



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 12:09 AM UTC

PDB ID : 1F4H / pdb_00001f4h
Title : E. COLI (LACZ) BETA-GALACTOSIDASE (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

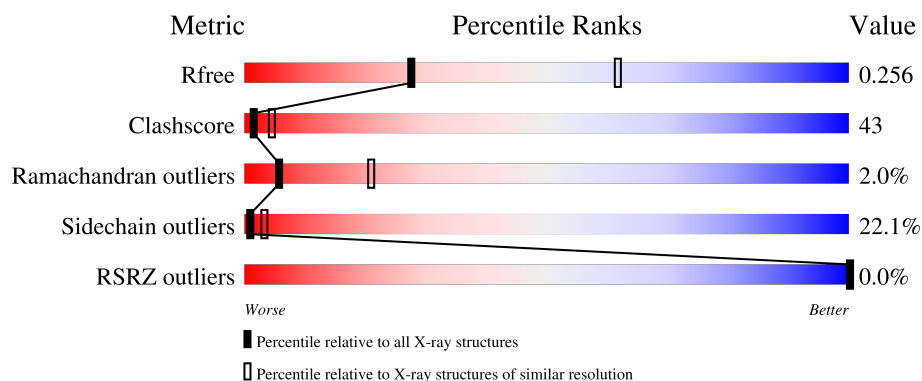
1 Overall quality at a glance

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	
1	B	1021	
1	C	1021	
1	D	1021	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 33805 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		

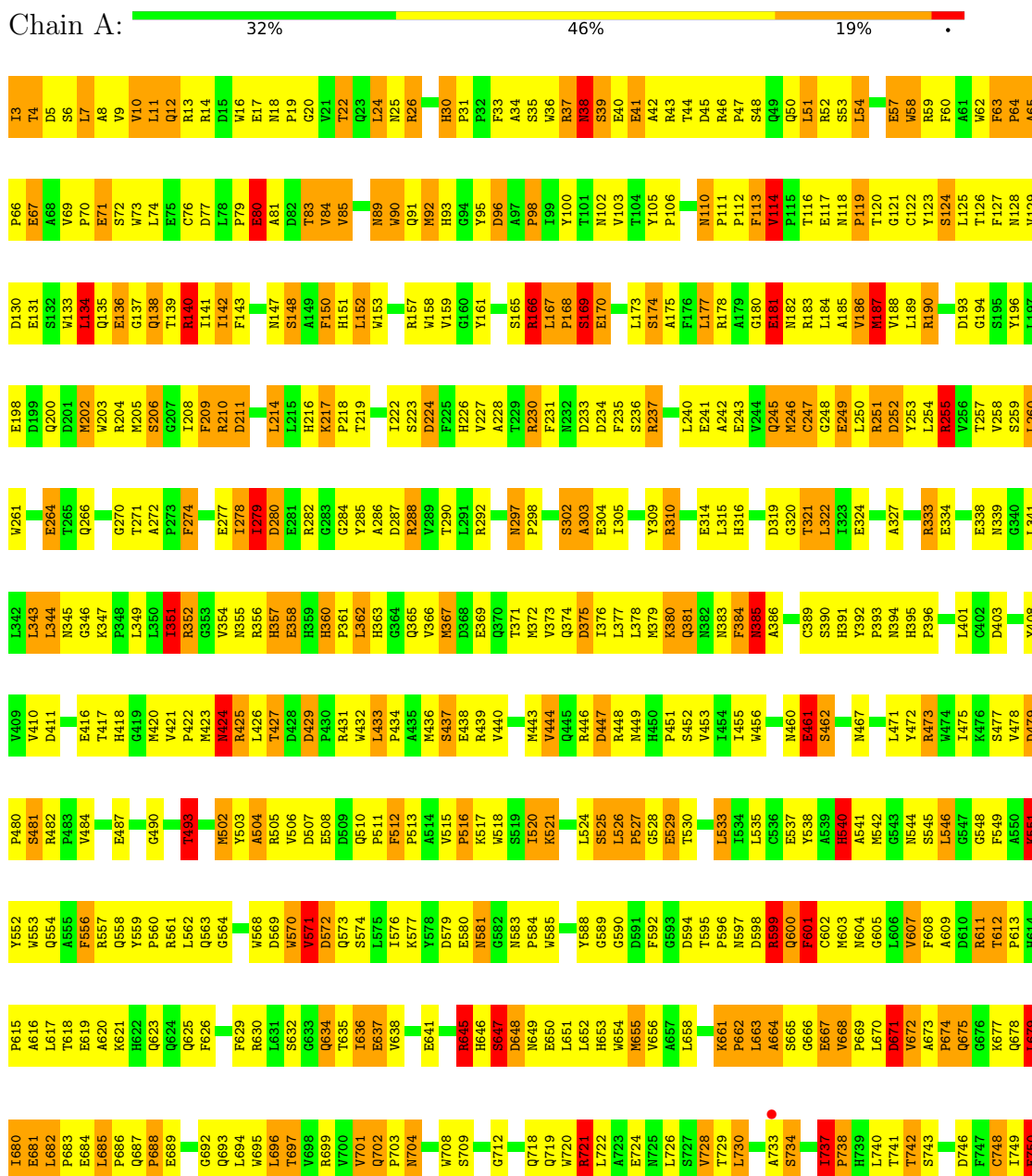
- Molecule 3 is water.

