wwPDB X-ray Structure Validation Summary Report (i)

Nov 1, 2023 – 02:43 PM JST

PDB ID	:	5I8I
Title	:	Crystal Structure of the K. lactis Urea Amidolyase
Authors	:	Zhao, J.; Xiang, S.
Deposited on	:	2016-02-19
Resolution	:	6.50 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:

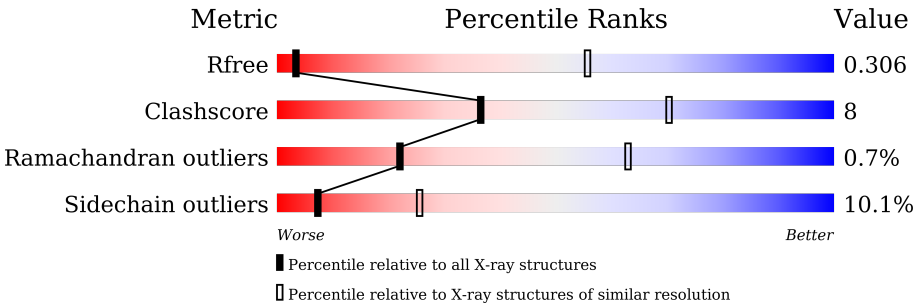
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 6.50 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	1000 (9.00-3.90)
Clashscore	141614	1064 (9.00-3.90)
Ramachandran outliers	138981	1012 (9.00-3.88)
Sidechain outliers	138945	1010 (9.00-3.84)

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%

Mol	Chain	Length	Quality of chain
1	A	1829	70% 18% • 9%
1	B	1829	69% 17% • 9%
1	C	1829	70% 17% • 9%
1	D	1829	70% 17% • 9%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 51722 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 Urea Amidolyase.

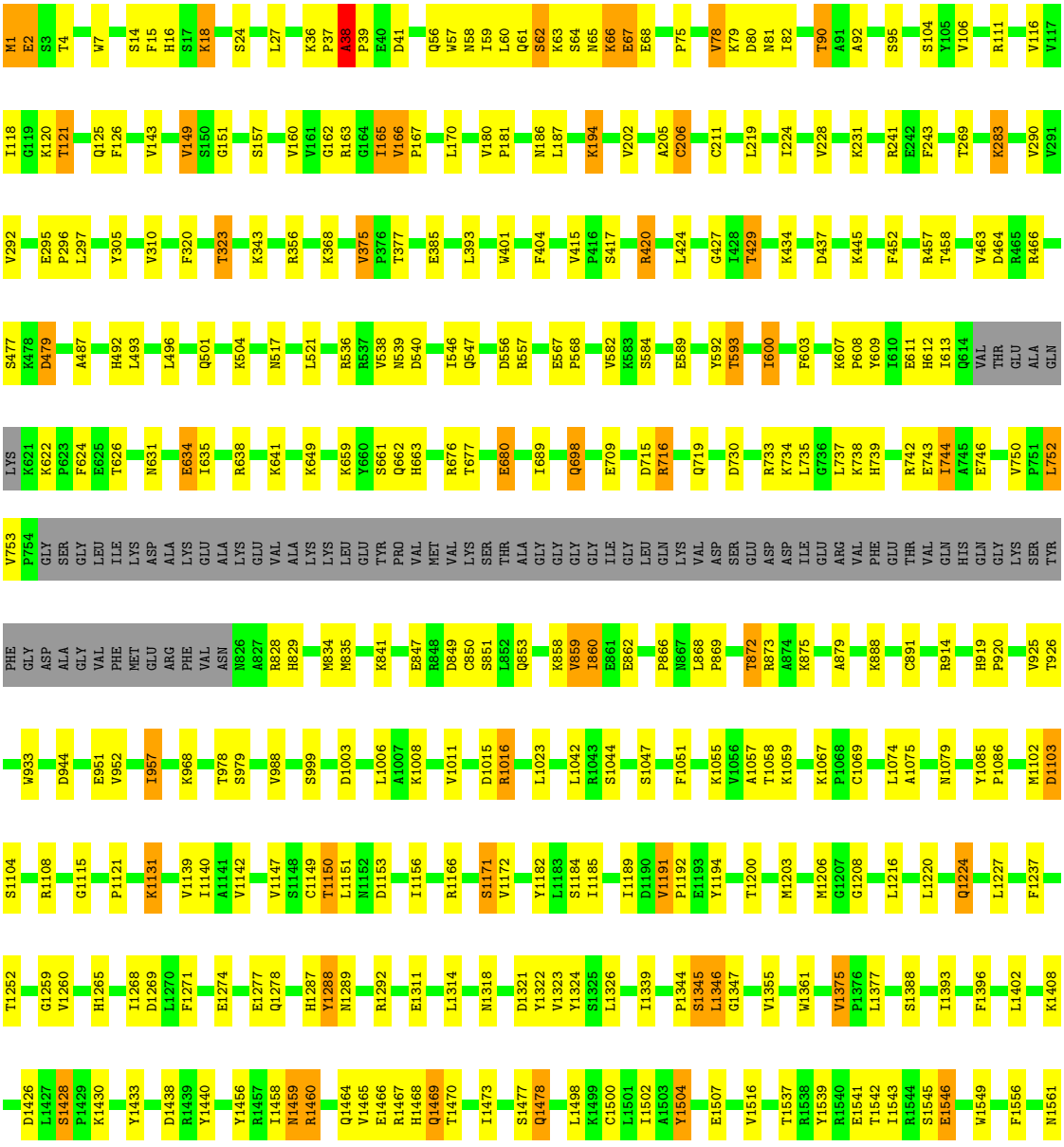
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1660	Total	C	N	O	S	0	0	0
			12939	8253	2183	2456	47			
1	B	1658	Total	C	N	O	S	0	0	0
			12922	8243	2181	2452	46			
1	C	1660	Total	C	N	O	S	0	0	0
			12939	8253	2183	2456	47			
1	D	1658	Total	C	N	O	S	0	0	0
			12922	8243	2181	2452	46			

