



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 05:08 AM GMT

PDB ID : 1RZR
Title : crystal structure of transcriptional regulator-phosphoprotein-DNAcomplex
Authors : Schumacher, M.A.; Allen, G.S.; Brennan, R.G.
Deposited on : 2003-12-27
Resolution : 2.80 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

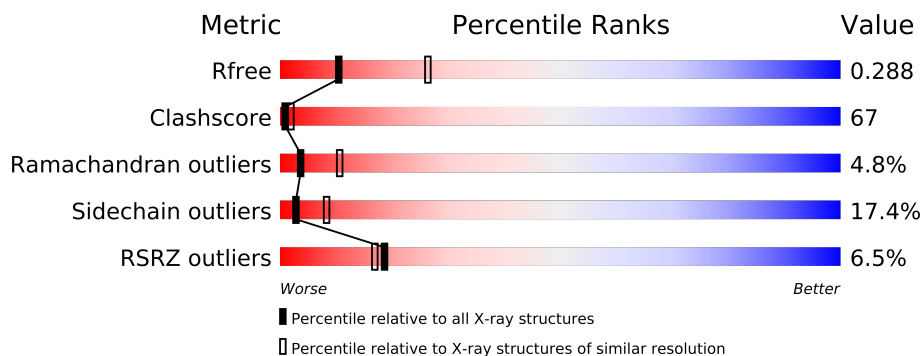
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance


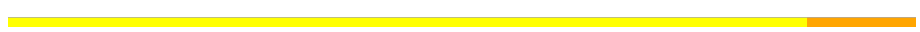
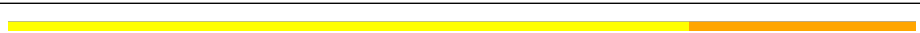
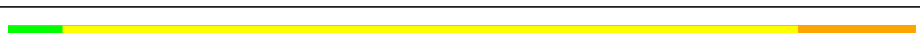
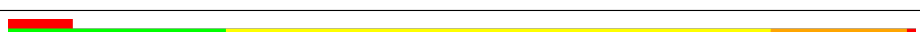


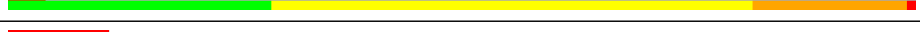
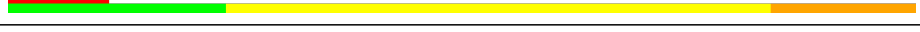
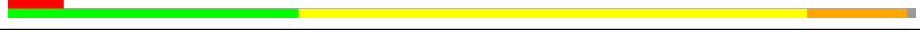

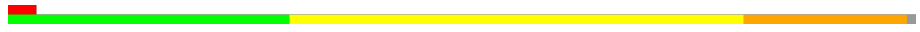
The reported resolution of this entry is 2.80 Å.

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}	66092	1799 (2.80-2.80)
Clashscore	79885	2295 (2.80-2.80)
Ramachandran outliers	78287	2252 (2.80-2.80)
Sidechain outliers	78261	2254 (2.80-2.80)
RSRZ outliers	66119	1802 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	E	16	
1	H	16	
2	B	16	
2	R	16	
3	A	332	
3	C	332	
3	G	332	
4	D	332	
5	L	88	
5	S	88	
5	T	88	
5	Y	88	

The following table lists non-polymeric compounds that are outliers for geometric or electron-

density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
6	SO4	A	946	-	X
6	SO4	C	846	-	X
6	SO4	C	947	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 14194 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA chain called 5'-D(*CP*TP*GP*AP*AP*AP*GP*CP*GP*CP*TP*AP*AP*CP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	16	Total	C	N	O	P	0	0	0
			327	156	66	90	15			
1	E	16	Total	C	N	O	P	0	0	0
			327	156	66	90	15			

