



wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 08:13 PM GMT

PDB ID : 1K32
Title : Crystal structure of the tricorn protease
Authors : Brandstetter, H.; Kim, J.-S.; Groll, M.; Huber, R.
Deposited on : 2001-10-01
Resolution : 2.00 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

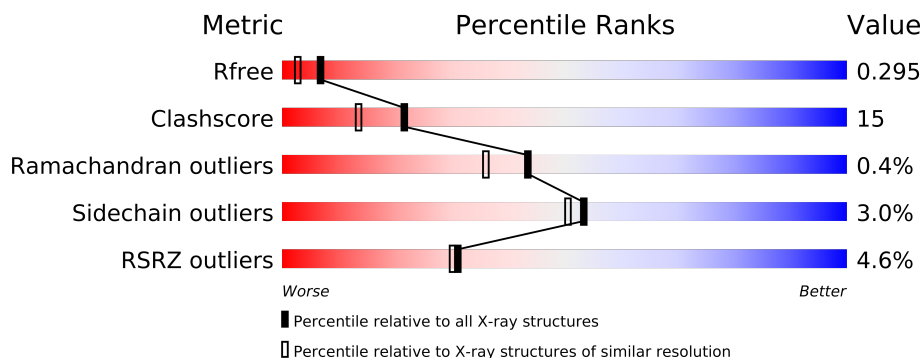
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.00 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}	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	1045	
1	B	1045	
1	C	1045	
1	D	1045	
1	E	1045	
1	F	1045	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 51456 atoms, of which 0 are hydrogen and 0 are deuterium.

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 tricorn protease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1023	Total	C	N	O	S	95	0	0
			8177	5196	1402	1551	28			
1	B	1023	Total	C	N	O	S	95	0	0
			8177	5196	1402	1551	28			
1	C	1023	Total	C	N	O	S	95	0	0
			8177	5196	1402	1551	28			
1	D	1023	Total	C	N	O	S	95	0	0
			8177	5196	1402	1551	28			
1	E	1023	Total	C	N	O	S	95	0	0
			8177	5196	1402	1551	28			
1	F	1023	Total	C	N	O	S	95	0	0
			8177	5196	1402	1551	28			

- Molecule 2 is water.

