



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:48 AM GMT

PDB ID : 2IJZ
Title : Crystal structure of aminopeptidase
Authors : Min, T.; Burley, S.K.; Shapiro, L.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2006-10-02
Resolution : 3.00 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
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

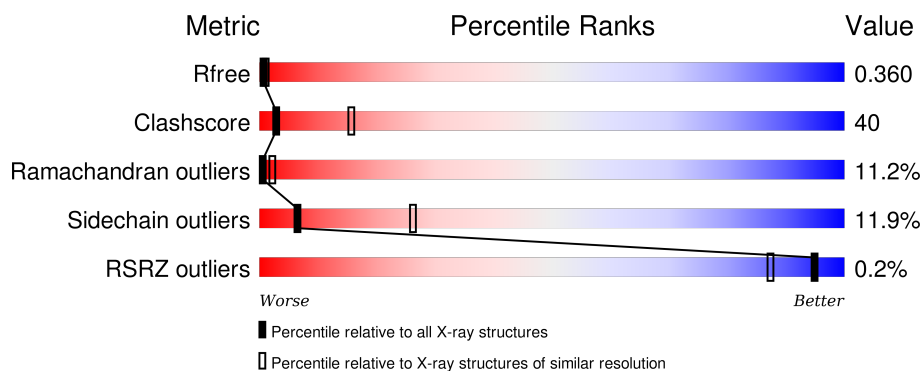
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

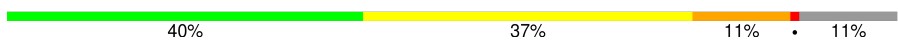
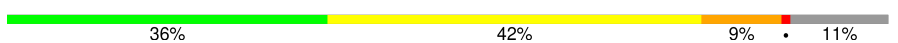


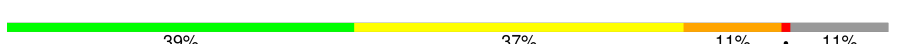
The reported resolution of this entry is 3.00 Å.

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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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 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	428	
1	B	428	
1	C	428	
1	D	428	
1	E	428	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	428	<div><div></div><div>40%38%9%11%</div></div>
1	G	428	<div><div></div><div>41%37%9%11%</div></div>
1	H	428	<div><div></div><div>39%39%10%11%</div></div>
1	I	428	<div><div></div><div>39%36%12%11%</div></div>
1	J	428	<div><div></div><div>42%36%10%11%</div></div>
1	K	428	<div><div></div><div>36%41%10%11%</div></div>
1	L	428	<div><div>%</div><div>39%38%10%11%</div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 37617 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 Probable M18-family aminopeptidase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	B	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	C	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	D	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	E	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	F	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	G	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	H	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	I	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	J	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	K	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			
1	L	379	Total	C	N	O	S	0	0	0
			2870	1795	519	546	10			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
B	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
C	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
D	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
E	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
G	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
H	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
I	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
J	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
K	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3
L	154	ASN	ALA	ENGINEERED	UNP Q9HYZ3

