



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 12:08 am BST

PDB ID : 1NJ8
Title : Crystal Structure of Prolyl-tRNA Synthetase from Methanocaldococcus
janaschii
Authors : Kamtekar, S.; Kennedy, W.D.; Wang, J.; Stathopoulos, C.; Soll, D.; Steitz,
T.A.
Deposited on : 2002-12-30
Resolution : 3.20 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 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	:	2.11
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)	:	2.11

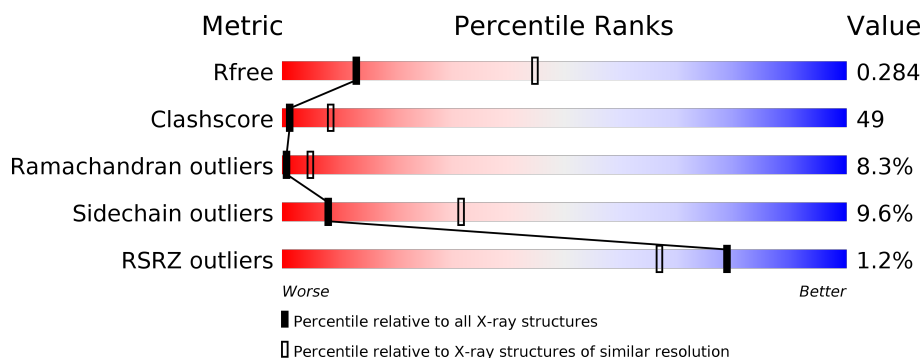
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 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	459	<div> <div>2%</div> <div> <div></div> <div>31%</div> <div>56%</div> <div>12%</div> <div>..</div> </div> </div>
1	B	459	<div> <div>%</div> <div> <div></div> <div>31%</div> <div>57%</div> <div>10%</div> <div>..</div> </div> </div>
1	C	459	<div> <div>%</div> <div> <div></div> <div>29%</div> <div>58%</div> <div>12%</div> <div>..</div> </div> </div>
1	D	459	<div> <div>2%</div> <div> <div></div> <div>31%</div> <div>58%</div> <div>10%</div> <div>.</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 15182 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 Proline-tRNA Synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	456	Total	C	N	O	S	0	0	0
			3772	2447	614	698	13			
1	B	456	Total	C	N	O	S	0	0	0
			3772	2447	614	698	13			
1	C	456	Total	C	N	O	S	0	0	0
			3772	2447	614	698	13			
1	D	456	Total	C	N	O	S	0	0	0
			3772	2447	614	698	13			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP Q58635
A	-2	SER	-	EXPRESSION TAG	UNP Q58635
A	-1	HIS	-	EXPRESSION TAG	UNP Q58635
A	0	MET	-	EXPRESSION TAG	UNP Q58635
A	1	LEU	MET	CLONING ARTIFACT	UNP Q58635
B	-3	GLY	-	EXPRESSION TAG	UNP Q58635
B	-2	SER	-	EXPRESSION TAG	UNP Q58635
B	-1	HIS	-	EXPRESSION TAG	UNP Q58635
B	0	MET	-	EXPRESSION TAG	UNP Q58635
B	1	LEU	MET	CLONING ARTIFACT	UNP Q58635
C	-3	GLY	-	EXPRESSION TAG	UNP Q58635
C	-2	SER	-	EXPRESSION TAG	UNP Q58635
C	-1	HIS	-	EXPRESSION TAG	UNP Q58635
C	0	MET	-	EXPRESSION TAG	UNP Q58635
C	1	LEU	MET	CLONING ARTIFACT	UNP Q58635
D	-3	GLY	-	EXPRESSION TAG	UNP Q58635
D	-2	SER	-	EXPRESSION TAG	UNP Q58635
D	-1	HIS	-	EXPRESSION TAG	UNP Q58635
D	0	MET	-	EXPRESSION TAG	UNP Q58635
D	1	LEU	MET	CLONING ARTIFACT	UNP Q58635

- Molecule 2 is water.

