



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:41 PM UTC

PDB ID : 1HQM / pdb_00001hqm
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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

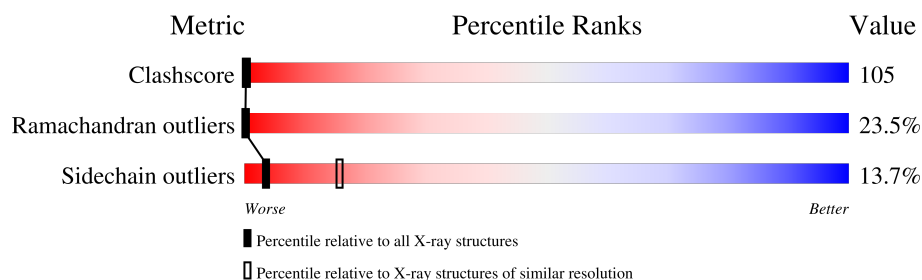
1 Overall quality at a glance

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	190562	1209 (3.32-3.28)
Ramachandran outliers	187476	1188 (3.32-3.28)
Sidechain outliers	187428	1187 (3.32-3.28)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	313	9% 38% 19% . 29%
1	B	313	12% 36% 21% . 27%
2	C	1119	13% 52% 31% . .
3	D	1265	15% 51% 24% . 7%
4	E	99	11% 65% 23% .

2 Entry composition

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 (CCD ID: MG) (formula: Mg).

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

- Molecule 6 is ZINC ION (CCD ID: ZN) (formula: Zn).

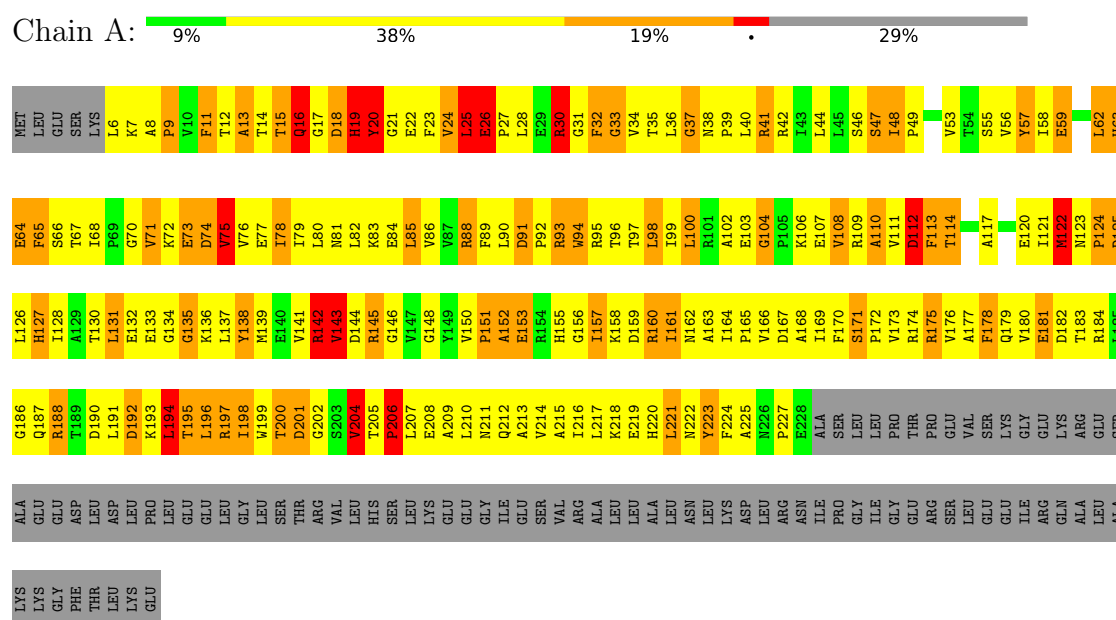
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	1	Total	Zn	0	0
			1	1		