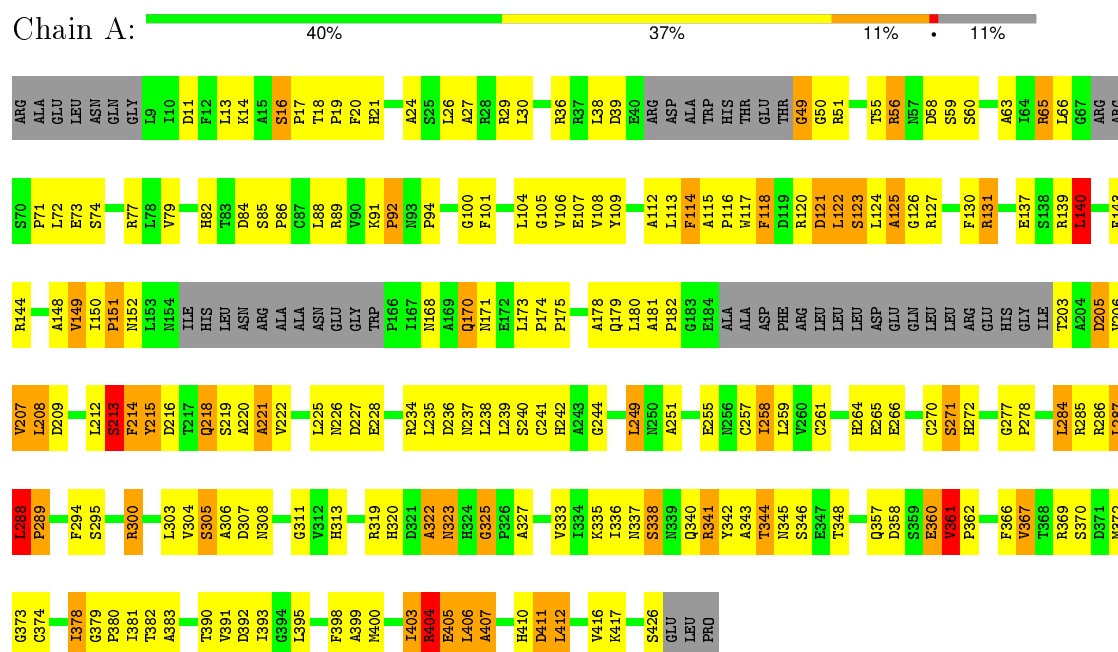
- Molecule 2 is water.

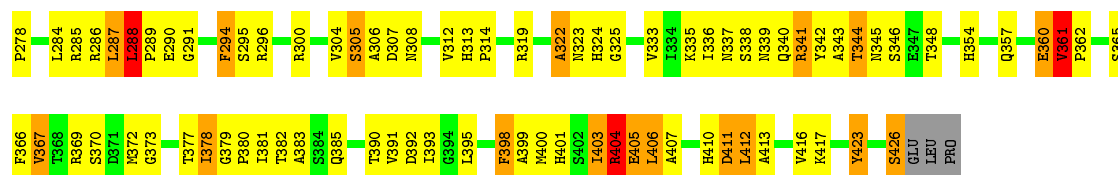
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	262	Total O 262 262	0	0
2	B	271	Total O 271 271	0	0
2	C	298	Total O 298 298	0	0
2	D	247	Total O 247 247	0	0
2	E	268	Total O 268 268	0	0
2	F	247	Total O 247 247	0	0
2	G	255	Total O 255 255	0	0
2	H	252	Total O 252 252	0	0
2	I	274	Total O 274 274	0	0
2	J	258	Total O 258 258	0	0
2	K	276	Total O 276 276	0	0
2	L	269	Total O 269 269	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 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.

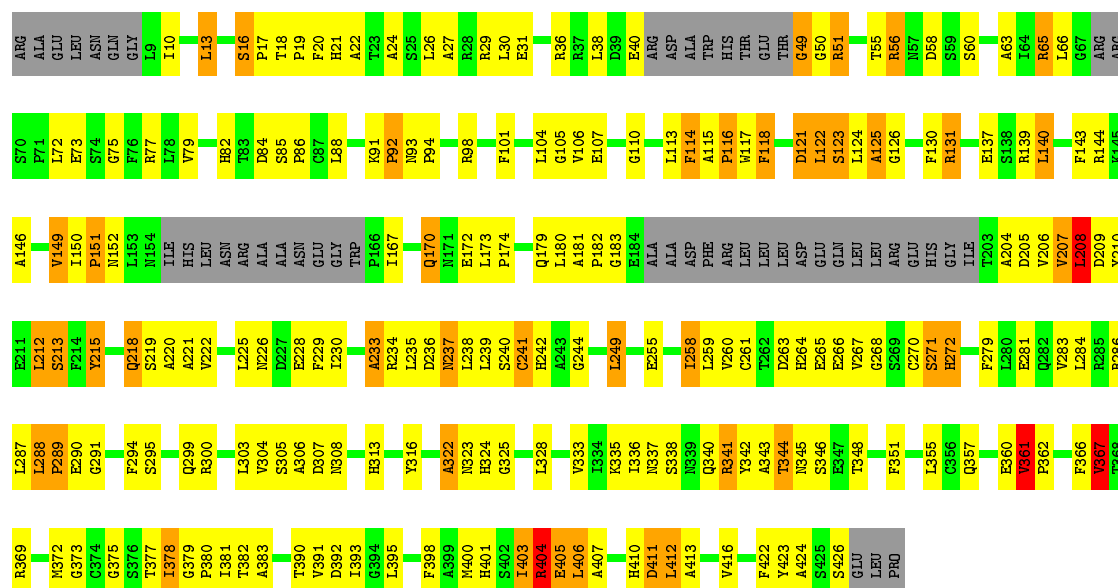
• Molecule 1: Probable M18-family aminopeptidase 2