- Molecule 2 is a DNA chain called 5'-D(*CP*TP*GP*TP*TP*AP*GP*CP*GP*CP*TP*TP*TP*CP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	R	16	Total	C	N	O	P	0	0	0
			323	156	54	98	15			
2	B	16	Total	C	N	O	P	0	0	0
			323	156	54	98	15			

- Molecule 3 is a protein called Glucose-resistance amylase regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	332	Total	C	N	O	Se	0	0	0
			2564	1610	438	506	10			
3	C	332	Total	C	N	O	Se	0	0	0
			2558	1606	437	505	10			
3	A	332	Total	C	N	O	Se	0	0	0
			2562	1608	437	507	10			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	1	MSE	MET	MODIFIED RESIDUE	UNP P46828
G	16	MSE	MET	MODIFIED RESIDUE	UNP P46828
G	88	MSE	MET	MODIFIED RESIDUE	UNP P46828
G	112	MSE	MET	MODIFIED RESIDUE	UNP P46828

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	123	MSE	MET	MODIFIED RESIDUE	UNP P46828
G	250	MSE	MET	MODIFIED RESIDUE	UNP P46828
G	282	MSE	MET	MODIFIED RESIDUE	UNP P46828
G	294	MSE	MET	MODIFIED RESIDUE	UNP P46828
G	302	MSE	MET	MODIFIED RESIDUE	UNP P46828
G	309	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	1	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	16	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	88	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	112	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	123	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	250	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	282	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	294	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	302	MSE	MET	MODIFIED RESIDUE	UNP P46828
C	309	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	1	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	16	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	88	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	112	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	123	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	250	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	282	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	294	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	302	MSE	MET	MODIFIED RESIDUE	UNP P46828
A	309	MSE	MET	MODIFIED RESIDUE	UNP P46828

- Molecule 4 is a protein called Glucose-resistance amylase regulator.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	332	Total	C	N	O	S	Se	0	0	0
			2558	1606	437	505	1	9			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	16	MSE	MET	MODIFIED RESIDUE	UNP P46828
D	88	MSE	MET	MODIFIED RESIDUE	UNP P46828
D	112	MSE	MET	MODIFIED RESIDUE	UNP P46828
D	123	MSE	MET	MODIFIED RESIDUE	UNP P46828
D	250	MSE	MET	MODIFIED RESIDUE	UNP P46828
D	282	MSE	MET	MODIFIED RESIDUE	UNP P46828

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	294	MSE	MET	MODIFIED RESIDUE	UNP P46828
D	302	MSE	MET	MODIFIED RESIDUE	UNP P46828
D	309	MSE	MET	MODIFIED RESIDUE	UNP P46828

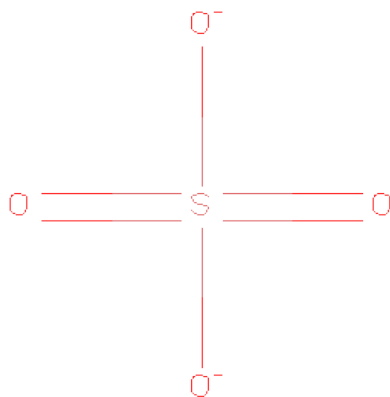
- Molecule 5 is a protein called Phosphocarrier protein HPr.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
5	T	87	Total	C	N	O	P	S	0	0	0
			632	386	104	138	1	3			
5	L	87	Total	C	N	O	P	S	0	0	0
			632	386	104	138	1	3			
5	Y	87	Total	C	N	O	P	S	0	0	0
			632	386	104	138	1	3			
5	S	87	Total	C	N	O	P	S	0	0	0
			632	386	104	138	1	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
T	46	SEP	SER	MODIFIED RESIDUE	UNP O69250
L	46	SEP	SER	MODIFIED RESIDUE	UNP O69250
Y	46	SEP	SER	MODIFIED RESIDUE	UNP O69250
S	46	SEP	SER	MODIFIED RESIDUE	UNP O69250

