



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2024 – 06:03 PM EST

PDB ID : 3L7P  
Title : Crystal structure of SMU.1657c, Putative nitrogen regulatory protein PII from streptococcus mutans  
Authors : Fan, X.-X.; Wang, K.-T.; Su, X.-D.  
Deposited on : 2009-12-29  
Resolution : 2.00 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

<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 : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

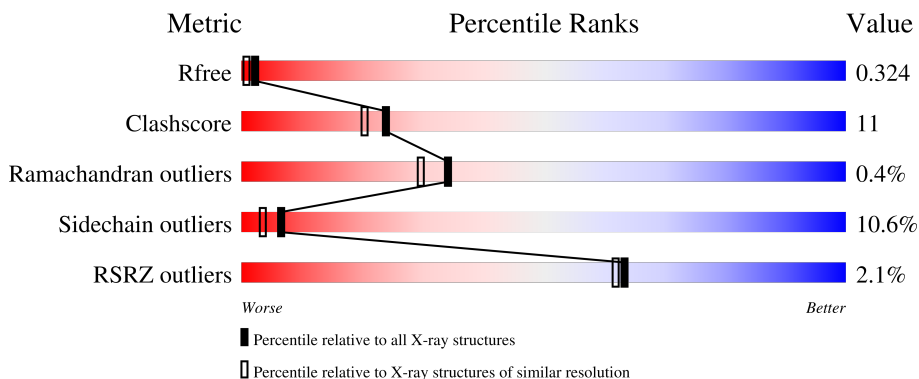
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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.



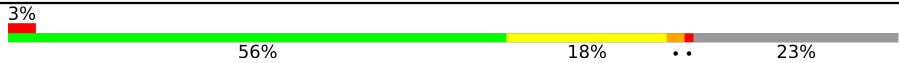
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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  $\geq 3$ , 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	115	 55% 16% 28% 3%
1	B	115	 53% 18% 23% 6%
1	C	115	 59% 19% 21% 1%
1	D	115	 57% 20% 22% 1%
1	E	115	 63% 9% 26% 2%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	115	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into four segments: a small red segment (3%), a large green segment (56%), a yellow segment (18%), and a grey segment (23%). There are two small black dots on the grey segment.</p>

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 4097 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 Putative nitrogen regulatory protein PII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	83	Total 625	C 403	N 103	O 116	S 3	0	0	0
1	B	88	Total 655	C 421	N 108	O 123	S 3	0	0	0
1	C	91	Total 685	C 436	N 115	O 131	S 3	0	0	0
1	D	90	Total 676	C 431	N 112	O 130	S 3	0	0	0
1	E	85	Total 636	C 409	N 105	O 119	S 3	0	0	0
1	F	88	Total 659	C 422	N 108	O 126	S 3	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q8DSV2
A	0	SER	-	expression tag	UNP Q8DSV2
B	-1	GLY	-	expression tag	UNP Q8DSV2
B	0	SER	-	expression tag	UNP Q8DSV2
C	-1	GLY	-	expression tag	UNP Q8DSV2
C	0	SER	-	expression tag	UNP Q8DSV2
D	-1	GLY	-	expression tag	UNP Q8DSV2
D	0	SER	-	expression tag	UNP Q8DSV2
E	-1	GLY	-	expression tag	UNP Q8DSV2
E	0	SER	-	expression tag	UNP Q8DSV2
F	-1	GLY	-	expression tag	UNP Q8DSV2
F	0	SER	-	expression tag	UNP Q8DSV2

- Molecule 2 is water.

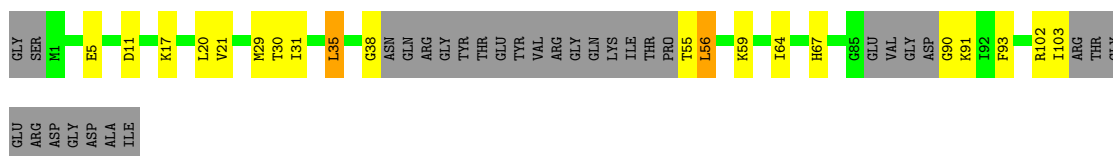
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	27	Total O 27 27	0	0
2	B	25	Total O 25 25	0	0
2	C	28	Total O 28 28	0	0
2	D	24	Total O 24 24	0	0
2	E	25	Total O 25 25	0	0
2	F	32	Total O 32 32	0	0

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

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.