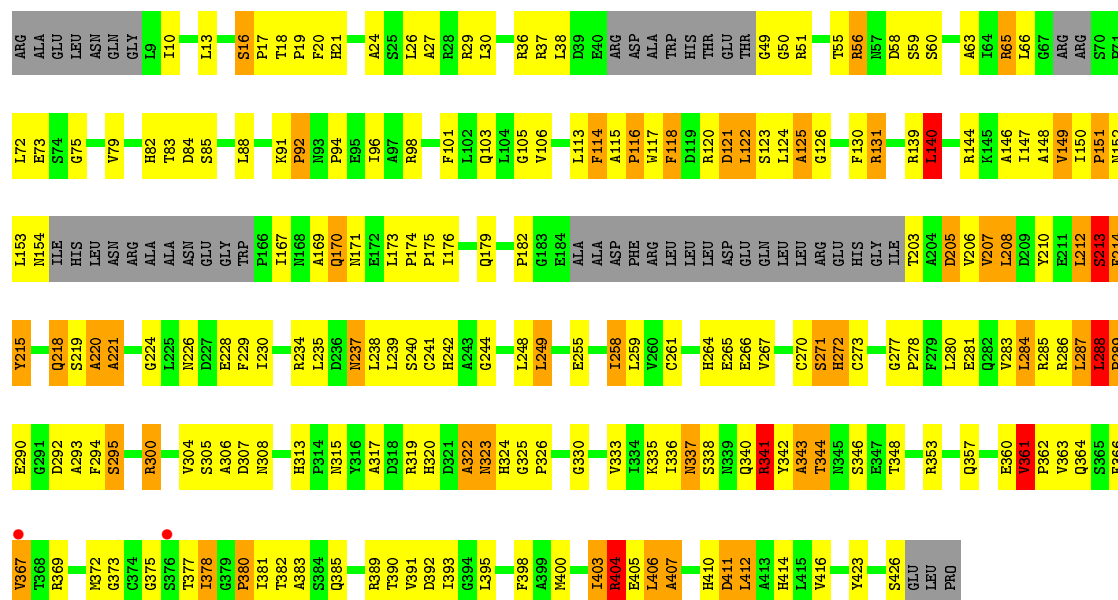
• Molecule 1: Probable M18-family aminopeptidase 2

Chain C: 39% 38% 10% 11%



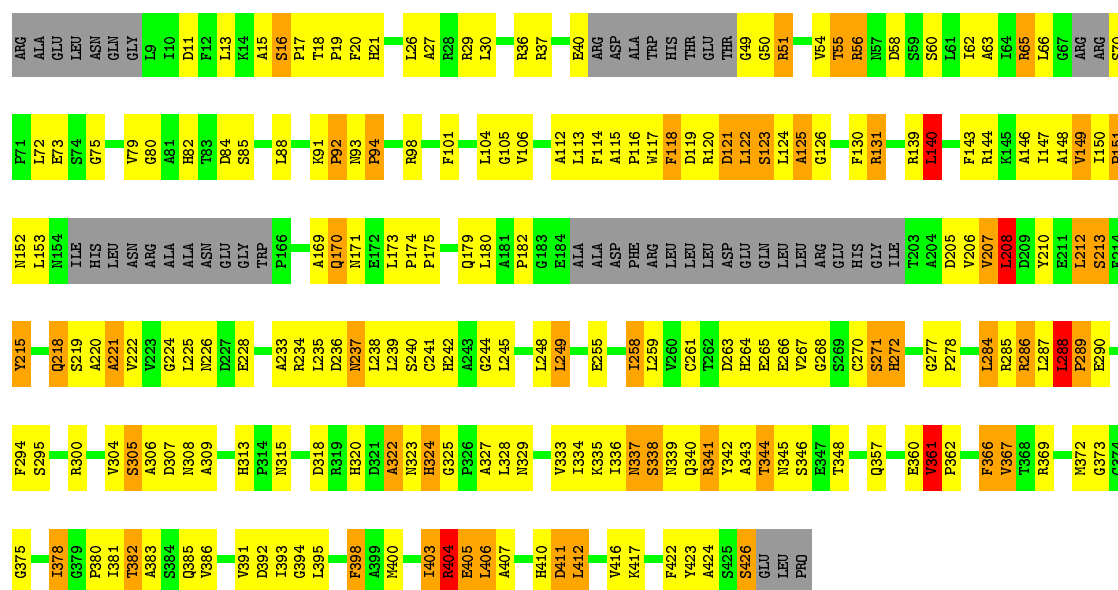
• Molecule 1: Probable M18-family aminopeptidase 2