- Molecule 6 is SULFATE ION (three-letter code: SO₄) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	C	1	Total O S 5 4 1	0	0
6	C	1	Total O S 5 4 1	0	0
6	G	1	Total O S 5 4 1	0	0
6	G	1	Total O S 5 4 1	0	0
6	C	1	Total O S 5 4 1	0	0

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total Mg 1 1	0	0
7	D	1	Total Mg 1 1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	17	Total O 17 17	0	0
8	B	1	Total O 1 1	0	0
8	C	13	Total O 13 13	0	0
8	D	12	Total O 12 12	0	0
8	E	2	Total O 2 2	0	0
8	G	10	Total O 10 10	0	0
8	H	1	Total O 1 1	0	0
8	L	4	Total O 4 4	0	0
8	R	1	Total O 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	S	7	Total 7	O 7	0	0
8	T	9	Total 9	O 9	0	0
8	Y	10	Total 10	O 10	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: 5'-D(*CP*TP*GP*AP*AP*AP*GP*CP*GP*CP*TP*AP*AP*CP*AP*G)-3',

Chain H: 

C700
T701
G702
A703
A704
A705
G706
C707
G708
C709
T710
A711
A712
C713
A714
G715

- Molecule 1: 5'-D(*CP*TP*GP*AP*AP*AP*GP*CP*GP*CP*TP*AP*AP*CP*AP*G)-3',

Chain E: 

C700
T701
G702
A703
A704
A705
G706
C707
G708
C709
T710
A711
A712
C713
A714
G715

- Molecule 2: 5'-D(*CP*TP*GP*TP*TP*AP*GP*CP*GP*CP*TP*TP*TP*CP*AP*G)-3',

Chain R: 

C700
T701
G702
T703
T704
A705
G706
G707
G708
C709
T710
T711
T712
C713
A714
G715

- Molecule 2: 5'-D(*CP*TP*GP*TP*TP*AP*GP*CP*GP*CP*TP*TP*TP*CP*AP*G)-3',

Chain B: 

C700
T701
G702
T703
T704
A705
G706
C707
G708
C709
T710
T711
T712
C713
A714
G715

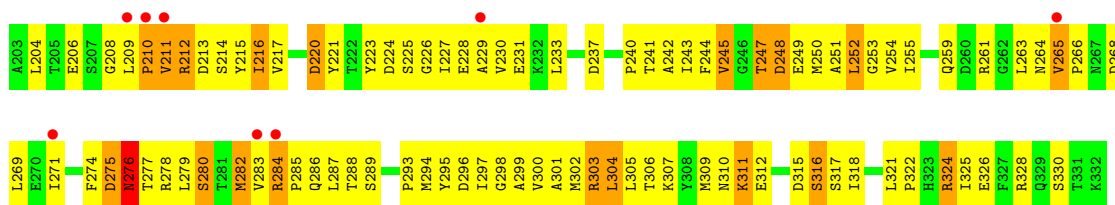
- Molecule 3: Glucose-resistance amylase regulator

Chain G: 

R1	N2	V3	T4	I5	Y6	D7	V8	A9	R10	E11	S15	R16	A17	T18	R21	V22	V23	N24	G25	N26	P27	N28	V29	R30	P31	S32	T33	R34	K35	R36	V37	L38	E39	T40	I41	E42	R43	L44	G45	Y46	R47	P48	N49	A50	V51	A52	R53	G54	L55	A56	T60	T61	T62	V63	G64	V65
----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