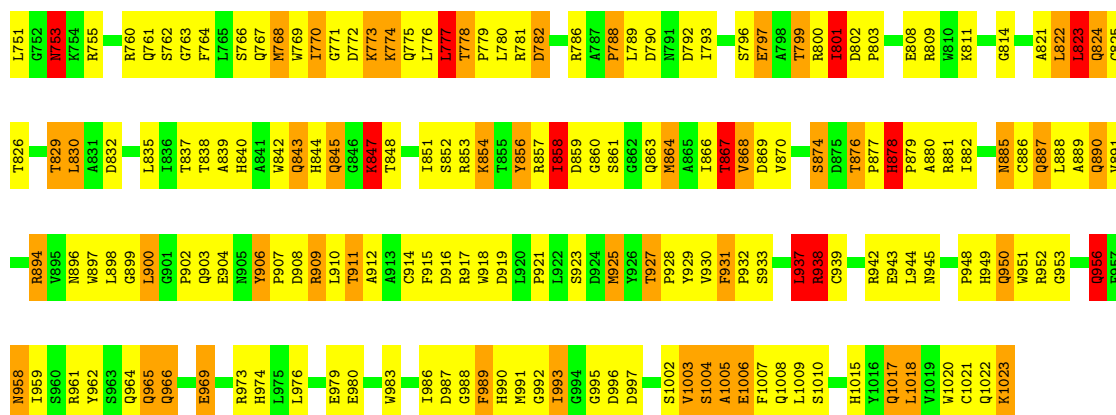
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	211	Total	O	0	0
			211	211		
3	B	202	Total	O	0	0
			202	202		
3	C	220	Total	O	0	0
			220	220		
3	D	212	Total	O	0	0
			212	212		

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.

