wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 09:06 AM GMT

PDB ID	:	3H8E
Title	:	Low pH native structure of leucine aminopeptidase from Pseudomonas putida
Authors	:	Kale, A.; Dijkstra, B.W.; Sonke, T.; Thunnissen, A.M.W.H
Deposited on	:	2009-04-29
Resolution	:	2.75 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

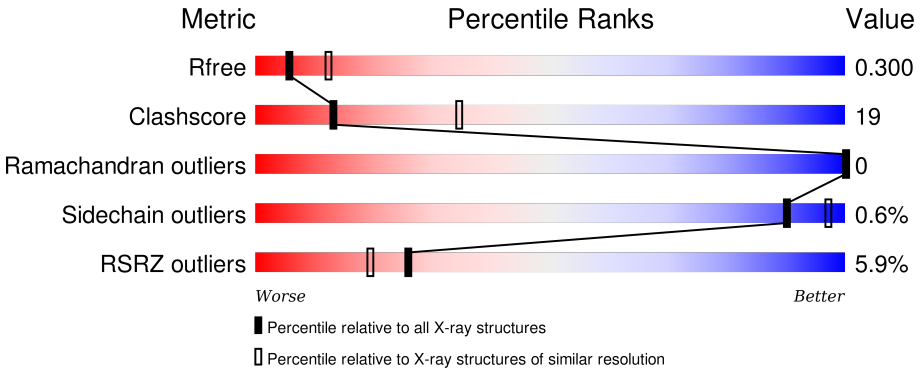
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 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	3340 (2.80-2.72)
Clashscore	102246	3829 (2.80-2.72)
Ramachandran outliers	100387	3767 (2.80-2.72)
Sidechain outliers	100360	3770 (2.80-2.72)
RSRZ outliers	91569	3352 (2.80-2.72)

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 <=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	497	6%, 74%, 18%, • 5%
1	B	497	6%, 71%, 21%, • 5%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 7006 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 Cytosol aminopeptidase.

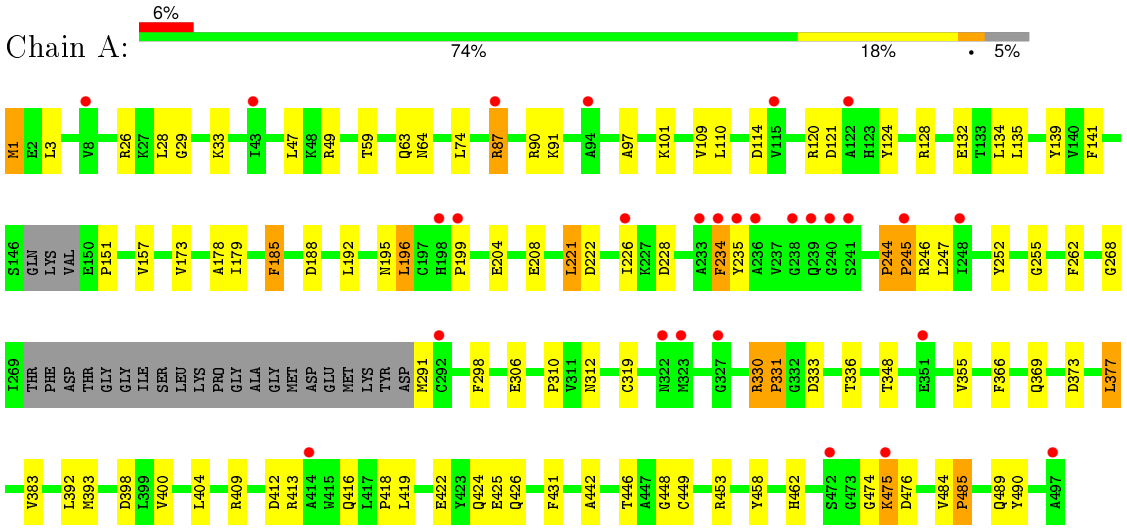
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	473	Total	C	N	O	S	0	0	0
			3503	2212	616	659	16			
1	B	473	Total	C	N	O	S	0	0	0
			3503	2212	616	659	16			

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 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 (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.