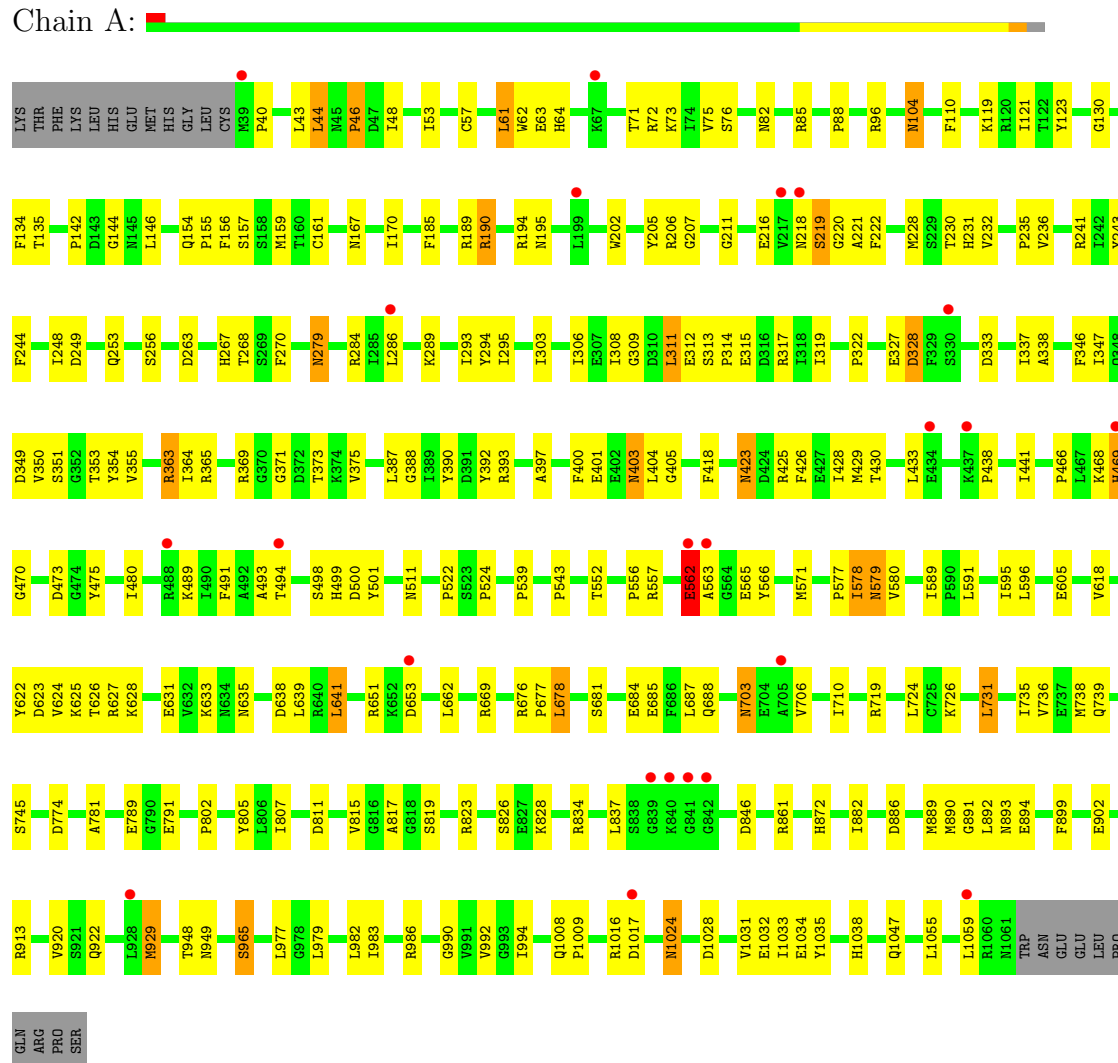
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	401	Total	O	0	0
			401	401		
2	B	395	Total	O	0	0
			395	395		
2	C	398	Total	O	0	0
			398	398		
2	D	401	Total	O	0	0
			401	401		
2	E	405	Total	O	0	0
			405	405		
2	F	394	Total	O	0	0
			394	394		

3 Residue-property plots

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.