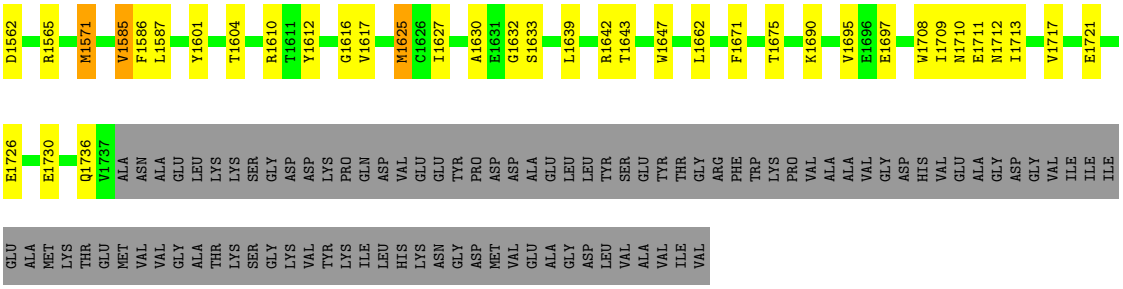
3 Residue-property plots (i)

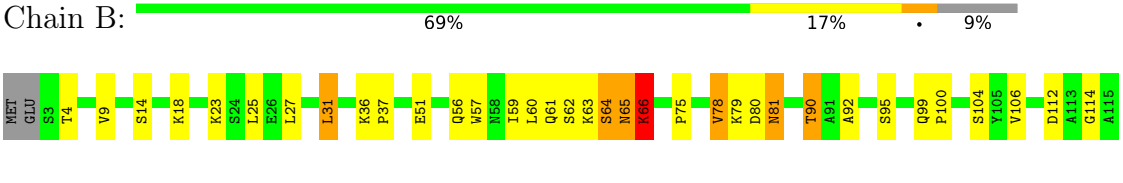
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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: Urea Amidolyase

Chain A:

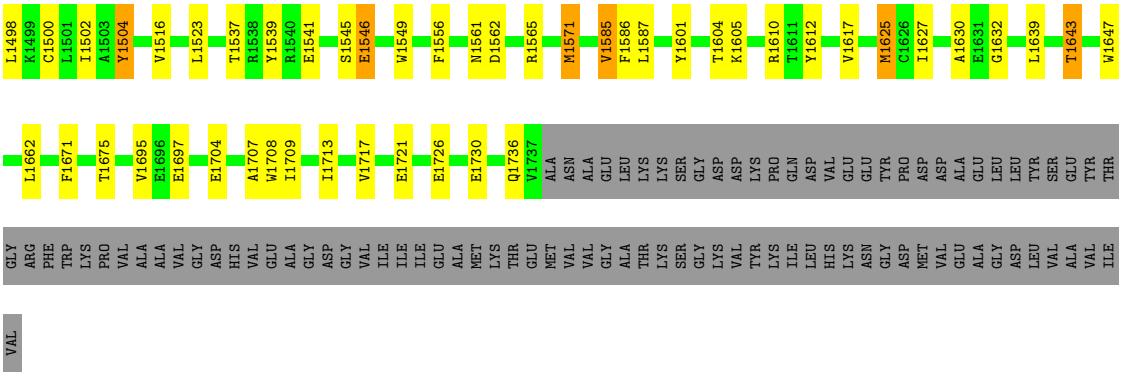






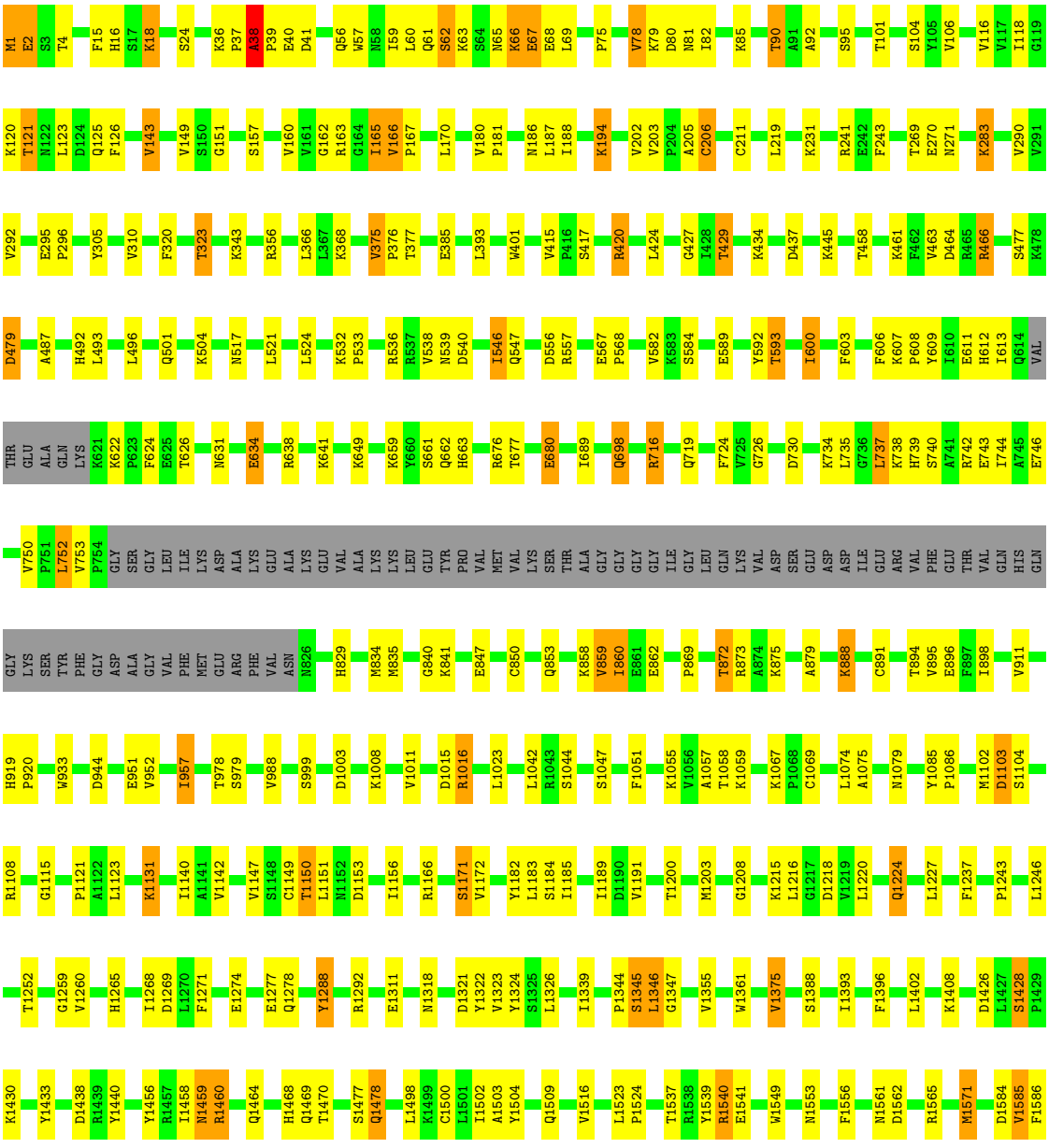
• Molecule 1: Urea Amidolyase

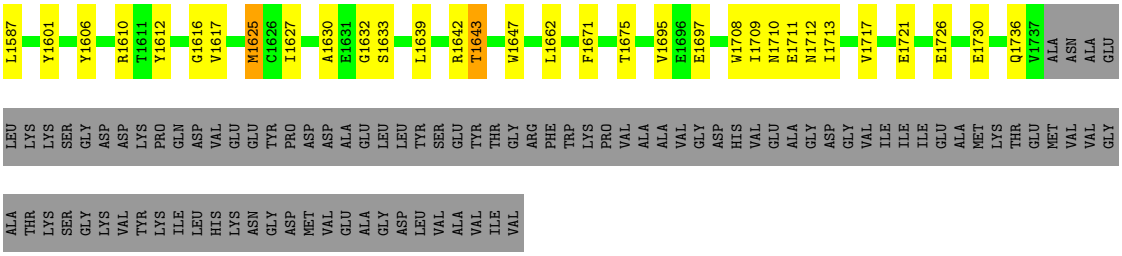
Chain B:



• Molecule 1: Urea Amidolyase

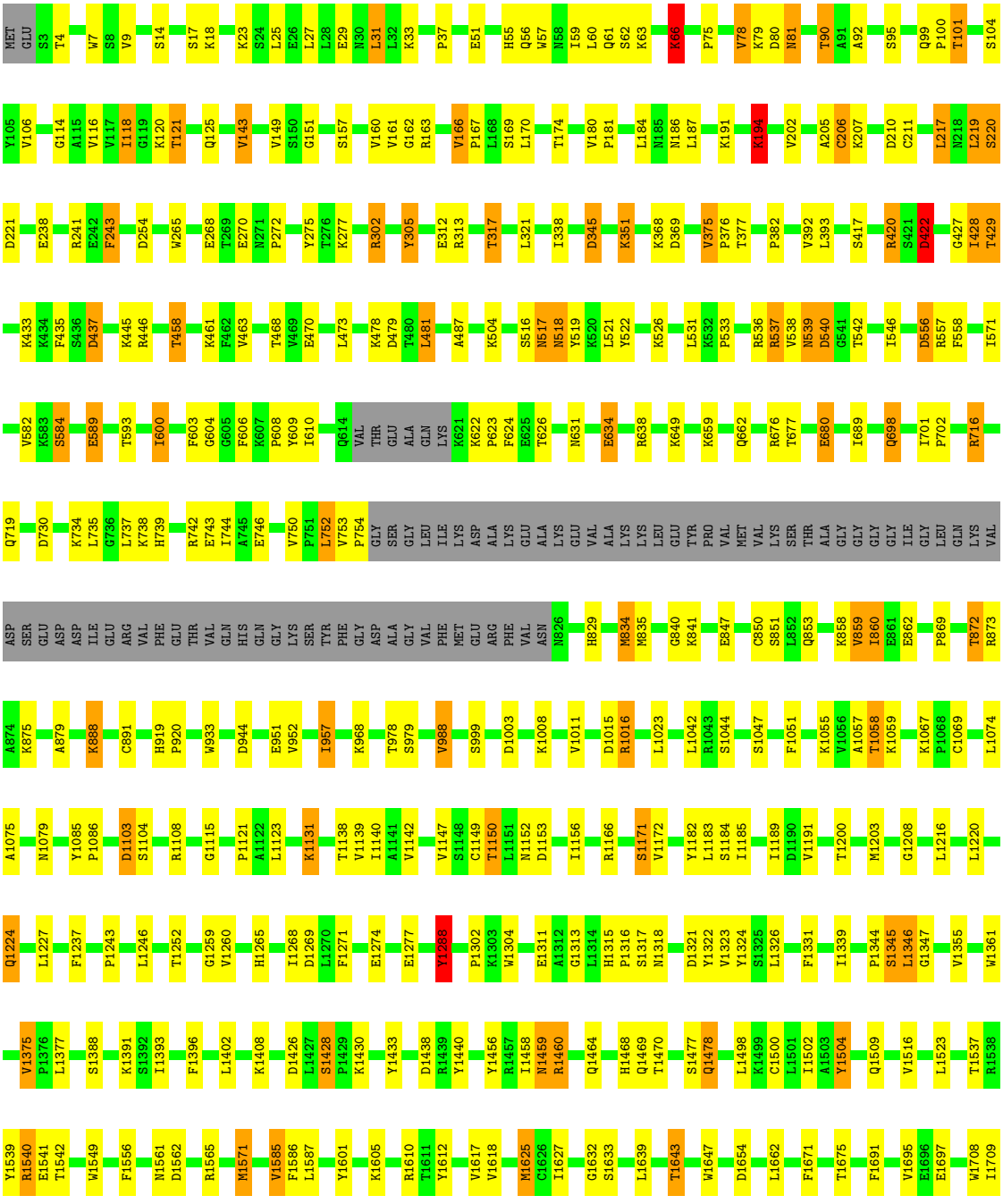
Chain C:

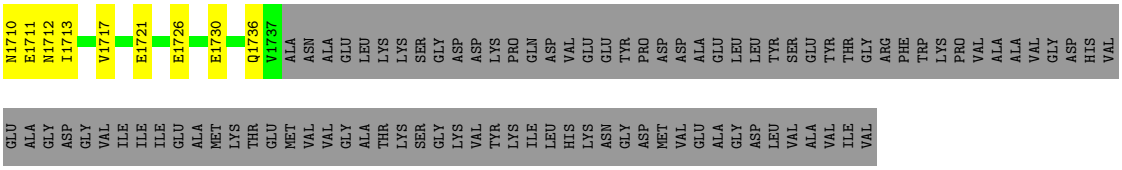




• Molecule 1: Urea Amidolyase

Chain D:




N1710
E1711
N1712
I1713
V1717
E1721
E1726
E1730
Q1736
V1737
ALA ASN ALA GLU LEU LYS LYS SER GLY ASP ASP LYS PRO GLN ASP VAL GLU GLU TYR PRO ASP ASP ALA GLU LEU LEU TYR SER GLU TYR THR GLY ARG PHE TRP LYS PRO VAL ALA ALA VAL GLY ASP HIS VAL
GLU ALA GLY ASP GLY VAL ILE ILE ILE GLU ALA MET LYS THR GLU MET VAL VAL GLY ALA THR LYS SER GLY LYS VAL TYR LYS ILE LEU HIS LYS ASN GLY ASP MET VAL GLU ALA GLY ASP LEU VAL ALA VAL ILE VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	105.74Å 181.94Å 549.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.98 – 6.50 29.98 – 6.50	Depositor EDS
% Data completeness (in resolution range)	95.8 (29.98-6.50) 96.0 (29.98-6.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	2.47 (at 6.58Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.278 , 0.302 0.281 , 0.306	Depositor DCC
R_{free} test set	1017 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	216.1	Xtriage
Anisotropy	0.897	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 238.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.78	EDS
Total number of atoms	51722	wwPDB-VP
Average B, all atoms (Å^2)	308.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% 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