Chain D: 40% 37% 11% 11%



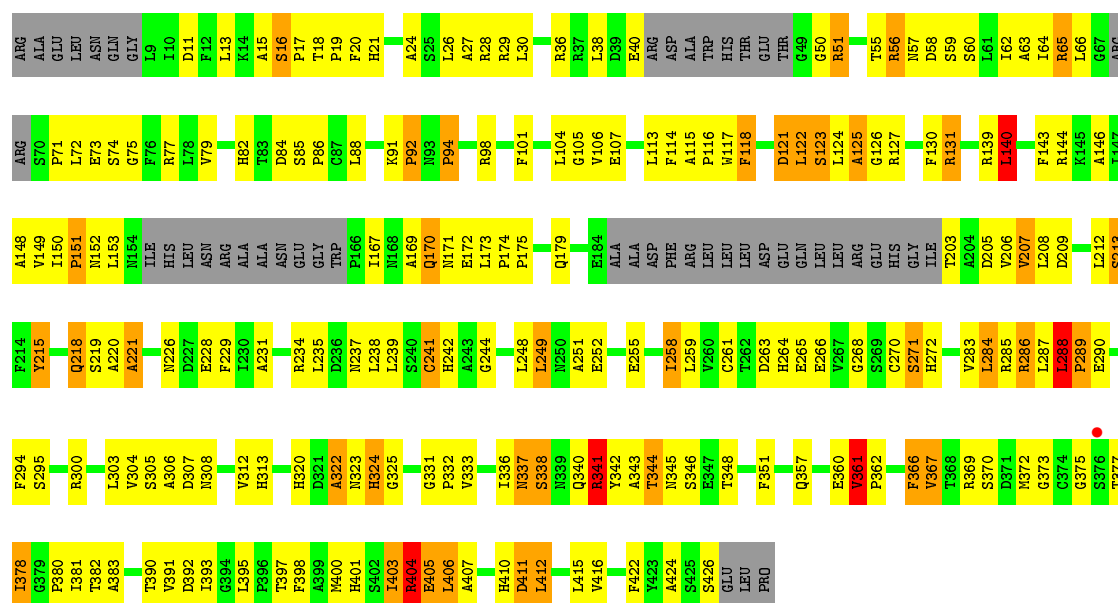
• Molecule 1: Probable M18-family aminopeptidase 2

Chain E: 



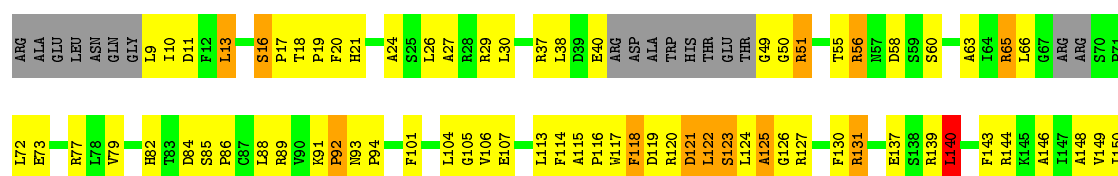
- Molecule 1: Probable M18-family aminopeptidase 2

