



wwPDB X-ray Structure Validation Summary Report

Mar 8, 2026 – 06:00 PM UTC

PDB ID : 1F4A / pdb_00001f4a
Title : E. COLI (LACZ) BETA-GALACTOSIDASE (NCS CONSTRAINED MONOMER-ORTHORHOMBIC)
Authors : Juers, D.H.; Jacobson, R.H.; Wigley, D.; Zhang, X.J.; Huber, R.E.; Tronrud, D.E.; Matthews, B.W.
Deposited on : 2000-06-07
Resolution : 2.80 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

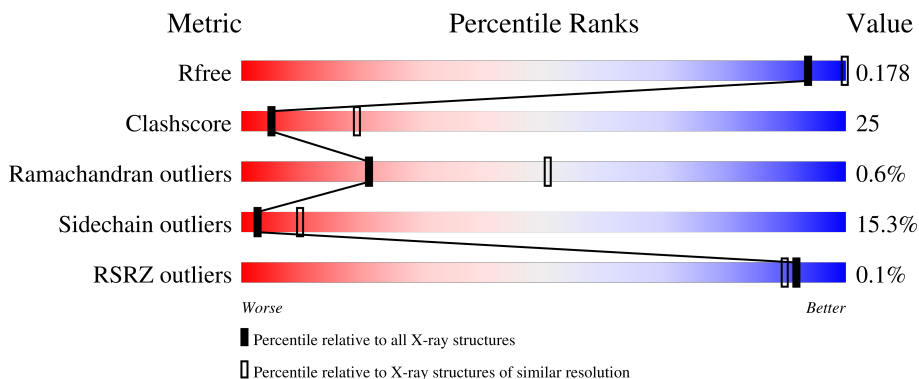
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	180053	3866 (2.80-2.80)
Clashscore	190562	4276 (2.80-2.80)
Ramachandran outliers	187476	4196 (2.80-2.80)
Sidechain outliers	187428	4198 (2.80-2.80)
RSRZ outliers	180081	3869 (2.80-2.80)

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	1021	 50% 36% 11% .
1	B	1021	 51% 36% 11% .
1	C	1021	 50% 37% 11% .
1	D	1021	 50% 36% 11% .

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 34424 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	1021	Total	C	N	O	S	84	7	0
			8238	5209	1466	1525	38			
1	B	1021	Total	C	N	O	S	84	7	0
			8238	5209	1466	1525	38			
1	C	1021	Total	C	N	O	S	84	7	0
			8238	5209	1466	1525	38			
1	D	1021	Total	C	N	O	S	84	7	0
			8238	5209	1466	1525	38			