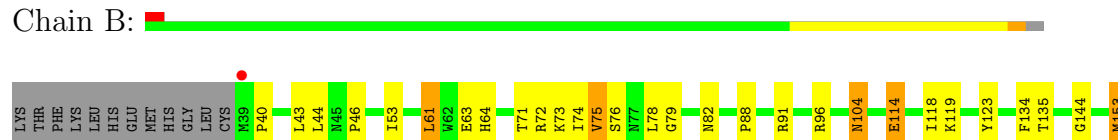
- Molecule 1: tricorn protease

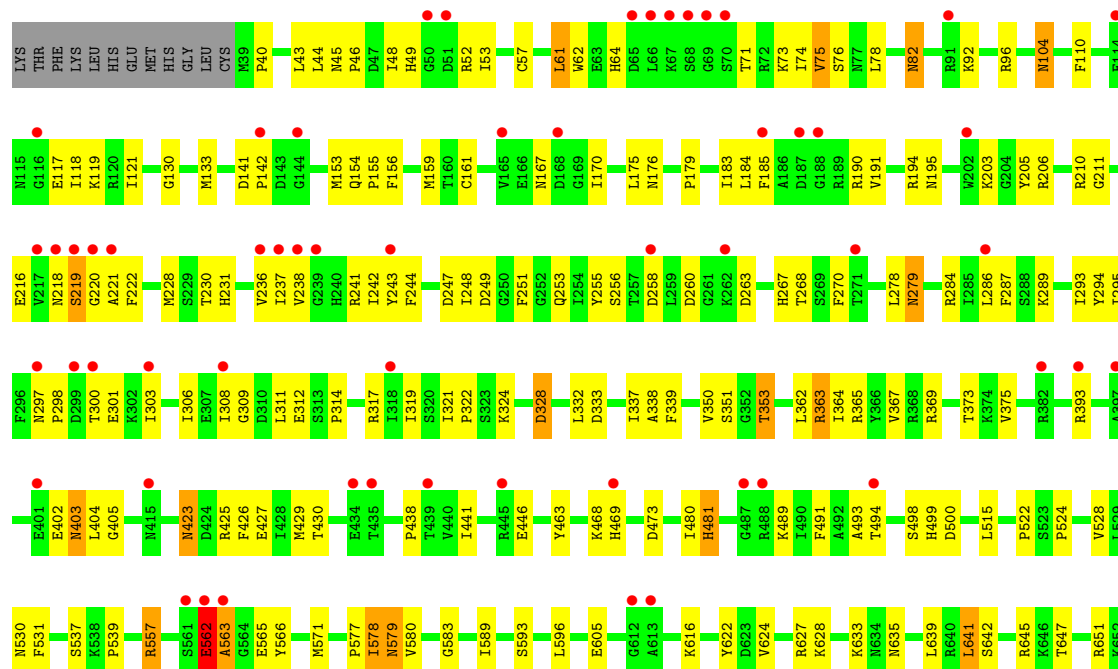
Chain A:

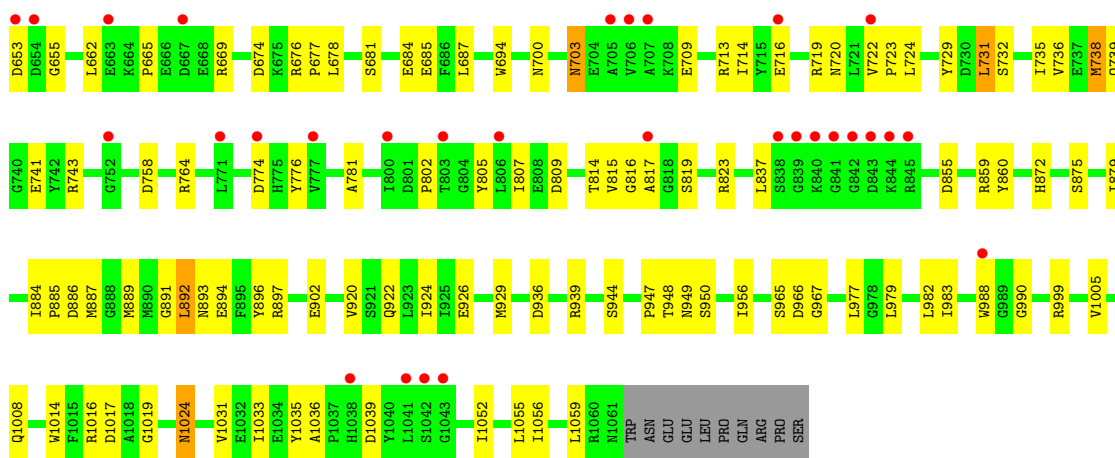


- Molecule 1: tricorn protease

Chain B:

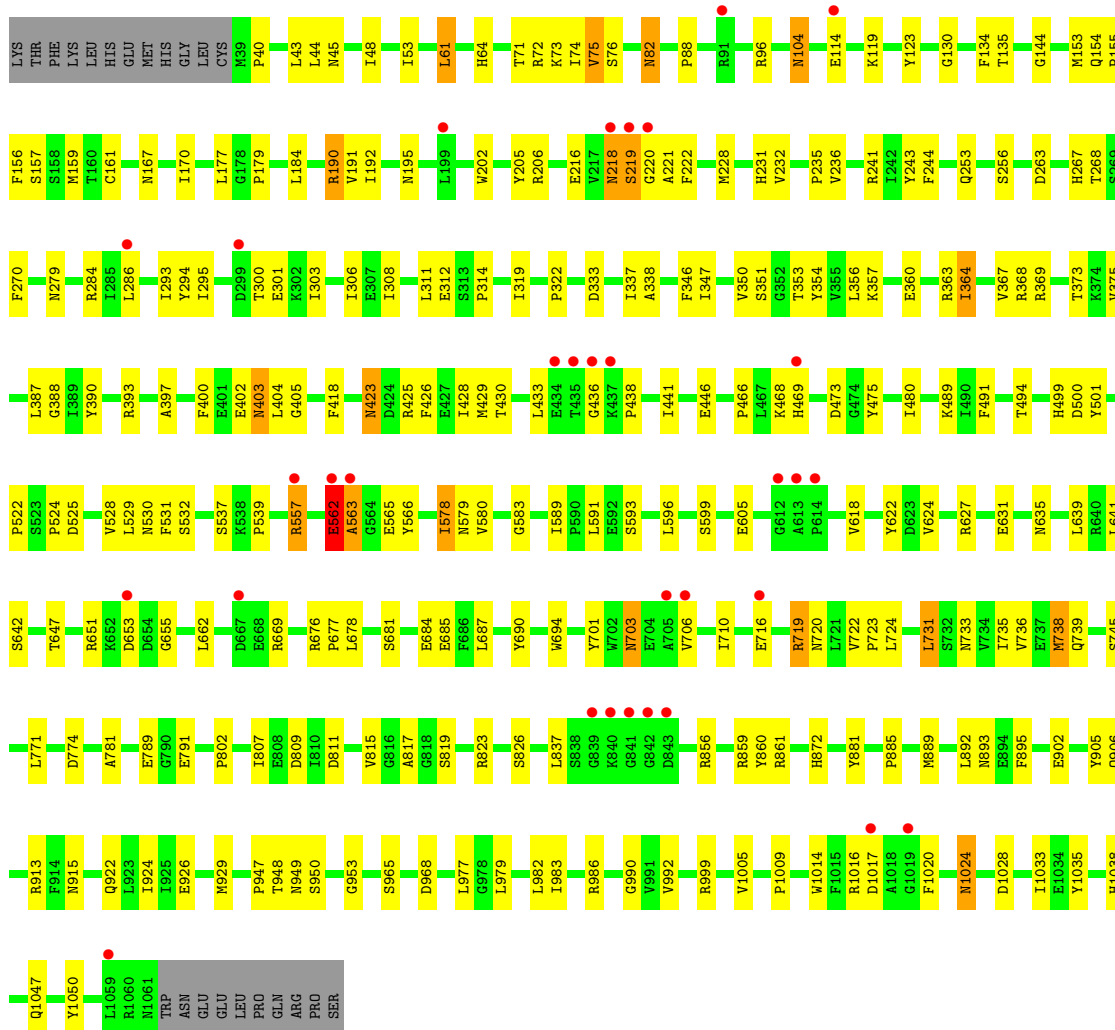






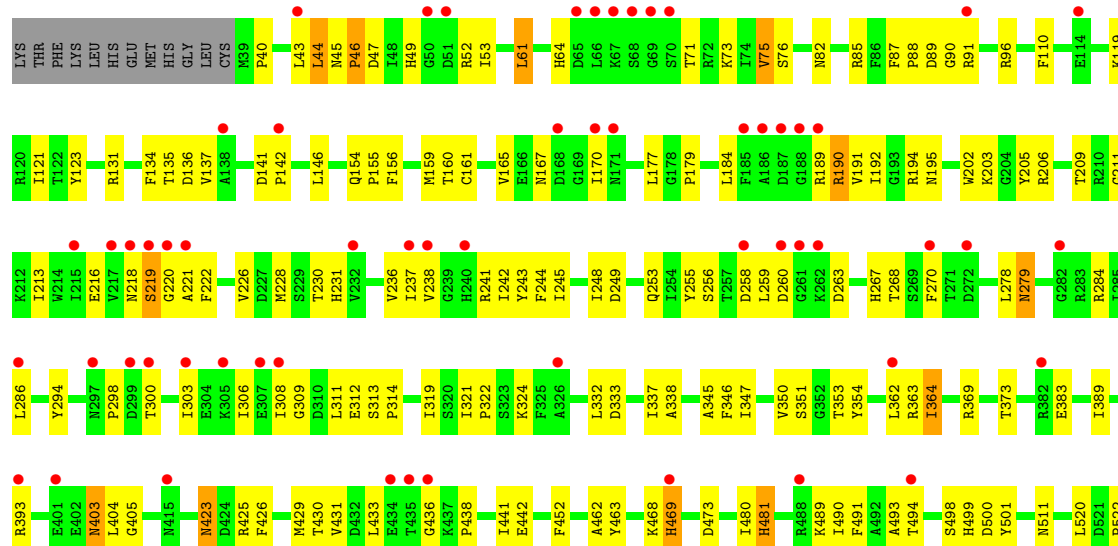
• Molecule 1: tricorn protease

Chain D:



• Molecule 1: tricorn protease

Chain E:



R1060	S523	K633	D730	G839	I966
N1061	P524	R634	L731	K840	A967
TRP		N635	S732	G841	I968
ASN	S537			G842	T959
GLU	K538	L639	I735	D843	
LEU	P539	R640	V736	K844	S965
PRO	P543	L641	E737		
GLN	P546	S642	M738	Y860	L977
ARG		T647	Q739	R861	G978
PRO	T552		G740	V864	L979
SER	R557	R651	S745	H872	L982
		K652	D756		I983
	S561	D653		I882	G984
E562	R561	G655	R761	I883	W988
A563	E562		C767	I884	G989
G564	A563	L662		P885	G990
E565	G564	D667	L771	M889	V991
Y566	E565	E668	D772	M890	V992
D567	Y566	R669	G773	G891	G993
	D567		D774	L892	I994
	M571	D674	H775	N893	
		Y776		E894	R999
	P577	P677	K780		V1005
I578	N579	L678	A781	R897	Q1008
V580	V580	S681	Y782	L898	P1009
		I682		F899	
G583		H683	E789	E902	R1016
		E684	C790	Y905	D1017
I889		L687	E791		F1020
P590	L591		I800	V912	
E592	S593	W694	D801	R913	N1024
S594	L595	N700	P802	F914	
L596			Y805	N915	V1031
E605			I807	V920	E1032
			D811	Q922	I1033
Y609				L923	E1034
			V815	I924	Y1035
G612	A613	K708	G816	E925	A1036
		E709	A817	E926	P1037
		S711	S818		H1038
K616	G617	E712	N820	M929	D1039
V618		E716			
			R823	D936	G1043
Y622			V824		K1044
D623		R719	L825	R939	D1045
V624		N720	S826	S944	F1046
		L721	F827		Q1047
K628		V722	K828	P947	I1048
		P723		T948	
E631		L724	L837	N949	I1052
V632		Y729	S838	S960	L1055
					I1056
					L1059

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	95.86Å 246.00Å 159.04Å 90.00° 105.30° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00 37.48 – 1.98	Depositor EDS
