



wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 07:38 pm BST

PDB ID : 1HQM
Title : CRYSTAL STRUCTURE OF THERMUS AQUATICUS CORE RNA POLYMERASE-INCLUDES COMPLETE STRUCTURE WITH SIDE-CHAINS (EXCEPT FOR DISORDERED REGIONS)-FURTHER REFINED FROM ORIGINAL DEPOSITION-CONTAINS ADDITIONAL SEQUENCE INFORMATION
Authors : Minakhin, L.; Bhagat, S.; Brunning, A.; Campbell, E.A.; Darst, S.A.; Ebright, R.H.; Severinov, K.
Deposited on : 2000-12-18
Resolution : 3.30 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

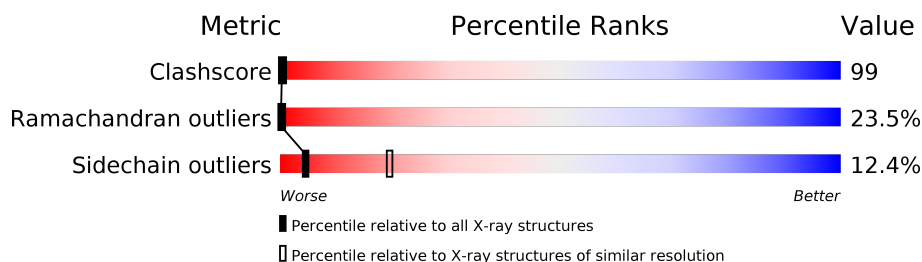
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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))
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)

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 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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	313	
1	B	313	
2	C	1119	
3	D	1265	
4	E	99	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 21254 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-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	S	0	0	0
			1750	1118	302	328	2			
1	B	229	Total	C	N	O	S	0	0	0
			1776	1135	305	334	2			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LYS	deletion	UNP Q9KWU8
A	93	ARG	MET	conflict	UNP Q9KWU8
A	94	TRP	ALA	conflict	UNP Q9KWU8
A	95	ARG	SER	conflict	UNP Q9KWU8
A	111	VAL	GLY	conflict	UNP Q9KWU8
B	?	-	LYS	deletion	UNP Q9KWU8
B	93	ARG	MET	conflict	UNP Q9KWU8
B	94	TRP	ALA	conflict	UNP Q9KWU8
B	95	ARG	SER	conflict	UNP Q9KWU8
B	111	VAL	GLY	conflict	UNP Q9KWU8

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1113	Total	C	N	O	S	12	0	0
			8508	5386	1514	1585	23			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	2	LYS	GLU	conflict	UNP Q9KWU7

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1175	Total	C	N	O	S	17	0	0
			8499	5328	1549	1595	27			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	119	PHE	SER	conflict	UNP Q9KWU6
D	863	THR	VAL	conflict	UNP Q9KWU6
D	866	THR	VAL	conflict	UNP Q9KWU6
D	876	ASN	SER	conflict	UNP Q9KWU6
D	947	ILE	-	insertion	UNP Q9KWU6
D	1010	ASN	LYS	conflict	UNP Q9KWU6
D	1117	LYS	ASN	conflict	UNP Q9KWU6
D	1389	PRO	ARG	conflict	UNP Q9KWU6

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	98	Total	C	N	O	S	0	0	0
			719	453	132	130	4			