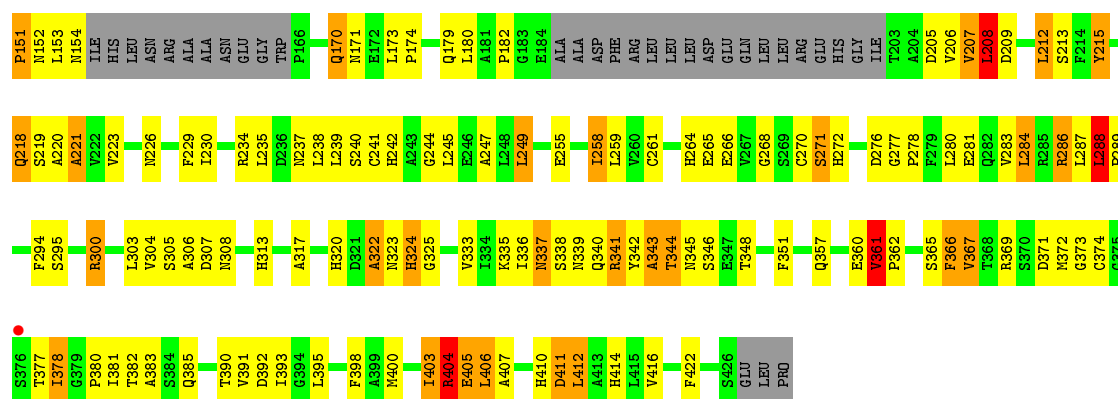
Chain F: 



- Molecule 1: Probable M18-family aminopeptidase 2

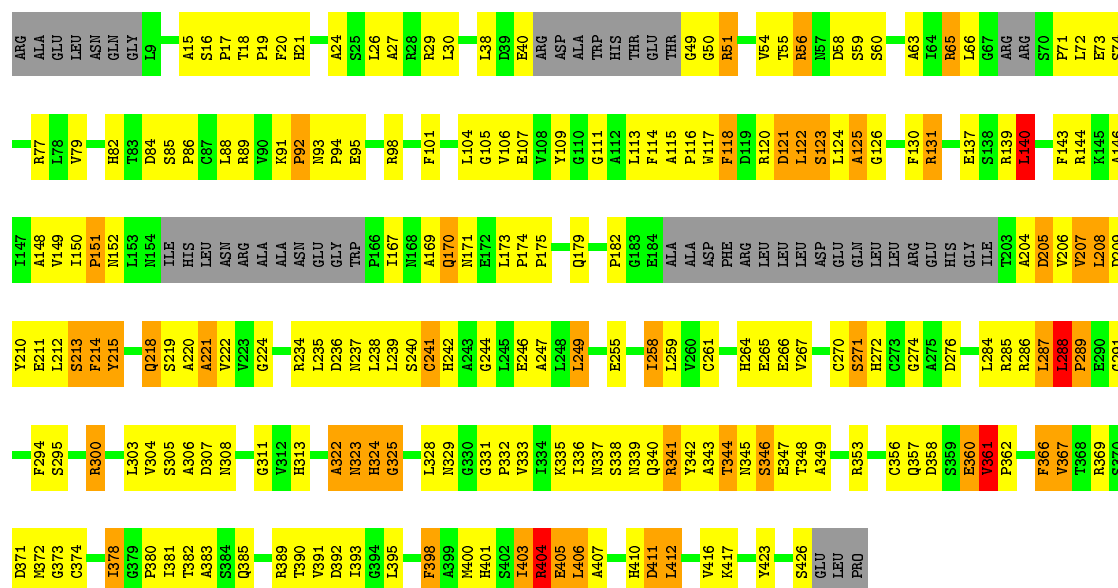
Chain G: 





