



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 20, 2026 – 03:05 AM UTC

PDB ID : 1HN1 / pdb_00001hn1
Title : E. COLI (LAC Z) BETA-GALACTOSIDASE (ORTHORHOMBIC)
Authors : Juers, D.H.; Matthews, B.W.
Deposited on : 2000-12-05
Resolution : 3.00 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)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

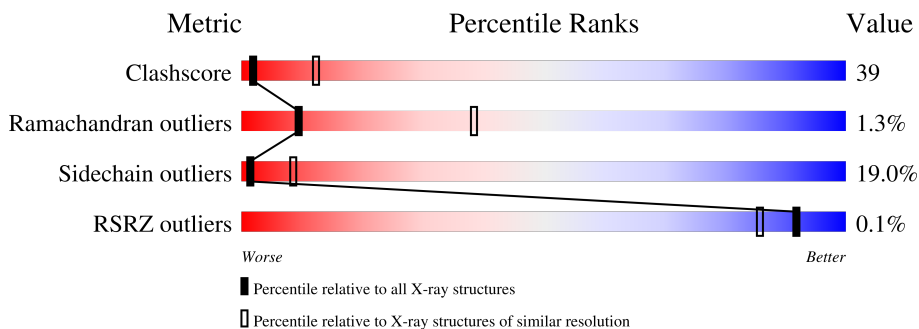
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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))
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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 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 $\leq 5\%$. 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	1023	30% 50% 17% ..
1	B	1023	29% 46% 20% ..
1	C	1023	30% 50% 16% ..
1	D	1023	33% 45% 18% ..

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 32954 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 BETA-GALACTOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1011	Total	C	N	O	S	0	2	0
			8136	5145	1443	1510	38			
1	B	1011	Total	C	N	O	S	0	2	0
			8136	5145	1443	1510	38			
1	C	1011	Total	C	N	O	S	0	2	0
			8136	5145	1443	1510	38			
1	D	1011	Total	C	N	O	S	0	1	0
			8130	5141	1441	1510	38			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP P00722
A	2	SER	-	cloning artifact	UNP P00722
A	3	HIS	-	cloning artifact	UNP P00722
A	4	MET	-	cloning artifact	UNP P00722
A	5	LEU	-	cloning artifact	UNP P00722
A	6	GLU	-	cloning artifact	UNP P00722
A	7	ASP	-	cloning artifact	UNP P00722
A	8	PRO	-	cloning artifact	UNP P00722
B	1	GLY	-	cloning artifact	UNP P00722
B	2	SER	-	cloning artifact	UNP P00722
B	3	HIS	-	cloning artifact	UNP P00722
B	4	MET	-	cloning artifact	UNP P00722
B	5	LEU	-	cloning artifact	UNP P00722
B	6	GLU	-	cloning artifact	UNP P00722
B	7	ASP	-	cloning artifact	UNP P00722
B	8	PRO	-	cloning artifact	UNP P00722
C	1	GLY	-	cloning artifact	UNP P00722
C	2	SER	-	cloning artifact	UNP P00722
C	3	HIS	-	cloning artifact	UNP P00722
C	4	MET	-	cloning artifact	UNP P00722
C	5	LEU	-	cloning artifact	UNP P00722

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	6	GLU	-	cloning artifact	UNP P00722
C	7	ASP	-	cloning artifact	UNP P00722
C	8	PRO	-	cloning artifact	UNP P00722
D	1	GLY	-	cloning artifact	UNP P00722
D	2	SER	-	cloning artifact	UNP P00722
D	3	HIS	-	cloning artifact	UNP P00722
D	4	MET	-	cloning artifact	UNP P00722
D	5	LEU	-	cloning artifact	UNP P00722
D	6	GLU	-	cloning artifact	UNP P00722
D	7	ASP	-	cloning artifact	UNP P00722
D	8	PRO	-	cloning artifact	UNP P00722

- Molecule 2 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

- Molecule 3 is SODIUM ION (CCD ID: NA) (formula: Na).

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	103	Total O 103 103	0	0

Continued on next page...

Continued from previous page...