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	1	Total	Mg	0	0
			1	1		

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

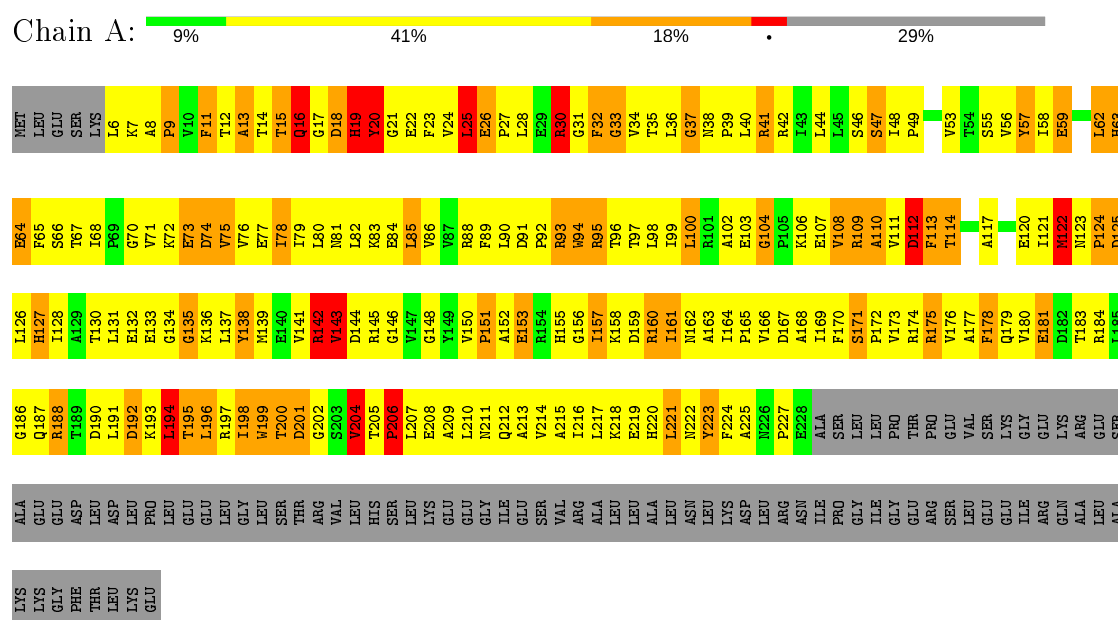
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	1	Total	Zn	0	0
			1	1		

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 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.


Note EDS was not executed.

- Molecule 1: DNA-directed RNA polymerase subunit alpha

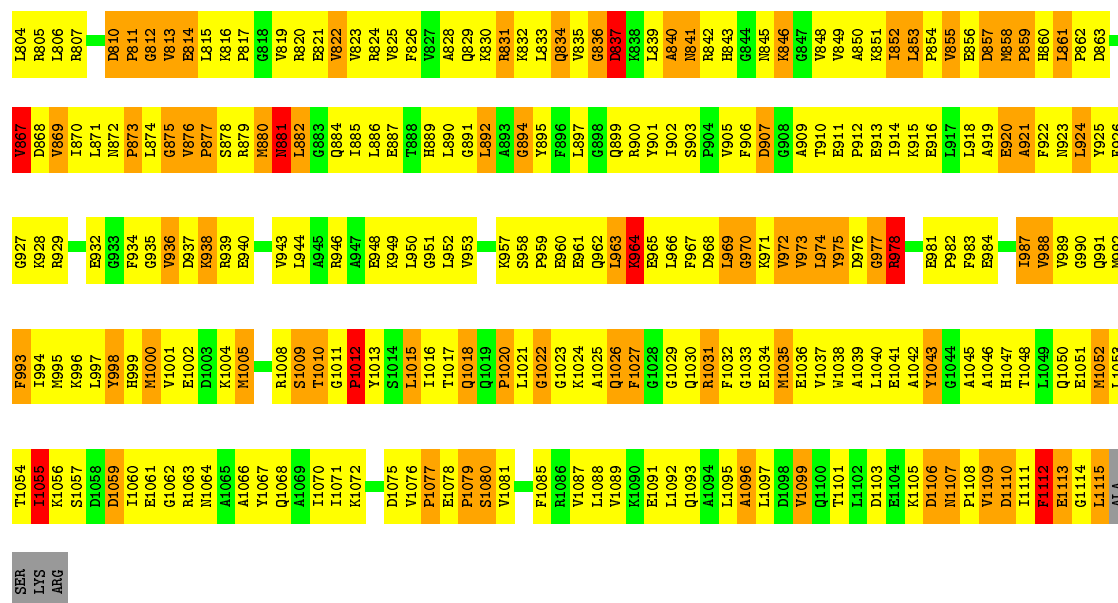


ASP	LEU	ASP	LEU	PRO	LEU	GLU	LEU	GLY	SER	THR	ARG	VAL	LEU	HIS	SER	LEU	LYS	GLU	GLY	ILE	GLU	SER	VAL	ANG	ALA	LEU	ALA	LEU	ASN	ASN	ILE	PRO	GLY	ILE	GLY	GLY	LEU	ANG	GLN	ALA	LEU	ALA	LYS	LYS	GLY
PHE	THR	LEU	LYS	GLU																																									