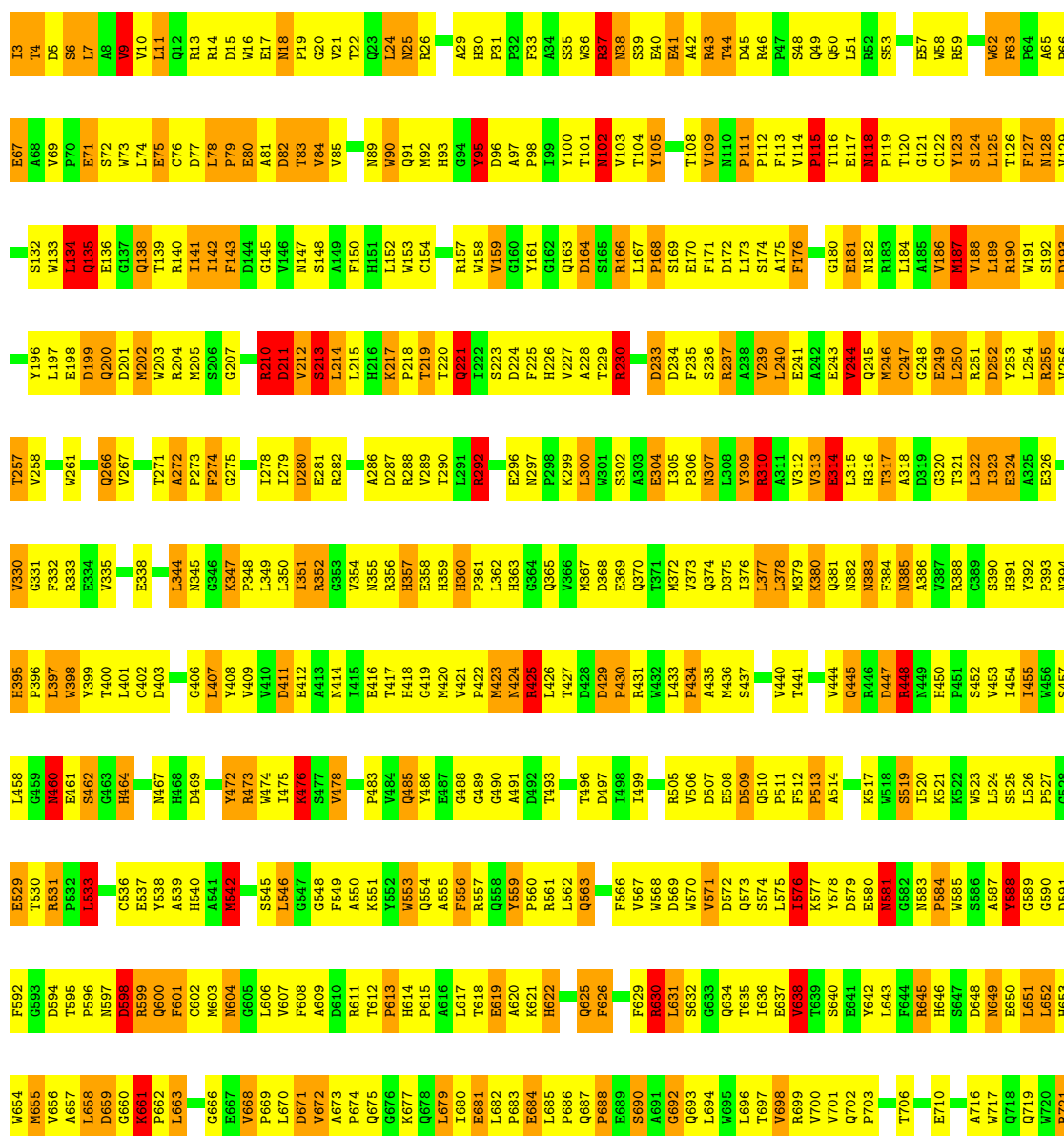
• Molecule 1: BETA-GALACTOSIDASE

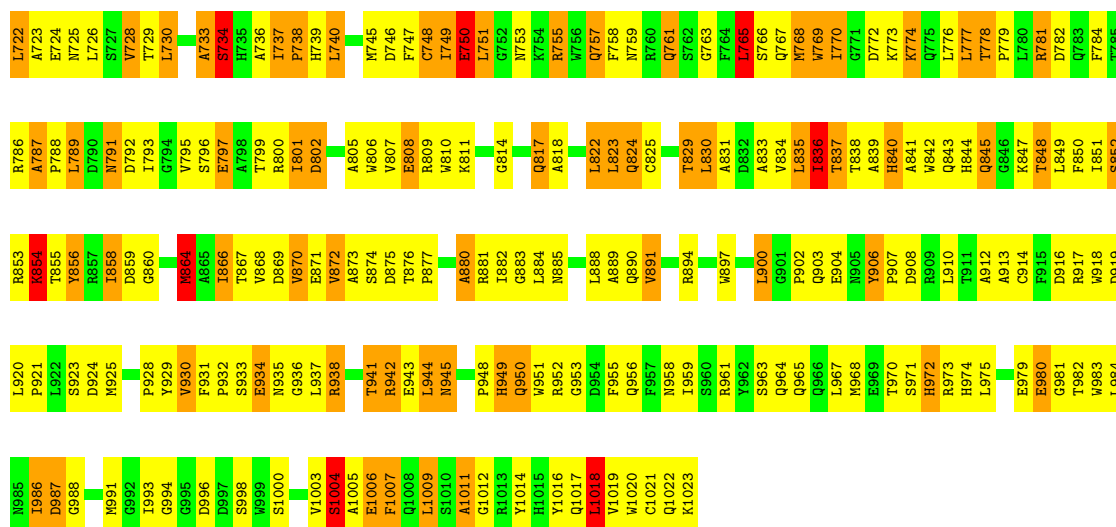