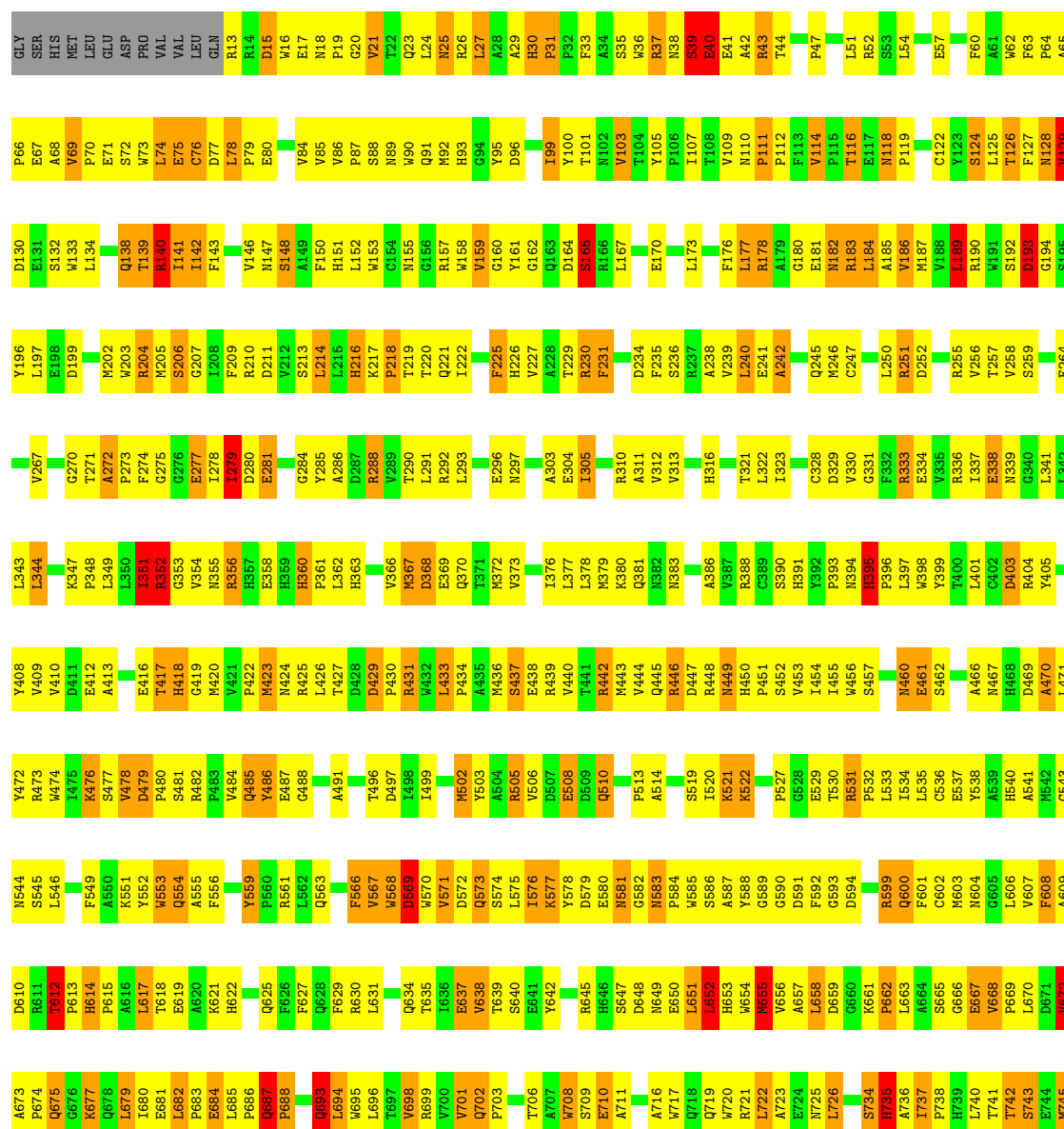
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	92	Total 92	O 92	0	0
4	C	108	Total 108	O 108	0	0
4	D	98	Total 98	O 98	0	0

