



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 15, 2026 – 12:20 AM UTC

PDB ID : 3BTU / pdb_00003btu
Title : Crystal structure of the super-repressor mutant of Gal80p from *Saccharomyces cerevisiae*; Gal80(S2) [E351K]
Authors : Kumar, P.R.; Joshua-Tor, L.
Deposited on : 2007-12-30
Resolution : 2.85 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

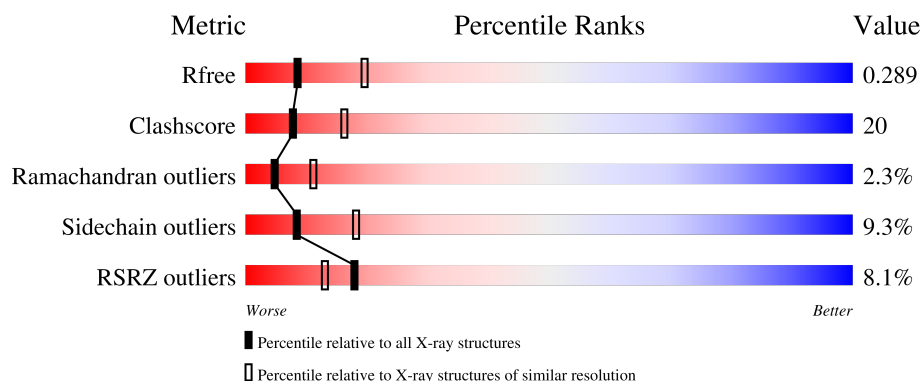
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.85 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	1407 (2.88-2.84)
Clashscore	190562	1446 (2.88-2.84)
Ramachandran outliers	187476	1406 (2.88-2.84)
Sidechain outliers	187428	1407 (2.88-2.84)
RSRZ outliers	180081	1408 (2.88-2.84)

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	438	
1	B	438	
1	C	438	
1	D	438	
1	E	438	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	438	14% 46% 36% 5% 12%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 18389 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 Galactose/lactose metabolism regulatory protein GAL80.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	392	Total	C	N	O	S	0	0	0
			3088	1987	518	572	11			
1	B	389	Total	C	N	O	S	0	0	0
			3065	1975	513	566	11			
1	C	388	Total	C	N	O	S	0	0	0
			3057	1969	511	566	11			
1	D	392	Total	C	N	O	S	0	0	0
			3087	1987	518	571	11			
1	E	387	Total	C	N	O	S	0	0	0
			3046	1960	512	563	11			
1	F	387	Total	C	N	O	S	0	0	0
			3046	1960	512	563	11			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP P04387
A	-1	SER	-	expression tag	UNP P04387
A	0	HIS	-	expression tag	UNP P04387
A	351	LYS	GLU	engineered mutation	UNP P04387
B	-2	GLY	-	expression tag	UNP P04387
B	-1	SER	-	expression tag	UNP P04387
B	0	HIS	-	expression tag	UNP P04387
B	351	LYS	GLU	engineered mutation	UNP P04387
C	-2	GLY	-	expression tag	UNP P04387
C	-1	SER	-	expression tag	UNP P04387
C	0	HIS	-	expression tag	UNP P04387
C	351	LYS	GLU	engineered mutation	UNP P04387
D	-2	GLY	-	expression tag	UNP P04387
D	-1	SER	-	expression tag	UNP P04387
D	0	HIS	-	expression tag	UNP P04387
D	351	LYS	GLU	engineered mutation	UNP P04387
E	-2	GLY	-	expression tag	UNP P04387

