



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 11:37 AM GMT

PDB ID : 3PJZ
Title : Crystal Structure of the Potassium Transporter TrkH from *Vibrio parahaemolyticus*
Authors : Cao, Y.; Jin, X.; Huang, H.; Levin, E.J.; Zhou, M.; New York Consortium on Membrane Protein Structure (NYCOMPS)
Deposited on : 2010-11-10
Resolution : 3.51 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
<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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

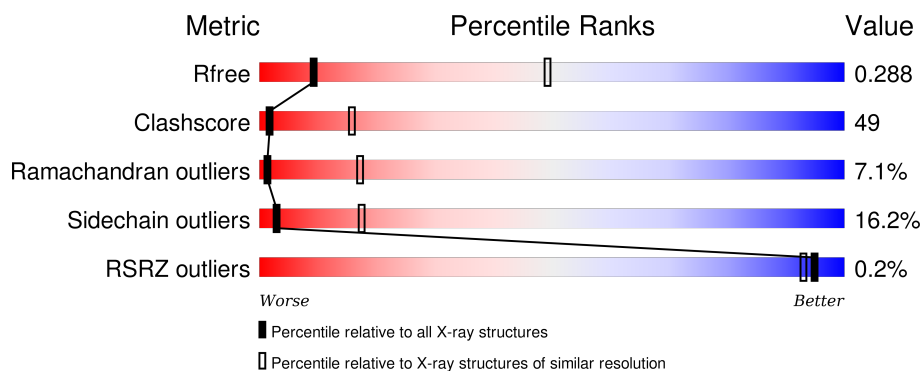
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.51 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}	91344	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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. 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	494	 32% 49% 13% 5%
1	B	494	 32% 50% 13% 5%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7242 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 Potassium uptake protein TrkH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	468	Total	C	N	O	S	0	0	0
			3620	2421	574	607	18			
1	B	468	Total	C	N	O	S	0	0	0
			3620	2421	574	607	18			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	486	ALA	-	EXPRESSION TAG	UNP Q87TN7
A	487	ALA	-	EXPRESSION TAG	UNP Q87TN7
A	488	ALA	-	EXPRESSION TAG	UNP Q87TN7
A	489	GLU	-	EXPRESSION TAG	UNP Q87TN7
A	490	ASN	-	EXPRESSION TAG	UNP Q87TN7
A	491	LEU	-	EXPRESSION TAG	UNP Q87TN7
A	492	TYR	-	EXPRESSION TAG	UNP Q87TN7
A	493	PHE	-	EXPRESSION TAG	UNP Q87TN7
A	494	GLN	-	EXPRESSION TAG	UNP Q87TN7
B	486	ALA	-	EXPRESSION TAG	UNP Q87TN7
B	487	ALA	-	EXPRESSION TAG	UNP Q87TN7
B	488	ALA	-	EXPRESSION TAG	UNP Q87TN7
B	489	GLU	-	EXPRESSION TAG	UNP Q87TN7
B	490	ASN	-	EXPRESSION TAG	UNP Q87TN7
B	491	LEU	-	EXPRESSION TAG	UNP Q87TN7
B	492	TYR	-	EXPRESSION TAG	UNP Q87TN7
B	493	PHE	-	EXPRESSION TAG	UNP Q87TN7
B	494	GLN	-	EXPRESSION TAG	UNP Q87TN7

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	K	0	0
			1	1		

Continued on next page...

