



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 05:30 am GMT

PDB ID : 2AXU  
Title : Structure of PrgX  
Authors : Shi, K.; Brown, C.K.; Gu, Z.Y.; Kozlowicz, B.K.; Dunny, G.M.; Ohlendorf, D.H.; Earhart, C.A.  
Deposited on : 2005-09-06  
Resolution : 2.90 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

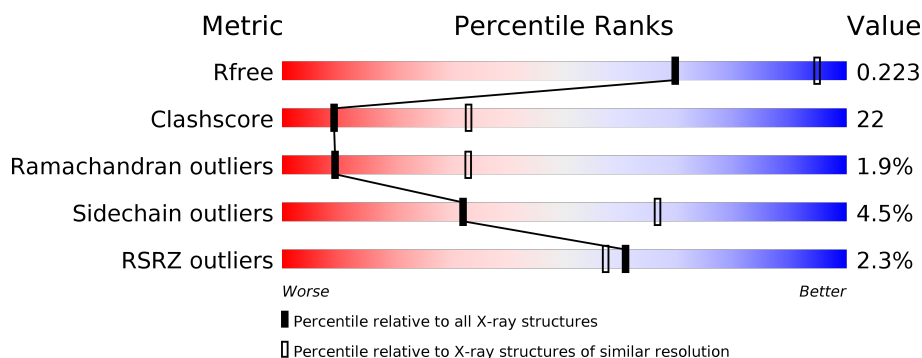
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
$R_{free}$	100719	1586 (2.90-2.90)
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	317	
1	B	317	
1	C	317	
1	D	317	
1	E	317	
1	F	317	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	317	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>2%</div><div>50%</div><div>38%</div><div>7%</div><div></div></div>
1	H	317	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>%</div><div>54%</div><div>37%</div><div></div><div>6%</div></div>
1	I	317	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>6%</div><div>49%</div><div>41%</div><div></div><div>5%</div></div>
1	J	317	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>4%</div><div>53%</div><div>38%</div><div></div><div>7%</div></div>
1	K	317	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>3%</div><div>56%</div><div>36%</div><div></div><div>6%</div></div>
1	L	317	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>%</div><div>54%</div><div>37%</div><div></div><div>6%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 30147 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 PrgX.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	Se	0	0	0
			2507	1630	402	469	1	5			
1	B	300	Total	C	N	O	S	Se	0	0	0
			2486	1616	399	466	1	4			
1	C	299	Total	C	N	O	S	Se	0	0	0
			2475	1607	398	465	1	4			
1	D	298	Total	C	N	O	S	Se	0	0	0
			2468	1602	397	464	1	4			
1	E	300	Total	C	N	O	S	Se	0	0	0
			2480	1610	399	466	1	4			
1	F	301	Total	C	N	O	S	Se	0	0	0
			2488	1616	400	467	1	4			
1	G	303	Total	C	N	O	S	Se	0	0	0
			2508	1630	402	471	1	4			
1	H	298	Total	C	N	O	S	Se	0	0	0
			2468	1602	397	464	1	4			
1	I	300	Total	C	N	O	S	Se	0	0	0
			2486	1616	399	466	1	4			
1	J	296	Total	C	N	O	S	Se	0	0	0
			2447	1587	394	461	1	4			
1	K	299	Total	C	N	O	S	Se	0	0	0
			2471	1604	397	465	1	4			
1	L	299	Total	C	N	O	S	Se	0	0	0
			2479	1611	398	465	1	4			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
A	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
A	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
A	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
A	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
B	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
B	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
B	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
B	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
C	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
C	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
C	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
C	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
D	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
D	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
D	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
D	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
E	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
E	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
E	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
E	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
E	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
F	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
F	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
F	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
F	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
F	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
G	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
G	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
G	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
G	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
G	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
H	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
H	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
H	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
H	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
H	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
I	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
I	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
I	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
I	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
I	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
J	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
J	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
J	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
J	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
J	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
K	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
K	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
K	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
K	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
K	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114
L	1	MSE	MET	MODIFIED RESIDUE	UNP Q04114
L	27	MSE	MET	MODIFIED RESIDUE	UNP Q04114
L	67	MSE	MET	MODIFIED RESIDUE	UNP Q04114
L	175	MSE	MET	MODIFIED RESIDUE	UNP Q04114
L	203	MSE	MET	MODIFIED RESIDUE	UNP Q04114