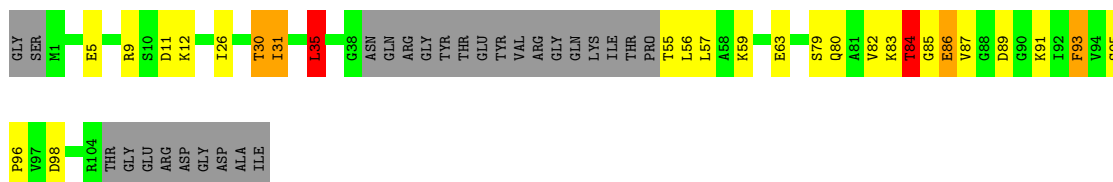
- Molecule 1: Putative nitrogen regulatory protein PII

Chain A: 



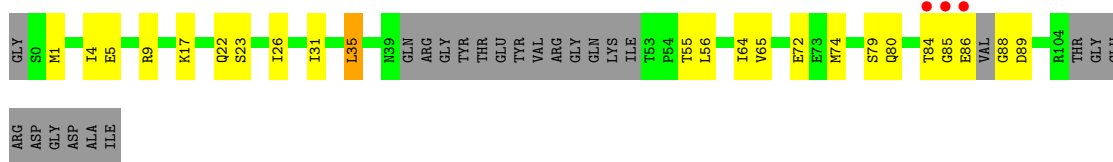
- Molecule 1: Putative nitrogen regulatory protein PII

Chain B: 



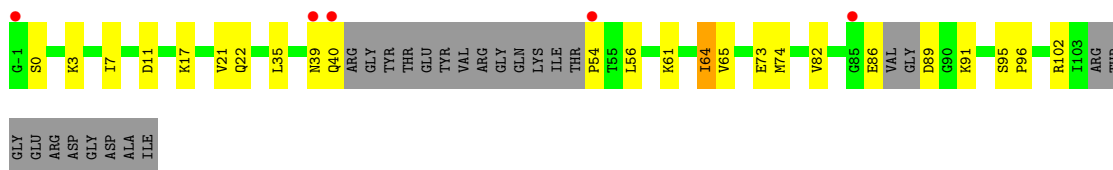
- Molecule 1: Putative nitrogen regulatory protein PII

Chain C: 



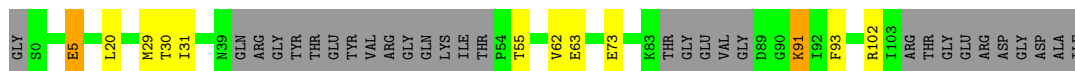
- Molecule 1: Putative nitrogen regulatory protein PII

Chain D: 



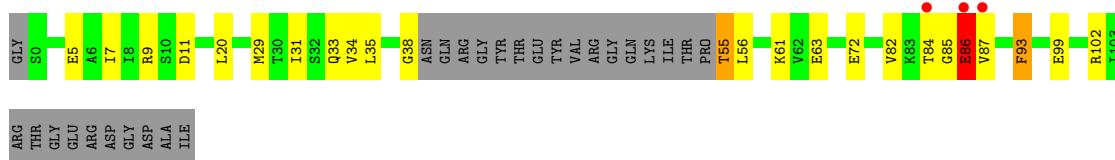
- Molecule 1: Putative nitrogen regulatory protein PII

Chain E:  63% 9% 26%



- Molecule 1: Putative nitrogen regulatory protein PII

Chain F:  3% 56% 18% 23%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.80Å 60.80Å 147.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.90 – 2.00 46.90 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.90-2.00) 99.7 (46.90-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.56 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.258 , 0.327 0.261 , 0.324	Depositor DCC
$R_{free}$ test set	1792 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.4	Xtrriage
Anisotropy	0.014	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 42.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.039 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	4097	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.38 % 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 6.8817e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	1.01	0/627	0.99	2/838 (0.2%)
1	B	0.97	0/658	1.04	4/882 (0.5%)
1	C	0.99	1/688 (0.1%)	0.97	0/921
1	D	1.03	1/679 (0.1%)	1.01	1/908 (0.1%)
1	E	1.07	1/639 (0.2%)	0.98	2/856 (0.2%)
1	F	1.03	0/662	1.01	3/887 (0.3%)
All	All	1.02	3/3953 (0.1%)	1.00	12/5292 (0.2%)

