPRESSION TAG	UNP P0A6U3
A	-3	ARG	-	EXPRESSION TAG	UNP P0A6U3
A	-2	GLY	-	EXPRESSION TAG	UNP P0A6U3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	EXPRESSION TAG	UNP P0A6U3
A	0	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-21	MET	-	EXPRESSION TAG	UNP P0A6U3
B	-20	GLY	-	EXPRESSION TAG	UNP P0A6U3
B	-19	SER	-	EXPRESSION TAG	UNP P0A6U3
B	-18	SER	-	EXPRESSION TAG	UNP P0A6U3
B	-17	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-16	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-15	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-14	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-13	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-12	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-11	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-10	HIS	-	EXPRESSION TAG	UNP P0A6U3
B	-9	SER	-	EXPRESSION TAG	UNP P0A6U3
B	-8	SER	-	EXPRESSION TAG	UNP P0A6U3
B	-7	GLY	-	EXPRESSION TAG	UNP P0A6U3
B	-6	LEU	-	EXPRESSION TAG	UNP P0A6U3
B	-5	VAL	-	EXPRESSION TAG	UNP P0A6U3
B	-4	PRO	-	EXPRESSION TAG	UNP P0A6U3
B	-3	ARG	-	EXPRESSION TAG	UNP P0A6U3
B	-2	GLY	-	EXPRESSION TAG	UNP P0A6U3
B	-1	SER	-	EXPRESSION TAG	UNP P0A6U3
B	0	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-21	MET	-	EXPRESSION TAG	UNP P0A6U3
C	-20	GLY	-	EXPRESSION TAG	UNP P0A6U3
C	-19	SER	-	EXPRESSION TAG	UNP P0A6U3
C	-18	SER	-	EXPRESSION TAG	UNP P0A6U3
C	-17	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-16	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-15	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-14	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-13	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-12	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-11	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-10	HIS	-	EXPRESSION TAG	UNP P0A6U3
C	-9	SER	-	EXPRESSION TAG	UNP P0A6U3
C	-8	SER	-	EXPRESSION TAG	UNP P0A6U3
C	-7	GLY	-	EXPRESSION TAG	UNP P0A6U3
C	-6	LEU	-	EXPRESSION TAG	UNP P0A6U3
C	-5	VAL	-	EXPRESSION TAG	UNP P0A6U3
C	-4	PRO	-	EXPRESSION TAG	UNP P0A6U3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	ARG	-	EXPRESSION TAG	UNP P0A6U3
C	-2	GLY	-	EXPRESSION TAG	UNP P0A6U3
C	-1	SER	-	EXPRESSION TAG	UNP P0A6U3
C	0	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-21	MET	-	EXPRESSION TAG	UNP P0A6U3
D	-20	GLY	-	EXPRESSION TAG	UNP P0A6U3
D	-19	SER	-	EXPRESSION TAG	UNP P0A6U3
D	-18	SER	-	EXPRESSION TAG	UNP P0A6U3
D	-17	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-16	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-15	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-14	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-13	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-12	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-11	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-10	HIS	-	EXPRESSION TAG	UNP P0A6U3
D	-9	SER	-	EXPRESSION TAG	UNP P0A6U3
D	-8	SER	-	EXPRESSION TAG	UNP P0A6U3
D	-7	GLY	-	EXPRESSION TAG	UNP P0A6U3
D	-6	LEU	-	EXPRESSION TAG	UNP P0A6U3
D	-5	VAL	-	EXPRESSION TAG	UNP P0A6U3
D	-4	PRO	-	EXPRESSION TAG	UNP P0A6U3
D	-3	ARG	-	EXPRESSION TAG	UNP P0A6U3
D	-2	GLY	-	EXPRESSION TAG	UNP P0A6U3
D	-1	SER	-	EXPRESSION TAG	UNP P0A6U3
D	0	HIS	-	EXPRESSION TAG	UNP P0A6U3

- Molecule 2 is water.