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

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

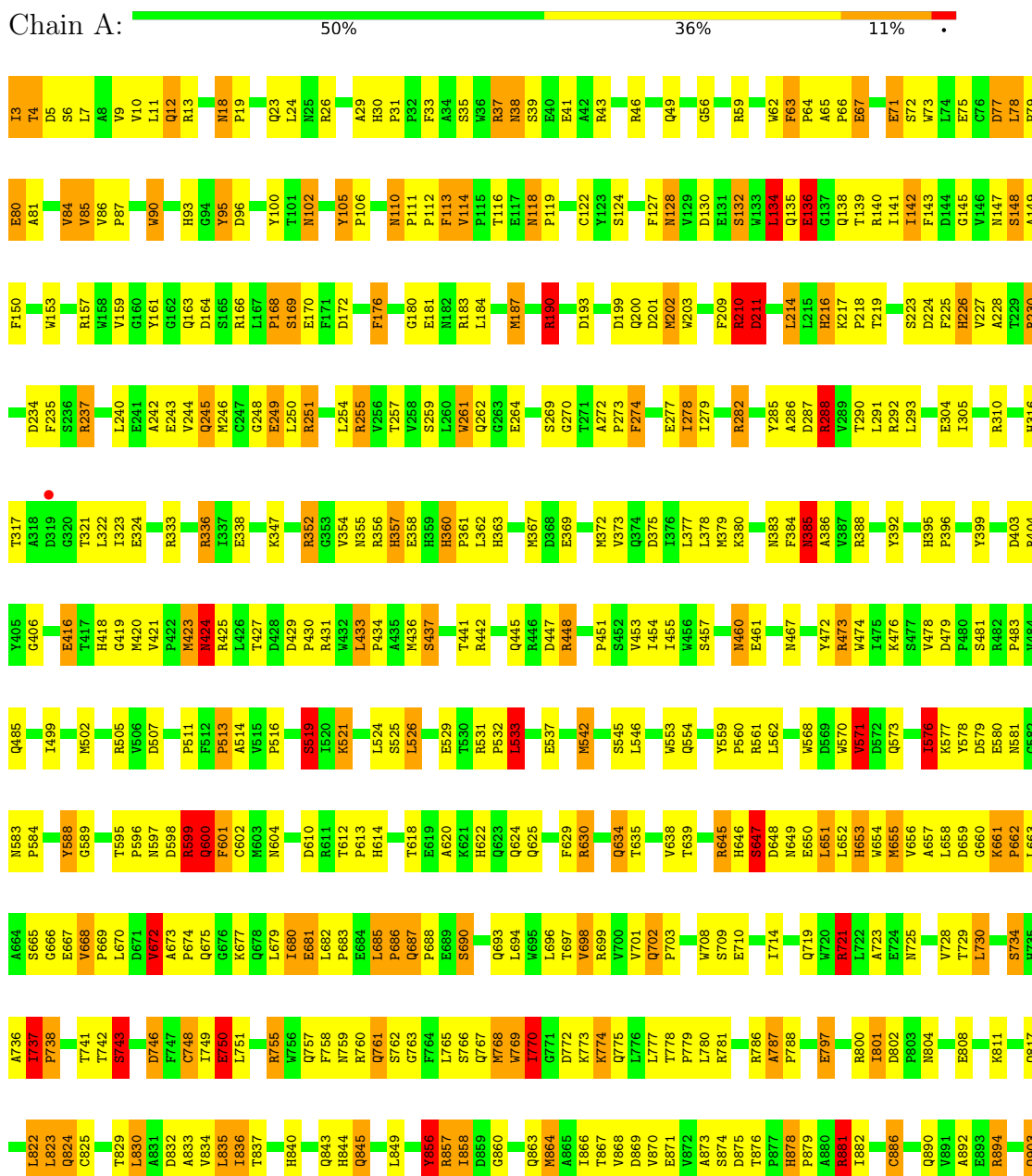
- Molecule 3 is water.

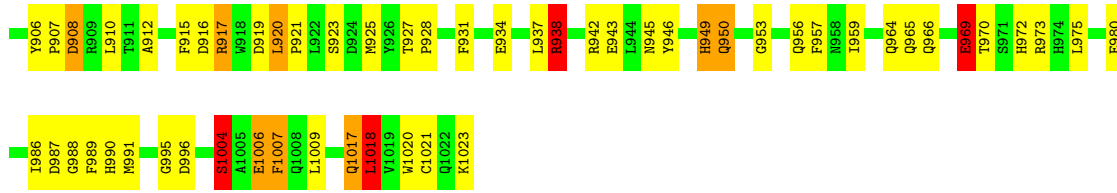
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	365	Total	O	0	0
			365	365		
3	B	366	Total	O	0	0
			366	366		
3	C	367	Total	O	0	0
			367	367		
3	D	366	Total	O	0	0
			366	366		