● Molecule 2: DNA-directed RNA polymerase subunit beta

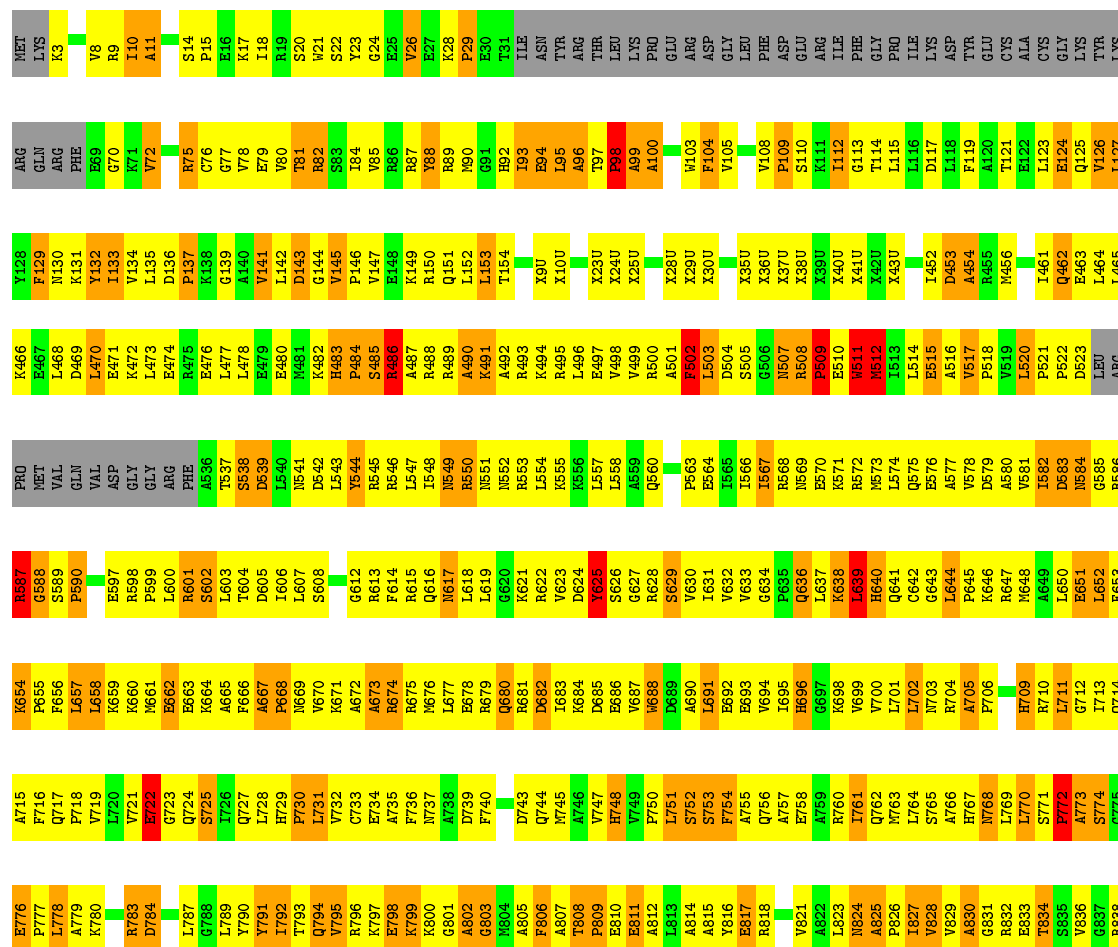
Chain C:  15% 56% 26% ..

