wwPDB X-ray Structure Validation Summary Report (i)

Nov 10, 2024 – 02:31 am GMT

PDB ID	:	1E2Y
Title	:	Tryparedoxin peroxidase from Crithidia fasciculata
Authors	:	Alphey, M.S.; Bond, C.S.; Hunter, W.N.
Deposited on	:	2000-05-30
Resolution	:	3.20 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 (1)) were used in the production of this report:

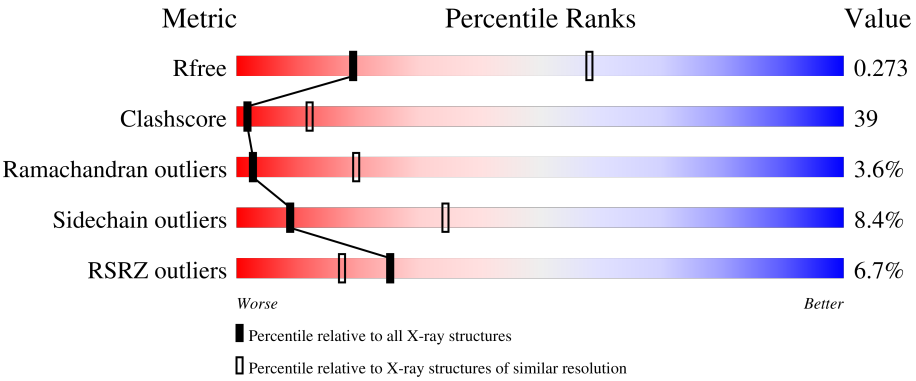
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 3.20 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}	164625	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (3.20-3.20)

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	188	3% 40% 44% • 11%
1	B	188	4% 43% 44% 5% 9%
1	C	188	6% 40% 41% 9% • 10%
1	D	188	8% 40% 43% 5% 12%
1	E	188	14% 40% 41% 6% • 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	188	13% 40% 37% 7% 14%
1	G	188	4% 44% 43% 5% 8%
1	H	188	% 39% 44% 5% 10%
1	I	188	3% 43% 41% 6% 10%
1	J	188	% 41% 39% 5% 12%

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	CL	F	1180	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 13211 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 TRYPAREDOXIN PEROXIDASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	167	Total	C	N	O	S	Se	0	0	0
			1308	838	216	245	3	6			
1	B	171	Total	C	N	O	S	Se	0	0	0
			1341	858	221	252	4	6			
1	C	170	Total	C	N	O	S	Se	0	0	0
			1333	854	219	250	4	6			
1	D	166	Total	C	N	O	S	Se	0	0	0
			1304	836	215	244	3	6			
1	E	165	Total	C	N	O	S	Se	0	0	0
			1299	833	214	243	3	6			
1	F	161	Total	C	N	O	S	Se	0	0	0
			1267	815	208	235	3	6			
1	G	173	Total	C	N	O	S	Se	0	0	0
			1364	875	225	254	4	6			
1	H	169	Total	C	N	O	S	Se	0	0	0
			1326	849	218	249	4	6			
1	I	170	Total	C	N	O	S	Se	0	0	0
			1333	854	219	250	4	6			
1	J	166	Total	C	N	O	S	Se	0	0	0
			1304	836	215	244	3	6			

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Cl	0	0
			1	1		
2	B	1	Total	Cl	0	0
			1	1		
2	C	1	Total	Cl	0	0
			1	1		
2	D	1	Total	Cl	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	E	1	Total Cl 1 1	0	0
2	F	1	Total Cl 1 1	0	0
2	G	1	Total Cl 1 1	0	0
2	H	1	Total Cl 1 1	0	0
2	I	1	Total Cl 1 1	0	0
2	J	1	Total Cl 1 1	0	0

- Molecule 3 is water.