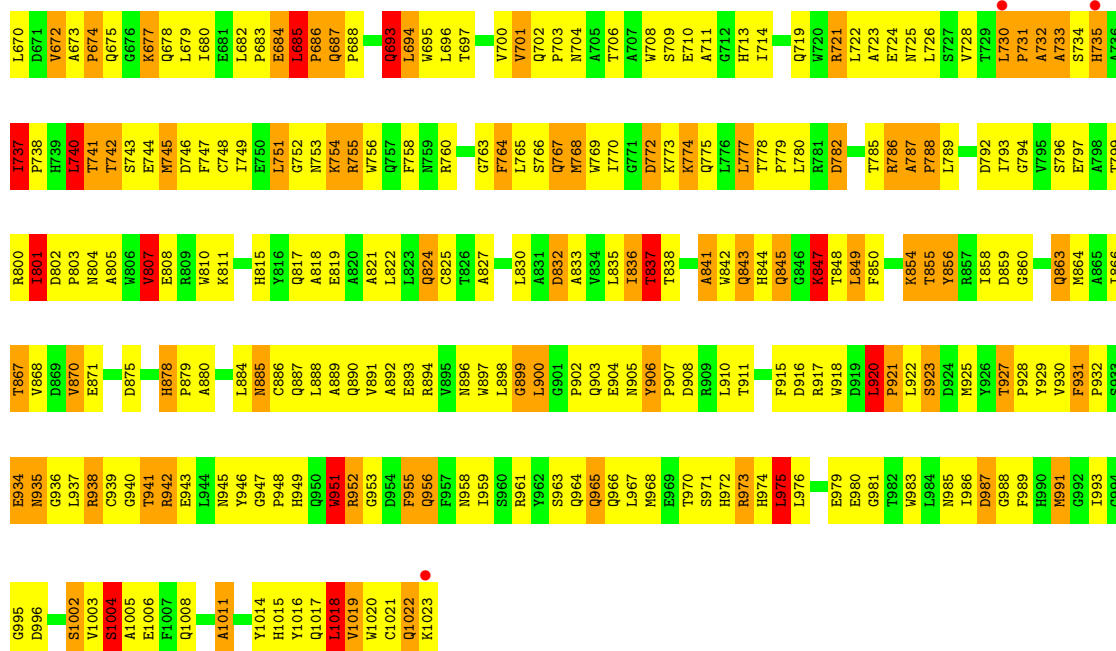
3 Residue-property plots

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 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: BETA-GALACTOSIDASE

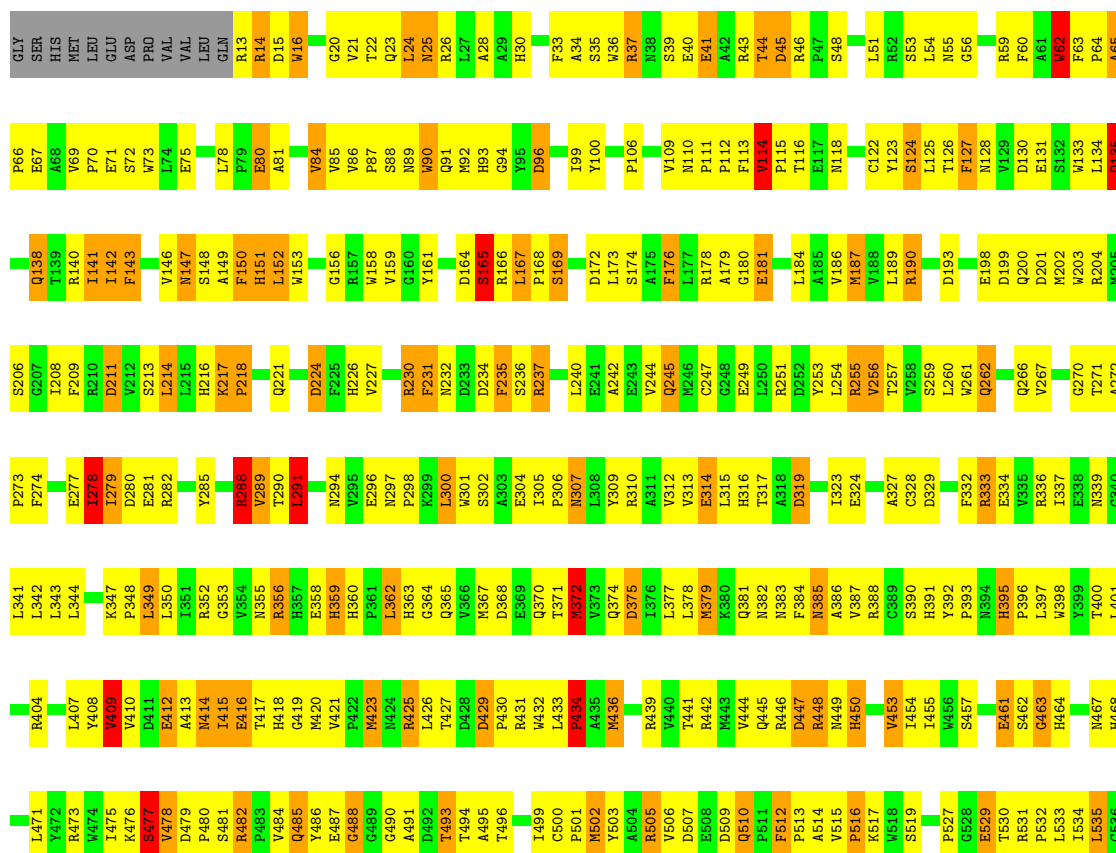
Chain A: 

