



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 10, 2024 – 12:28 PM EST

PDB ID : 2P5T
Title : Molecular and structural characterization of the PezAT chromosomal toxin-antitoxin system of the human pathogen *Streptococcus pneumoniae*
Authors : Loll, B.; Meinhart, A.
Deposited on : 2007-03-16
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 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

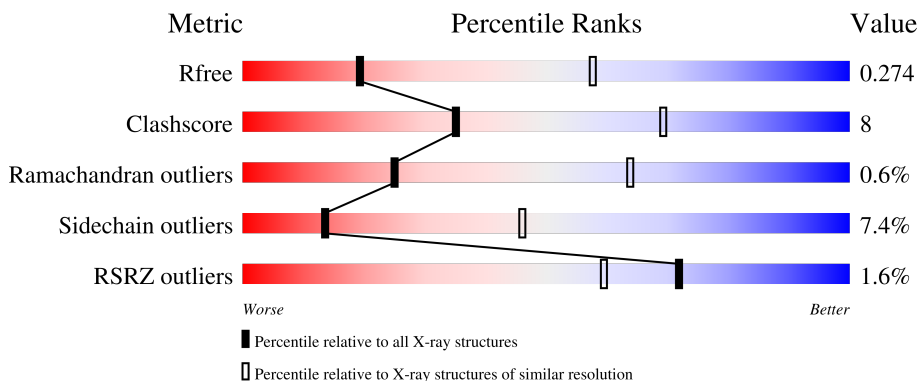
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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.



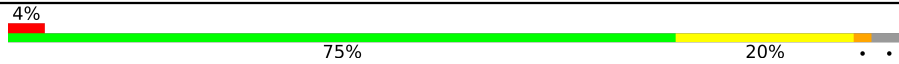

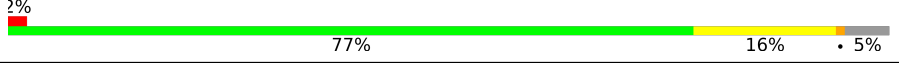

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 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 $\leq 5\%$. 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	X	33	
2	A	158	
2	C	158	
2	E	158	
2	G	158	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	B	253	
3	D	253	
3	F	253	
3	H	253	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 11093 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 fragment of PezA helix-turn-helix motif.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	X	33	132	66	33	33	0	0	0

- Molecule 2 is a protein called Putative transcriptional regulator PezA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	92	759	478	122	156	3	0	0	0
2	C	93	767	483	123	157	4	0	0	0
2	E	95	784	493	126	161	4	0	0	0
2	G	93	767	483	123	157	4	0	0	0

- Molecule 3 is a protein called PezT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	244	1979	1253	342	381	3	0	0	0
3	D	247	2002	1268	345	385	4	0	0	0
3	F	240	1944	1232	332	377	3	0	0	0
3	H	242	1959	1242	334	379	4	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	109	GLY	ARG	conflict	UNP Q97QZ1
B	228	PHE	LEU	conflict	UNP Q97QZ1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	109	GLY	ARG	conflict	UNP Q97QZ1
D	228	PHE	LEU	conflict	UNP Q97QZ1
F	109	GLY	ARG	conflict	UNP Q97QZ1
F	228	PHE	LEU	conflict	UNP Q97QZ1
H	109	GLY	ARG	conflict	UNP Q97QZ1
H	228	PHE	LEU	conflict	UNP Q97QZ1

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: fragment of PezA helix-turn-helix motif

Chain X:  100%

There are no outlier residues recorded for this chain.

- Molecule 2: Putative transcriptional regulator PezA

Chain A:  44% 13% 42%

MET ILE GLY LYS ASN LYS ILE LYS SER LEU ARG LYS THR HIS ASP LEU THR GLN LEU LEU PHE ALA ILE VAL GLY ILE SER ARG ASN SER SER LEU SER ARG TYR GLU ASN GLY THR SER SER VAL THR GLU LEU ILE ASP ILE ILE CYS GLN LYS PHE ASN VAL SER TYR VAL ASP ILE

VAL GLY GLU ASP MET L67 M66 P68 E80 L75 L79 E80 K83 S91 R95 D106 P110 L113 L116 L121 M125 I126 I135 S139 G140 D143 I151 R155 M156 V157 A158