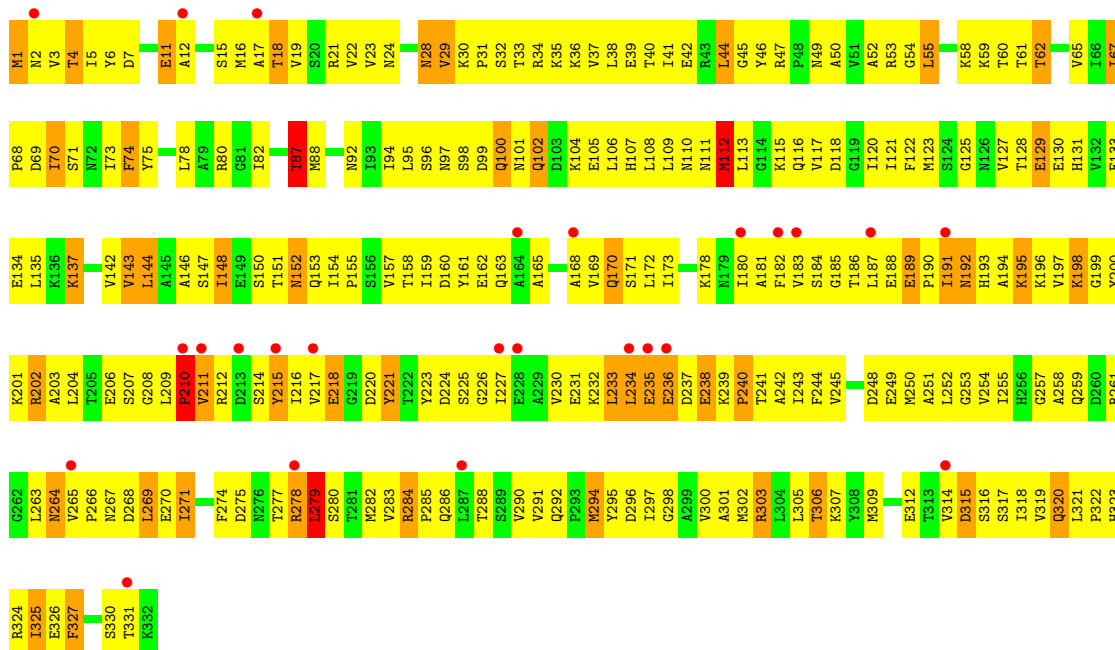
I66	I67	P68	D69	S70	N71	N72	I73	F74	A75	A76	E77	L78	A79	R80	G81	E82	E83	D84	I85	M88	Y89	K90	Y91	N92	I93	I94	L95	S96	N97	S98	D99	Q100	N101	Q102	D103	K104	L106	L107	L108	L109	N110	M111	M112	L113	G114	K115	Q116	V117	D118	K195	G119	I120	I121	F122	V127	E129
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

E130	R131	V132	L135	K136	P139	V140	P141	V142	V143	L144	A145	A146	S147	I148	E149	S150	T151	N152	Q153	I154	P155	S156	V157	T158	I159	E162	Q163	F166	V169	I173	H177	I180	V183	S184	G185	T186	L187	I191	N192	H193	A194	K195	K196	V197	K198	G199	Y200	K201	R202
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------



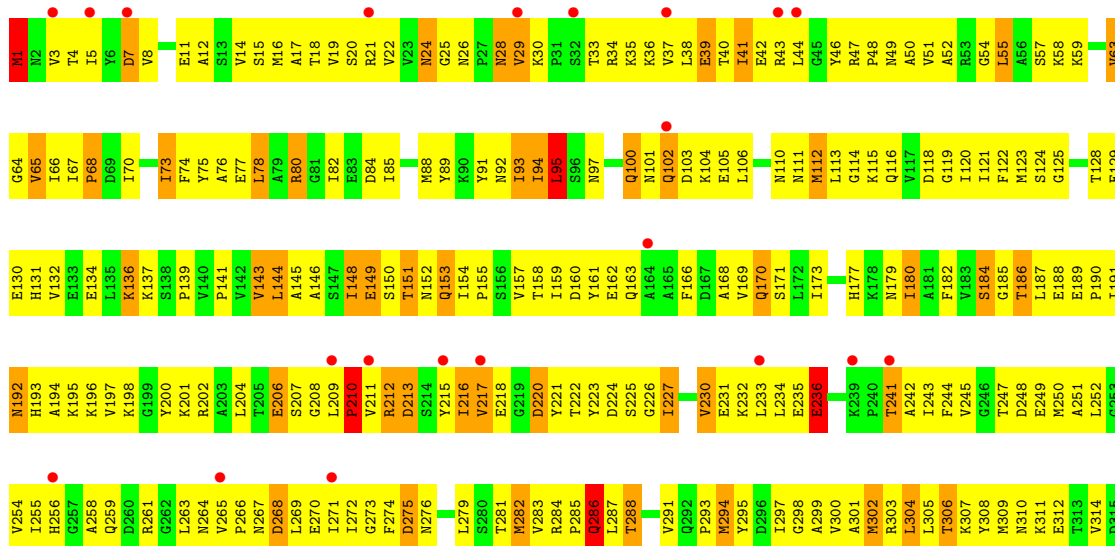
• Molecule 3: Glucose-resistance amylase regulator