3 Residue-property plots

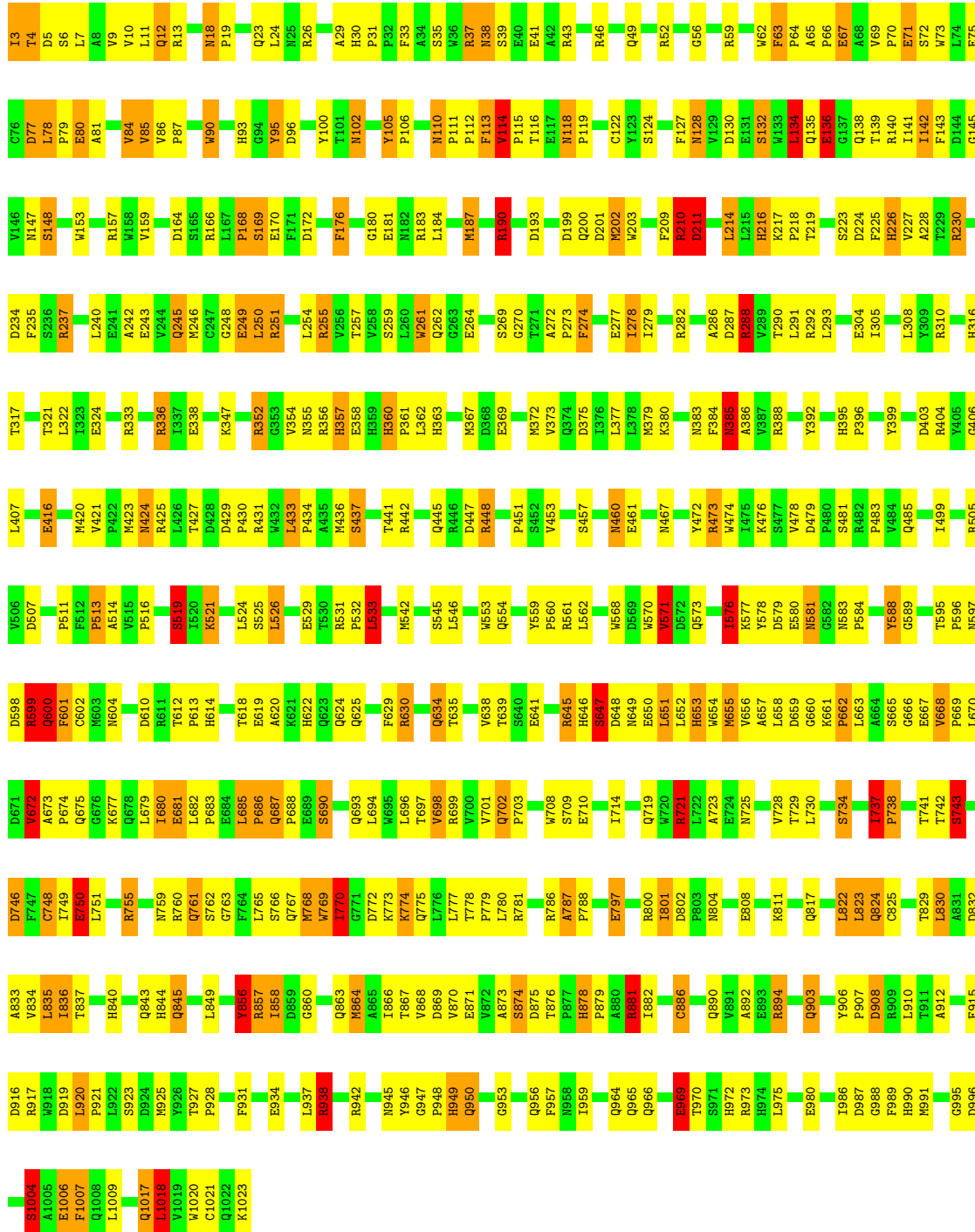
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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