% Data completeness (in resolution range)	78.6 (20.00-2.00) 81.6 (37.48-1.98)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 1.98Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.245 , 0.264 0.258 , 0.295	Depositor DCC
R_{free} test set	20086 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	25.8	Xtriage
Anisotropy	0.345	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 42.0	EDS
Estimated twinning fraction	0.054 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	9 of 400057 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	51456	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0/8367	0.66	1/11311 (0.0%)
1	B	0.40	0/8367	0.66	2/11311 (0.0%)
1	C	0.36	0/8367	0.62	1/11311 (0.0%)
1	D	0.39	0/8367	0.65	1/11311 (0.0%)
1	E	0.38	0/8367	0.65	1/11311 (0.0%)
1	F	0.36	0/8367	0.62	1/11311 (0.0%)
All	All	0.38	0/50202	0.64	7/67866 (0.0%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	364	ILE	N-CA-C	-6.20	94.25	111.00
1	A	364	ILE	N-CA-C	-5.63	95.81	111.00
1	E	364	ILE	N-CA-C	-5.41	96.39	111.00
1	B	79	GLY	N-CA-C	-5.28	99.89	113.10
1	D	364	ILE	N-CA-C	-5.23	96.87	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	8177	0	8003	256	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	8177	0	8003	265	0
1	C	8177	0	8003	272	0
1	D	8177	0	8003	226	0
1	E	8177	0	8003	249	0
1	F	8177	0	8003	275	0
2	A	401	0	0	12	0
2	B	395	0	0	8	0
2	C	398	0	0	18	0
2	D	401	0	0	9	0
2	E	405	0	0	10	0
2	F	394	0	0	10	0
All	All	51456	0	48018	1453	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 15.