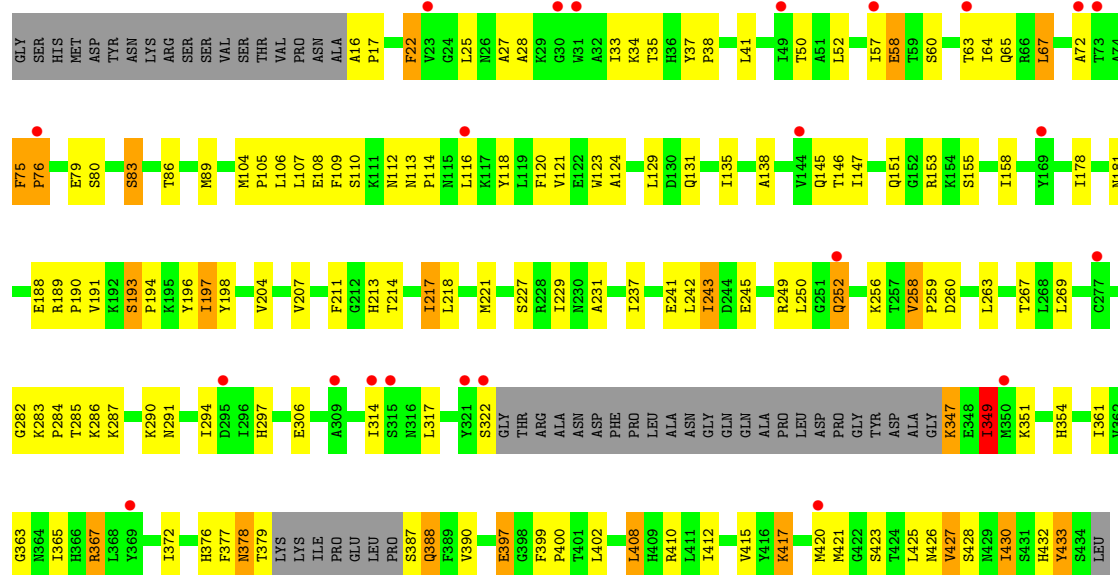
Continued on next page...

Continued from previous page...

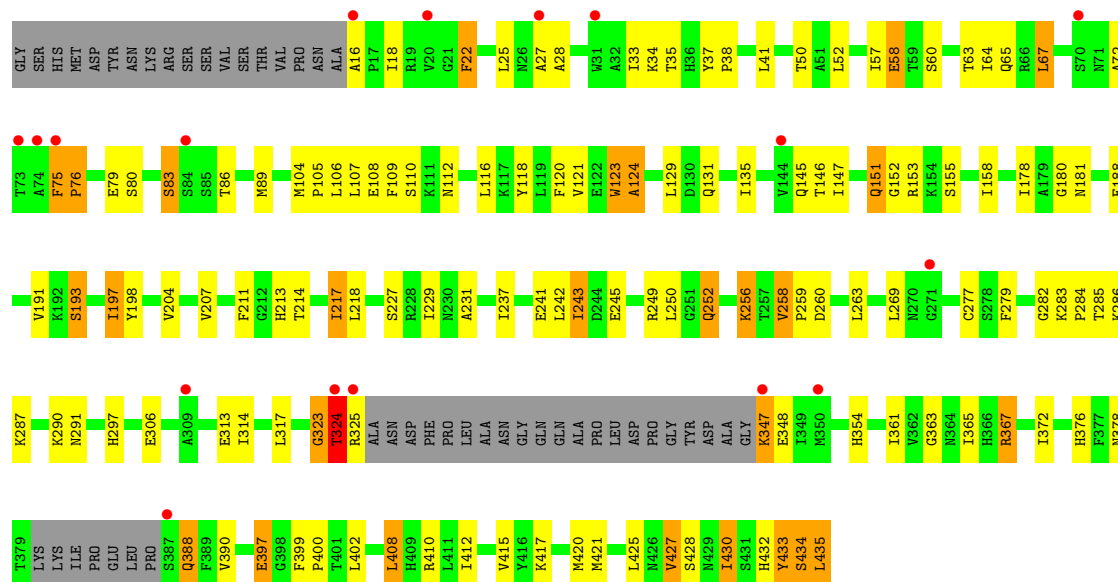
Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	SER	-	expression tag	UNP P04387
E	0	HIS	-	expression tag	UNP P04387
E	351	LYS	GLU	engineered mutation	UNP P04387
F	-2	GLY	-	expression tag	UNP P04387
F	-1	SER	-	expression tag	UNP P04387
F	0	HIS	-	expression tag	UNP P04387
F	351	LYS	GLU	engineered mutation	UNP P04387