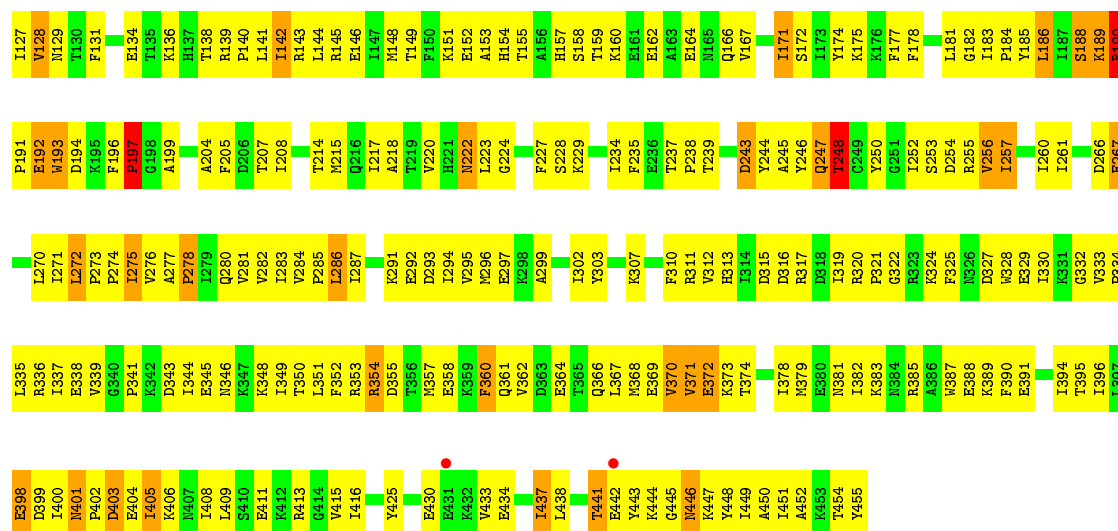
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	35	Total 35	O 35	0	0
2	B	16	Total 16	O 16	0	0
2	C	25	Total 25	O 25	0	0
2	D	18	Total 18	O 18	0	0

3 Residue-property plots

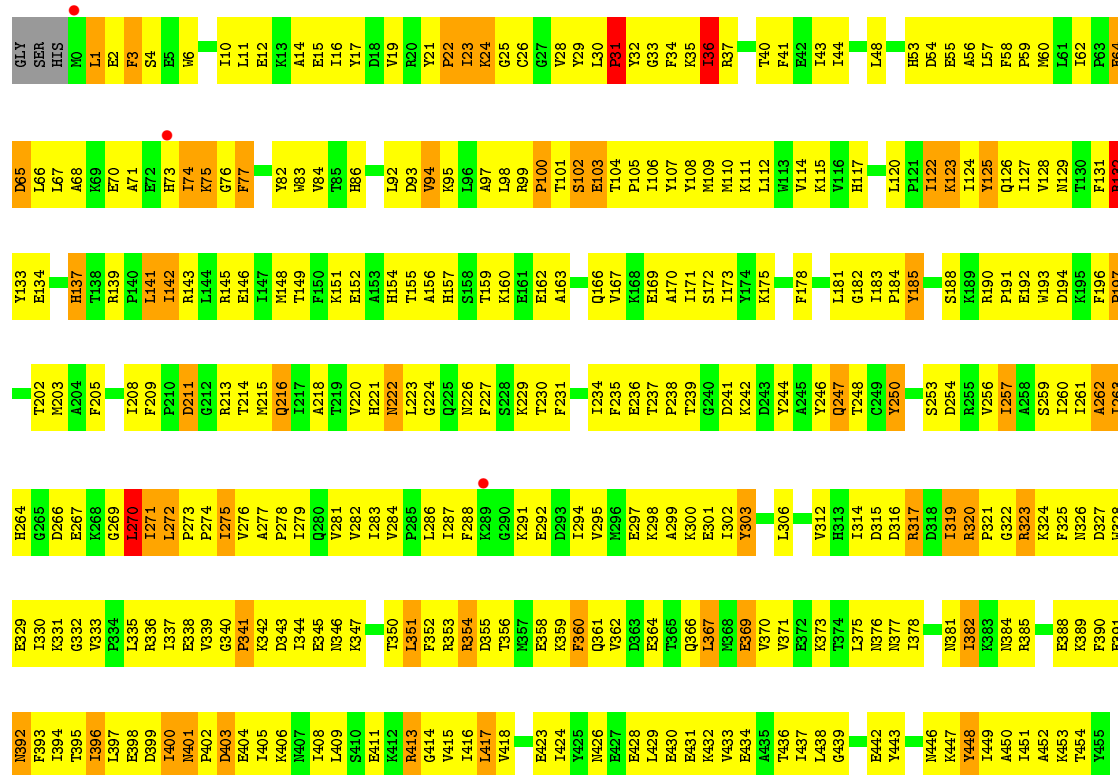
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Proline-tRNA Synthetase