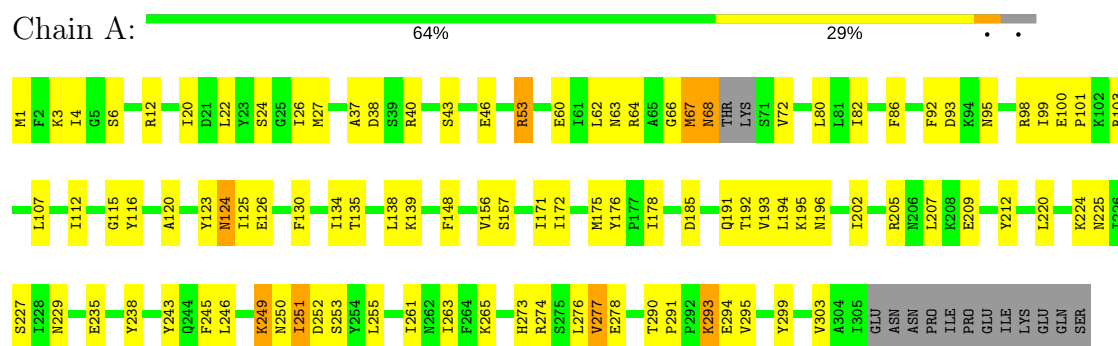
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	37	Total O 37 37	0	0
2	B	45	Total O 45 45	0	0
2	C	25	Total O 25 25	0	0
2	D	29	Total O 29 29	0	0
2	E	33	Total O 33 33	0	0
2	F	42	Total O 42 42	0	0
2	G	28	Total O 28 28	0	0
2	H	36	Total O 36 36	0	0
2	I	23	Total O 23 23	0	0
2	J	28	Total O 28 28	0	0
2	K	28	Total O 28 28	0	0
2	L	30	Total O 30 30	0	0