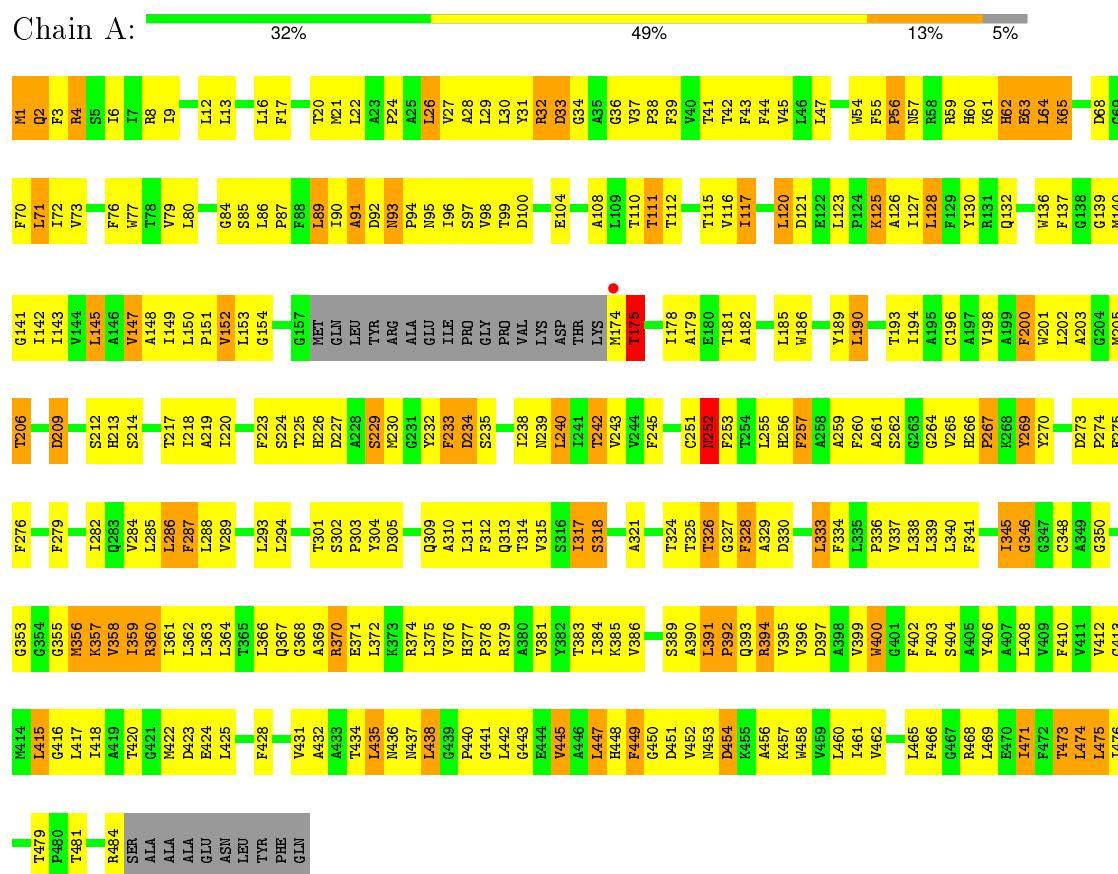
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	K	0	0
			1	1		

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 errors 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 ($\text{RSRZ} > 2$). 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: Potassium uptake protein TrkH