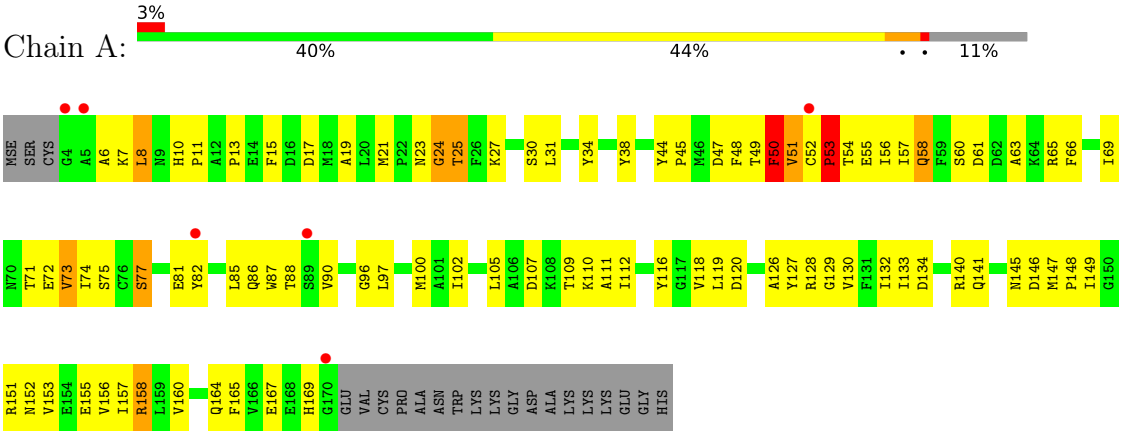
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	6	Total O 6 6	0	0
3	B	2	Total O 2 2	0	0
3	G	3	Total O 3 3	0	0
3	H	2	Total O 2 2	0	0
3	I	5	Total O 5 5	0	0
3	J	4	Total O 4 4	0	0

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.

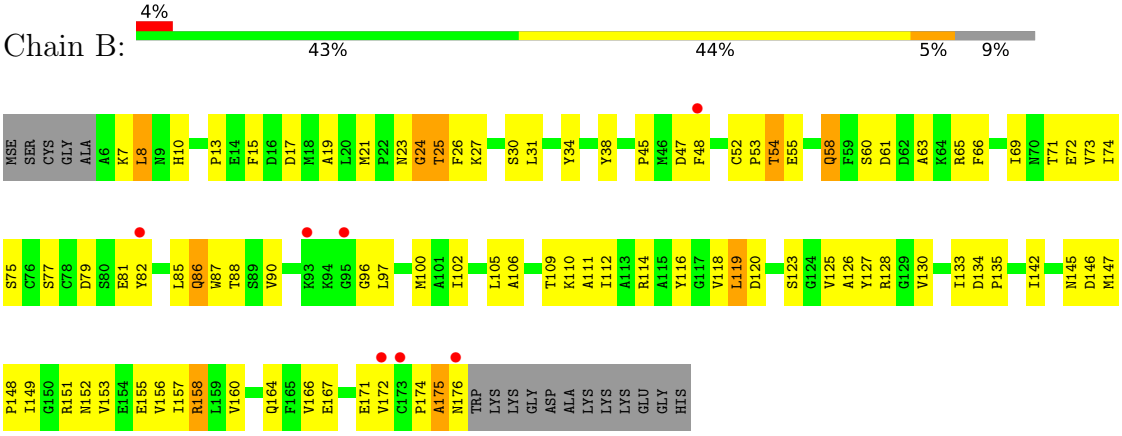
• Molecule 1: TRYPAREDOXIN PEROXIDASE

Chain A:



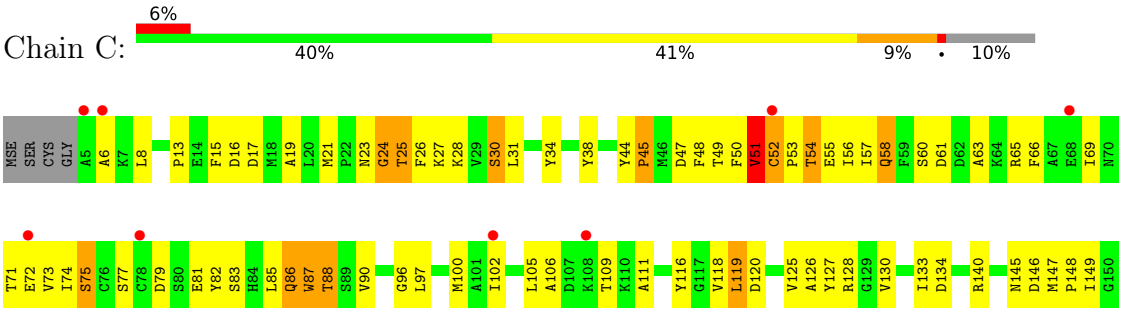
• Molecule 1: TRYPAREDOXIN PEROXIDASE

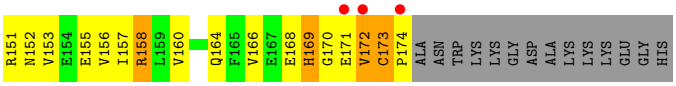
Chain B:



• Molecule 1: TRYPAREDOXIN PEROXIDASE

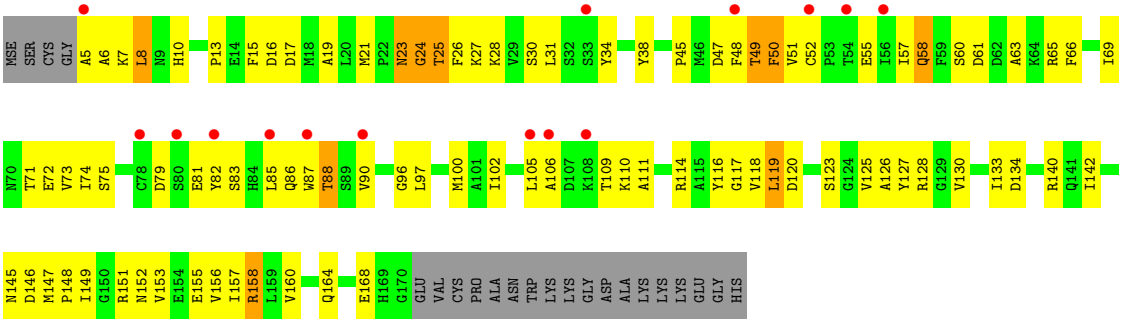
Chain C:





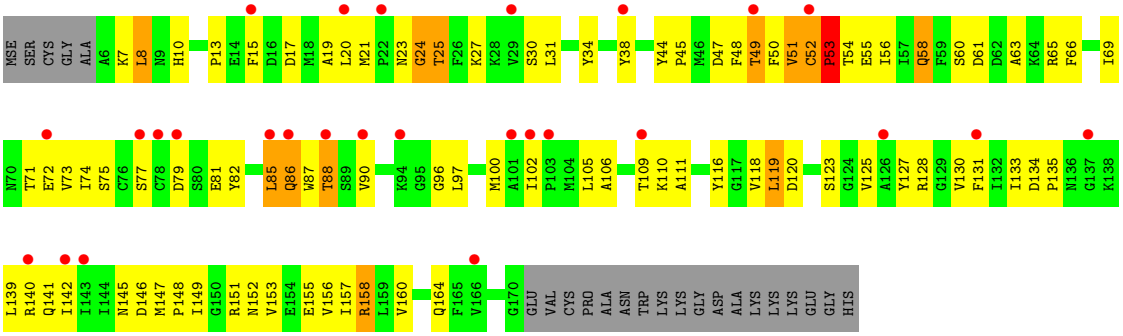
• Molecule 1: TRYPAREDOXIN PEROXIDASE

Chain D:



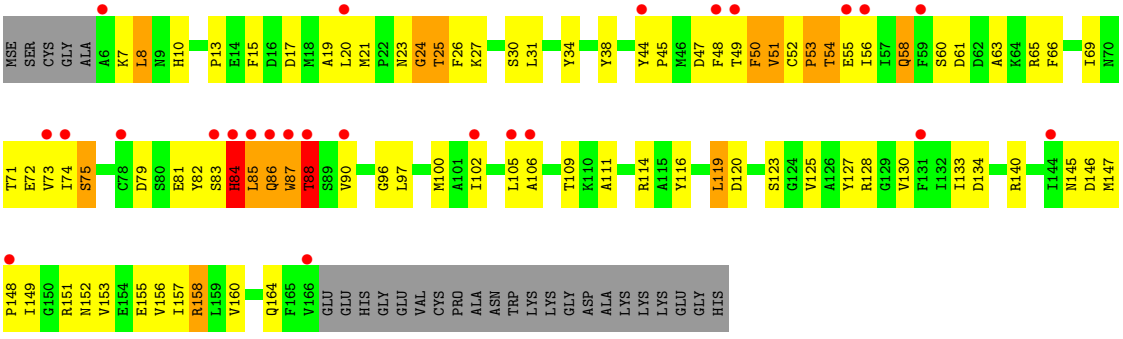
• Molecule 1: TRYPAREDOXIN PEROXIDASE

Chain E:



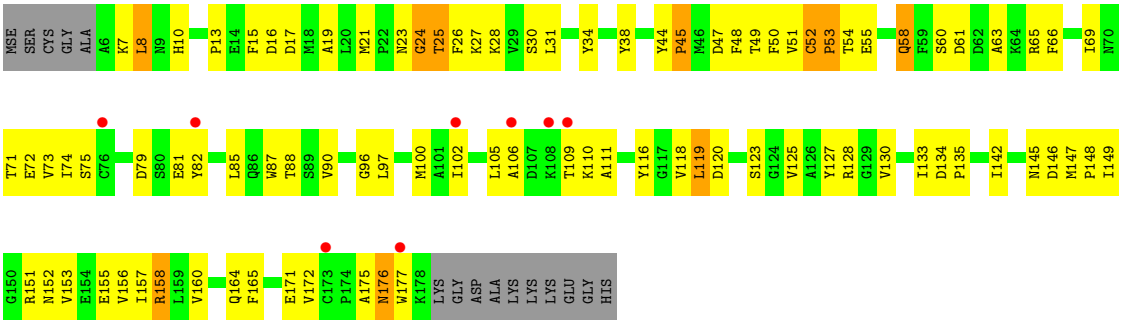
• Molecule 1: TRYPAREDOXIN PEROXIDASE

Chain F:



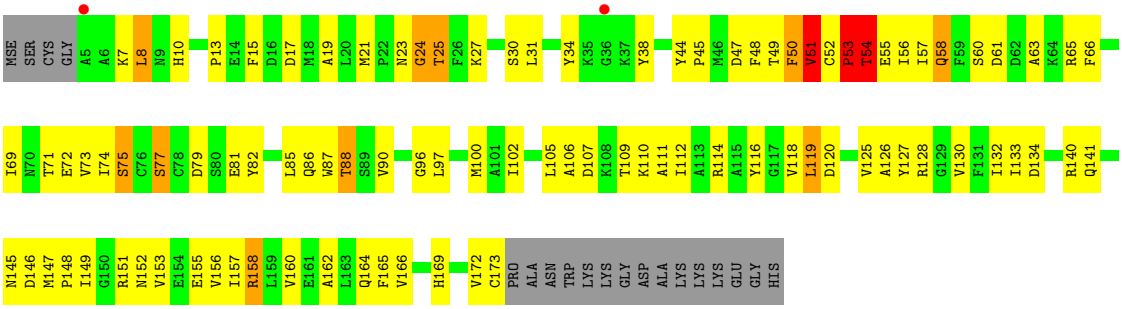
• Molecule 1: TRYPAREDOXIN PEROXIDASE

Chain G:



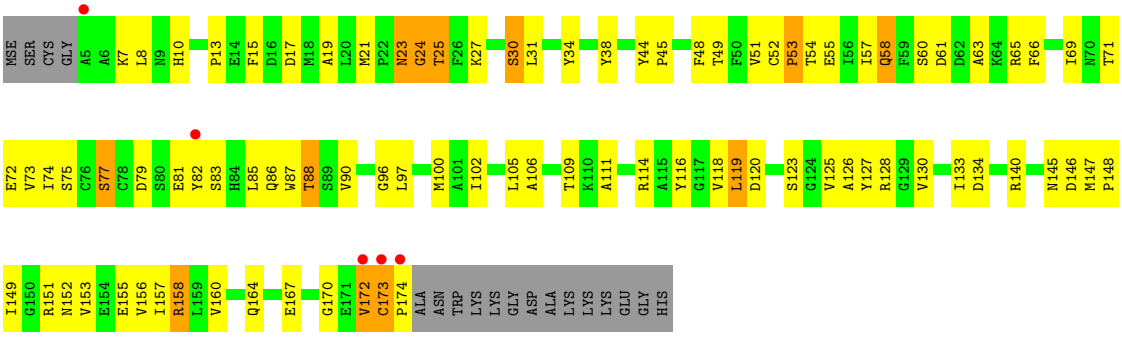
• Molecule 1: TRYPAREDOXIN PEROXIDASE

Chain H:



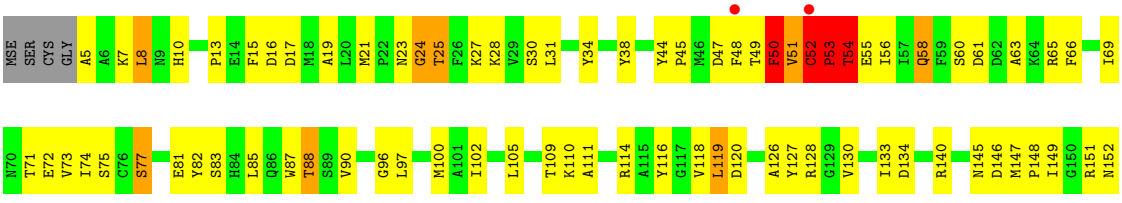
• Molecule 1: TRYPAREDOXIN PEROXIDASE

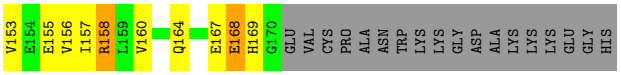
Chain I:



• Molecule 1: TRYPAREDOXIN PEROXIDASE

Chain J:




V153
E154
E155
V156
I157
R158
L159
V160
Q164
E167
E168
H169
G170
GLU
VAL
CYS
PRO
ALA
ASN
TRP
LYS
LYS
GLY
ASP
ALA
LYS
LYS
LYS
GLU
GLY
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	123.80Å 97.20Å 133.80Å 90.00° 93.10° 90.00°	Depositor
Resolution (Å)	21.00 – 3.20 21.00 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.0 (21.00-3.20) 98.9 (21.00-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 3.22Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.273 , 0.286 0.258 , 0.273	Depositor DCC
R_{free} test set	2625 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	84.5	Xtriage
Anisotropy	0.393	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 56.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	13211	wwPDB-VP
Average B, all atoms (Å^2)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% 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

Bond lengths and bond angles in the following residue types are not validated in this section: CL

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	0.57	0/1331	0.73	2/1789 (0.1%)
1	B	0.56	0/1365	0.67	0/1837
1	C	0.46	0/1357	0.67	1/1826 (0.1%)
1	D	0.44	0/1327	0.66	0/1784
1	E	0.52	1/1322 (0.1%)	0.70	3/1777 (0.2%)
1	F	0.56	1/1289 (0.1%)	0.73	1/1733 (0.1%)
1	G	0.47	0/1390	0.65	0/1871
1	H	0.57	1/1349 (0.1%)	0.72	1/1814 (0.1%)
1	I	0.55	0/1357	0.69	0/1826
1	J	0.58	0/1327	0.81	4/1784 (0.2%)
All	All	0.53	3/13414 (0.0%)	0.70	12/18041 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	53	PRO	CA-C	-8.39	1.36	1.52
1	H	53	PRO	CA-C	-7.01	1.38	1.52
1	F	88	THR	CA-CB	5.30	1.67	1.53

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	J	53	PRO	CA-N-CD	-9.38	98.36	111.50
1	J	52	CYS	C-N-CD	-8.34	102.24	120.60
1	E	53	PRO	CA-C-N	-7.49	100.71	117.20
1	E	53	PRO	CA-N-CD	-6.96	101.76	111.50
1	E	53	PRO	O-C-N	5.68	131.79	122.70

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	1308	0	1296	100	0
1	B	1341	0	1326	99	0
1	C	1333	0	1320	106	0
1	D	1304	0	1293	93	0
1	E	1299	0	1288	116	0
1	F	1267	0	1266	126	0
1	G	1364	0	1349	110	0
1	H	1326	0	1313	121	0
1	I	1333	0	1320	106	0
1	J	1304	0	1293	101	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
3	A	6	0	0	0	0
3	B	2	0	0	0	0
3	G	3	0	0	0	0
3	H	2	0	0	0	0
3	I	5	0	0	0	0
3	J	4	0	0	0	0
All	All	13211	0	13064	1032	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 39.

