wwPDB X-ray Structure Validation Summary Report (i)

Apr 29, 2024 – 12:31 pm BST

PDB ID	:	2IUS
Title	:	E. coli FtsK motor domain
Authors	:	Massey, T.H.; Mercogliano, C.P.; Yates, J.; Sherratt, D.J.; Lowe, J.
Deposited on	:	2006-06-07
Resolution	:	2.70 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:

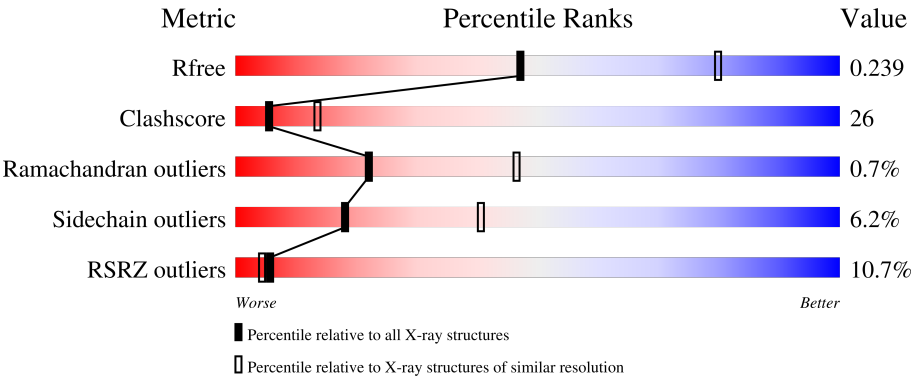
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.70 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.



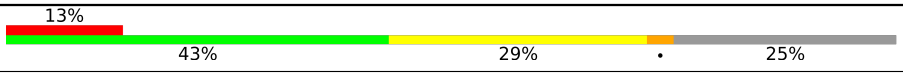
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	512	5% 43% 31% · 23%
1	B	512	5% 46% 27% · 23%
1	C	512	10% 44% 30% · 23%
1	D	512	6% 45% 30% · 23%
1	E	512	10% 43% 28% · 26%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	512	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 18569 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 DNA TRANSLOCASE FTSK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	S	0	0	1
			3029	1930	532	551	16			
1	B	393	Total	C	N	O	S	0	0	1
			3029	1930	532	551	16			
1	C	393	Total	C	N	O	S	0	0	1
			3029	1930	532	551	16			
1	D	394	Total	C	N	O	S	0	0	1
			3036	1935	533	552	16			
1	E	381	Total	C	N	O	S	0	0	1
			2941	1877	512	536	16			
1	F	382	Total	C	N	O	S	0	0	1
			2952	1883	516	537	16			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	997	ALA	LYS	engineered mutation	UNP P46889
B	997	ALA	LYS	engineered mutation	UNP P46889
C	997	ALA	LYS	engineered mutation	UNP P46889
D	997	ALA	LYS	engineered mutation	UNP P46889
E	997	ALA	LYS	engineered mutation	UNP P46889
F	997	ALA	LYS	engineered mutation	UNP P46889

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	91	Total	O	0	0
			91	91		
2	B	94	Total	O	0	0
			94	94		
2	C	80	Total	O	0	0
			80	80		

Continued on next page...

Continued from previous page...