I742	Y682	G618	N556	I496	R432	K371	I310	P247	V186	T122	G62
V743	Y682	R619	R567	A497	T433	L372	F311	K248	N187	E123	G63
R744	N683	L620	A558	A498	H434	V373	A312	K249	K188	D124	L64
G745	F684	G611	I560	A499	Y435	N374	L313	K250	R189	V65	
I746	B685	G112	N560	R500	G436	S375	T314	K251	K190	F127	L66
A747	D686	H623	G561	T601	R437	A376	A315	K252	F191	I128	D67
E748	A687	P624	S562	P602	I438	P377	G316	A253	P192	I129	F68
V749	I688	R627	N562	L503	C439	L378	P317	L254	P193	N130	L69
K750	Y689	A654	N564	E504	P440	E379	P318	A255	V194	E70	R10
P751	I690	Y628	G565	G505	V441	A380	G319	Y256	L195	Y71	E11
G752	S691	A629	T566	D606	E442	A381	H320	L257	L196	R72	V12
V753	S692	R630	G567	R507	T443	L382	E321	P258	R197	I134	R72
E754	B693	S631	I508	I508	P444	R383	V322	G259	R198	V135	I13
L755	L694	N632	A509	A509	F384	D323	D323	L260	V199	G74	P14
V756	L695	G633	P570	T510	G446	F385	D324	L261	L190	V137	L15
G757	K696	G634	L571	D611	A447	F386	D325	A262	G201	N138	P16
R758	E697	R635	I572	R512	N448	S387	D326	D263	Y202	Q139	P17
V759	D698	A636	R573	V513	I449	R388	H327	P264	R78	I140	L18
S760	F700	P637	A574	V514		S389	L328	K265	D203	H141	T19
Y761	I700	Q645	Q575	A515	T453	Q390	G329	R266	E205	S143	E20
K762	T701	R640	A576	R516	L454	L391	N330	Y267	T206	R142	D81
G763	S702	P641	P577	R517	L455	S392	R331	D268	L207	G145	Q22
E764	I703	R642	V578	R518	A456	Q393	R332	L269	V208	V146	V23
Q765	H704	V643	V579	G519	A457	F394	L333	G270	R209	I147	E24
V766	I705	R644	N580	E520	Y458	K395	R334	E271	E210	Y147	S25
P767	E706	V645	T581	P521	A459	D396	T335	A272	L211	F148	R36
S768	R707	G646	G582	V522	R460	E397	V336	G273	S212	T149	D87
P769	Y708	Q647	L583	I523	V461	T398	G337	R274	A213	D151	K27
E770	E709	R648	E584	E524	D462	N399	E338	Y275	Y214	P152	A29
I771	I710	V649	E585	A525	A463	P400	L339		G215	Q90	L30
R772	E711	R650	R586	P526	L464	L401	N340	E278	D216	A153	Q31
L773	A712	K651	V587	E527	G465	S402	A341	R408	R154	R92	A32
R774	R713	G652	V588	E528	F466	S403	D342	L281	P155	G166	D33
V775	D714	D653	V529	V529	I467	L404	Q343	G282	V218	L162	E40
T776	T715	L654	E530	E530	R468	R405	F344	V283	Q219	I163	N41
S776	T716	L655	P531	P531	T469	H406	R345	G284	G220	Y158	P36
I777	L717	A656	N593	M532	P470	K407	V346	L285	L221	I159	E37
	G718	D657	L595	D533	Y471	R408	G347	S286	D222	A160	R38
K781	F719	G658	Y596	V534	R472	R409	L348	E224	E224	S161	Q39
A782	E720	P659	A597	S535	A273	I410	A349	T289	A225	E224	L101
R783	R721	A660	E598	P536	R474	S411	R350	L290	V226	P164	H02
D784	I722	S661	E599	K337	K475	A412	L351	V291	L227	L165	Q43
V785	T723	B662	D601	Q538		L413	A352	R292	A228	P166	T44
K786	R724	E663	G601	V539	V479	L413	R353	F293	M229	K167	Q45
D787	T725	G664	E602	F540	T480	P415	G354	E294	R230	R168	A46
T788	I726	F665	V603	S541	E491	R356	V355	D295	P231	G169	L107
S789	P727	L666	V604	L542	R604		R356	G296	R232	E106	A47
V790	H728	A667	K605	I543	V483	T419	E357	E297	E233	P170	F48
R791	L729	L668	V606	T544	V484	R420	R358	F298	E233	M171	K109
V792	S730		D607	N545	Y485	E421	K359	K299	A234	I172	E50
F793	E731	N671	G608	L546	M486	R422	V360	D300	M235	D173	T51
T794	A732	V672	T609	I547	T487	A423	N361	E301	V236	L174	E112
G795	R733	L673	R610	P648	A488	G424	G362	V302	R237	E175	P53
E796	L734	V674	I611	F549	S489	R425	S363	F303	L238	F114	L54
V797	R735	A675	A612	L550	E490	D426				E177	E55
	T736		V613	E551	E491	V427	T366	P305	L241		E56
Y800	D737	L676	R614	E552	D492	R428	L367	T306	R243	V181	G57
V801	L738	P677	G615	D553	R493	D429	T368	L307	P244	T182	L118
E802	E739	R679	E616	E554	Y494	V430	P369	R308	G245	T183	P119
		R680	E617	E555	T495	R431	A370	R208	D246	M184	L120



- Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain D: 17% 54% 20% 7%



- Molecule 4: DNA-directed RNA polymerase subunit omega

[illegible]

4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	200.76Å 200.76Å 292.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 3.30	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-3.30)	Depositor
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 0.9	Depositor
R, R_{free}	0.300 , 0.360	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	21254	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	0.42	0/1786	0.77	0/2434
1	B	0.39	0/1812	0.74	0/2471
2	C	0.42	0/8672	0.78	5/11752 (0.0%)
3	D	0.42	0/8437	0.78	14/11443 (0.1%)
4	E	0.35	0/730	0.65	0/991
All	All	0.41	0/21437	0.77	19/29091 (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	1
2	C	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	834	THR	N-CA-C	-6.40	93.72	111.00
3	D	137	PRO	N-CA-CB	6.22	110.77	103.30
2	C	580	MET	N-CA-C	6.19	127.70	111.00
2	C	836	GLY	N-CA-C	-6.04	97.99	113.10
2	C	329	GLY	N-CA-C	-5.97	98.17	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	138	TYR	Sidechain
2	C	975	TYR	Sidechain

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	1750	0	1759	403	0
1	B	1776	0	1776	323	0
2	C	8508	0	8418	1886	0
3	D	8499	0	7993	1651	0
4	E	719	0	685	125	0
5	D	1	0	0	0	0
6	D	1	0	0	0	0
All	All	21254	0	20631	4161	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 99.

The worst 5 of 4161 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)
3:D:1020:PRO:HB2	3:D:1023:VAL:HB	1.20	1.18
2:C:508:ILE:H	2:C:508:ILE:HD13	1.10	1.15
2:C:438:ILE:HG21	2:C:470:PRO:HB3	1.22	1.15
2:C:605:LYS:HG2	2:C:606:VAL:H	1.05	1.14
2:C:262:ALA:HB1	2:C:266:ARG:HD2	1.23	1.14

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	221/313 (71%)	98 (44%)	68 (31%)	55 (25%)	0	0
1	B	227/313 (72%)	109 (48%)	61 (27%)	57 (25%)	0	0
2	C	1111/1119 (99%)	559 (50%)	300 (27%)	252 (23%)	0	0
3	D	1127/1265 (89%)	543 (48%)	319 (28%)	265 (24%)	0	0
4	E	96/99 (97%)	49 (51%)	22 (23%)	25 (26%)	0	0
All	All	2782/3109 (90%)	1358 (49%)	770 (28%)	654 (24%)	0	0

5 of 654 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	26	GLU
1	A	59	GLU
1	A	64	GLU
1	A	73	GLU
1	A	75	VAL

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	190/271 (70%)	163 (86%)	27 (14%)	3	15
1	B	191/271 (70%)	171 (90%)	20 (10%)	7	25
2	C	869/936 (93%)	747 (86%)	122 (14%)	3	16
3	D	782/1036 (76%)	693 (89%)	89 (11%)	5	22
4	E	67/88 (76%)	64 (96%)	3 (4%)	27	58
All	All	2099/2602 (81%)	1838 (88%)	261 (12%)	4	19

5 of 261 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	672	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	873	PRO
3	D	1197	THR
2	C	680	ASP
2	C	758	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 68 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	845	ASN
2	C	1026	GLN
3	D	1368	HIS
2	C	860	HIS
2	C	889	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	D	3

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	155:ASP	C	2(U):UNK	N	55.17
1	D	46(U):UNK	C	452:ILE	N	46.65
1	D	10(U):UNK	C	20(U):UNK	N	14.79

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.