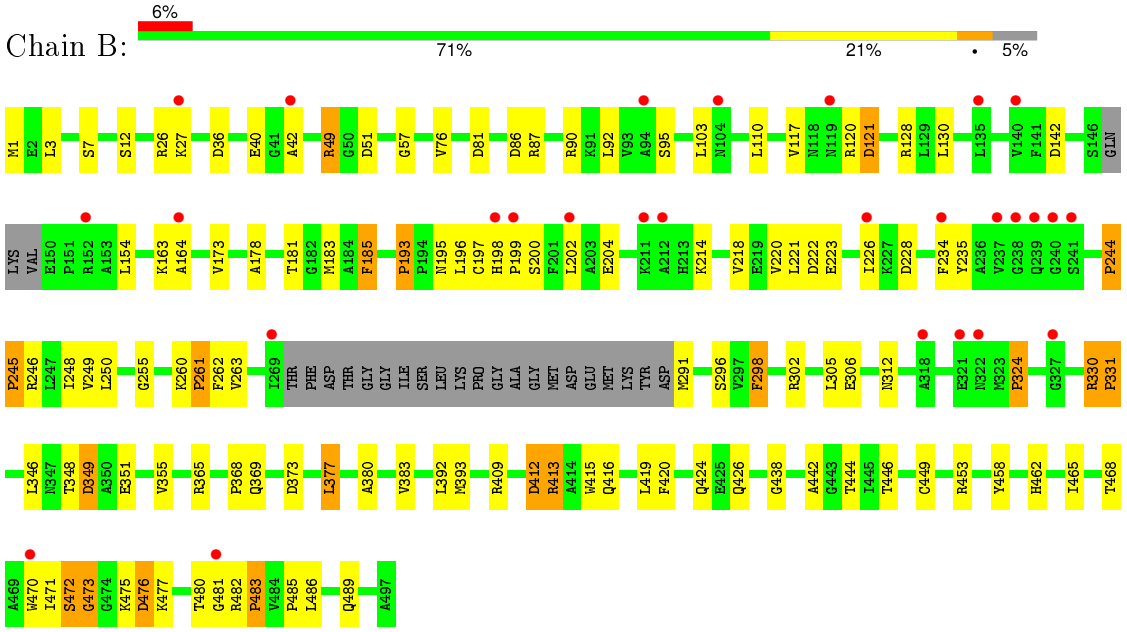
- Molecule 1: Cytosol aminopeptidase

Chain A:



- Molecule 1: Cytosol aminopeptidase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α, β, γ	116.95Å 116.95Å 137.90Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	58.00 – 2.75 58.47 – 2.75	Depositor EDS
% Data completeness (in resolution range)	100.0 (58.00-2.75) 100.0 (58.47-2.75)	Depositor EDS
R_{merge}	0.35	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.11 (at 2.77Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.212 , 0.267 0.290 , 0.300	Depositor DCC
R_{free} test set	1399 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	41.2	Xtriage
Anisotropy	0.360	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , -15.8	EDS
Estimated twinning fraction	0.479 for h,-h-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Outliers	0 of 27831 reflections	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	7006	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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	1.38	9/3553 (0.3%)	1.34	39/4805 (0.8%)
1	B	1.35	16/3553 (0.5%)	1.31	40/4805 (0.8%)
All	All	1.37	25/7106 (0.4%)	1.32	79/9610 (0.8%)

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	A	0	2
1	B	0	3
All	All	0	5

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	235	TYR	CD1-CE1	10.94	1.55	1.39
1	A	310	PRO	N-CD	-10.55	1.33	1.47
1	B	245	PRO	N-CD	-7.34	1.37	1.47
1	B	324	PRO	N-CD	-7.14	1.37	1.47
1	B	235	TYR	CD1-CE1	7.13	1.50	1.39

The worst 5 of 79 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	246	ARG	NE-CZ-NH1	11.27	125.94	120.30
1	A	310	PRO	CA-N-CD	-9.63	98.02	111.50
1	A	235	TYR	CG-CD1-CE1	-9.49	113.71	121.30
1	B	185	PHE	CB-CG-CD1	9.33	127.33	120.80
1	B	413	ARG	NE-CZ-NH2	-9.23	115.68	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	474	GLY	Peptide
1	A	475	LYS	Peptide
1	B	472	SER	Peptide
1	B	473	GLY	Peptide
1	B	476	ASP	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	3503	0	3600	123	11
1	B	3503	0	3600	152	12
All	All	7006	0	7200	268	12

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 19.

The worst 5 of 268 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:226:ILE:HD13	1:A:234:PHE:CE2	1.28	1.65
1:B:226:ILE:HD13	1:B:234:PHE:CE2	1.10	1.62
1:B:226:ILE:CD1	1:B:234:PHE:HE2	1.13	1.52
1:A:185:PHE:CE2	1:A:298:PHE:HB3	1.44	1.51
1:B:473:GLY:CA	1:B:477:LYS:HG3	1.38	1.49

The worst 5 of 12 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:336:THR:CB	1:B:214:LYS:CE[4_444]	0.56	1.64
1:B:163:LYS:CE	1:B:164:ALA:O[2_445]	1.27	0.93
1:A:336:THR:CB	1:B:214:LYS:NZ[4_444]	1.56	0.64
1:A:336:THR:OG1	1:B:214:LYS:CE[4_444]	1.60	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:63:GLN:NE2	1:B:49:ARG:NH2[2_455]	1.64	0.56

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	467/497 (94%)	455 (97%)	12 (3%)	0	100	100
1	B	467/497 (94%)	454 (97%)	13 (3%)	0	100	100
All	All	934/994 (94%)	909 (97%)	25 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	359/378 (95%)	357 (99%)	2 (1%)	90	97
1	B	359/378 (95%)	357 (99%)	2 (1%)	90	97
All	All	718/756 (95%)	714 (99%)	4 (1%)	90	97

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	330	ARG
1	A	377	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	330	ARG
1	B	377	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	213	HIS
1	B	489	GLN
1	B	424	GLN
1	A	462	HIS
1	B	312	ASN

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

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

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^{th} 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	473/497 (95%)	0.49	28 (5%)	26	19	5, 10, 25, 45	0
1	B	473/497 (95%)	0.53	28 (5%)	26	19	6, 12, 28, 51	0
All	All	946/994 (95%)	0.51	56 (5%)	26	19	5, 11, 27, 51	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	269	ILE	4.3
1	A	239	GLN	4.0
1	A	241	SER	3.8
1	A	323	MET	3.8
1	B	94	ALA	3.8

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

There are no carbohydrates 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.