- Molecule 2: Putative transcriptional regulator PezA

Chain C:  42% 16% 41%

MET ILE GLY LYS ASN LYS ILE LYS SER LEU ARG LYS THR HIS ASP LEU THR GLN LEU LEU PHE ALA ILE VAL GLY ILE SER ARG ASN SER SER LEU SER ARG TYR GLU ASN GLY THR SER SER VAL THR GLU LEU ILE ASP ILE ILE CYS GLN LYS PHE ASN VAL SER TYR VAL ASP ILE

VAL GLY GLU ASP MET L67 M66 I79 E80 R92 Y96 I104 S108 N109 P110 W111 D120 L121 M125 I126 T131 F132 D133 E134 I135 E136 R137 Y138 Y141 M148 L149 E150 K154 R156 A158


- Molecule 2: Putative transcriptional regulator PezA

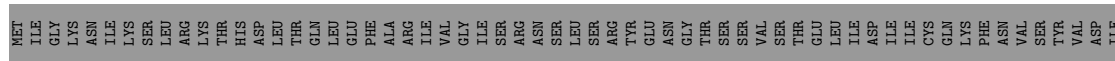
Chain E:  % 43% 15% 40%

MET ILE GLY LYS ASN LYS ILE LYS SER LEU ARG LYS THR HIS ASP LEU THR GLN LEU LEU PHE ALA ILE VAL GLY ILE SER ARG ASN SER SER LEU SER ARG TYR GLU ASN GLY THR SER SER VAL THR GLU LEU ILE ASP ILE ILE CYS GLN LYS PHE ASN VAL SER TYR VAL ASP ILE

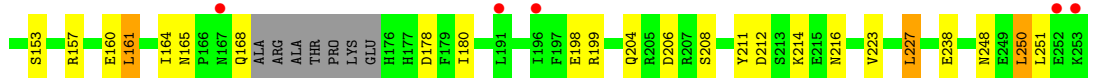
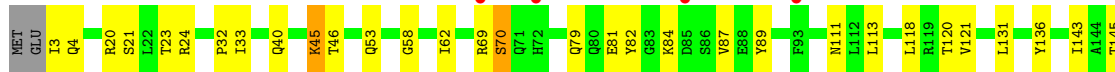
VAL GLY GLU D94 L67 N68 E71 K78 V82 L90 Y96 Q97 P110 W111 S115 L118 L121 I122 I125 N126 I127 Y129 V129 D133 E134 I135 Y138 S139 G140 M148 L149 E150 I151 S152 A158

- Molecule 2: Putative transcriptional regulator PezA

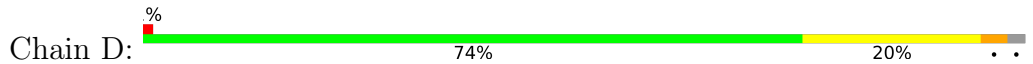
Chain G:  % 43% 14% 41%



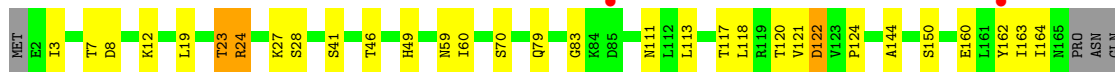
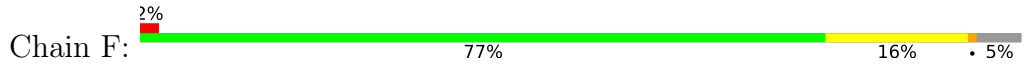
• Molecule 3: PezT



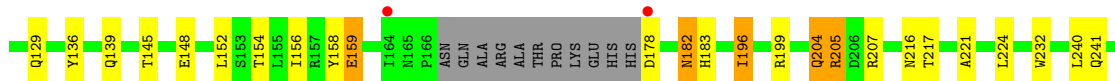
• Molecule 3: PezT

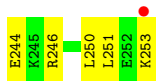


• Molecule 3: PezT



• Molecule 3: PezT





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	80.52Å 102.86Å 254.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.67 – 3.20 47.68 – 3.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.67-3.20) 98.5 (47.68-3.20)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.00 (at 3.19Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.214 , 0.277 0.211 , 0.274	Depositor DCC
R_{free} test set	1726 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	90.3	Xtrriage
Anisotropy	0.014	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 84.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11093	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% 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
2	A	0.38	0/770	0.54	0/1039
2	C	0.39	0/778	0.55	0/1049
2	E	0.44	0/795	0.56	0/1071
2	G	0.41	0/778	0.55	0/1049
3	B	0.38	0/2010	0.52	0/2703
3	D	0.39	0/2034	0.55	0/2736
3	F	0.38	0/1972	0.54	0/2650
3	H	0.43	0/1988	0.57	0/2672
All	All	0.40	0/11125	0.55	0/14969