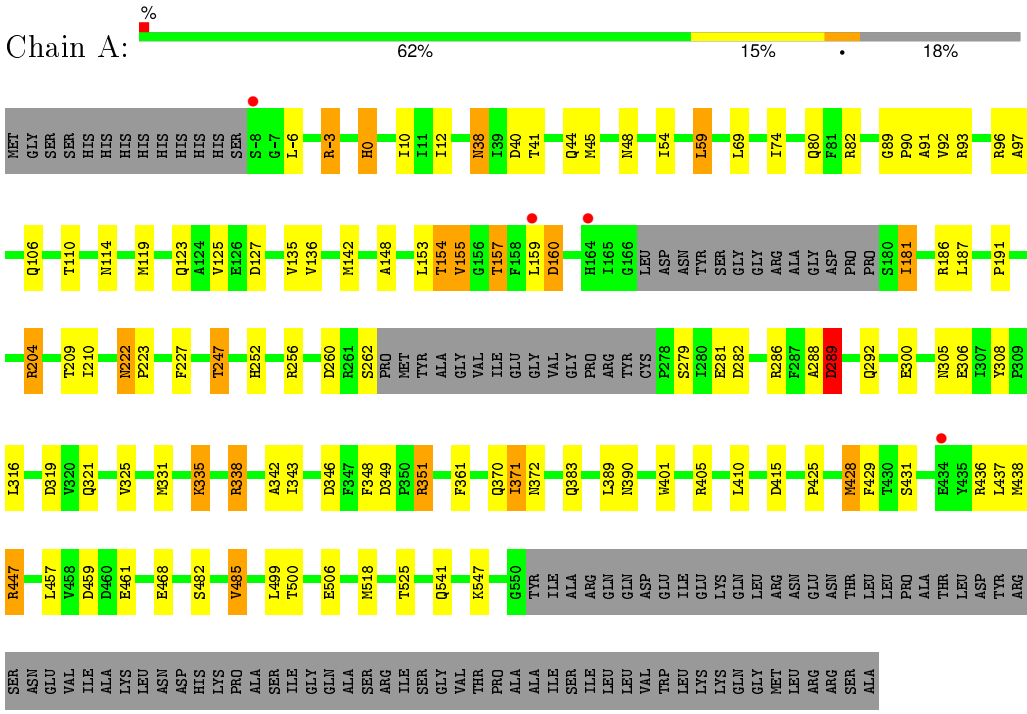
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	147	Total O 147 147	0	0
2	B	114	Total O 114 114	0	0
2	C	41	Total O 41 41	0	0
2	D	30	Total O 30 30	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

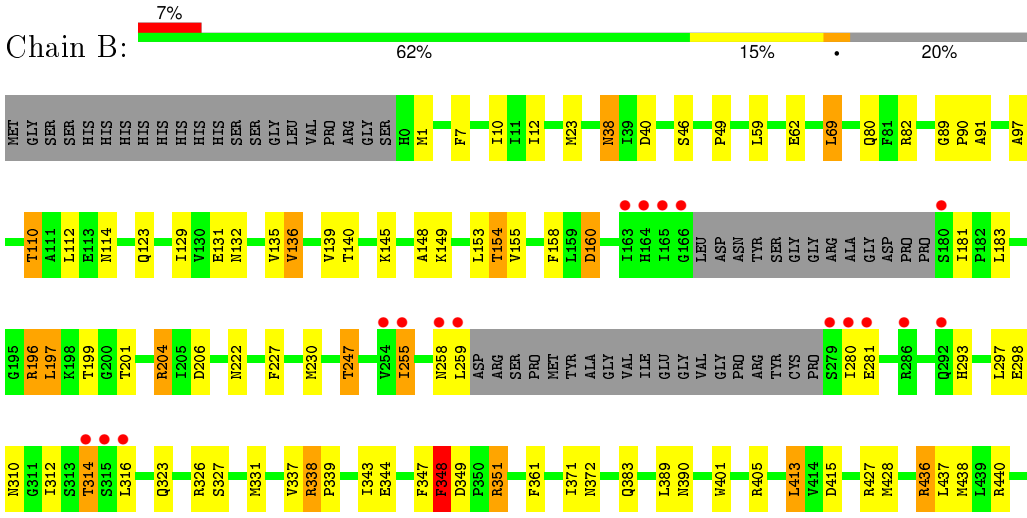
- Molecule 1: tRNA uridine 5-carboxymethylaminomethyl modification enzyme gidA