Chain C:



• Molecule 3: Glucose-resistance amylase regulator

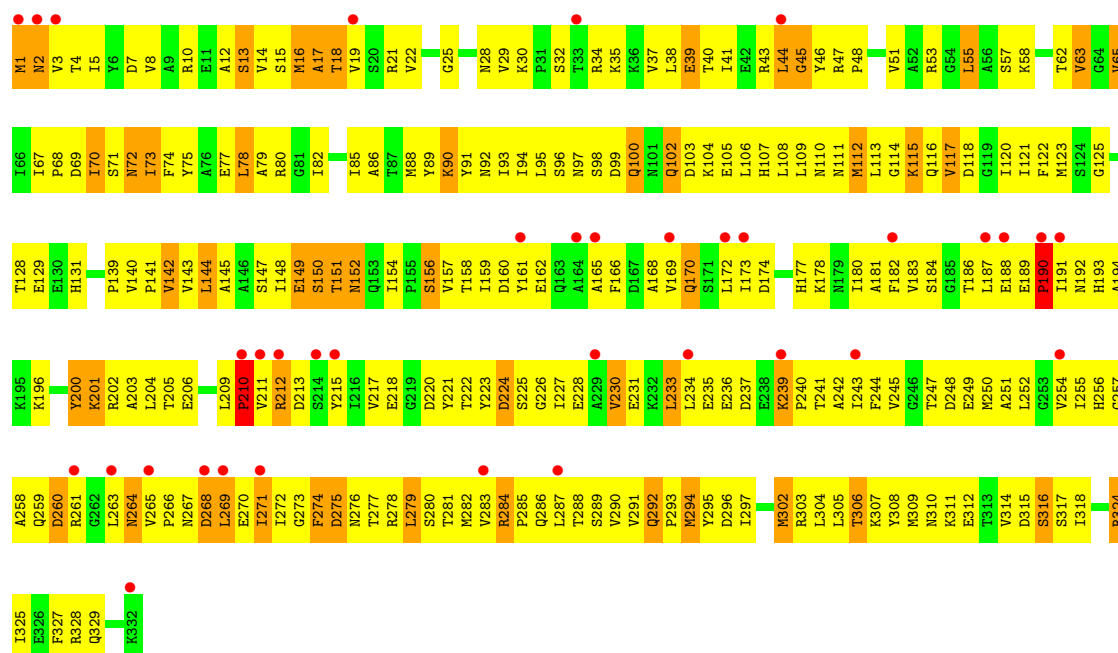
Chain A:

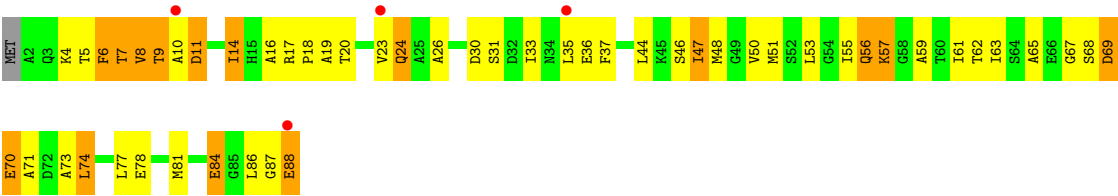




• Molecule 4: Glucose-resistance amylase regulator

Chain D:





● Molecule 5: Phosphocarrier protein HPr

Chain S:



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	105.71Å 109.24Å 117.81Å 90.00° 90.05° 90.00°	Depositor
Resolution (Å)	78.70 – 2.80 78.71 – 2.80	Depositor EDS
% Data completeness (in resolution range)	86.8 (78.70-2.80) 86.6 (78.71-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.93 (at 2.82Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.242 , 0.288 0.246 , 0.288	Depositor DCC
R_{free} test set	5845 reflections (10.17%)	DCC
Wilson B-factor (Å ²)	81.1	Xtriage
Anisotropy	0.221	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 75.7	EDS
Estimated twinning fraction	0.037 for -k,-h,-l 0.038 for k,h,-l 0.349 for h,-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	2 of 57480 reflections (0.003%)	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	14194	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.32% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: MG, SO4, SEP

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	E	0.75	0/368	0.89	0/566
1	H	1.03	3/368 (0.8%)	0.94	1/566 (0.2%)
2	B	0.70	0/360	0.95	0/554
2	R	0.74	0/360	1.00	0/554
3	A	0.73	7/2590 (0.3%)	0.85	3/3492 (0.1%)
3	C	0.68	3/2586 (0.1%)	0.85	3/3486 (0.1%)
3	G	0.73	4/2593 (0.2%)	0.93	8/3498 (0.2%)
4	D	0.66	3/2586 (0.1%)	0.83	1/3486 (0.0%)
5	L	0.51	0/625	0.75	0/839
5	S	0.54	0/625	0.76	0/839
5	T	0.53	0/625	0.75	0/839
5	Y	0.56	0/625	0.87	1/839 (0.1%)
All	All	0.69	20/14311 (0.1%)	0.86	17/19558 (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	E	0	2
1	H	0	1
2	B	0	4
2	R	0	2
4	D	0	1
5	Y	0	1
All	All	0	11

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1	MSE	SE-CE	11.43	2.62	1.95
3	A	129	GLU	CD-OE1	9.85	1.36	1.25
4	D	129	GLU	CD-OE1	9.60	1.36	1.25
3	C	129	GLU	CD-OE1	8.89	1.35	1.25
3	C	129	GLU	CD-OE2	8.75	1.35	1.25

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	711	DA	C8-N9-C4	-9.02	102.19	105.80
3	A	236	GLU	N-CA-C	-7.22	91.49	111.00
3	A	95	LEU	CA-CB-CG	6.44	130.10	115.30
3	G	3	VAL	N-CA-C	6.43	128.36	111.00
3	C	1	MSE	CB-CG-SE	-6.39	93.53	112.70

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	702	DG	Sidechain
1	E	703	DA	Sidechain
1	H	700	DC	Sidechain
2	R	712	DT	Sidechain
2	R	715	DG	Sidechain

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	327	0	180	52	0
1	H	327	0	180	56	0
2	B	323	0	184	43	0
2	R	323	0	184	57	0
3	A	2562	0	2587	389	0
3	C	2558	0	2584	386	1
3	G	2564	0	2594	374	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	2558	0	2584	376	0
5	L	632	0	624	70	0
5	S	632	0	624	69	0
5	T	632	0	624	79	0
5	Y	632	0	624	69	0
6	A	10	0	0	1	0
6	C	15	0	0	1	0
6	G	10	0	0	0	0
7	A	1	0	0	0	0
7	D	1	0	0	0	0
8	A	17	0	0	0	0
8	B	1	0	0	0	0
8	C	13	0	0	0	0
8	D	12	0	0	0	0
8	E	2	0	0	0	0
8	G	10	0	0	0	0
8	H	1	0	0	0	0
8	L	4	0	0	0	0
8	R	1	0	0	0	0
8	S	7	0	0	0	0
8	T	9	0	0	0	0
8	Y	10	0	0	0	0
All	All	14194	0	13573	1861	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 67.