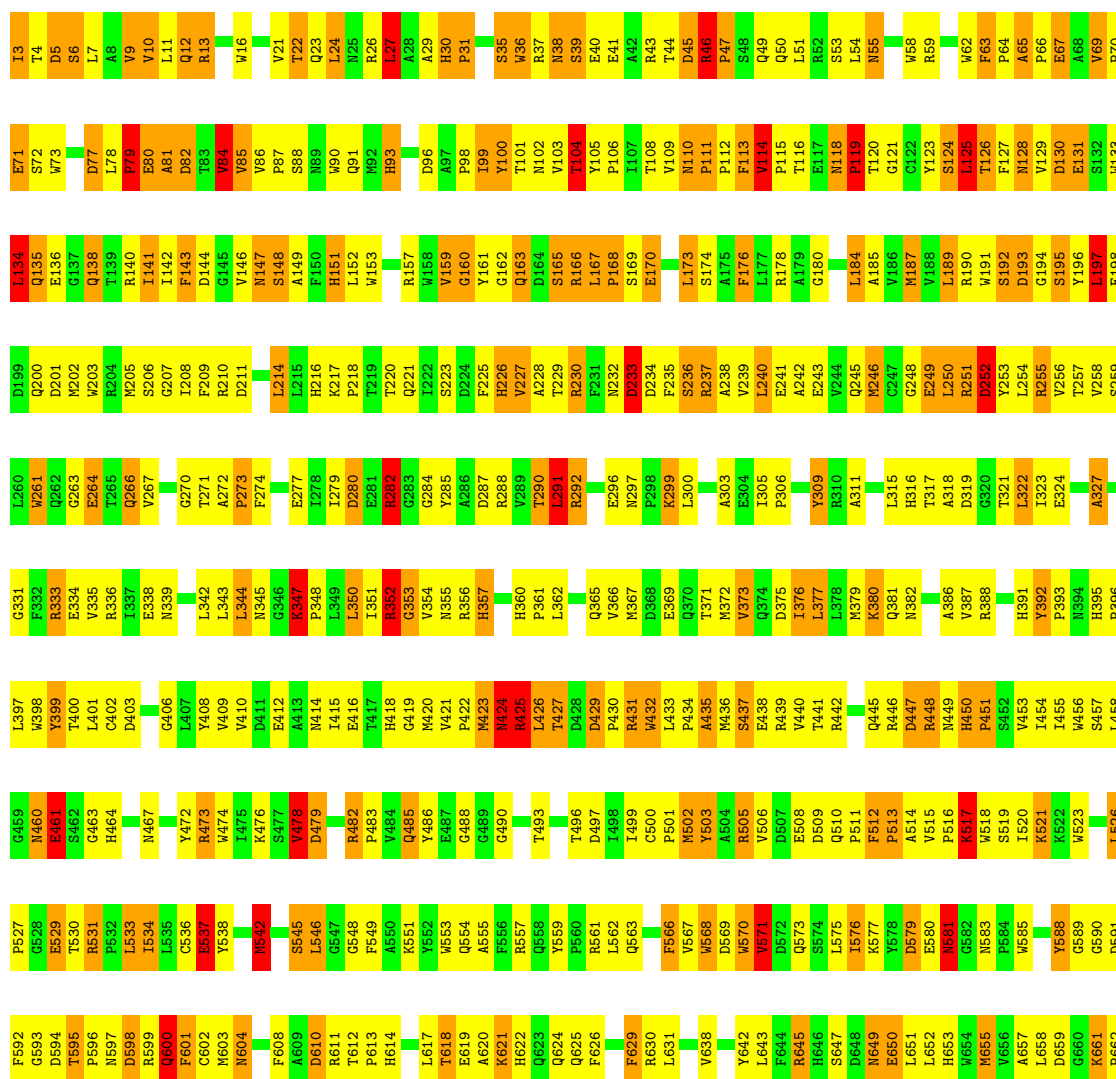
• Molecule 1: BETA-GALACTOSIDASE