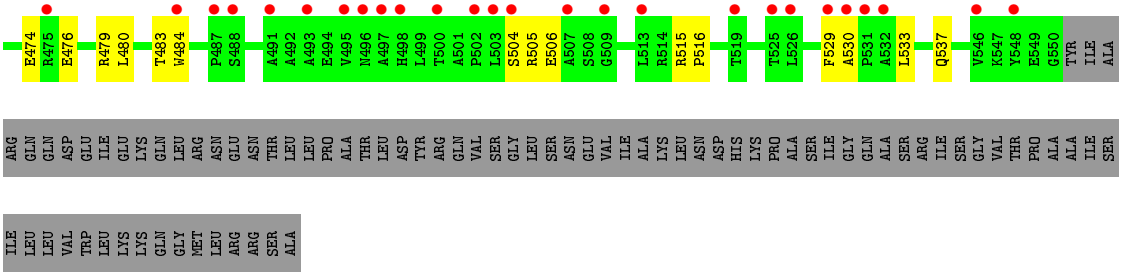
Chain A:



- Molecule 1: tRNA uridine 5-carboxymethylaminomethyl modification enzyme gidA

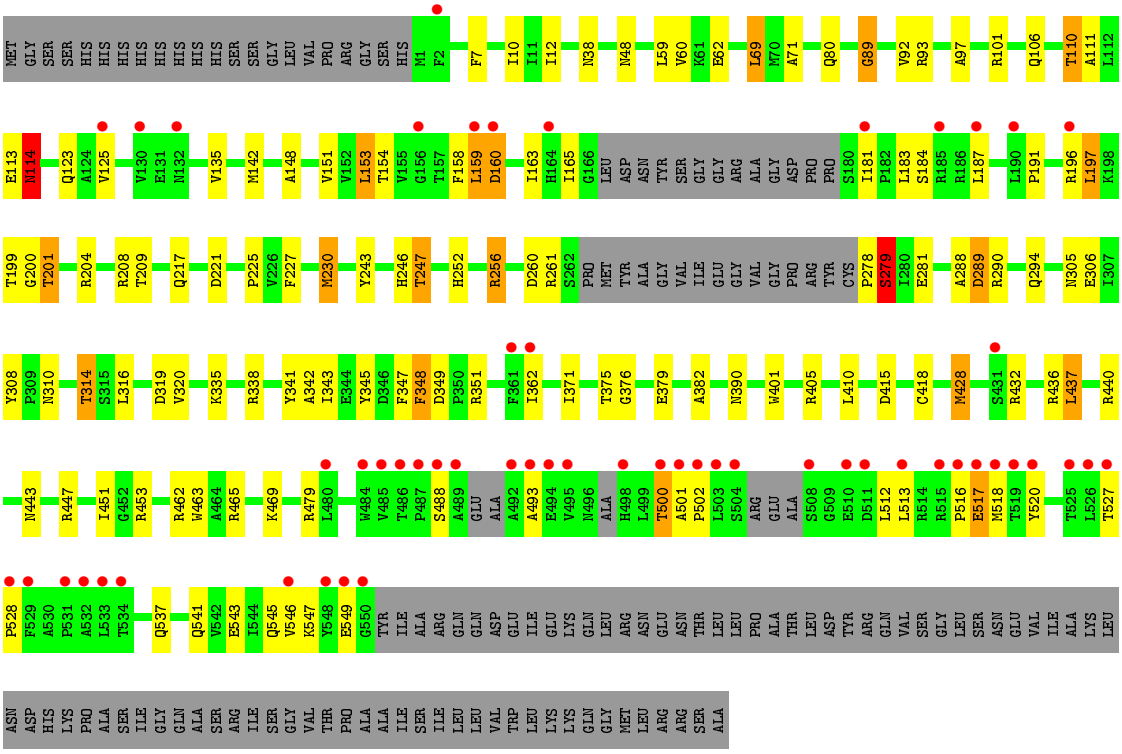
Chain B:





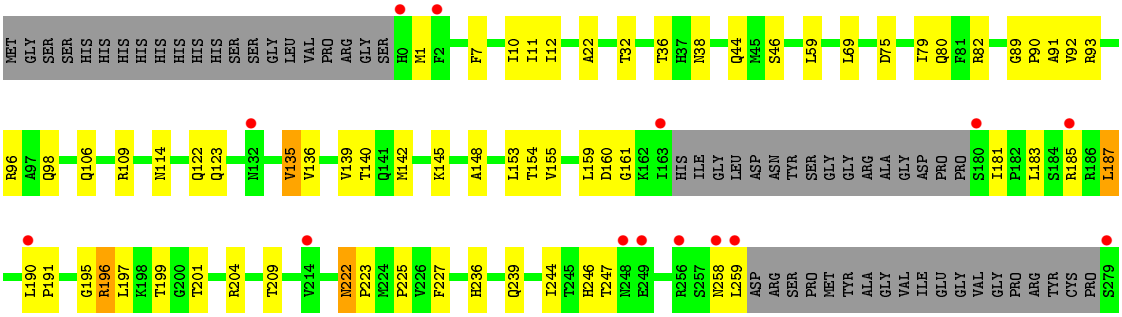
• Molecule 1: tRNA uridine 5-carboxymethylaminomethyl modification enzyme gidA

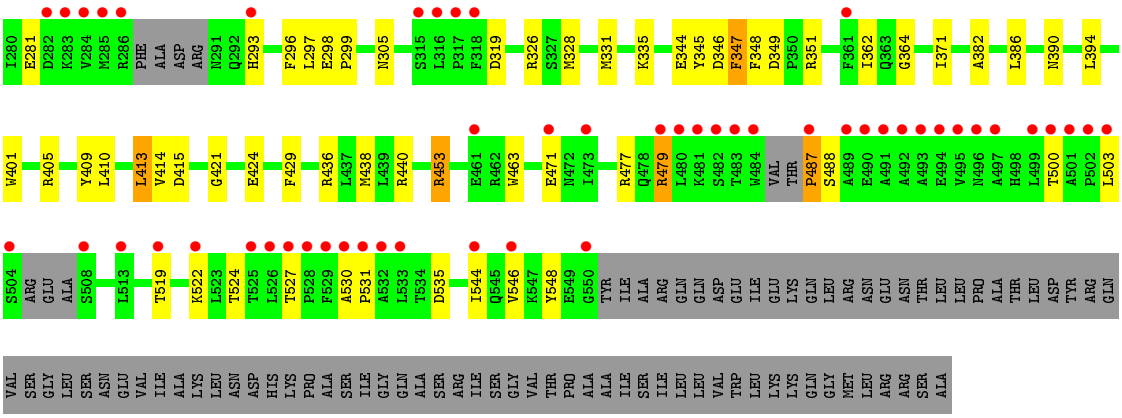
Chain C:



• Molecule 1: tRNA uridine 5-carboxymethylaminomethyl modification enzyme gidA

Chain D:




I280 E281 D282 K283 V284 M285 R286 PHE ALA ASP ARG N291 Q292 H293 F296 L297 E298 P299 N305 S315 L316 P317 F318 D319 R326 S327 M328 M331 K335 E344 Y345 D346 F347 F348 D349 P350 R351 F361 I362 Q363 G364 I371 A382 L386 N390 L394
W401 R405 Y409 L410 L413 V414 D415 G421 E424 F429 R436 L437 M438 L439 R440 R453 E461 R462 W463 E471 N472 I473 R477 Q478 R479 L480 K481 S482 T483 W484 VAL THR P487 S488 A489 E490 A491 A492 A493 E494 V495 N496 A497 H498 L499 T500 A501 P502 L503
S504 ARG GLU ALA S508 L513 T519 K522 L523 T524 T525 L526 T527 P528 F529 A530 P531 A532 L533 T534 D535 I544 Q545 V546 K547 Y548 E549 G550 TYR ILE ALA ARG GLN GLN ASP GLU ILE GLU LYS GLN LEU ARG ASN GLU ASN THR LEU LEU PRO ALA THR LEU ASP TYR ARG GLN
VAL SER GLY LEU SER ASN GLU VAL ILE ALA LYS LEU ASN ASP HIS LYS PRO ALA SER ILE GLY GLN ALA SER ARG ILE SER GLY VAL THR PRO ALA ALA ILE SER ILE LEU LEU VAL TRP LEU LYS LYS GLN GLY MET LEU ARG ARG SER ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	85.90Å 144.33Å 147.75Å 90.00° 106.63° 90.00°	Depositor
Resolution (Å)	141.42 – 2.41 47.75 – 2.41	Depositor EDS
% Data completeness (in resolution range)	97.4 (141.42-2.41) 97.4 (47.75-2.41)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.08 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, R_{free}	0.201 , 0.238 0.201 , 0.237	Depositor DCC
R_{free} test set	6511 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	41.1	Xtriage
Anisotropy	0.531	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 52.7	EDS
Estimated twinning fraction	0.014 for h,-k,-h-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 129227 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	15980	wwPDB-VP
Average B, all atoms (Å^2)	54.0	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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.78	2/4174 (0.0%)	0.87	13/5651 (0.2%)
1	B	0.68	0/3970	0.85	14/5391 (0.3%)
1	C	0.54	1/3949 (0.0%)	0.85	10/5354 (0.2%)
1	D	0.47	0/3843	0.84	6/5213 (0.1%)
All	All	0.63	3/15936 (0.0%)	0.85	43/21609 (0.2%)