W136	W201	D273	A349	L415	T479
F137	L202	P274	G350	G416	P490
G138	A203	E275	G353	L417	T481
G139	G204	F276	G353	L418	F482
M140	M205		M356	A419	F483
G141	T206	F279	K357	T420	F484
I142			K358	G491	F484
I143	D209		V358	M422	SER
V144		I282	I359	D423	ALA
L145	S212	Q283	R360	E424	ALA
A146	H213	L285	I361	L425	ALA
V147	S214	L286	L362		GLU
A148		F287	L363	F428	ASN
I149	T217	L288	L364		LEU
L150	I218	V289	T365	V431	TYR
P151	A219	C290	G367	A432	PHE
V152	I220		Q368	T434	GLN
L153		L293	G368	L435	
G154		L294	A369	M436	
	F223		R370	M437	
G157	S224	T301	E371	L438	
MET	T225	S302	L372	G439	
GLN	H226	P303	R373	P440	
LEU	D227	Y304	R374	G441	
	A228	D305	L375	L442	
TVR	S229		V376	G443	
ARG	M230			E444	
	G231	Q309	P378	V445	
ALA	Y232	A310	T383	A446	
GLU	F233	L311	I384	L447	
ILE	D234	F312	K385	H448	
PRO	S235	Q313	V386	F449	
GLY		T314		G450	
PRO	N239	V315		V451	
VAL	I240	S316	S389	M452	
LYS	I241	I317	A390	M453	
ASP	T242	S318	I391	D454	
THR	V243		P392	F455	
	V244	A321	Q393	A456	
M174	F245		R394	K457	
I178		T324	V395	M458	
C251		T325	V396	V459	
A179	R252	T326	D397	L460	
F180		G327	A398	I461	
T181	L255	F328	V399	V462	
A182	H256	A329	W400	S463	
	F257	D330		M464	
L185	A258		G401	L465	
W186	A259	L333	F403	F466	
	F260	F334	S404	G467	
Y189	A261	L335	A405	R468	
L190	S262	P336	Y406	L469	
	G263	V337	A407	E470	
T193	G264	L340	V409	F471	
I194	V265	F341	F410	T472	
A195	H266		V411	T473	
C196	P267	I345	V412	L474	
A197	K268	G346	C413	L475	
V198	Y269	G347	M414		
A199	Y270	C348			
F200					

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	91.49Å 97.41Å 195.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.99 – 3.51 44.54 – 3.51	Depositor EDS
% Data completeness (in resolution range)	85.7 (14.99-3.51) 85.6 (44.54-3.51)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	18.32 (at 3.48Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.249 , 0.299 0.237 , 0.288	Depositor DCC
R_{free} test set	1007 reflections (5.21%)	DCC
Wilson B-factor (Å ²)	63.6	Xtriage
Anisotropy	0.083	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 47.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	2 of 19351 reflections (0.010%)	Xtriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	7242	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: K

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	0/3723	0.61	0/5073
1	B	0.43	0/3723	0.60	0/5073
All	All	0.44	0/7446	0.60	0/10146

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	3620	0	3696	367	0
1	B	3620	0	3696	370	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
All	All	7242	0	7392	712	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 49.

The worst 5 of 712 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:136:TRP:HE1	1:A:193:THR:HG21	0.97	1.12
1:B:136:TRP:HE1	1:B:193:THR:HG21	0.95	1.06
1:B:93:ASN:HB2	1:B:94:PRO:HD3	1.40	1.04
1:B:132:GLN:HG3	1:B:212:SER:HB2	1.38	1.04
1:A:93:ASN:HB2	1:A:94:PRO:HD3	1.41	1.00

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	464/494 (94%)	337 (73%)	93 (20%)	34 (7%)	1	16
1	B	464/494 (94%)	338 (73%)	94 (20%)	32 (7%)	1	18
All	All	928/988 (94%)	675 (73%)	187 (20%)	66 (7%)	1	17

5 of 66 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	HIS
1	A	91	ALA
1	A	175	THR
1	A	267	PRO
1	A	329	ALA

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	380/401 (95%)	318 (84%)	62 (16%)	3	17
1	B	380/401 (95%)	319 (84%)	61 (16%)	3	18
All	All	760/802 (95%)	637 (84%)	123 (16%)	3	17

5 of 123 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	454	ASP
1	B	33	ASP
1	B	435	LEU
1	A	474	LEU
1	B	1	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	437	ASN
1	B	60	HIS
1	B	437	ASN
1	A	453	ASN
1	B	2	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	468/494 (94%)	-0.56	1 (0%) 95 93	17, 53, 112, 179	0
1	B	468/494 (94%)	-0.50	1 (0%) 95 93	18, 53, 115, 178	0
All	All	936/988 (94%)	-0.53	2 (0%) 95 93	17, 53, 113, 179	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	174	MET	4.2
1	B	484	ARG	2.3

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	K	A	501	1/1	0.97	0.09	-3.28	41,41,41,41	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	K	B	501	1/1	0.97	0.04	-4.13	49,49,49,49	0

6.5 Other polymers [i](#)

There are no such residues in this entry.