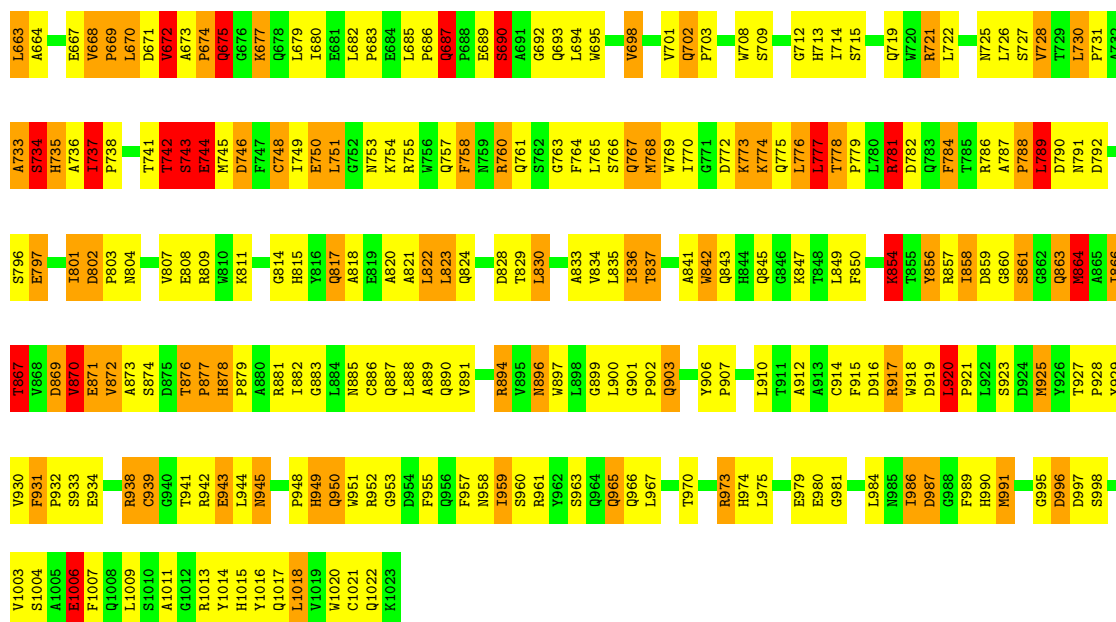
Chain B: 26% 49% 21% •





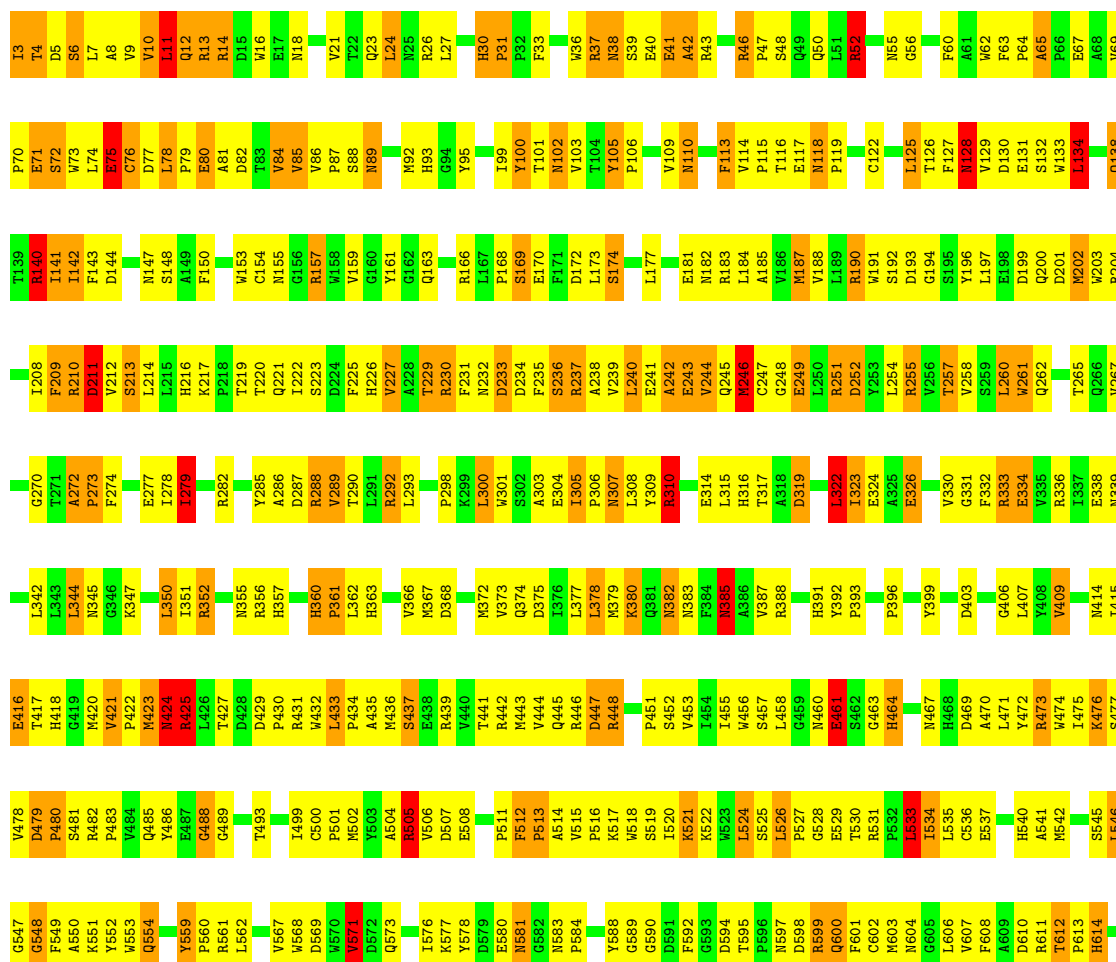
• Molecule 1: BETA-GALACTOSIDASE





• Molecule 1: BETA-GALACTOSIDASE

Chain D: 33% 45% 19% .



