W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 28, 2023 – 01:24 PM EDT

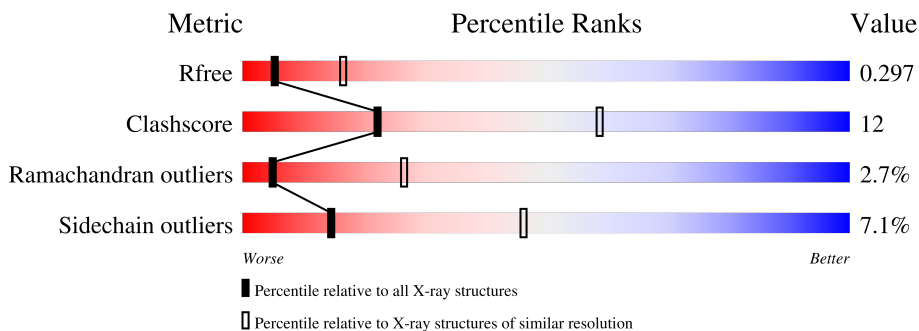|		
PDB ID	:	3K1P
Title	:	Crystal Structure of full-length BenM E226K mutant
Authors	:	Ruangprasert, A.; Momany, C.; Neidle, E.L.; Craven, S.H.
Deposited on	:	2009-09-28
Resolution	:	3.00 Å(reported)

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	Percentile Ranks	Value
Rfree		0.297
Clashscore		12
Ramachandran outliers		2.7%
Sidechain outliers		7.1%

Worse   *Better*

▮ Percentile relative to all X-ray structures

⯐ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (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 <=5%

Mol	Chain	Length	Quality of chain
1	A	312	65% 28% · ·
1	B	312	70% 25% · ·

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 5235 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 HTH-type transcriptional regulator benM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	1	0
			2424	1554	418	443	9			
1	B	303	Total	C	N	O	S	0	3	0
			2438	1562	422	445	9			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	226	LYS	GLU	engineered mutation	UNP O68014
A	305	LEU	-	expression tag	UNP O68014
A	306	GLU	-	expression tag	UNP O68014
A	307	HIS	-	expression tag	UNP O68014
A	308	HIS	-	expression tag	UNP O68014
A	309	HIS	-	expression tag	UNP O68014
A	310	HIS	-	expression tag	UNP O68014
A	311	HIS	-	expression tag	UNP O68014
A	312	HIS	-	expression tag	UNP O68014
B	226	LYS	GLU	engineered mutation	UNP O68014
B	305	LEU	-	expression tag	UNP O68014
B	306	GLU	-	expression tag	UNP O68014
B	307	HIS	-	expression tag	UNP O68014
B	308	HIS	-	expression tag	UNP O68014
B	309	HIS	-	expression tag	UNP O68014
B	310	HIS	-	expression tag	UNP O68014
B	311	HIS	-	expression tag	UNP O68014
B	312	HIS	-	expression tag	UNP O68014

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	196	Total	O	0	0
			196	196		

Continued on next page...

Continued from previous page...