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

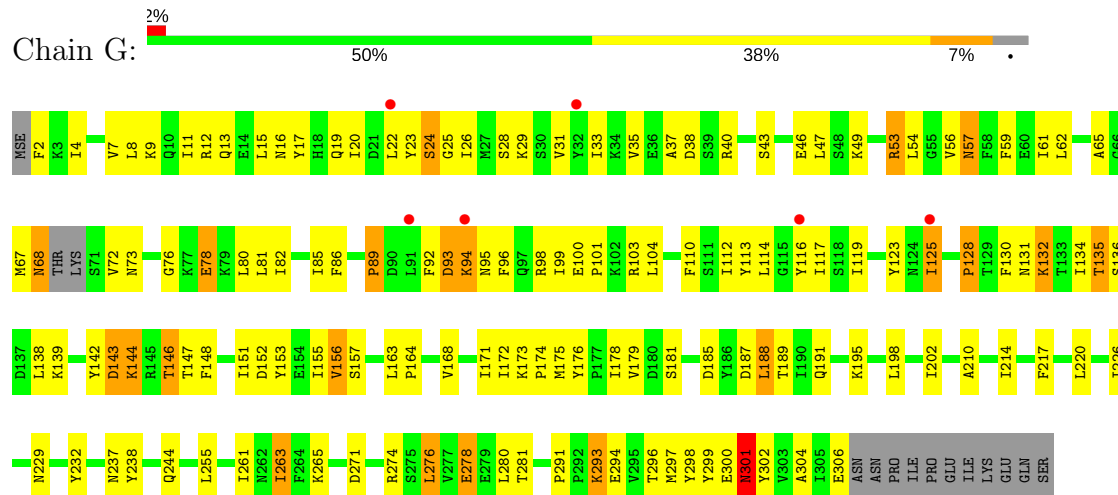
#### • Molecule 1: PrgX



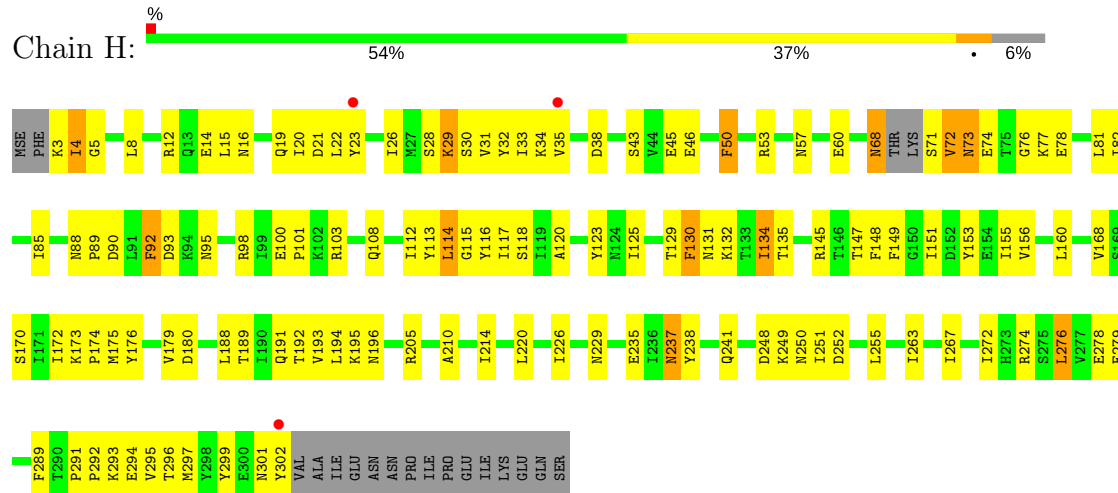




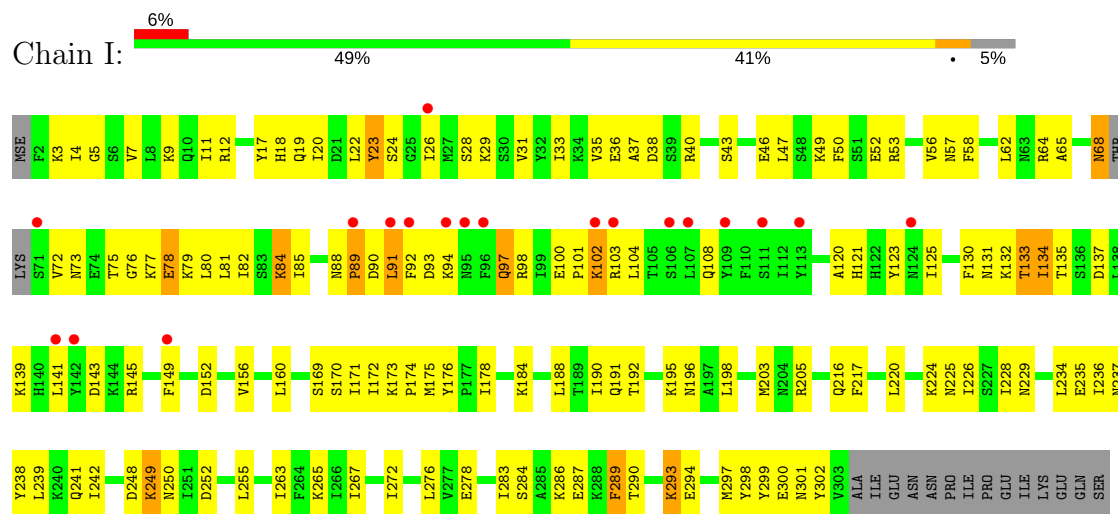
- Molecule 1: PrgX



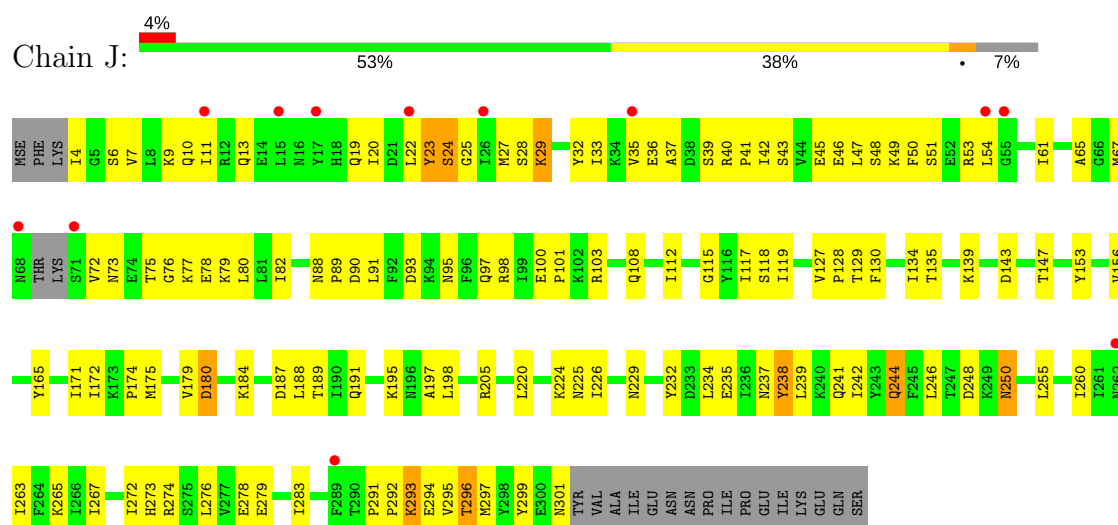
- Molecule 1: PrgX



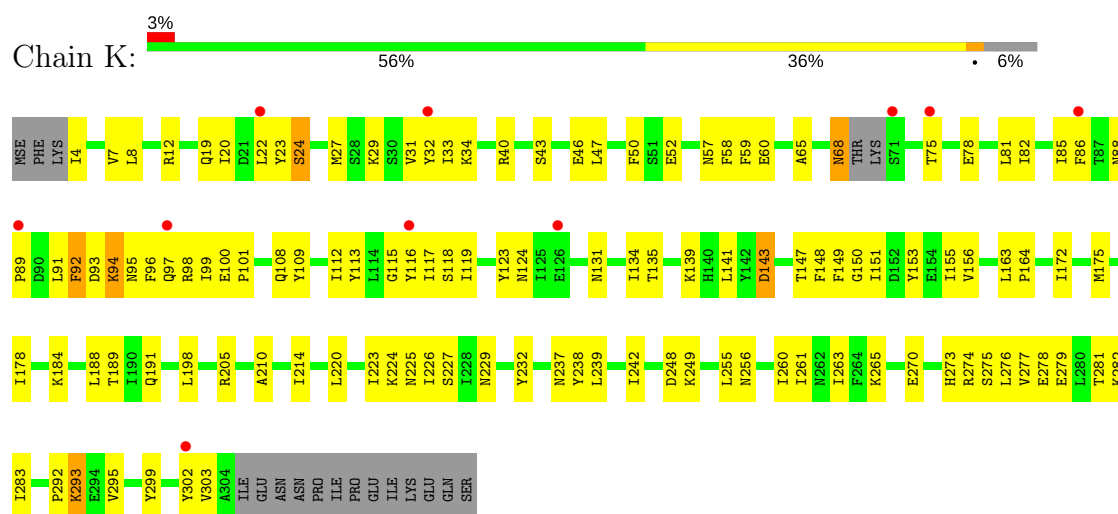
- Molecule 1: PrgX



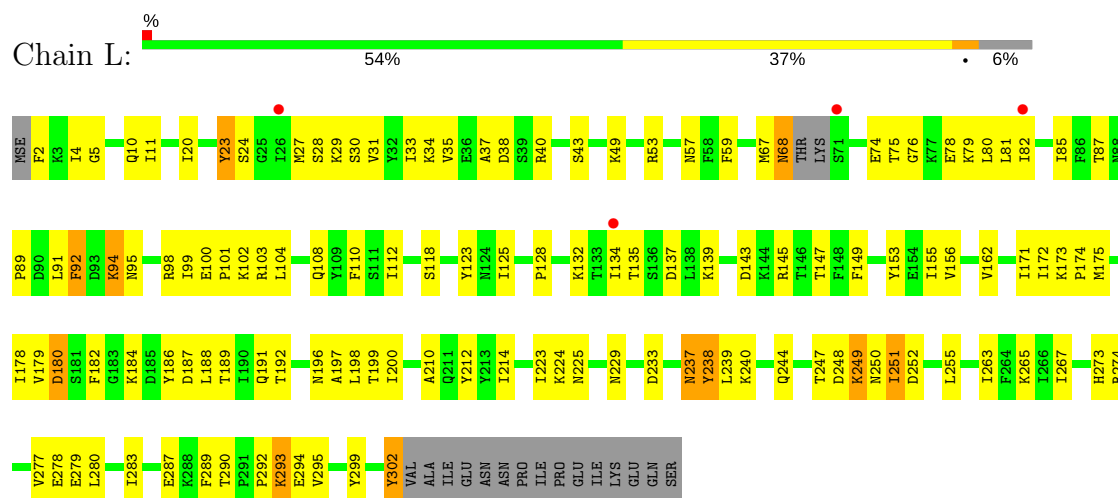
- Molecule 1: PrgX



• Molecule 1: PrgX



• Molecule 1: PrgX



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.65Å 134.72Å 195.93Å 90.00° 100.30° 90.00°	Depositor
Resolution (Å)	15.00 – 2.90 49.53 – 2.89	Depositor EDS
% Data completeness (in resolution range)	(Not available) (15.00-2.90) 99.1 (49.53-2.89)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.50 (at 2.91Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.199 , 0.266 0.197 , 0.223	Depositor DCC