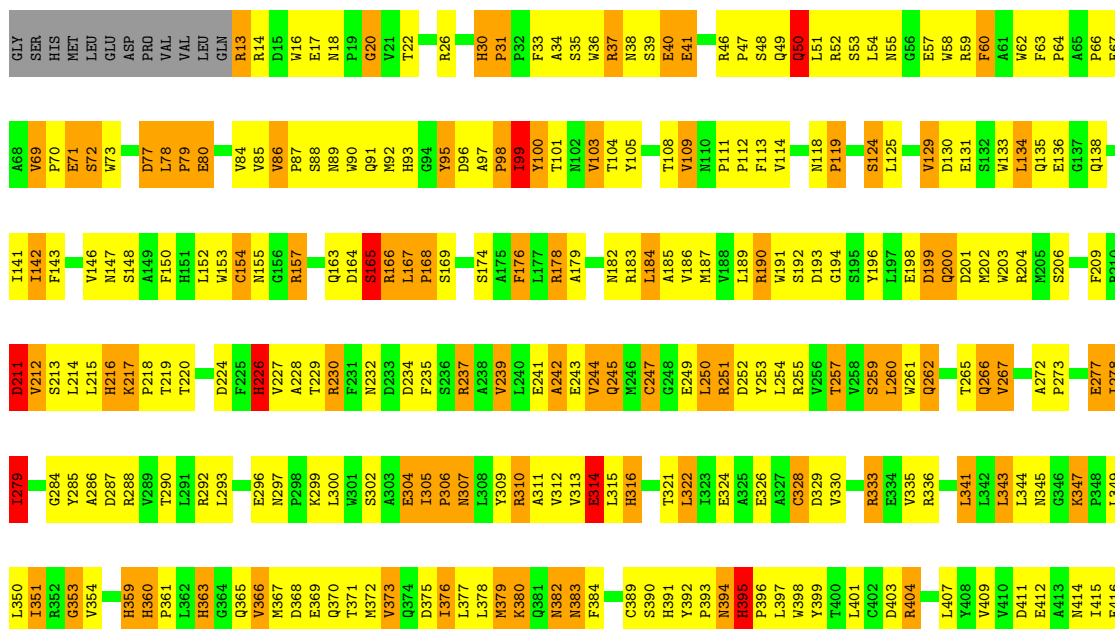




• Molecule 1: BETA-GALACTOSIDASE

Chain C: 30% 50% 16% ..