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	B	1	1
1	C	0	2
1	D	0	3
All	All	1	6

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	0	HIS	C-N	14.28	1.66	1.34
1	C	500	THR	C-N	-6.51	1.19	1.34
1	A	506	GLU	CB-CG	-5.92	1.40	1.52

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	487	PRO	O-C-N	-35.11	66.53	122.70
1	C	500	THR	O-C-N	-32.11	71.33	122.70
1	D	487	PRO	CA-C-N	15.40	151.09	117.20
1	A	447	ARG	NE-CZ-NH2	-9.57	115.52	120.30
1	C	500	THR	CA-C-N	-8.81	97.82	117.20

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	348	PHE	CA

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	89	GLY	Peptide
1	C	500	THR	Mainchain
1	C	89	GLY	Peptide
1	D	487	PRO	Mainchain
1	D	89	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	4098	0	4061	95	0
1	B	3895	0	3741	71	0
1	C	3884	0	3757	78	0
1	D	3771	0	3604	73	0
2	A	147	0	0	5	0
2	B	114	0	0	2	0
2	C	41	0	0	2	0
2	D	30	0	0	2	0
All	All	15980	0	15163	309	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 10.

The worst 5 of 309 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:230:MET:HE2	1:C:230:MET:HA	1.25	1.12
1:D:154:THR:HG21	2:D:647:HOH:O	1.51	1.10
1:B:49:PRO:HB3	1:B:101:ARG:NH1	1.67	1.09
1:C:288:ALA:O	1:C:289:ASP:HB2	1.52	1.03

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:90:PRO:HB2	1:D:440:ARG:HD2	1.39	1.02

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	526/651 (81%)	507 (96%)	14 (3%)	5 (1%)	19	27
1	B	514/651 (79%)	476 (93%)	33 (6%)	5 (1%)	19	27
1	C	504/651 (77%)	461 (92%)	28 (6%)	15 (3%)	5	4
1	D	496/651 (76%)	466 (94%)	22 (4%)	8 (2%)	12	15
All	All	2040/2604 (78%)	1910 (94%)	97 (5%)	33 (2%)	12	15

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	160	ASP
1	B	348	PHE
1	B	437	LEU
1	B	504	SER
1	C	279	SER

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	429/535 (80%)	389 (91%)	40 (9%)	11	16
1	B	383/535 (72%)	349 (91%)	34 (9%)	12	17
1	C	389/535 (73%)	361 (93%)	28 (7%)	18	27
1	D	365/535 (68%)	339 (93%)	26 (7%)	18	28
All	All	1566/2140 (73%)	1438 (92%)	128 (8%)	14	21

5 of 128 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	204	ARG
1	B	413	LEU
1	D	227	PHE
1	B	227	PHE
1	B	314	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 52 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	323	GLN
1	C	123	GLN
1	D	310	ASN
1	B	383	GLN
1	C	38	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 carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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	531/651 (81%)	0.07	4 (0%) 87 87	20, 34, 63, 86	0
1	B	519/651 (79%)	0.43	43 (8%) 14 13	23, 38, 125, 136	0
1	C	516/651 (79%)	0.71	56 (10%) 7 7	36, 58, 105, 148	4 (0%)
1	D	507/651 (77%)	0.75	66 (13%) 5 4	40, 61, 128, 160	1 (0%)
All	All	2073/2604 (79%)	0.49	169 (8%) 14 14	20, 49, 114, 160	5 (0%)

The worst 5 of 169 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	504	SER	28.3
1	C	503	LEU	22.1
1	C	502	PRO	15.0
1	D	259	LEU	8.7
1	D	526	LEU	7.6

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.