$R_{free}$ test set	10465 reflections (10.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	52.1	Xtriage
Anisotropy	0.357	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 70.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	30147	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.11% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

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	0.31	0/2554	0.54	0/3442
1	B	0.31	0/2533	0.55	0/3414
1	C	0.29	0/2521	0.52	0/3398
1	D	0.30	0/2514	0.54	0/3388
1	E	0.30	0/2526	0.53	0/3405
1	F	0.30	0/2534	0.56	1/3416 (0.0%)
1	G	0.42	0/2555	0.59	0/3444
1	H	0.31	0/2514	0.55	0/3388
1	I	0.29	0/2533	0.51	0/3414
1	J	0.29	0/2492	0.53	0/3359
1	K	0.30	0/2517	0.53	0/3394
1	L	0.30	0/2526	0.53	0/3404
All	All	0.31	0/30319	0.54	1/40866 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	302	TYR	N-CA-C	-5.26	96.79	111.00

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	2507	0	2543	89	0
1	B	2486	0	2515	109	0
1	C	2475	0	2506	118	0
1	D	2468	0	2497	91	0
1	E	2480	0	2511	96	0
1	F	2488	0	2522	125	0
1	G	2508	0	2537	153	0
1	H	2468	0	2497	132	0
1	I	2486	0	2515	144	0
1	J	2447	0	2475	115	0
1	K	2471	0	2498	99	0
1	L	2479	0	2506	115	0
2	A	37	0	0	11	0
2	B	45	0	0	6	0
2	C	25	0	0	11	0
2	D	29	0	0	6	0
2	E	33	0	0	6	0
2	F	42	0	0	6	0
2	G	28	0	0	19	0
2	H	36	0	0	5	0
2	I	23	0	0	10	0
2	J	28	0	0	5	0
2	K	28	0	0	12	0
2	L	30	0	0	6	0
All	All	30147	0	30122	1323	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 22.