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 [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	X	132	0	2	0	0
2	A	759	0	740	20	0
2	C	767	0	749	22	0
2	E	784	0	766	18	0
2	G	767	0	749	24	0
3	B	1979	0	1997	31	0
3	D	2002	0	2023	37	0
3	F	1944	0	1968	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	1959	0	1987	42	0
All	All	11093	0	10981	180	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 8.

The worst 5 of 180 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)
2:G:104:ILE:HD11	3:H:17:ARG:HB2	1.51	0.89
3:D:147:PRO:HB2	3:H:148:GLU:HG2	1.58	0.85
2:G:79:ILE:HD13	3:H:50:ARG:HD2	1.61	0.80
3:H:40:GLN:HB2	3:H:43:ALA:HB2	1.65	0.79
2:C:131:THR:HG23	2:C:134:GLU:HB2	1.66	0.77

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	
2	A	90/158 (57%)	84 (93%)	5 (6%)	1 (1%)	14	51
2	C	91/158 (58%)	87 (96%)	3 (3%)	1 (1%)	14	51
2	E	93/158 (59%)	88 (95%)	5 (5%)	0	100	100
2	G	91/158 (58%)	87 (96%)	4 (4%)	0	100	100
3	B	240/253 (95%)	218 (91%)	21 (9%)	1 (0%)	34	69
3	D	243/253 (96%)	226 (93%)	15 (6%)	2 (1%)	19	58
3	F	236/253 (93%)	221 (94%)	13 (6%)	2 (1%)	19	58
3	H	238/253 (94%)	224 (94%)	13 (6%)	1 (0%)	34	69
All	All	1322/1644 (80%)	1235 (93%)	79 (6%)	8 (1%)	25	64

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	164	ILE
3	F	24	ARG
2	A	106	ASP
2	C	104	ILE
3	D	177	HIS

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
2	A	86/147 (58%)	83 (96%)	3 (4%)	36 69
2	C	87/147 (59%)	85 (98%)	2 (2%)	50 78
2	E	89/147 (60%)	83 (93%)	6 (7%)	16 50
2	G	87/147 (59%)	81 (93%)	6 (7%)	15 49
3	B	219/226 (97%)	203 (93%)	16 (7%)	14 46
3	D	222/226 (98%)	204 (92%)	18 (8%)	11 42
3	F	215/226 (95%)	198 (92%)	17 (8%)	12 43
3	H	217/226 (96%)	194 (89%)	23 (11%)	6 27
All	All	1222/1492 (82%)	1131 (93%)	91 (7%)	13 46

5 of 91 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	162	TYR
3	H	70	SER
3	F	179	PHE
2	G	104	ILE
3	H	120	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	204	GLN
3	H	129	GLN
3	H	241	GLN
3	H	183	HIS
3	H	49	HIS

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 [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	X	0/33	-	-	-	-
2	A	92/158 (58%)	-0.07	0 100 100	78, 88, 97, 104	0
2	C	93/158 (58%)	-0.13	0 100 100	79, 90, 105, 112	0
2	E	95/158 (60%)	-0.15	1 (1%) 80 69	79, 89, 107, 114	0
2	G	93/158 (58%)	-0.10	2 (2%) 62 48	79, 89, 108, 111	0
3	B	244/253 (96%)	0.05	9 (3%) 41 26	78, 91, 106, 115	0
3	D	247/253 (97%)	-0.15	3 (1%) 79 67	79, 89, 105, 114	0
3	F	240/253 (94%)	-0.09	4 (1%) 70 57	74, 89, 114, 123	0
3	H	242/253 (95%)	-0.06	3 (1%) 79 67	75, 90, 112, 130	0
All	All	1346/1677 (80%)	-0.08	22 (1%) 72 59	74, 90, 107, 130	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	157	VAL	5.7
2	E	64	ASP	5.1
3	H	164	ILE	4.1
3	B	253	LYS	4.0
2	G	158	ALA	4.0

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.