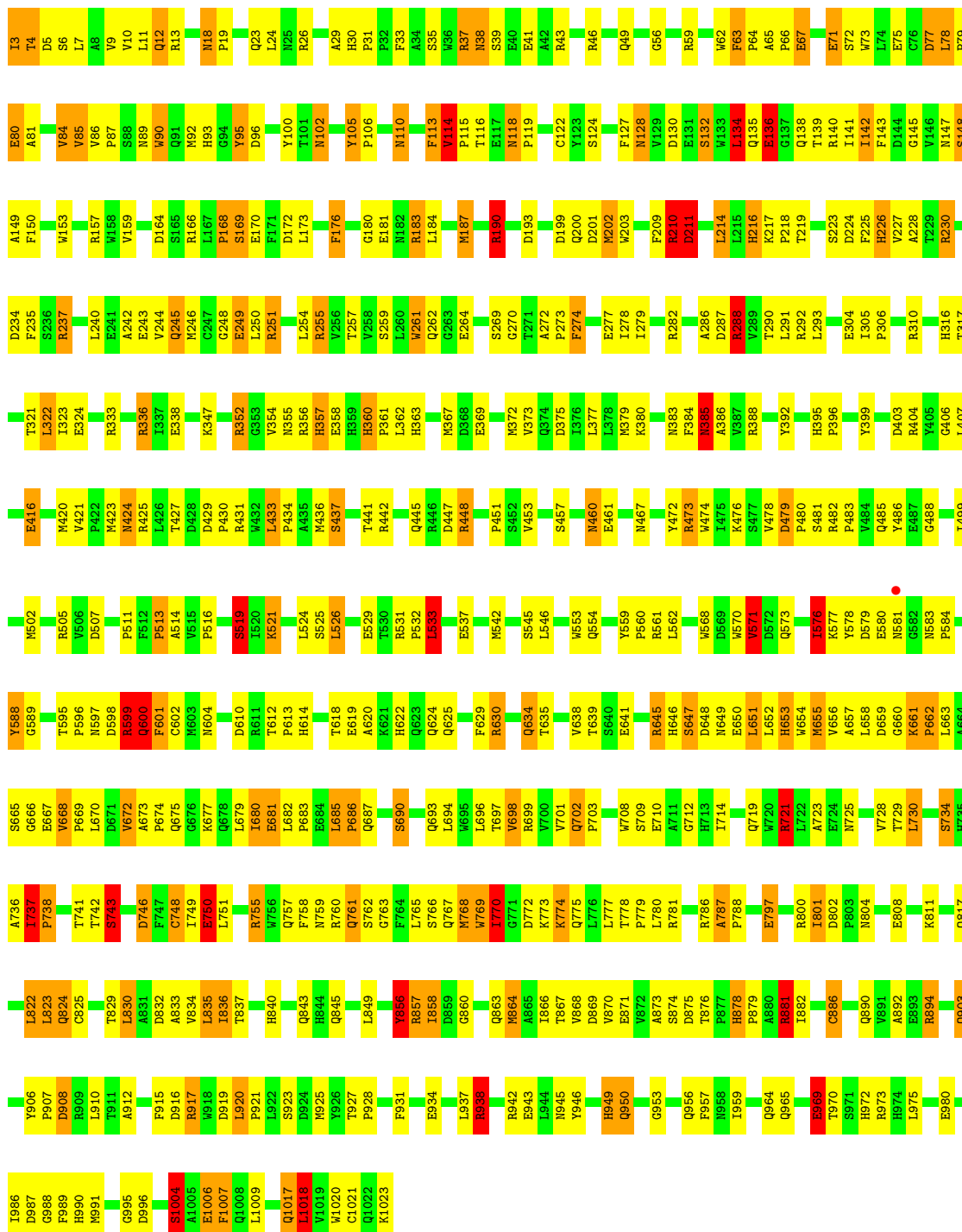


● Molecule 1: BETA-GALACTOSIDASE



● Molecule 1: BETA-GALACTOSIDASE

Chain C:  50% 37% 11%



● Molecule 1: BETA-GALACTOSIDASE

Chain D:  50% 36% 11%

F989	H990	M991	G995	D996	S1004	A1005	E1006	F1007	Q1008	L1009	Q1017	L1018	V1019	W1020	C1021	Q1022	K1023																																						
Y906	P907	R908	R909	L910	T911	A912	F915	D916	R917	V918	L919	L920	L921	S923	T927	P928	F931	E934	L937	R938	R942	E943	L944	N945	Y946	H949	Q950	G953	Q956	I959	Q964	Q965	Q966	E969	T970	S971	H972	R973	L975	E980	T986	D987	G988												
L822	L823	Q824	C825	T829	L830	A831	D832	A833	V834	L835	L836	L837	H840	Q843	H844	Q845	L849	V856	R857	I858	G860	Q863	E864	A865	I866	T867	V868	D869	V870	E871	V872	A873	S874	D875	T876	R877	H878	P879	A880	R881	I882	C886	Q890	V891	T729	L730	K811	Q817							
P738	T741	T742	S743	E744	M745	D746	F747	C748	I749	E750	L751	M755	V756	Q757	F758	N759	Q761	Q762	G763	F764	L765	S766	Q767	Q768	Q769	Q770	Q771	D772	K773	Q774	Q775	L776	T777	T778	P779	L780	R781	R782	T783	R784	R785	R786	R787	P788	E797	R800	I801	D802	C806	Q809	V891	T729	L730	K811	Q817
G666	E667	P668	L670	D671	R672	A673	F601	C602	M603	N604	D610	H611	T612	P613	H614	T618	E619	A620	H621	H622	Q623	Q624	Q625	F629	R630	Q634	T635	V638	T639	E641	R645	H646	G647	D648	N649	E650	L651	L652	H653	M654	M655	V656	A657	L658	D659	E580	G660	N581	X661	P662	L663	A664	S665		
I3	T4	D6	S6	L7	A8	V9	V10	L11	Q12	R13	M18	P19	Q23	L24	M25	R26	A29	H30	P31	F32	F33	A34	S35	W36	R37	N38	S39	E40	E41	A42	R43	R46	Q49	G56	N128	V129	D130	E131	S132	W133	L134	A65	A66	E67	A68	V69	T139	R140	E71	S72	W73	L74	E75	D77	
L78	P79	E80	A81	V84	V85	V86	P87	S88	N89	W90	Q91	M92	H93	G94	Y95	D96	Y100	T101	N102	Y105	P106	M110	P111	P112	F113	V114	P115	T116	E117	N118	G122	Y123	S124	F127	N129	D130	E131	S132	W133	L134	Q135	E136	G137	Q138	T139	R140	E71	S72	W73	L74	E75	D77			
N147	S148	W153	R157	M158	V159	D164	S165	N166	Q167	P168	S169	E170	F171	D172	F176	G180	E181	S259	W260	R261	Q262	E263	G264	T257	S269	G270	T271	A272	P273	F274	E277	M202	W203	F209	R210	D211	L214	L215	K217	P218	L291	R292	L293	N297	E304	I305	V227	A228	F229	R230	D234				