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	0.40	1/13231 (0.0%)	0.70	4/17964 (0.0%)
1	B	0.51	2/13214 (0.0%)	0.86	23/17942 (0.1%)
1	C	0.39	1/13231 (0.0%)	0.69	5/17964 (0.0%)
1	D	0.49	3/13214 (0.0%)	0.83	21/17942 (0.1%)
All	All	0.45	7/52890 (0.0%)	0.77	53/71812 (0.1%)

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	2
1	B	0	2
1	C	0	3
1	D	0	6
All	All	0	13

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1546	GLU	CG-CD	-7.32	1.41	1.51
1	C	1288	TYR	CD2-CE2	6.77	1.49	1.39
1	B	1546	GLU	CG-CD	-5.72	1.43	1.51
1	D	17	SER	CA-CB	5.61	1.61	1.52
1	D	1288	TYR	CD2-CE2	5.39	1.47	1.39

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	556	ASP	N-CA-CB	-8.99	94.41	110.60
1	B	556	ASP	N-CA-CB	-8.84	94.69	110.60
1	A	1504	TYR	CB-CG-CD1	-8.11	116.14	121.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	1504	TYR	CB-CG-CD1	-7.93	116.24	121.00
1	D	422	ASP	N-CA-CB	-7.68	96.78	110.60

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1469	GLN	Mainchain
1	A	539	ASN	Peptide
1	B	1469	GLN	Mainchain
1	B	539	ASN	Peptide
1	C	539	ASN	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	12939	0	12896	226	0
1	B	12922	0	12878	218	1
1	C	12939	0	12896	203	1
1	D	12922	0	12878	215	0
All	All	51722	0	51548	813	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 8.

The worst 5 of 813 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:59:ILE:O	1:B:62:SER:OG	1.78	1.01
1:B:79:LYS:HA	1:B:121:THR:HG22	1.45	0.97
1:C:1509:GLN:NE2	1:D:101:THR:O	1.97	0.97
1:D:79:LYS:HA	1:D:121:THR:HG22	1.45	0.95
1:A:493:LEU:HB2	1:A:496:LEU:HD12	1.47	0.94

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1288:TYR:OH	1:C:1541:GLU:OE2[2_455]	1.97	0.23

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	1654/1829 (90%)	1540 (93%)	101 (6%)	13 (1%)	19	60
1	B	1652/1829 (90%)	1544 (94%)	97 (6%)	11 (1%)	22	63
1	C	1654/1829 (90%)	1539 (93%)	102 (6%)	13 (1%)	19	60
1	D	1652/1829 (90%)	1544 (94%)	97 (6%)	11 (1%)	22	63
All	All	6612/7316 (90%)	6167 (93%)	397 (6%)	48 (1%)	22	63

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	38	ALA
1	A	1166	ARG
1	A	1711	GLU
1	B	1166	ARG
1	C	38	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	1413/1547 (91%)	1275 (90%)	138 (10%)	8	27
1	B	1411/1547 (91%)	1264 (90%)	147 (10%)	7	25
1	C	1413/1547 (91%)	1274 (90%)	139 (10%)	8	27
1	D	1411/1547 (91%)	1263 (90%)	148 (10%)	7	24
All	All	5648/6188 (91%)	5076 (90%)	572 (10%)	7	25

5 of 572 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	428	ILE
1	D	538	VAL
1	D	420	ARG
1	D	1015	ASP
1	B	600	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	65	ASN
1	D	1459	ASN
1	C	829	HIS
1	D	1509	GLN
1	D	829	HIS

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)

There are no ligands in this entry.

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.