3 Residue-property plots

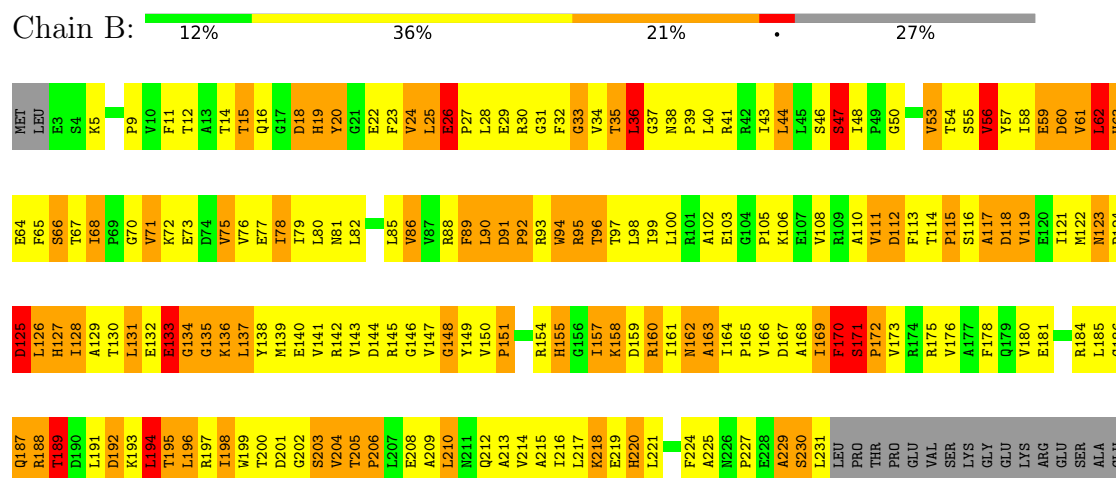
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded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

Note EDS was not executed.

- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 1: DNA-directed RNA polymerase subunit alpha



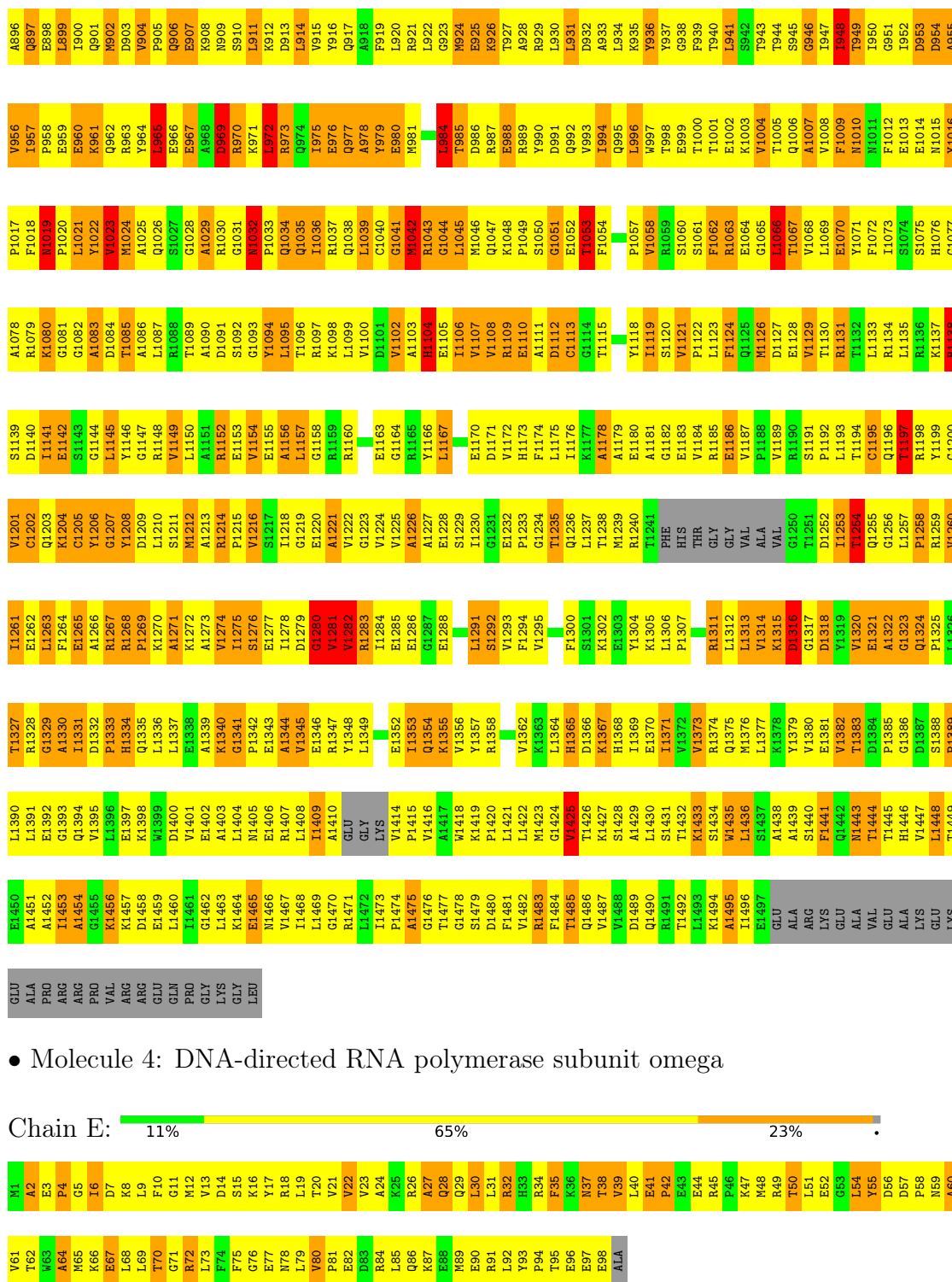
GLY	ASP	LEU	ASP	LEU	PRO	LEU	LEU	GLU	GLU	GLY	LEU	SER	THR	ARG	VAL	LEU	HIS	SER	LEU	LYS	GLY	GLU	GLU	GLY	ILE	GLU	SER	VAL	ARG	ASN	LEU	LYS	ASP	LEU	ARG	ASN	ILE	PRO	GLY	ILE	GLY	GLU	ARG	GLY	ILE	ARG	GLN	ALA	LEU	ALA	LYS	LYS
GLY	PHE	THR	LEU	LEU	LYS	GLU																																														

● Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C: 13% 52% 31%

T799	D738	P678	Y615	A555	T495	H431	A370	Y309	P248	Y186	T122	G62
V801	E739	P679	G616	M556	I496	R432	K371	L310	K249	M187	E123	G63
G802	I742	D680	D617	R557	D681	T433	L372	F311	K250	K188	D124	L64
R803	T743	G681	G618	A558	Q498	H434	R373	A312	D251	K189	F127	V65
L804	R744	G682	R619	L559	A499	Y435	R374	L313	K252	K190	F128	F66
R805	R745	M683	L620	M560	R500	Q436	S375	T314	A253	F191	I128	D67
L806	E746	E685	G1U	S562	P502	T438	P377	A315	L254	P192	I129	F68
R807	D686	D687	H623	M563	L503	C439	L378	F317	Y256	F194	G131	L69
G748	E747	E687	P624	M564	E504	P440	E379	P318	L257	L195		E70
	L688	L689	G627	Q565	G505	V441	A380	G319	F258	L196	R134	Y71
R811	R750	V689	R627	T566	D506	E442	A381	H320	K259	L197	V135	V12
G812	P751	L690	Y628	Q567	R507	T443	L382	E321	L260	R198	I136	I13
R813	G752	S691	A629	A568	L508	P444	R383	V322	L261	F199	V137	G74
L814	D753	E692	R630	V569	A509	E445	E384	D323	A282	L200	G138	D75
L815	T754	E693	S631	P570	T510	G446	F385	D324	D263	G201	Q139	P76
K816	L755	L694	N632	L571	D511	A447	F386	I325	P264	Y202	Q139	P77
R817	V756	L695	Q633	I572	R512	M448	S387	D326	K285	D203	I140	F78
G818	E757	K696	G634	R573	V513	L449		H327	K286	Q204	H141	S79
	R758	L697	T635	A574	V514		R388	L328	E272	E210	F148	Q80
E820	T759	D698	A636	Q575	R516	T453	Q390	G330	D268	E205	S143	D81
R821	S760	F699	F637	A576	R516	S454	L391	N330	L269	T206	F144	E82
R822	P761	Y700	P577	L455	R517	L455	S392	R331	G270	V208	V146	C83
E823	G762	T702	R640	V578	R518	A456	Q393	R332	E271	R209	Y147	E24
R824	P763	L703	P641	V579	E519	A457	F394	I333	A272	E210	F148	S25
E825	E764	H704	R642	M580	E520	Y458	D396	L334	K273	L211	T149	R86
R826	Q765	T705	G643	T581	P521	A459	R396	T335	R274	D87	P150	K28
E827	E766	I706	V645	G582	V522	R460	E397	V336	Y275	D151	F150	L88
R828	P767	E706	G646	L583	I523	V461	T398	G337	K276	G215	P152	Y90
Q829	S768	R707	Q646	E584	V524	D462	N399	E338	A277	D216	A183	Q91
R830	P769	Y708	Q647	E585	A525	A463	P400	L339	E278	L217	R154	A92
R831	E770	E709	R648	R586	P526	L464	L401	M340		V218	P155	A92
L832	E771	T710	V649	V587	E527	G465	S402	A341	L281	Q219	G156	P93
R833	R772	H711	K650	V588	E528	F466	S403	D342	G282	Q220	R157	L94
L834	L773	A712	K651	R589	V529	I467	L404	Q343	Y283	L221	Y158	P35
Q835	L774	R713	G652	D590	E530	A468	R495	F344	K284	L222	F159	A96
R836	R775	D714	D653	S591	F531	T469	H406	R345	L285	D223	I159	E37
E837	S776	T715	L654	L592	M532	P470	K407	R346	E224	K230	P166	L98
R838	T777	K716	L655	A593	D533	Y471	R408	G347	G287	A225	S161	Q99
L839	R778	L717	A656	A594	V534	R472	R409	L348	R288	V226	I163	L100
E840	G779	G718	D657	L595	S535	R473	T410	A349	T289	L227	P164	H102
R841	R780	F719	G658	Y586	P536	V474	S411	R350	L290	A228	L165	V42
R842	K781	E720	P659	A597	K537	K475	A412	L351	V291	K229	G103	G43
H843	E782	R721	A660	E598	Q538		L413	A352	R292	R230	P167	I104
G844	R783	T722	G661	E599	V539	V479	G414	R353	F293	P231	K167	I105
L845	T784	T723	E662	D600	F540	T480	P415	G354	E294	E232	R168	G106
K846	V785	R724	E663	G601	S541	E481	G416	V355	D295	E233	G169	L107
G847	K786	R725	G664	B602	L542	E482		R356	G296	K234	P170	L108
R848	T787	I726	P665	V603	M543	V483	T419	E357	E297	K235	W171	K109
H849	E788	P727	L666	V604	T544	V484	R420	R358	F298	V236	I172	E110
A850	G789	H728	A667	K605	Y485	Y485	R421	M359	K299	G236	D173	D111
K851	L790	L729	L668	V606	L546	M486	R422	V360	D300	L238	L174	K212
L852	R791	S730		D607	I547	T487	A423	M361	F301		E175	V113
E853	V792	E731	M671	G608	P548	A488	G424	G382	V302	L241	V176	F114
R854	P793	A732	P672	T609	F549	S489	G425	S363	F303	L242	E177	E55
E855	R794	A733	L673	R610	L550	E490	F426		L304	R243		E56
E856	E795	L734	P674	T611	E551	E491	V427	T366	P305	P244	W181	H17
D857	E796	R735	A675	A612	H552	D492	R428	T367	G245	G245	V182	P118
M858	G797	D736	L676	V613	D553	R493	D429	T368	L307	D246	T183	P119
	E798	L727	V677	V614	D554	Y104	V420	D430	P308	D247	M184	L120
											G60	V21





- Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E: 11% 65% 23%

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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.53	0/1786	1.11	8/2434 (0.3%)
1	B	0.50	1/1812 (0.1%)	1.12	17/2471 (0.7%)
2	C	0.54	0/8672	1.20	93/11752 (0.8%)
3	D	0.54	1/8437 (0.0%)	1.20	99/11443 (0.9%)
4	E	0.46	0/730	1.03	3/991 (0.3%)
All	All	0.53	2/21437 (0.0%)	1.18	220/29091 (0.8%)

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

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	231	LEU	CB-CG	-5.56	1.42	1.53
3	D	949	THR	CA-CB	5.24	1.58	1.53

The worst 5 of 220 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	834	THR	N-CA-C	-14.67	93.72	114.12
3	D	1070	GLU	N-CA-C	-13.62	96.94	113.15
2	C	736	ASP	N-CA-C	-12.93	97.54	113.18
3	D	567	ILE	N-CA-C	-12.76	101.58	113.71

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	836	GLY	N-CA-C	-12.71	97.99	111.21

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	422	0
1	B	1776	0	1776	337	0
2	C	8508	0	8418	1971	0
3	D	8499	0	7993	1744	0
4	E	719	0	685	138	0
5	D	1	0	0	0	0
6	D	1	0	0	0	0
All	All	21254	0	20631	4381	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 105.

The worst 5 of 4381 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%)	161 (85%)	29 (15%)	3	13
1	B	191/271 (70%)	169 (88%)	22 (12%)	5	21
2	C	869/936 (93%)	740 (85%)	129 (15%)	3	13
3	D	782/1036 (76%)	678 (87%)	104 (13%)	4	17

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	E	67/88 (76%)	63 (94%)	4 (6%)	17	46
All	All	2099/2602 (81%)	1811 (86%)	288 (14%)	3	16

5 of 288 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	985	THR
4	E	32	ARG
3	D	1035	GLN
3	D	1167	LEU
2	C	434	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 71 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	768	ASN
3	D	917	GLN
3	D	1354	GLN
2	C	633	GLN
2	C	632	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 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.