F235	S236	R237	L240	E241	A242	E243	V244	M245	M246	Q247	G248	E249	L250	R251	L254	R255	T256	V258	S259	W260	Q262	E263	G264	T257	S269	G270	T271	A272	P273	F274	E277	M202	W203	F209	R210	D211	L214	L215	K217	P218	L291	R292	L293	N297	E304	I305	V227	A228	F229	R230	D234				
H316	T317	A318	D319	G320	T321	L322	I323	E324	R333	R336	T337	E338	F347	G352	G353	V354	N355	R356	H357	E358	H359	H360	P361	L362	H363	K367	P368	E369	M372	V373	G374	D375	I376	L377	L378	N379	K380	N383	F384	R385	A386	Y387	R388	Y392	H395	P396	Y399	D403							
R404	L407	F416	T417	H418	C419	M420	P421	P422	M423	N424	R425	L426	T427	D428	D429	P430	R431	Y432	L433	P434	A435	M436	S437	T441	R442	Q445	R448	P451	S452	V453	L454	I455	S457	M460	E461	M467	Y472	R473	A474	I475	R476	S477	V478	D479	P480	S481	R482	P483	V484						
Q485	I499	R505	W506	D507	P511	F512	P513	P514	A514	W515	P516	S519	I520	K521	L524	S525	L526	E529	T530	R531	P532	L533	M542	S545	L546	M553	O554	Y559	P560	R561	L562	M568	O569	M570	V571	D572	O573	R473	A474	I475	R476	S477	V478	D479	P480	S481	R482	P483	V484						
G589	T595	P596	M597	D598	R599	Q600	F601	C602	M603	N604	D610	H611	T612	P613	H614	T618	E619	A620	H621	H622	Q623	Q624	Q625	F629	R630	Q634	T635	V638	T639	E641	R645	H646	G647	D648	N649	E650	L651	L652	H653	M654	M655	V656	A657	L658	D659	E580	G660	N581	X661	P662	L663	A664	S665		
G666	E667	P668	L670	D671	R672	A673	F601	C602	M603	N604	D610	H611	T612	P613	H614	T618	E619	A620	H621	H622	Q623	Q624	Q625	F629	R630	Q634	T635	V638	T639	E641	R645	H646	G647	D648	N649	E650	L651	L652	H653	M654	M655	V656	A657	L658	D659	E580	G660	N581	X661	P662	L663	A664	S665		
P738	T741	T742	S743	E744	M745	D746	F747	C748	I749	E750	L751	M755	V756	Q757	F758	N759	Q761	Q762	G763	F764	L765	S766	Q767	Q768	Q769	Q770	Q771	D772	K773	Q774	Q775	L776	T777	T778	P779	L780	R781	R782	T783	R784	R785	R786	R787	P788	E797	R800	I801	D802	C806	Q809	V891	T729	L730	K811	Q817
L822	L823	Q824	C825	T829	L830	A831	D832	A833	V834	L835	L836	L837	H840	Q843	H844	Q845	L849	V856	R857	I858	G860	Q863	E864	A865	I866	T867	V868	D869	V870	E871	V872	A873	S874	D875	T876	R877	H878	P879	A880	R881	I882	C886	Q890	V891	T729	L730	K811	Q817							
Y906	P907	R908	R909	L910	T911	A912	F915	D916	R917	V918	L919	L920	L921	S923	T927	P928	F931	E934	L937	R938	R942	E943	L944	N945	Y946	H949	Q950	G953	Q956	I959	Q964	Q965	Q966	E969	T970	S971	H972	R973	L975	E980	T986	D987	G988												