L617	P883	F758	T826	R894	R961	L617
T618	E684	N759	A827	V895	R961	T618
E619	L685	R760	D828	N896	Q964	E619
A620	P686	T621	T829	W897	Q965	A620
K621	Q687	S762	L830		Q966	K621
Q624	P688	G763		P902	L967	Q624
Q625	E689	F764	V834	Q903	M968	Q625
F626	S690	L765	L835	E904	E969	F626
	A691	S766	L836	T970	T970	
R630	G692	Q767	T837	Y906	S971	R630
L831	Q693	M768	T838	P907	H972	L831
S632	L694	W769		D908	R973	S632
G633	W695	L770	H844	R909	H974	G633
Q634	L696	G771	Q845	L910		Q634
T635	G697	D772	G846	T911	E979	T635
T636	K773	K774	K847	A912	E980	T636
E637	R699	T774	T848	A913	G981	E637
V638	V700	Q775	L849		T982	V638
T639	V701	L776	F850	W918	Q983	T639
	Q702	L777	L851	D919	L984	
L643	P703	T778	S852	L920	N985	L643
F644	N704	P779	R853	P921	I986	F644
R645	A705	L780	K854	L922	D987	R645
H646	T706	R781	T855	S923	G988	H646
S647	A707	D782	Y856	D924	F989	S647
D648	W708	Q783	R857	M925	H990	D648
N649	S709	F784	L858	Y926	M991	N649
E650	A716	T785	D859	T927	G992	E650
L651	W717	R786	G860	P928	I993	L651
L652		A787	S861	Y929		L652
H653		P788	Q862	V930	D996	H653
W654	R721		F931	F931	D997	W654
M655	E724	L793	M864	P932	S998	M655
V656		G794		S933	W999	V656
A657	L730		T867	R935		A657
L658	P731	E797	V868	N934	S1002	L658
	A732	R798	G366	G366	V1003	
K661	A733	T799	V870	S1004	A1004	K661
P662	S734	R800	E871	R938	A1005	P662
L663		S801	V872		E1006	L663
A664	T737	D802	A873	T941		A664
S665	P738	P803	S874	R942	L1009	S665
G666	H739	F603	D875	E943	S1010	G666
E667	L740	N904	T876	L944	A1011	E667
V668	T741	E808	P877	N945		V668
P669	T742	E809	H878	Y946	Y1014	P669
L670	S743	R879	P879	G947		L670
D671	E744	W810	A880	P948	Q1017	D671
V672	M745	K911	R881	H949	L1018	V672
A673	D746	G814	L882	Q950	V1019	A673
P674	F747	H815	G883	W951	W1020	P674
Q675	G748	Y816	L884	R952	C1021	Q675
G676	I749		S885	G953	Q1022	G676
K677	E750	E919	C886	D954	K1023	K677
Q678	L751		Q887	F955		Q678
L679		L822	L888	Q956		L679
E680	R755	L823	A889	P957	N958	E680
L681	W756	Q824	Q890	T957	T959	L681
L682	C757	C925				L682

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)	Xtriage
Refinement program	TNT 5E	Depositor
R, R_{free}	0.137 , 0.279 0.125 , 0.256	Depositor DCC
R_{free} test set	1590 reflections (1.50%)	wwPDB-VP
Wilson B-factor (Å ²)	40.7	Xtriage
Anisotropy	0.148	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 156.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	33805	wwPDB-VP
Average B, all atoms (Å ²)	32.0	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 41.86 % 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 ⓘ

5.1 Standard geometry ⓘ

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	1.38	20/8515 (0.2%)	1.97	235/11615 (2.0%)
1	B	1.34	11/8515 (0.1%)	2.01	273/11615 (2.4%)
1	C	1.34	19/8515 (0.2%)	1.97	248/11615 (2.1%)
1	D	1.36	20/8515 (0.2%)	1.99	250/11615 (2.2%)
All	All	1.35	70/34060 (0.2%)	1.99	1006/46460 (2.2%)

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	4	0
1	B	1	0
1	C	4	1
1	D	3	0
All	All	12	1

The worst 5 of 70 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	479	ASP	C-N	-7.97	1.27	1.33
1	D	749	ILE	N-CA	-6.92	1.38	1.46
1	A	258	VAL	CA-C	-6.38	1.45	1.52
1	D	679	LEU	CA-C	-6.37	1.45	1.52
1	A	376	ILE	CA-CB	-6.34	1.47	1.54