The worst 5 of 1861 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:C:1:MSE:CE	3:C:1:MSE:SE	2.20	1.40
4:D:1:MET:SD	4:D:1:MET:CE	2.15	1.34
3:G:139:PRO:CG	3:C:1:MSE:HE2	1.66	1.26
3:G:73:ILE:CG2	3:C:278:ARG:HH22	1.53	1.22
3:G:73:ILE:HG22	3:C:278:ARG:NH2	1.54	1.21

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:G:198:LYS:NZ	3:C:133:GLU:OE1[2_556]	2.17	0.03

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	
3	A	330/332 (99%)	259 (78%)	56 (17%)	15 (4%)	4	12
3	C	330/332 (99%)	251 (76%)	62 (19%)	17 (5%)	3	9
3	G	330/332 (99%)	271 (82%)	40 (12%)	19 (6%)	3	7
4	D	330/332 (99%)	246 (74%)	63 (19%)	21 (6%)	2	5
5	L	84/88 (96%)	71 (84%)	12 (14%)	1 (1%)	19	54
5	S	84/88 (96%)	76 (90%)	7 (8%)	1 (1%)	19	54
5	T	84/88 (96%)	72 (86%)	8 (10%)	4 (5%)	4	10
5	Y	84/88 (96%)	73 (87%)	10 (12%)	1 (1%)	19	54
All	All	1656/1680 (99%)	1319 (80%)	258 (16%)	79 (5%)	4	10

5 of 79 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	100	GLN
3	G	102	GLN
3	G	141	PRO
3	G	150	SER
3	G	151	THR

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	
3	A	286/279 (102%)	240 (84%)	46 (16%)	3	10
3	C	285/279 (102%)	241 (85%)	44 (15%)	4	12
3	G	287/279 (103%)	235 (82%)	52 (18%)	2	7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	D	285/280 (102%)	236 (83%)	49 (17%)	3	8
5	L	66/67 (98%)	54 (82%)	12 (18%)	2	7
5	S	66/67 (98%)	53 (80%)	13 (20%)	2	6
5	T	66/67 (98%)	51 (77%)	15 (23%)	1	3
5	Y	66/67 (98%)	52 (79%)	14 (21%)	1	4
All	All	1407/1385 (102%)	1162 (83%)	245 (17%)	3	8

5 of 245 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	102	GLN
3	A	304	LEU
5	Y	56	GLN
3	A	144	LEU
3	A	207	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	97	ASN
3	A	193	HIS
5	L	56	GLN
3	A	152	ASN
3	A	256	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z > 2$	Counts	RMSZ	# $ Z > 2$
5	SEP	L	46	5	9,9,10	6.36	2 (22%)	10,12,14	1.52	1 (10%)
5	SEP	S	46	5	9,9,10	6.25	3 (33%)	10,12,14	1.53	1 (10%)
5	SEP	T	46	5	9,9,10	6.35	3 (33%)	10,12,14	1.69	4 (40%)
5	SEP	Y	46	5	9,9,10	5.94	3 (33%)	10,12,14	1.43	2 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	SEP	L	46	5	-	0/6/8/10	0/0/0/0
5	SEP	S	46	5	-	0/6/8/10	0/0/0/0
5	SEP	T	46	5	-	0/6/8/10	0/0/0/0
5	SEP	Y	46	5	-	0/6/8/10	0/0/0/0

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	L	46	SEP	O-C	18.72	1.24	1.11
5	T	46	SEP	O-C	18.54	1.24	1.11
5	S	46	SEP	O-C	18.20	1.23	1.11
5	Y	46	SEP	O-C	17.31	1.23	1.11
5	S	46	SEP	CA-C	3.28	1.54	1.48