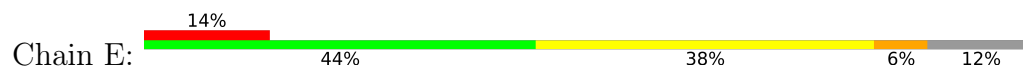
- Molecule 1: Galactose/lactose metabolism regulatory protein GAL80

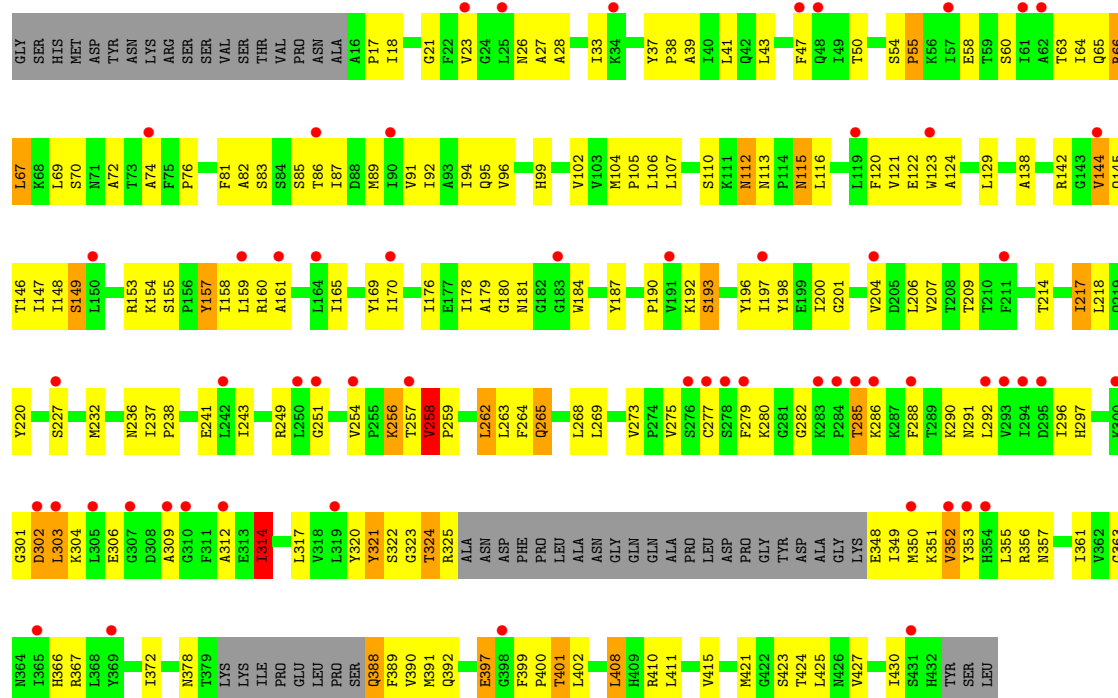


- Molecule 1: Galactose/lactose metabolism regulatory protein GAL80



- Molecule 1: Galactose/lactose metabolism regulatory protein GAL80





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	495.32Å 84.86Å 66.46Å 90.00° 98.90° 90.00°	Depositor
Resolution (Å)	50.00 – 2.85 50.00 – 2.85	Depositor EDS
% Data completeness (in resolution range)	85.6 (50.00-2.85) 85.6 (50.00-2.85)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.33 (at 2.86Å)	Xtriage
Refinement program	REFMAC 5.2.0019, PHENIX	Depositor
R, R_{free}	0.228 , 0.278 0.249 , 0.289	Depositor DCC
R_{free} test set	5498 reflections (8.48%)	wwPDB-VP
Wilson B-factor (Å ²)	66.4	Xtriage
Anisotropy	0.629	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 72.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.040 for -h-2*k,l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	18389	wwPDB-VP
Average B, all atoms (Å ²)	109.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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	0.43	0/3155	0.80	6/4270 (0.1%)
1	B	0.43	0/3132	0.82	4/4240 (0.1%)
1	C	0.43	0/3124	0.82	6/4230 (0.1%)
1	D	0.45	0/3154	0.82	8/4270 (0.2%)
1	E	0.36	0/3112	0.77	2/4214 (0.0%)
1	F	0.35	0/3112	0.75	2/4214 (0.0%)
All	All	0.41	0/18789	0.80	28/25438 (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	D	0	1

There are no bond length outliers.