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	153.40Å 173.40Å 204.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.80 25.00 – 2.82	Depositor EDS
% Data completeness (in resolution range)	88.0 (25.00-2.80) 80.3 (25.00-2.82)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 2.80Å)	Xtrriage
Refinement program	TNT 5E	Depositor
R, R_{free}	0.167 , 0.198 0.149 , 0.178	Depositor DCC
R_{free} test set	1590 reflections (1.50%)	wwPDB-VP
Wilson B-factor (Å ²)	40.7	Xtrriage
Anisotropy	0.148	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 117.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	34424	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 41.86 % 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 2.2179e-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

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	1.34	17/8515 (0.2%)	1.82	169/11615 (1.5%)
1	B	1.34	17/8515 (0.2%)	1.82	167/11615 (1.4%)
1	C	1.34	17/8515 (0.2%)	1.82	166/11615 (1.4%)
1	D	1.34	18/8515 (0.2%)	1.82	167/11615 (1.4%)
All	All	1.34	69/34060 (0.2%)	1.82	669/46460 (1.4%)

The worst 5 of 69 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	479	ASP	C-N	-9.15	1.25	1.34
1	D	479	ASP	C-N	-9.14	1.25	1.34
1	A	479	ASP	C-N	-9.14	1.25	1.34
1	C	479	ASP	C-N	-9.05	1.25	1.34
1	D	249	GLU	CA-C	-8.01	1.42	1.52