F957	A892	Q824	E750	E684	L617	Q554	R482	T417
N958	E893	C325	L751	P886	T618	A555	P483	H418
I959	R894	T326		P686	E619	F556	V484	G419
S960	N895		R755	Q687	A620	Q485	Y486	M420
R961	N896		W756		K621	Q558	Y486	V421
	W897	T829	Q757	G692	H622	Y559	E487	P422
Q964	L898	L830		Q893		P560	G488	N423
	G899	A831	N759	L694	F626	R561	G489	M424
S970	L900	D832	R760	W695	P627	L562	G490	R425
T971	G901	A833	Q761	L696				L426
R972	P902	W834	S762	T697	R630	G565	A491	T427
R973	Q903	L835		R698	L631	F566	T493	D428
H974	E904	I836	L765	R699	S632	V567	T494	D429
L975	N905	T837	S766	V700	G633	W568		P430
L976	Y906	T838	Q767	W701	Q634	D569	D497	R431
H977	P907	H839	W768	Q702	T635	W570	I498	N432
A978	D908	H840	W769	P703	L636	V571	I499	L433
E979	R909	A841	I770	P703	E637	D572	C500	P434
E980	L910	W842		T706	R638	Q573	P501	A435
G981	T911	Q843	K773	A707	T639		M502	M436
T982	A912	H844	K774	W708	E641	I576	S437	S437
W983	A913	Q845	Q775	S709	E641	K577	Y504	E438
L984	C914	G846	L776	E710	L642	V578	R505	R439
N985		K347	L777	A711	L643	D579		V440
	R917			G712	F644	E580	E508	T441
I986	W918	F850	P779	H713	H646	N581	P511	R442
G988	D919	L780	R781	I714	S647	P584	F512	N443
F989	L920	R853			D648	N583		
H990	P921	K854	W717		N649	W585		R446
N991		T855	Q718			I585	V515	D447
G992	Y926	Y856	W719	Q718	E650	S586	P516	R448
I993	T927	R857	L720	W720	L651	A587	K517	N449
G994	P928	L858	A787	R721	L652	Y588		
G995	Y929	D859	P788	L722	H653	G589		
D996	W930	G860	L789	A723	W654	G590	I520	S452
D997	F931		D790	E724	M655	D591	K521	V453
S998	P932	T866			V656	F592	K522	I454
	S933	T867	G794	W725	A657	D592	W523	I455
V1003	E934	V868	L726	N725	L726	G593	S525	W456
S1004	N935	D869	S727	W728	L658	D594	S525	S457
A1005	G936	W870	S796	T729	D659	T595		
E1006	L937	E871	E797	W729	G660	P596	P527	N460
F1007	R938	W872	A798	L730	K661	N597	Q528	E461
	C939		T799	P731	P662	D598	E529	S462
A1011	G940	D875	R800		L663	R599	T530	G463
G1012	T941	T876	I801	S734	A664	Q600	R531	H464
R1013	R342	H878	D802	H735	S665	F601	P532	
Y1014		H878	P803	A736	G666	G602	L533	N467
H1015	N945	P878	N804	I737	E667	M603	I534	
Y1016	Y946	P879	A805	P738		N604		A470
Q1017	G947	R880	W806	P738	A673	E537		L471
L1018	N948	R881	W907	H739	L606			Y472
V1019	H949	I882	E808	L740	G543			R473
W1020	Q950	G883	R809	T742	N544	G543		N474
C1021	Y951	L884	S743	T742	G676	F608		I475
	Q887		H815	S743	K677	A609		
Q1022	R952	Q887	E744	E744	Q678	D610		K476
G953	G954	Q887	Y816	W745	L679	R611		S477
K1023	N954	A889	Q817	D746	I680	T612		V478
	F955	Q890	Q817	F747	E681	P613	K551	D479
	Q956	W891	A820	C748	L682	H614	Y552	P480
					D682		I453	S481

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	153.90Å 171.40Å 204.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.00 15.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	94.0 (15.00-3.00) 88.5 (15.00-3.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 3.00Å)	Xtriage
Refinement program	TNT	Depositor
R, R_{free}	0.148 , 0.299 0.134 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	31.4	Xtriage
Anisotropy	0.253	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 172.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	32954	wwPDB-VP
Average B, all atoms (Å ²)	42.0	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 41.90 % 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 2.2078e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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, NA

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	1.32	9/8387 (0.1%)	1.92	204/11442 (1.8%)
1	B	1.29	16/8387 (0.2%)	1.92	208/11442 (1.8%)
1	C	1.29	8/8387 (0.1%)	1.91	211/11442 (1.8%)
1	D	1.30	15/8376 (0.2%)	1.95	221/11427 (1.9%)
All	All	1.30	48/33537 (0.1%)	1.92	844/45753 (1.8%)

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	B	1	0
1	D	1	0
All	All	2	0

The worst 5 of 48 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	306	PRO	N-CA	-7.15	1.41	1.47
1	B	479	ASP	C-N	-6.46	1.27	1.33
1	D	949	HIS	CA-C	-6.35	1.44	1.52
1	B	329	ASP	CA-C	-5.95	1.45	1.52
1	B	987	ASP	CA-C	-5.93	1.45	1.52

The worst 5 of 844 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	685	LEU	C-N-CD	-12.56	73.50	125.00
1	B	730	LEU	C-N-CD	-12.38	74.26	125.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	949	HIS	CA-CB-CG	-11.46	102.34	113.80
1	D	927	THR	CA-C-N	11.07	133.68	119.84
1	D	927	THR	C-N-CA	11.07	133.68	119.84