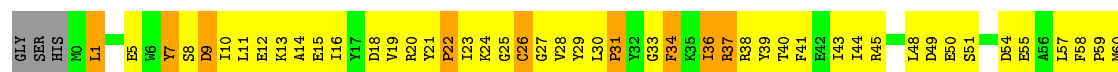




• Molecule 1: Proline-tRNA Synthetase



• Molecule 1: Proline-tRNA Synthetase



T395	T396	T397	E398	D399	I400	M401	P402	D403	E404	I405	K406	M407	I408	L409	E410	E411	K412	R413	G414	V415	I416		Y425		L426	E430	E431	K432	V433	E434	A435	T436	L437	L438		T441	E442	Y443	K444	G445	M446	K447	Y448	I449	A450	I451	A452	K453	T454	Y455										
V333	P334	L335	R336	I337	E338	V339	G340	P341	K342	D343	I344	F345	I346	K347	K348	T349	V350	L351	F352	R353	R354	D355	T356	K357	E358	K359	F360	Q361	V362	D363	E364	T365	Q366	L367	K368	K369	V370	V371	E372	K373	T374		I378	K379	E380	I381	I382	K383	M384	K385	K386	W387	E388	K389	F390	E391	N392	F393	I394	
A262		D266	E267		L270	I271	L272	P273	P274	I275	V276	A277	P278	I279	Q280	V281	V282	I283	V284	V285	L286	I287		K291	E292	I293	I294	V295	V296	E297		T302	Y303		K307		F310	R311	V312	R313	I314	D315	D316	K317	D318	E380	I381	I382	K383	M384	K385	K386	W387	E388	K389	F390	E391	N392	F393	I394
R190	P191	E192	W193	D194	K195	F196	P197	G198	A199		T202	W203	A204	P205	D206	T207	L144	R145	E146	M148	T149		F150	K151	E152	A153	H154	T155	A156	H157	S158	T159	K160	E161	A162	A163	E164	N165	Q166	V167		I171	S172	I173	Y174	K175	K176	F177	F178		L181	G182	I183	P184	Y185	L186	S187	K188	K189	
L61	I62	P63	E64	D65	L66	L67	A68	K69	E70	A71	E72	H73	I74	K75	F76	G77	F78		Y82	W83	V84	T85	H86		Q91	L92	D93	V94	K95	L96	A97	L98	R99	P100	T101	S102	E103	T104	P105	I106	Y107	Y108	M109	M110	K111	L112	W113	V114	K115	V116	H117	T118	D119	L120	P121	I122	K123	I124	Y125	