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	S	46	SEP	O3P-P-O1P	3.71	122.57	110.44
5	T	46	SEP	O3P-P-O1P	3.31	121.27	110.44
5	L	46	SEP	O3P-P-O1P	3.19	120.88	110.44
5	Y	46	SEP	O3P-P-O1P	2.93	120.02	110.44
5	T	46	SEP	C-CA-N	2.47	116.29	113.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
6	SO4	A	599	-	4,4,4	0.37	0	6,6,6	0.13	0
6	SO4	A	946	-	4,4,4	0.65	0	6,6,6	0.58	0
6	SO4	C	346	-	4,4,4	0.31	0	6,6,6	0.12	0
6	SO4	C	846	-	4,4,4	0.28	0	6,6,6	0.23	0
6	SO4	C	947	-	4,4,4	0.19	0	6,6,6	0.20	0
6	SO4	G	646	-	4,4,4	0.32	0	6,6,6	0.12	0
6	SO4	G	647	-	4,4,4	0.21	0	6,6,6	0.17	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	SO4	A	599	-	-	0/0/0/0	0/0/0/0
6	SO4	A	946	-	-	0/0/0/0	0/0/0/0
6	SO4	C	346	-	-	0/0/0/0	0/0/0/0
6	SO4	C	846	-	-	0/0/0/0	0/0/0/0
6	SO4	C	947	-	-	0/0/0/0	0/0/0/0
6	SO4	G	646	-	-	0/0/0/0	0/0/0/0
6	SO4	G	647	-	-	0/0/0/0	0/0/0/0

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.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9	
1	E	16/16 (100%)	-0.25	0	100	100	49, 75, 90, 91	0
1	H	16/16 (100%)	-0.19	0	100	100	52, 78, 93, 111	0
2	B	16/16 (100%)	-0.27	0	100	100	54, 79, 95, 104	0
2	R	16/16 (100%)	-0.17	0	100	100	47, 77, 94, 94	0
3	A	332/332 (100%)	0.45	23 (6%)	17	15	34, 79, 119, 141	0
3	C	332/332 (100%)	0.58	25 (7%)	14	12	33, 85, 119, 135	0
3	G	332/332 (100%)	0.43	12 (3%)	41	41	28, 70, 107, 150	0
4	D	332/332 (100%)	0.55	36 (10%)	6	5	33, 90, 129, 148	0
5	L	87/88 (98%)	0.34	5 (5%)	23	23	43, 79, 102, 127	0
5	S	87/88 (98%)	0.25	4 (4%)	31	31	43, 73, 96, 109	0
5	T	87/88 (98%)	0.32	3 (3%)	43	44	54, 75, 96, 103	0
5	Y	87/88 (98%)	0.15	4 (4%)	31	31	54, 75, 89, 111	0
All	All	1740/1744 (99%)	0.43	112 (6%)	18	17	28, 79, 118, 150	0

The worst 5 of 112 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	44	LEU	7.0
3	C	210	PRO	5.0
4	D	211	VAL	4.9
3	G	265	VAL	4.5
4	D	215	TYR	4.5

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
5	SEP	S	46	10/11	0.20	0.29	48,58,66,69	0
5	SEP	T	46	10/11	0.18	-0.36	54,66,74,75	0
5	SEP	L	46	10/11	0.20	-0.48	49,59,66,72	0
5	SEP	Y	46	10/11	0.18	-0.58	61,66,80,84	0

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	SO4	C	947	5/5	0.50	6.46	133,136,138,141	0
6	SO4	C	846	5/5	0.37	4.22	94,98,99,102	0
6	SO4	A	946	5/5	0.28	2.35	110,111,114,115	0
6	SO4	A	599	5/5	0.27	1.70	109,114,116,116	0
6	SO4	G	646	5/5	0.33	-0.04	139,141,143,146	0
6	SO4	C	346	5/5	0.17	-0.82	113,119,120,122	0
6	SO4	G	647	5/5	0.14	-1.14	120,123,124,127	0
7	MG	D	754	1/1	0.09	-2.38	74,74,74,74	0
7	MG	A	704	1/1	0.07	-2.42	61,61,61,61	0

6.5 Other polymers ⓘ

There are no such residues in this entry.