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	347	LYS	N-CA-C	-10.23	82.37	111.00
1	B	347	LYS	N-CA-C	-9.98	83.05	111.00
1	D	323	GLY	N-CA-C	8.04	121.29	110.43
1	C	379	THR	N-CA-C	-6.41	93.06	111.00
1	F	258	VAL	CA-C-N	5.95	125.89	119.76

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	123	TRP	Peptide

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	3088	0	3101	112	0
1	B	3065	0	3080	111	0
1	C	3057	0	3067	116	0
1	D	3087	0	3101	109	0
1	E	3046	0	3058	169	0
1	F	3046	0	3058	160	0
All	All	18389	0	18465	734	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 20.

The worst 5 of 734 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:347:LYS:N	1:D:348:GLU:CA	1.70	1.34
1:E:155:SER:HB3	1:E:158:ILE:HD13	1.41	1.03
1:B:347:LYS:HD2	1:B:347:LYS:O	1.59	0.99
1:C:349:ILE:HD11	1:C:351:LYS:HE3	1.48	0.94
1:D:347:LYS:N	1:D:348:GLU:HA	0.77	0.92

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	386/438 (88%)	356 (92%)	24 (6%)	6 (2%)	7 17

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	383/438 (87%)	353 (92%)	26 (7%)	4 (1%)	12	25
1	C	382/438 (87%)	350 (92%)	25 (6%)	7 (2%)	6	15
1	D	386/438 (88%)	352 (91%)	27 (7%)	7 (2%)	6	15
1	E	381/438 (87%)	304 (80%)	65 (17%)	12 (3%)	3	7
1	F	381/438 (87%)	298 (78%)	67 (18%)	16 (4%)	2	3
All	All	2299/2628 (88%)	2013 (88%)	234 (10%)	52 (2%)	5	11

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	433	TYR
1	C	349	ILE
1	C	378	ASN
1	C	433	TYR
1	D	124	ALA

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	338/375 (90%)	305 (90%)	33 (10%)	7	16
1	B	335/375 (89%)	302 (90%)	33 (10%)	7	16
1	C	335/375 (89%)	305 (91%)	30 (9%)	9	19
1	D	338/375 (90%)	306 (90%)	32 (10%)	8	17
1	E	333/375 (89%)	300 (90%)	33 (10%)	7	16
1	F	333/375 (89%)	307 (92%)	26 (8%)	11	25
All	All	2012/2250 (89%)	1825 (91%)	187 (9%)	8	18

5 of 187 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	258	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	256	LYS
1	D	367	ARG
1	E	70	SER
1	E	355	LEU

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	D	112	ASN
1	F	364	ASN
1	D	357	ASN
1	F	357	ASN
1	F	131	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](#)

There are no ligands in this entry.

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	392/438 (89%)	0.45	15 (3%)	44 35	51, 88, 138, 216	0
1	B	389/438 (88%)	0.43	11 (2%)	55 46	53, 89, 131, 169	0
1	C	388/438 (88%)	0.58	23 (5%)	28 21	54, 93, 163, 296	0
1	D	392/438 (89%)	0.43	17 (4%)	40 31	20, 86, 131, 183	0
1	E	387/438 (88%)	1.16	62 (16%)	5 4	80, 129, 203, 281	0
1	F	387/438 (88%)	1.23	60 (15%)	5 4	74, 136, 215, 350	0
All	All	2335/2628 (88%)	0.71	188 (8%)	18 13	20, 102, 181, 350	0

The worst 5 of 188 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	325	ARG	5.5
1	F	47	PHE	4.8
1	C	57	ILE	4.3
1	A	30	GLY	4.2
1	F	365	ILE	4.1

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.