The worst 5 of 1006 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	385	ASN	N-CA-CB	-12.32	92.58	110.81
1	A	385	ASN	CB-CA-C	-12.19	89.17	109.65
1	B	385	ASN	CB-CA-C	-11.71	88.75	109.24
1	D	479	ASP	CA-C-N	11.21	132.15	120.04
1	D	479	ASP	C-N-CA	11.21	132.15	120.04

5 of 12 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	166	ARG	CA
1	A	187	MET	CA
1	A	903[A]	GLN	CA
1	A	903[B]	GLN	CA
1	B	118	ASN	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	352	ARG	Mainchain

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	8238	0	7824	673	0
1	B	8238	0	7824	743	0
1	C	8238	0	7824	718	0
1	D	8238	0	7823	658	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	211	0	0	13	0
3	B	202	0	0	12	0
3	C	220	0	0	24	0
3	D	212	0	0	13	0
All	All	33805	0	31295	2736	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 43.

The worst 5 of 2736 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:A:173:LEU:HB3	1:A:177:LEU:HD21	1.21	1.17
1:A:427:THR:HA	1:A:436:MET:HE1	1.24	1.15
1:B:734:SER:HB3	1:B:860:GLY:HA3	1.24	1.12
1:C:427:THR:HA	1:C:436:MET:HE1	1.17	1.11
1:A:777:LEU:HD21	1:A:889:ALA:HB2	1.28	1.10

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%)	923 (90%)	86 (8%)	17 (2%)	7	25
1	B	1026/1021 (100%)	917 (89%)	89 (9%)	20 (2%)	6	22
1	C	1026/1021 (100%)	918 (90%)	82 (8%)	26 (2%)	4	16
1	D	1026/1021 (100%)	921 (90%)	85 (8%)	20 (2%)	6	22
All	All	4104/4084 (100%)	3679 (90%)	342 (8%)	83 (2%)	6	21

5 of 83 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	647	SER
1	A	733	ALA
1	A	734	SER
1	A	1005	ALA
1	B	79	PRO

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%)	688 (78%)	192 (22%)	1	3
1	B	880/873 (101%)	676 (77%)	204 (23%)	1	3
1	C	880/873 (101%)	680 (77%)	200 (23%)	1	3
1	D	880/873 (101%)	703 (80%)	177 (20%)	1	4
All	All	3520/3492 (101%)	2747 (78%)	773 (22%)	1	3

5 of 773 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	324	GLU
1	C	874	SER
1	C	431[A]	ARG
1	C	322	LEU
1	C	702	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 111 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	138	GLN
1	D	965	GLN
1	C	597	ASN
1	D	949	HIS
1	D	614	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 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%)	-0.90	1 (0%) 92 90	2, 25, 64, 100	26 (2%)
1	B	1021/1021 (100%)	-0.78	0 100 100	4, 30, 70, 100	26 (2%)
1	C	1021/1021 (100%)	-0.86	0 100 100	6, 27, 63, 100	26 (2%)
1	D	1021/1021 (100%)	-0.91	0 100 100	3, 25, 64, 100	26 (2%)
All	All	4084/4084 (100%)	-0.87	1 (0%) 100 100	2, 27, 66, 100	104 (2%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	733	ALA	2.3

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(Å ²)	Q<0.9
2	MG	B	3002	1/1	0.95	0.06	36,36,36,36	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	MG	D	3001	1/1	0.96	0.04	40,40,40,40	0
2	MG	A	3001	1/1	0.97	0.05	28,28,28,28	0
2	MG	C	3002	1/1	0.97	0.04	31,31,31,31	0
2	MG	B	3001	1/1	0.97	0.04	20,20,20,20	0
2	MG	A	3002	1/1	0.98	0.07	31,31,31,31	0
2	MG	C	3001	1/1	0.99	0.02	15,15,15,15	0
2	MG	D	3002	1/1	0.99	0.05	25,25,25,25	0

6.5 Other polymers [i](#)

There are no such residues in this entry.