The worst 5 of 1453 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:155:PRO:HG2	1:F:159:MET:HE1	1.33	1.09
1:B:155:PRO:HG2	1:B:159:MET:HE1	1.36	1.07
1:C:155:PRO:HG2	1:C:159:MET:HE1	1.37	1.06
1:D:155:PRO:HG2	1:D:159:MET:HE1	1.35	1.03
1:B:983:ILE:HG23	1:B:1033:ILE:HD13	1.41	1.02

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	1021/1045 (98%)	979 (96%)	38 (4%)	4 (0%)	43	36
1	B	1021/1045 (98%)	977 (96%)	41 (4%)	3 (0%)	50	44
1	C	1021/1045 (98%)	961 (94%)	53 (5%)	7 (1%)	30	20
1	D	1021/1045 (98%)	979 (96%)	38 (4%)	4 (0%)	43	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	1021/1045 (98%)	975 (96%)	43 (4%)	3 (0%)	50	44
1	F	1021/1045 (98%)	971 (95%)	44 (4%)	6 (1%)	33	24
All	All	6126/6270 (98%)	5842 (95%)	257 (4%)	27 (0%)	43	36

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	328	ASP
1	B	328	ASP
1	C	219	SER
1	C	579	ASN
1	D	219	SER

5.3.2 Protein sidechains ⓘ

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	883/904 (98%)	857 (97%)	26 (3%)	55	52
1	B	883/904 (98%)	853 (97%)	30 (3%)	49	45
1	C	883/904 (98%)	860 (97%)	23 (3%)	59	58
1	D	883/904 (98%)	856 (97%)	27 (3%)	52	49
1	E	883/904 (98%)	857 (97%)	26 (3%)	55	52
1	F	883/904 (98%)	857 (97%)	26 (3%)	55	52
All	All	5298/5424 (98%)	5140 (97%)	158 (3%)	53	50

5 of 158 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	687	LEU
1	D	363	ARG
1	F	578	ILE
1	C	719	ARG
1	D	75	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 193 such

sidechains are listed below:

Mol	Chain	Res	Type
1	C	949	ASN
1	D	569	ASN
1	F	635	ASN
1	C	1024	ASN
1	D	267	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

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, 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	1022/1045 (97%)	0.18	23 (2%)	57 57	14, 26, 43, 56	19 (1%)
1	B	1022/1045 (97%)	0.18	21 (2%)	60 61	16, 26, 43, 55	19 (1%)
1	C	1022/1045 (97%)	0.62	87 (8%)	11 10	20, 33, 46, 56	19 (1%)
1	D	1022/1045 (97%)	0.22	32 (3%)	47 46	18, 28, 45, 54	19 (1%)
1	E	1022/1045 (97%)	0.20	26 (2%)	54 54	17, 28, 45, 55	19 (1%)
1	F	1022/1045 (97%)	0.58	91 (8%)	10 9	21, 32, 46, 55	19 (1%)
All	All	6132/6270 (97%)	0.33	280 (4%)	31 30	14, 30, 45, 56	114 (1%)

The worst 5 of 280 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	841	GLY	9.3
1	C	219	SER	8.6
1	F	219	SER	8.2
1	F	220	GLY	7.9
1	E	219	SER	7.1

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 ⓘ

There are no such residues in this entry.