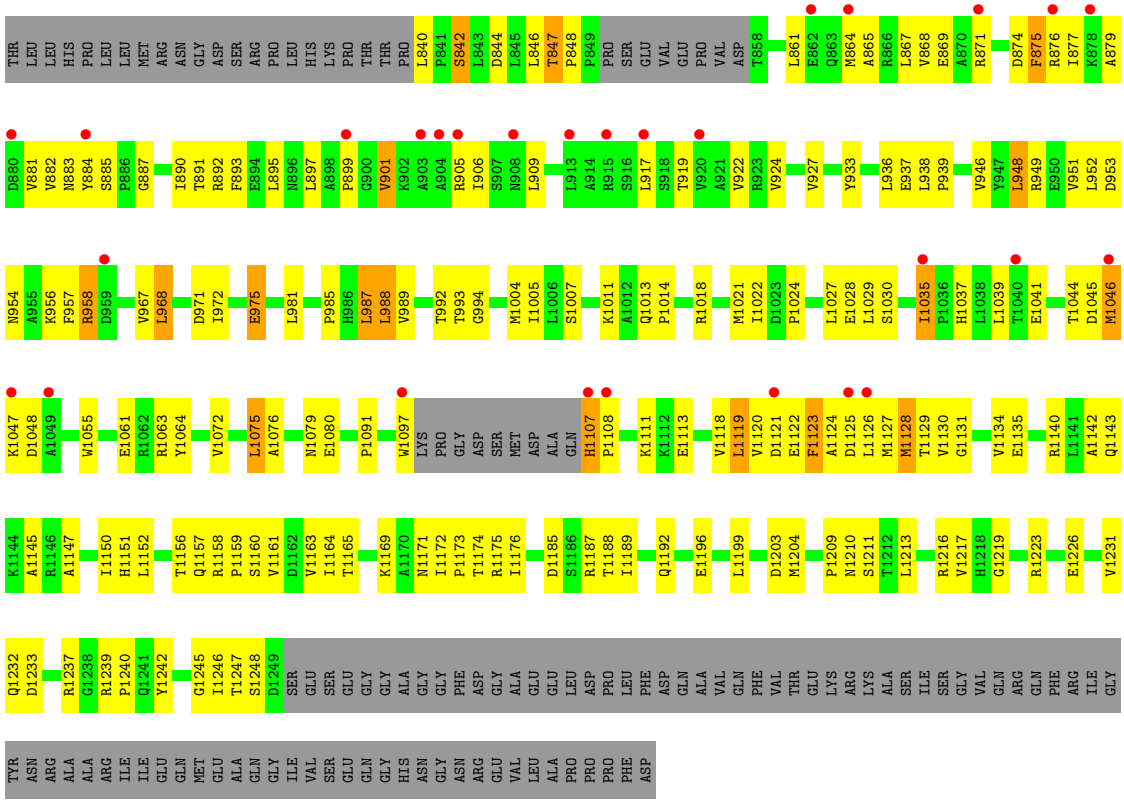
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	116	Total O 116 116	0	0
2	E	99	Total O 99 99	0	0
2	F	73	Total O 73 73	0	0

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.

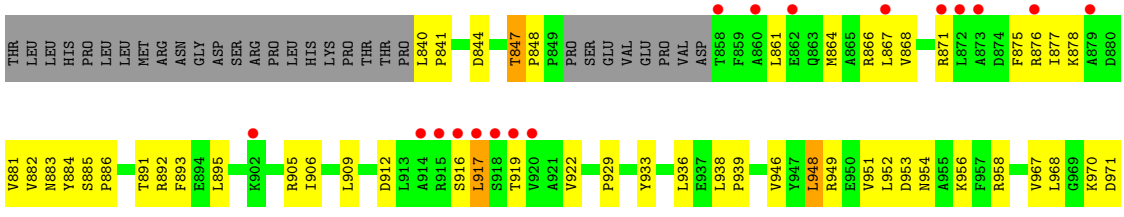
• Molecule 1: DNA TRANSLOCASE FTSK

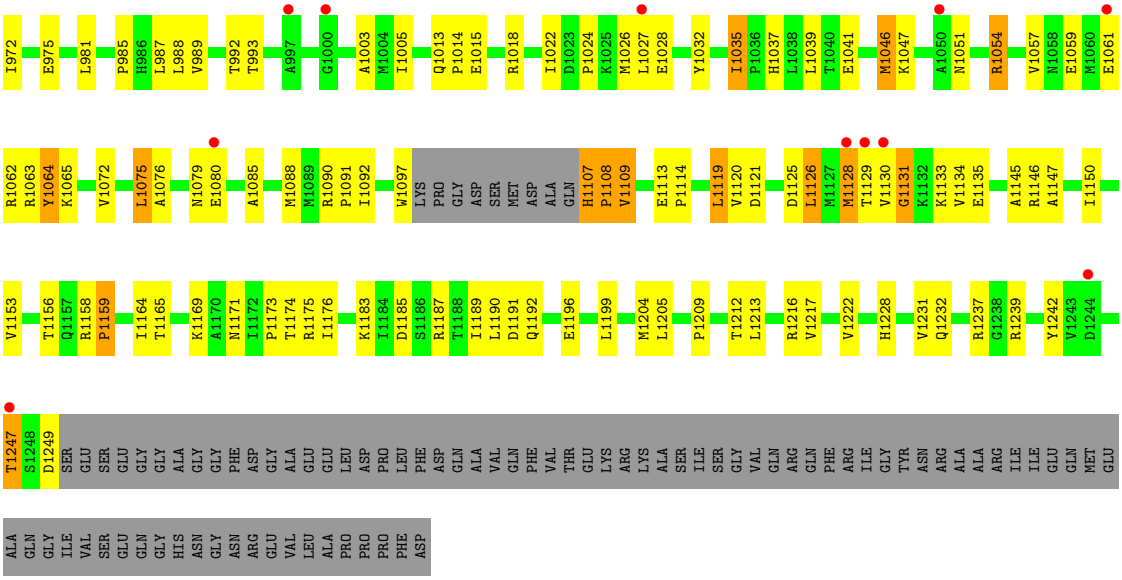
Chain A:



• Molecule 1: DNA TRANSLOCASE FTSK

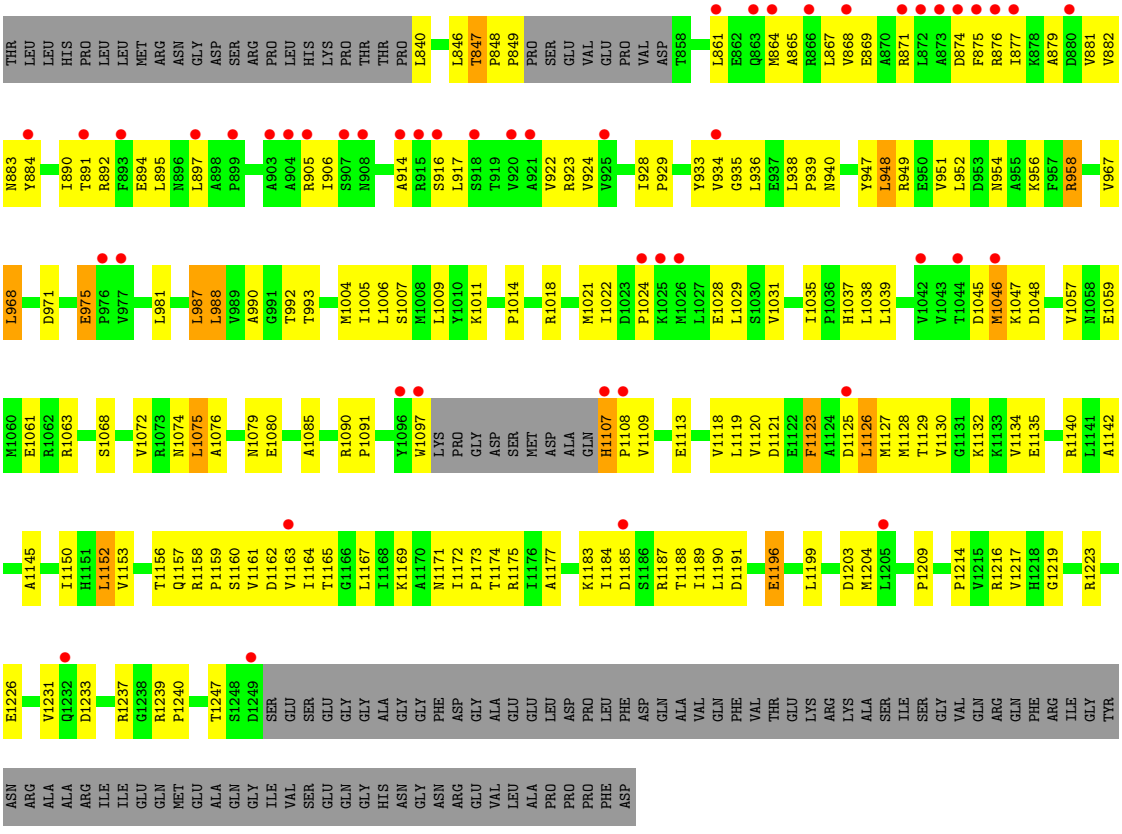
Chain B:





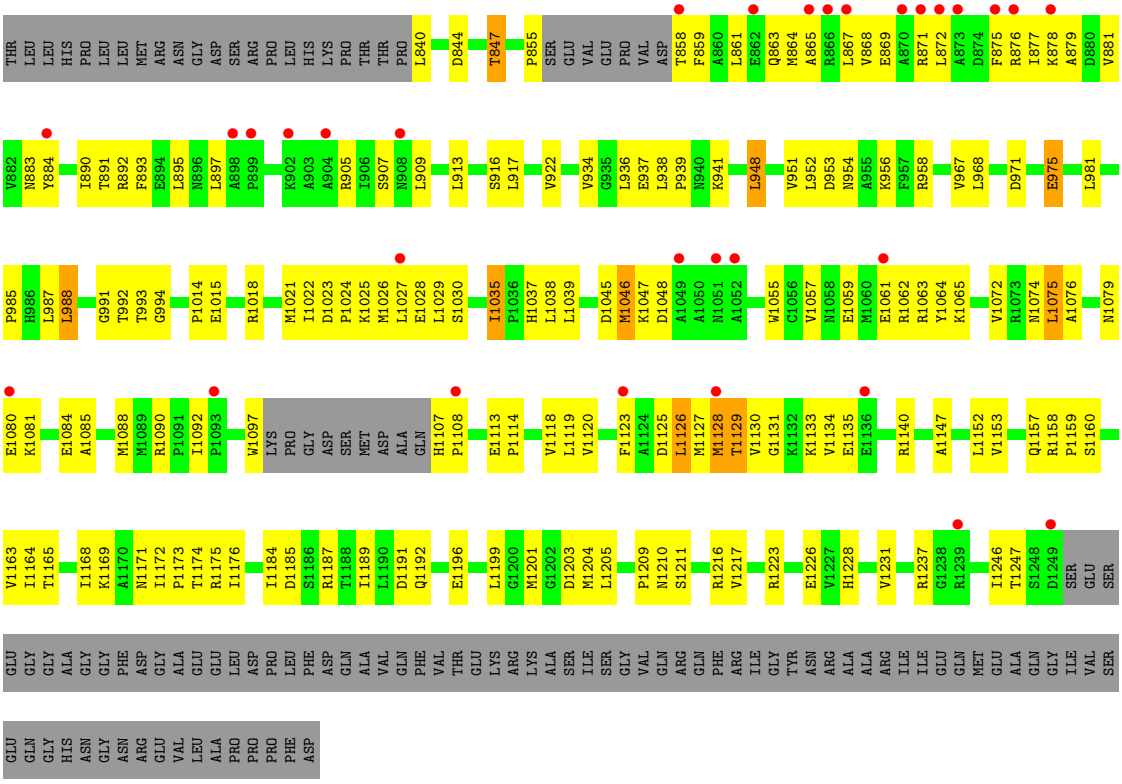
• Molecule 1: DNA TRANSLOCASE FTSK

Chain C:



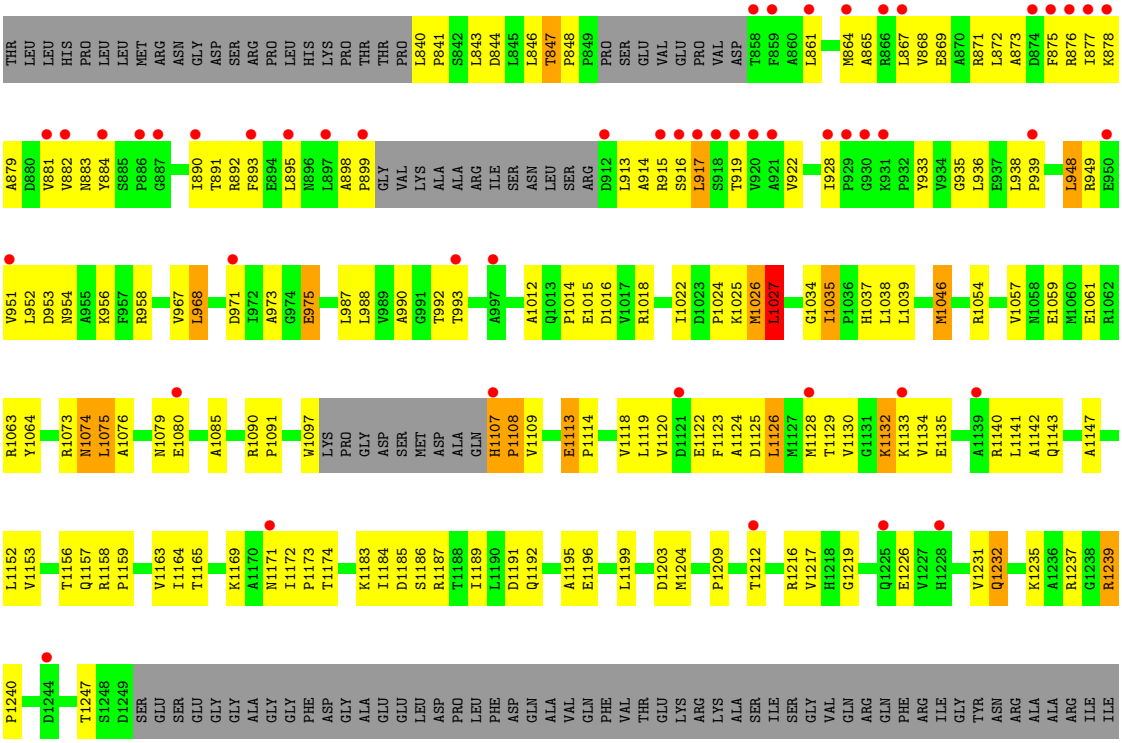
• Molecule 1: DNA TRANSLOCASE FTSK

Chain D:



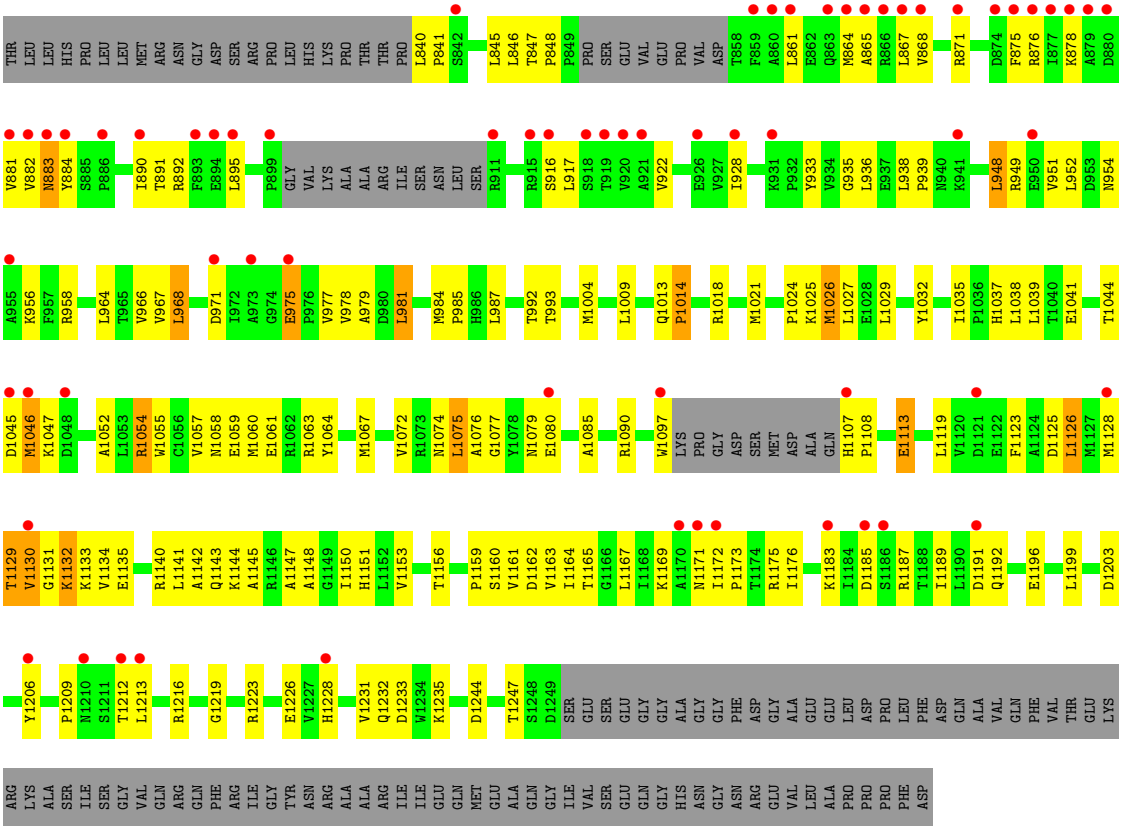
• Molecule 1: DNA TRANSLOCASE FTSK

Chain E:



• Molecule 1: DNA TRANSLOCASE FTSK

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	97.60Å 117.20Å 132.80Å 90.00° 100.50° 90.00°	Depositor
Resolution (Å)	100.00 – 2.70 130.58 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (100.00-2.70) 99.7 (130.58-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.53 (at 2.69Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.246 , 0.298 0.243 , 0.239	Depositor DCC
R_{free} test set	4056 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	45.5	Xtriage
Anisotropy	0.266	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 63.1	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	0.90	EDS
Total number of atoms	18569	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.21 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.8823e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.36	0/3083	0.66	0/4186
1	B	0.39	0/3083	0.66	0/4186
1	C	0.35	0/3083	0.63	0/4186
1	D	0.38	0/3090	0.65	0/4194
1	E	0.38	0/2994	0.67	0/4066
1	F	0.35	0/3005	0.63	0/4080
All	All	0.37	0/18338	0.65	0/24898

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

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	3029	0	3146	167	0
1	B	3029	0	3146	157	3
1	C	3029	0	3146	163	0
1	D	3036	0	3153	166	0
1	E	2941	0	3046	159	3
1	F	2952	0	3059	146	0
2	A	91	0	0	18	0
2	B	94	0	0	27	0
2	C	80	0	0	20	0
2	D	116	0	0	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	99	0	0	22	0
2	F	73	0	0	25	0
All	All	18569	0	18696	948	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 26.

The worst 5 of 948 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:E:1046:MET:HG2	1:E:1129:THR:HG21	1.20	1.10
1:D:1159:PRO:HB3	1:D:1189:ILE:HD11	1.38	1.06
1:E:1159:PRO:HB3	1:E:1189:ILE:HD11	1.34	1.05
1:B:1159:PRO:HB3	1:B:1189:ILE:HD11	1.33	1.05
1:F:1035:ILE:HG22	1:F:1038:LEU:HG	1.45	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:B:919:THR:N	1:E:919:THR:N[1_465]	1.90	0.30
1:B:919:THR:O	1:E:919:THR:O[1_465]	2.09	0.11
1:B:917:LEU:O	1:E:919:THR:CG2[1_465]	2.18	0.02

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	387/512 (76%)	355 (92%)	29 (8%)	3 (1%)	19 43
1	B	387/512 (76%)	360 (93%)	23 (6%)	4 (1%)	15 37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	387/512 (76%)	354 (92%)	33 (8%)	0	100	100
1	D	387/512 (76%)	355 (92%)	30 (8%)	2 (0%)	29	54
1	E	373/512 (73%)	347 (93%)	22 (6%)	4 (1%)	14	34
1	F	374/512 (73%)	347 (93%)	24 (6%)	3 (1%)	19	43
All	All	2295/3072 (75%)	2118 (92%)	161 (7%)	16 (1%)	22	46

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1131	GLY
1	D	1129	THR
1	E	1026	MET
1	E	1027	LEU
1	E	1108	PRO

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	330/429 (77%)	310 (94%)	20 (6%)	18	41
1	B	330/429 (77%)	310 (94%)	20 (6%)	18	41
1	C	330/429 (77%)	309 (94%)	21 (6%)	17	39
1	D	331/429 (77%)	314 (95%)	17 (5%)	24	50
1	E	321/429 (75%)	300 (94%)	21 (6%)	17	38
1	F	322/429 (75%)	299 (93%)	23 (7%)	14	34
All	All	1964/2574 (76%)	1842 (94%)	122 (6%)	18	40

5 of 122 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1128	MET
1	F	1054	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	1046	MET
1	F	1046	MET
1	F	1196	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 83 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	954	ASN
1	F	954	ASN
1	E	1051	ASN
1	E	1192	GLN
1	F	1107	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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	849:PRO	C	855:PRO	N	13.90

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	393/512 (76%)	0.67	28 (7%) 16 14	18, 43, 85, 100	0
1	B	393/512 (76%)	0.63	28 (7%) 16 14	14, 37, 77, 94	0
1	C	393/512 (76%)	1.00	49 (12%) 3 3	21, 49, 106, 120	0
1	D	394/512 (76%)	0.68	31 (7%) 12 10	16, 37, 76, 103	0
1	E	381/512 (74%)	1.01	50 (13%) 3 2	20, 45, 106, 128	0
1	F	382/512 (74%)	1.11	65 (17%) 1 1	19, 51, 118, 138	0
All	All	2336/3072 (76%)	0.85	251 (10%) 6 4	14, 43, 96, 138	0

The worst 5 of 251 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	918	SER	13.5
1	C	904	ALA	10.7
1	E	920	VAL	10.2
1	E	877	ILE	9.3
1	F	874	ASP	8.7

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 monosaccharides 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.