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	37	ARG	CA
1	D	951	TRP	CA

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	8136	0	7723	623	0
1	B	8136	0	7723	697	0
1	C	8136	0	7723	582	0
1	D	8130	0	7720	593	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	1	0	0	0	0
4	A	103	0	0	2	0
4	B	92	0	0	9	0
4	C	108	0	0	8	0
4	D	98	0	0	6	0
All	All	32954	0	30889	2463	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 2463 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:D:502:MET:HB2	1:D:537:GLU:HB2	1.25	1.18
1:D:734:SER:HB3	1:D:860:GLY:HA3	1.20	1.11
1:B:18:ASN:HD22	1:B:21:VAL:HG23	1.09	1.10
1:B:737:ILE:HD12	1:B:738:PRO:HD2	1.26	1.09
1:A:737:ILE:HG13	1:A:832:ASP:HA	1.35	1.08

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	1011/1023 (99%)	902 (89%)	99 (10%)	10 (1%)	12	45
1	B	1011/1023 (99%)	901 (89%)	96 (10%)	14 (1%)	9	36
1	C	1011/1023 (99%)	904 (89%)	93 (9%)	14 (1%)	9	36
1	D	1010/1023 (99%)	904 (90%)	93 (9%)	13 (1%)	9	38
All	All	4043/4092 (99%)	3611 (89%)	381 (9%)	51 (1%)	9	38

5 of 51 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	129	VAL
1	A	425	ARG
1	A	688	PRO
1	B	201	ASP
1	B	647	SER

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	866/875 (99%)	696 (80%)	170 (20%)	1	8
1	B	866/875 (99%)	684 (79%)	182 (21%)	1	6
1	C	866/875 (99%)	717 (83%)	149 (17%)	2	11
1	D	865/875 (99%)	708 (82%)	157 (18%)	2	9
All	All	3463/3500 (99%)	2805 (81%)	658 (19%)	1	8

5 of 658 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	783	GLN
1	D	508	GLU
1	C	861	SER
1	C	781	ARG
1	D	125	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 102 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	262	GLN
1	C	844	HIS
1	D	985	ASN
1	C	381	GLN
1	C	702	GLN

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 15 ligands modelled in this entry, 15 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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	1011/1023 (98%)	-0.89	0	100 100	10, 35, 92, 211	2 (0%)
1	B	1011/1023 (98%)	-0.82	3 (0%)	90 79	8, 38, 101, 212	2 (0%)
1	C	1011/1023 (98%)	-0.91	2 (0%)	91 83	14, 34, 90, 214	2 (0%)
1	D	1011/1023 (98%)	-0.91	1 (0%)	92 86	5, 34, 93, 210	1 (0%)
All	All	4044/4092 (98%)	-0.88	6 (0%)	92 86	5, 35, 94, 214	7 (0%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	734	SER	2.9
1	C	734	SER	2.5
1	B	1023	LYS	2.3
1	C	800	ARG	2.3
1	B	730	LEU	2.2

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no oligosaccharides 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, 95th 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($\text{\AA}^2$)	Q<0.9
3	NA	A	3101	1/1	0.95	0.08	52,52,52,52	0
3	NA	B	3102	1/1	0.96	0.04	32,32,32,32	0
3	NA	C	3101	1/1	0.96	0.08	49,49,49,49	0
2	MG	B	3002	1/1	0.97	0.06	30,30,30,30	0
2	MG	C	3001	1/1	0.97	0.04	25,25,25,25	0
3	NA	B	3103	1/1	0.97	0.05	45,45,45,45	0
2	MG	D	3001	1/1	0.97	0.11	26,26,26,26	0
3	NA	A	3102	1/1	0.98	0.05	24,24,24,24	0
2	MG	A	3001	1/1	0.98	0.09	29,29,29,29	0
3	NA	C	3102	1/1	0.98	0.03	24,24,24,24	0
2	MG	B	3001	1/1	0.99	0.03	13,13,13,13	0
2	MG	A	3002	1/1	0.99	0.04	29,29,29,29	0
2	MG	D	3002	1/1	0.99	0.07	30,30,30,30	0
3	NA	D	3102	1/1	0.99	0.03	47,47,47,47	0
2	MG	C	3002	1/1	1.00	0.07	25,25,25,25	0

6.5 Other polymers [i](#)

There are no such residues in this entry.