The worst 5 of 1032 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:F:83:SER:HA	1:G:48:PHE:HD2	1.09	1.16

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:52:CYS:H	1:H:53:PRO:HD2	1.01	1.12
1:F:83:SER:HA	1:G:48:PHE:CD2	1.85	1.11
1:H:107:ASP:HB3	1:H:112:ILE:CD1	1.82	1.10
1:J:38:TYR:HB2	1:J:71:THR:HG22	1.34	1.09

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	165/188 (88%)	139 (84%)	21 (13%)	5 (3%)	3	23
1	B	169/188 (90%)	143 (85%)	22 (13%)	4 (2%)	5	29
1	C	168/188 (89%)	136 (81%)	25 (15%)	7 (4%)	2	17
1	D	164/188 (87%)	134 (82%)	26 (16%)	4 (2%)	5	29
1	E	163/188 (87%)	135 (83%)	19 (12%)	9 (6%)	1	11
1	F	159/188 (85%)	132 (83%)	21 (13%)	6 (4%)	2	18
1	G	171/188 (91%)	142 (83%)	25 (15%)	4 (2%)	5	29
1	H	167/188 (89%)	138 (83%)	23 (14%)	6 (4%)	3	20
1	I	168/188 (89%)	143 (85%)	19 (11%)	6 (4%)	3	20
1	J	164/188 (87%)	139 (85%)	17 (10%)	8 (5%)	2	14
All	All	1658/1880 (88%)	1381 (83%)	218 (13%)	59 (4%)	3	20

5 of 59 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	51	VAL
1	C	173	CYS
1	E	49	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	53	PRO
1	F	54	THR

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	142/152 (93%)	131 (92%)	11 (8%)	10 39
1	B	147/152 (97%)	136 (92%)	11 (8%)	11 40
1	C	146/152 (96%)	132 (90%)	14 (10%)	7 28
1	D	142/152 (93%)	129 (91%)	13 (9%)	7 29
1	E	142/152 (93%)	132 (93%)	10 (7%)	12 42
1	F	139/152 (91%)	126 (91%)	13 (9%)	7 28
1	G	149/152 (98%)	139 (93%)	10 (7%)	13 44
1	H	145/152 (95%)	132 (91%)	13 (9%)	8 30
1	I	146/152 (96%)	135 (92%)	11 (8%)	11 40
1	J	142/152 (93%)	127 (89%)	15 (11%)	5 24
All	All	1440/1520 (95%)	1319 (92%)	121 (8%)	9 34

5 of 121 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	119	LEU
1	J	50	PHE
1	F	158	ARG
1	J	30	SER
1	J	119	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	152	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	J	10	HIS
1	F	164	GLN
1	I	169	HIS
1	J	152	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 10 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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

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	161/188 (85%)	-0.15	6 (3%) 45 31	10, 40, 90, 100	0
1	B	165/188 (87%)	0.03	7 (4%) 41 28	16, 56, 100, 100	0
1	C	164/188 (87%)	0.60	11 (6%) 25 17	34, 90, 100, 100	0
1	D	160/188 (85%)	0.81	15 (9%) 15 11	53, 96, 100, 100	0
1	E	159/188 (84%)	1.13	27 (16%) 5 4	55, 100, 100, 100	0
1	F	155/188 (82%)	1.10	25 (16%) 5 4	53, 100, 100, 100	0
1	G	167/188 (88%)	0.43	8 (4%) 36 25	28, 85, 100, 100	0
1	H	163/188 (86%)	0.03	2 (1%) 76 61	22, 56, 98, 100	0
1	I	164/188 (87%)	-0.11	5 (3%) 52 37	20, 52, 100, 100	0
1	J	160/188 (85%)	-0.18	2 (1%) 74 60	12, 41, 100, 100	0
All	All	1618/1880 (86%)	0.36	108 (6%) 25 17	10, 77, 100, 100	0

The worst 5 of 108 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	59	PHE	5.2
1	E	102	ILE	4.9
1	I	173	CYS	4.9
1	E	78	CYS	4.8
1	C	171	GLU	4.7

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 (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, 95^{th} 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	CL	F	1180	1/1	0.57	0.40	138,138,138,138	0
2	CL	D	1180	1/1	0.66	0.35	110,110,110,110	0
2	CL	E	1180	1/1	0.81	0.51	127,127,127,127	0
2	CL	C	1180	1/1	0.87	0.34	101,101,101,101	0
2	CL	G	1180	1/1	0.90	0.35	87,87,87,87	0
2	CL	B	1180	1/1	0.94	0.32	78,78,78,78	0
2	CL	H	1180	1/1	0.96	0.28	74,74,74,74	0
2	CL	I	1180	1/1	0.96	0.25	54,54,54,54	0
2	CL	A	1180	1/1	0.98	0.32	66,66,66,66	0
2	CL	J	1180	1/1	0.98	0.30	60,60,60,60	0

6.5 Other polymers (i)

There are no such residues in this entry.