The worst 5 of 1323 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:156:VAL:HG21	1:B:175:MSE:HE1	1.21	1.19
1:I:93:ASP:HB3	1:I:97:GLN:HE22	1.21	1.03
1:H:4:ILE:HG21	1:H:38:ASP:HA	1.38	1.01
1:C:263:ILE:HG13	1:D:263:ILE:HG13	1.42	1.01
1:B:20:ILE:HD12	1:B:20:ILE:H	1.24	0.99

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	299/317 (94%)	273 (91%)	23 (8%)	3 (1%)	18	51
1	B	296/317 (93%)	279 (94%)	15 (5%)	2 (1%)	25	60
1	C	295/317 (93%)	251 (85%)	32 (11%)	12 (4%)	3	13
1	D	294/317 (93%)	275 (94%)	14 (5%)	5 (2%)	11	36
1	E	296/317 (93%)	271 (92%)	20 (7%)	5 (2%)	11	36
1	F	297/317 (94%)	266 (90%)	23 (8%)	8 (3%)	6	23
1	G	299/317 (94%)	257 (86%)	32 (11%)	10 (3%)	4	18
1	H	294/317 (93%)	268 (91%)	18 (6%)	8 (3%)	6	23
1	I	296/317 (93%)	257 (87%)	33 (11%)	6 (2%)	9	31
1	J	292/317 (92%)	259 (89%)	30 (10%)	3 (1%)	18	51
1	K	295/317 (93%)	274 (93%)	17 (6%)	4 (1%)	13	41
1	L	295/317 (93%)	269 (91%)	24 (8%)	2 (1%)	25	60
All	All	3548/3804 (93%)	3199 (90%)	281 (8%)	68 (2%)	9	33

5 of 68 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	124	ASN
1	C	72	VAL
1	C	88	ASN
1	C	249	LYS
1	D	88	ASN

#### 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	285/294 (97%)	273 (96%)	12 (4%)	34	69
1	B	283/294 (96%)	273 (96%)	10 (4%)	41	75
1	C	282/294 (96%)	270 (96%)	12 (4%)	33	68
1	D	281/294 (96%)	270 (96%)	11 (4%)	37	72
1	E	282/294 (96%)	267 (95%)	15 (5%)	26	60
1	F	283/294 (96%)	272 (96%)	11 (4%)	37	72
1	G	285/294 (97%)	264 (93%)	21 (7%)	16	42
1	H	281/294 (96%)	270 (96%)	11 (4%)	37	72
1	I	283/294 (96%)	273 (96%)	10 (4%)	41	75
1	J	279/294 (95%)	265 (95%)	14 (5%)	28	62
1	K	281/294 (96%)	270 (96%)	11 (4%)	37	72
1	L	282/294 (96%)	269 (95%)	13 (5%)	31	65
All	All	3387/3528 (96%)	3236 (96%)	151 (4%)	32	66

5 of 151 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	279	GLU
1	G	189	THR
1	L	23	TYR
1	F	302	TYR
1	G	132	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 151 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	131	ASN
1	G	122	HIS
1	L	88	ASN
1	F	161	ASN
1	F	244	GLN

### 5.3.3 RNA ⓘ

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 carbohydrates 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 ⓘ

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	298/317 (94%)	-0.28	0 100 100	19, 42, 99, 133	0
1	B	296/317 (93%)	-0.25	1 (0%) 93 93	15, 42, 95, 193	0
1	C	295/317 (93%)	-0.02	12 (4%) 38 32	19, 60, 121, 160	0
1	D	294/317 (92%)	-0.07	6 (2%) 65 62	19, 56, 110, 189	0
1	E	296/317 (93%)	-0.20	3 (1%) 82 81	16, 47, 98, 132	0
1	F	297/317 (93%)	-0.11	5 (1%) 70 68	18, 53, 116, 167	0
1	G	299/317 (94%)	0.01	6 (2%) 65 62	23, 57, 121, 159	0
1	H	294/317 (92%)	-0.10	3 (1%) 82 81	18, 52, 111, 169	0
1	I	296/317 (93%)	0.10	19 (6%) 20 15	28, 66, 134, 169	0
1	J	292/317 (92%)	0.01	12 (4%) 38 32	24, 56, 119, 171	0
1	K	295/317 (93%)	-0.01	10 (3%) 46 39	17, 51, 112, 169	0
1	L	295/317 (93%)	-0.09	4 (1%) 75 74	16, 52, 103, 149	0
All	All	3547/3804 (93%)	-0.09	81 (2%) 61 57	15, 53, 114, 193	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	68	ASN	4.5
1	J	71	SER	4.2
1	E	71	SER	4.1
1	C	120	ALA	4.0
1	I	92	PHE	4.0

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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.