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	108.94Å 104.84Å 91.75Å 90.00° 93.63° 90.00°	Depositor
Resolution (Å)	20.00 – 3.20 29.81 – 3.16	Depositor EDS
% Data completeness (in resolution range)	87.0 (20.00-3.20) 84.9 (29.81-3.16)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	720.88 (at 3.18Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.231 , 0.301 0.223 , 0.284	Depositor DCC
R_{free} test set	2974 reflections (9.83%)	wwPDB-VP
Wilson B-factor (Å ²)	67.4	Xtriage
Anisotropy	0.553	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 54.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	15182	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% 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 [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	0.44	0/3860	0.71	3/5213 (0.1%)
1	B	0.40	0/3860	0.63	0/5213
1	C	0.44	0/3860	0.70	3/5213 (0.1%)
1	D	0.39	0/3860	0.63	0/5213
All	All	0.42	0/15440	0.67	6/20852 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	320	ARG	NE-CZ-NH2	-11.57	114.52	120.30
1	C	320	ARG	NE-CZ-NH1	-11.40	114.60	120.30
1	C	320	ARG	NE-CZ-NH2	11.27	125.94	120.30
1	A	320	ARG	NE-CZ-NH1	11.16	125.88	120.30
1	C	320	ARG	CD-NE-CZ	5.36	131.10	123.60

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	3772	0	3804	386	2
1	B	3772	0	3804	353	0
1	C	3772	0	3804	409	1
1	D	3772	0	3804	363	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	35	0	0	2	0
2	B	16	0	0	1	0
2	C	25	0	0	6	0
2	D	18	0	0	3	0
All	All	15182	0	15216	1472	2

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 49.

The worst 5 of 1472 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:275:ILE:HD12	1:B:275:ILE:H	1.06	1.14
1:D:275:ILE:H	1:D:275:ILE:HD12	1.08	1.11
1:B:48:LEU:HD11	1:B:151:LYS:HD2	1.32	1.10
1:D:48:LEU:HD11	1:D:151:LYS:HD2	1.34	1.07
1:A:337:ILE:HG12	1:A:351:LEU:HD12	1.36	1.02

All (2) 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	Interatomic distance (Å)	Clash overlap (Å)
1:A:323:ARG:NH1	1:C:323:ARG:NH1[2_646]	1.95	0.25
1:A:293:ASP:OD1	1:D:392:ASN:OD1[1_554]	2.10	0.10

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	454/459 (99%)	331 (73%)	84 (18%)	39 (9%)	1 4
1	B	454/459 (99%)	313 (69%)	104 (23%)	37 (8%)	1 5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	454/459 (99%)	330 (73%)	85 (19%)	39 (9%)	1	4
1	D	454/459 (99%)	312 (69%)	107 (24%)	35 (8%)	1	6
All	All	1816/1836 (99%)	1286 (71%)	380 (21%)	150 (8%)	1	5

5 of 150 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	36	ILE
1	A	74	ILE
1	A	75	LYS
1	A	76	GLY
1	A	102	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	412/414 (100%)	373 (90%)	39 (10%)	8	32
1	B	412/414 (100%)	373 (90%)	39 (10%)	8	32
1	C	412/414 (100%)	370 (90%)	42 (10%)	7	29
1	D	412/414 (100%)	374 (91%)	38 (9%)	9	33
All	All	1648/1656 (100%)	1490 (90%)	158 (10%)	8	32

5 of 158 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	360	PHE
1	C	137	HIS
1	D	303	TYR
1	B	398	GLU
1	C	31	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	401	ASN
1	C	216	GLN
1	D	366	GLN
1	C	46	ASN
1	C	222	ASN

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 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 [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	456/459 (99%)	-0.45	7 (1%) 73 61	10, 49, 112, 140	0
1	B	456/459 (99%)	-0.30	4 (0%) 84 75	17, 69, 129, 144	0
1	C	456/459 (99%)	-0.44	3 (0%) 87 81	11, 49, 111, 140	0
1	D	456/459 (99%)	-0.25	8 (1%) 68 55	16, 69, 129, 144	0
All	All	1824/1836 (99%)	-0.36	22 (1%) 79 67	10, 58, 122, 144	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	73	HIS	6.3
1	B	74	ILE	4.7
1	C	73	HIS	4.5
1	D	74	ILE	4.0
1	B	73	HIS	3.7

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 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.