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	F	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	5	GLU	CB-CG	-5.49	1.41	1.52
1	C	72	GLU	CG-CD	5.19	1.59	1.51
1	D	73	GLU	CB-CG	5.00	1.61	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	35	LEU	CB-CG-CD1	-6.97	99.15	111.00
1	F	11	ASP	CB-CG-OD2	-6.65	112.31	118.30
1	B	98	ASP	CB-CG-OD2	6.11	123.80	118.30
1	A	11	ASP	CB-CG-OD1	6.11	123.80	118.30
1	B	35	LEU	CB-CG-CD2	-5.75	101.23	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	85	GLY	Peptide
1	F	86	GLU	Peptide

## 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	625	0	682	13	0
1	B	655	0	705	33	0
1	C	685	0	733	10	0
1	D	676	0	722	18	0
1	E	636	0	682	4	0
1	F	659	0	710	21	0
2	A	27	0	0	4	0
2	B	25	0	0	3	0
2	C	28	0	0	1	0
2	D	24	0	0	2	0
2	E	25	0	0	0	0
2	F	32	0	0	3	0
All	All	4097	0	4234	91	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 11.

The worst 5 of 91 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:86:GLU:HG3	1:D:89:ASP:HB2	1.24	1.14
1:C:85:GLY:O	1:C:86:GLU:HB2	1.67	0.93
1:B:9:ARG:CZ	1:B:84:THR:HG22	2.05	0.86
1:D:86:GLU:HG3	1:D:89:ASP:CB	2.05	0.85
1:F:82:VAL:HG13	2:F:130:HOH:O	1.78	0.81

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	77/115 (67%)	76 (99%)	1 (1%)	0	100	100
1	B	84/115 (73%)	80 (95%)	3 (4%)	1 (1%)	13	7
1	C	85/115 (74%)	83 (98%)	1 (1%)	1 (1%)	13	7
1	D	84/115 (73%)	81 (96%)	3 (4%)	0	100	100
1	E	79/115 (69%)	79 (100%)	0	0	100	100
1	F	84/115 (73%)	80 (95%)	4 (5%)	0	100	100
All	All	493/690 (71%)	479 (97%)	12 (2%)	2 (0%)	34	30

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	84	THR
1	C	26	ILE

### 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	69/94 (73%)	63 (91%)	6 (9%)	10	6
1	B	71/94 (76%)	62 (87%)	9 (13%)	4	2
1	C	76/94 (81%)	66 (87%)	10 (13%)	4	2
1	D	75/94 (80%)	66 (88%)	9 (12%)	5	2
1	E	70/94 (74%)	64 (91%)	6 (9%)	10	6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	F	73/94 (78%)	67 (92%)	6 (8%)	11 7
All	All	434/564 (77%)	388 (89%)	46 (11%)	6 3

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	39	ASN
1	E	31	ILE
1	D	40	GLN
1	D	91	LYS
1	E	73	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	80	GLN
1	E	22	GLN
1	F	33	GLN
1	F	22	GLN
1	B	22	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

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 [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, 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	83/115 (72%)	-0.04	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	16, 26, 34, 41	0
1	B	88/115 (76%)	0.09	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	18, 31, 47, 57	0
1	C	91/115 (79%)	0.12	3 (3%) <span style="border: 1px solid red; padding: 2px;">46</span> <span style="border: 1px solid red; padding: 2px;">45</span>	18, 28, 44, 54	0
1	D	90/115 (78%)	0.15	5 (5%) <span style="border: 1px solid red; padding: 2px;">24</span> <span style="border: 1px solid red; padding: 2px;">23</span>	20, 26, 40, 55	0
1	E	85/115 (73%)	-0.02	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	19, 27, 36, 39	0
1	F	88/115 (76%)	0.09	3 (3%) <span style="border: 1px solid red; padding: 2px;">45</span> <span style="border: 1px solid red; padding: 2px;">44</span>	16, 25, 38, 53	0
All	All	525/690 (76%)	0.07	11 (2%) <span style="border: 1px solid blue; padding: 2px;">63</span> <span style="border: 1px solid blue; padding: 2px;">62</span>	16, 27, 43, 57	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	87	VAL	6.7
1	D	40	GLN	4.0
1	D	39	ASN	3.1
1	D	54	PRO	2.7
1	C	86	GLU	2.4

### 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 monosaccharides 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.