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	177	Total 177	O 177	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. 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. 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: HTH-type transcriptional regulator benM
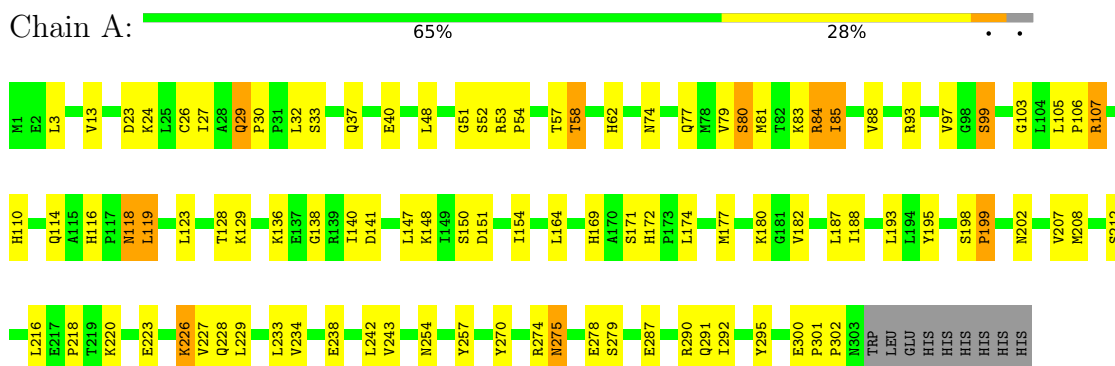
- Molecule 1: HTH-type transcriptional regulator benM
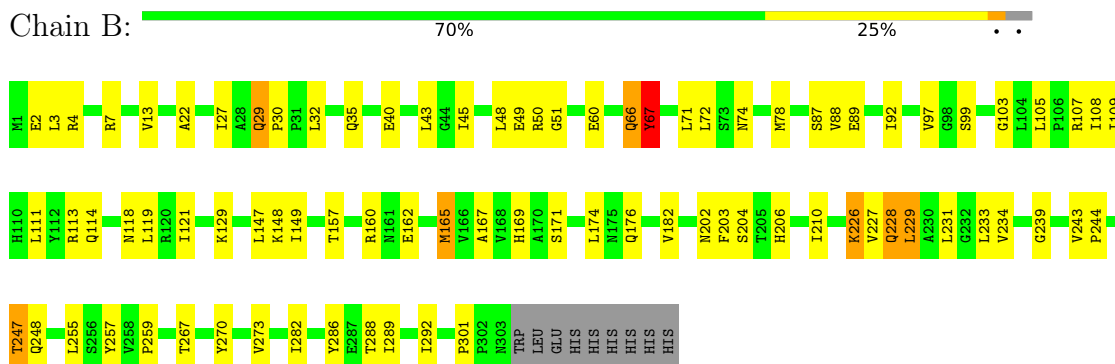
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	70.60Å 70.70Å 187.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.27 – 3.00 48.27 – 3.00	Depositor EDS
% Data completeness (in resolution range)	90.7 (48.27-3.00) 98.0 (48.27-3.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.14 (at 3.01Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.197 , 0.226 0.269 , 0.297	Depositor DCC
R_{free} test set	1011 reflections (5.21%)	wwPDB-VP
Wilson B-factor (Å^2)	61.4	Xtriage
Anisotropy	0.236	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 30.3	EDS
L-test for twinning[2]	$< |L| > = 0.52, < L^2 > = 0.36$	Xtriage
Estimated twinning fraction	0.167 for k,h,-l	Xtriage
Reported twinning fraction	0.598 for H, K, L 0.402 for K, H, -L	Depositor
Outliers	1 of 19392 reflections (0.005%)	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	5235	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.40	0/2477	0.57	0/3355
1	B	0.40	0/2494	0.56	0/3377
All	All	0.40	0/4971	0.56	0/6732

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2424	0	2486	65	0
1	B	2438	0	2503	53	0
2	A	196	0	0	14	0
2	B	177	0	0	6	0
All	All	5235	0	4989	115	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 12.

The worst 5 of 115 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:226:LYS:HD3	1:A:227:VAL:H	1.29	0.97

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:198:SER:HB2	1:A:199:PRO:HD2	1.47	0.94
1:A:118:ASN:HA	2:A:356:HOH:O	1.68	0.94
1:B:226:LYS:HD2	1:B:227:VAL:H	1.33	0.93
1:B:66:GLN:O	1:B:67:TYR:HB3	1.78	0.83

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	302/312 (97%)	268 (89%)	23 (8%)	11 (4%)	3	19
1	B	304/312 (97%)	265 (87%)	33 (11%)	6 (2%)	7	34
All	All	606/624 (97%)	533 (88%)	56 (9%)	17 (3%)	5	25

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	129	LYS
1	A	24	LYS
1	A	26	CYS
1	A	40	GLU
1	A	80	SER

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	268/276 (97%)	246 (92%)	22 (8%)	11	39
1	B	270/276 (98%)	253 (94%)	17 (6%)	18	51
All	All	538/552 (98%)	499 (93%)	39 (7%)	14	45

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	114	GLN
1	B	228	GLN
1	B	147	LEU
1	B	165	MET
1	B	247	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	29	GLN
1	B	35	GLN
1	B	169	HIS
1	B	66	GLN
1	B	74	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.