The worst 5 of 669 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	114	VAL	N-CA-CB	-14.09	102.11	111.83
1	C	114	VAL	N-CA-CB	-14.08	102.11	111.83
1	D	114	VAL	N-CA-CB	-14.08	102.12	111.83
1	A	114	VAL	N-CA-CB	-14.06	102.12	111.83
1	D	385	ASN	CB-CA-C	-12.05	89.51	109.03

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	8238	0	7824	423	0
1	B	8238	0	7824	397	0
1	C	8238	0	7824	400	0
1	D	8238	0	7824	402	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	365	0	0	10	0
3	B	366	0	0	10	0
3	C	367	0	0	10	0
3	D	366	0	0	10	0
All	All	34424	0	31296	1597	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 25.

The worst 5 of 1597 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:B:746:ASP:HA	1:B:760:ARG:HG3	1.34	1.10
1:A:746:ASP:HA	1:A:760:ARG:HG3	1.34	1.09
1:D:746:ASP:HA	1:D:760:ARG:HG3	1.34	1.09
1:B:427:THR:HA	1:B:436:MET:HE1	1.09	1.08
1:D:427:THR:HA	1:D:436:MET:HE1	1.09	1.07

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	1026/1021 (100%)	953 (93%)	67 (6%)	6 (1%)	21	51
1	B	1026/1021 (100%)	953 (93%)	67 (6%)	6 (1%)	21	51
1	C	1026/1021 (100%)	953 (93%)	67 (6%)	6 (1%)	21	51
1	D	1026/1021 (100%)	953 (93%)	67 (6%)	6 (1%)	21	51
All	All	4104/4084 (100%)	3812 (93%)	268 (6%)	24 (1%)	21	51

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	647	SER
1	B	647	SER
1	C	647	SER
1	D	647	SER
1	A	77	ASP

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	880/873 (101%)	746 (85%)	134 (15%)	3	10
1	B	880/873 (101%)	746 (85%)	134 (15%)	3	10
1	C	880/873 (101%)	746 (85%)	134 (15%)	3	10
1	D	880/873 (101%)	746 (85%)	134 (15%)	3	10
All	All	3520/3492 (101%)	2984 (85%)	536 (15%)	3	10

5 of 536 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	529	GLU
1	D	647	SER
1	D	526	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	874	SER
1	B	526	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 104 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	316	HIS
1	C	718	GLN
1	D	824	GLN
1	C	363	HIS
1	C	581	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 8 ligands modelled in this entry, 8 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	1021/1021 (100%)	-0.64	1 (0%) 92 90	8, 29, 69, 100	26 (2%)
1	B	1021/1021 (100%)	-0.61	0 100 100	8, 29, 69, 100	26 (2%)
1	C	1021/1021 (100%)	-0.70	1 (0%) 92 90	8, 29, 69, 100	26 (2%)
1	D	1021/1021 (100%)	-0.70	2 (0%) 91 88	8, 29, 69, 100	26 (2%)
All	All	4084/4084 (100%)	-0.66	4 (0%) 92 90	8, 29, 69, 100	104 (2%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	319	ASP	3.9
1	D	319	ASP	2.7
1	C	581	ASN	2.1
1	D	744	GLU	2.1

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(\AA^2)	Q<0.9
2	MG	A	3001	1/1	0.97	0.05	28,28,28,28	0
2	MG	C	3002	1/1	0.97	0.07	31,31,31,31	0
2	MG	D	3002	1/1	0.97	0.05	31,31,31,31	0
2	MG	B	3002	1/1	0.98	0.04	31,31,31,31	0
2	MG	C	3001	1/1	0.99	0.03	28,28,28,28	0
2	MG	B	3001	1/1	0.99	0.04	28,28,28,28	0
2	MG	D	3001	1/1	0.99	0.02	28,28,28,28	0
2	MG	A	3002	1/1	0.99	0.07	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.