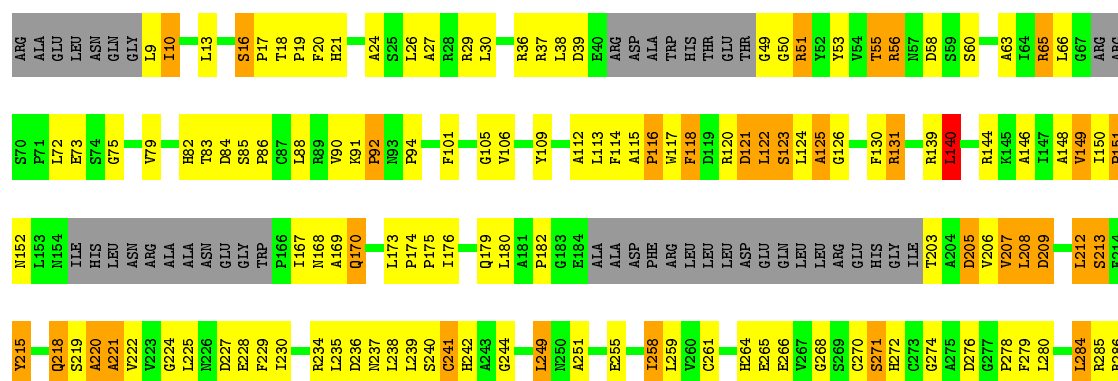
• Molecule 1: Probable M18-family aminopeptidase 2

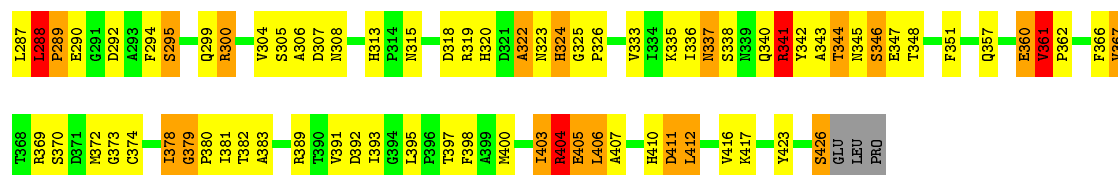
Chain H: 39% 39% 10% 11%



• Molecule 1: Probable M18-family aminopeptidase 2

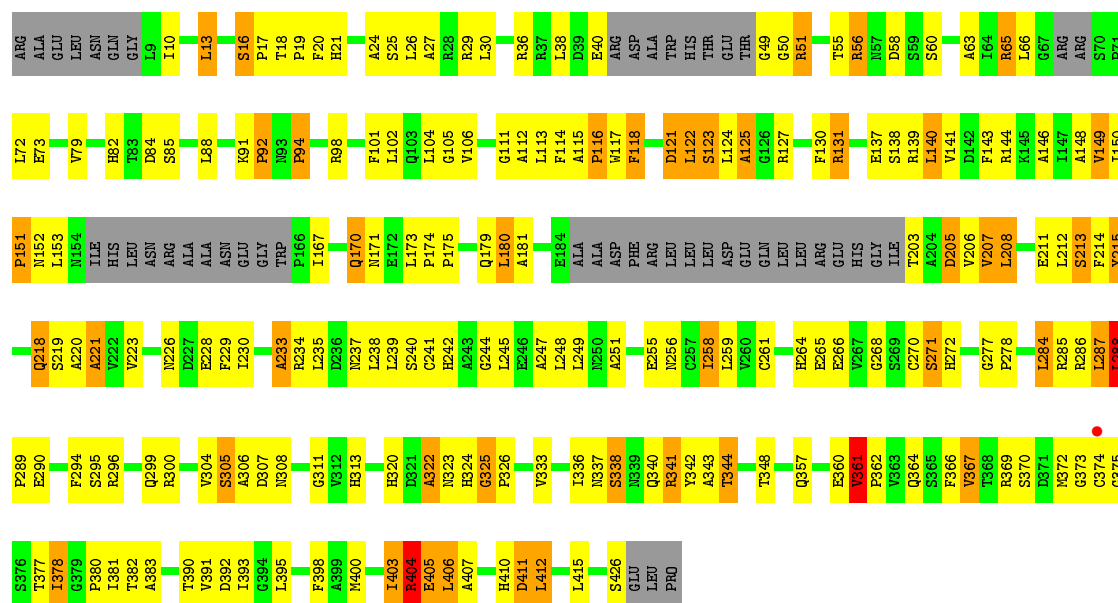
Chain I: 39% 36% 12% 11%





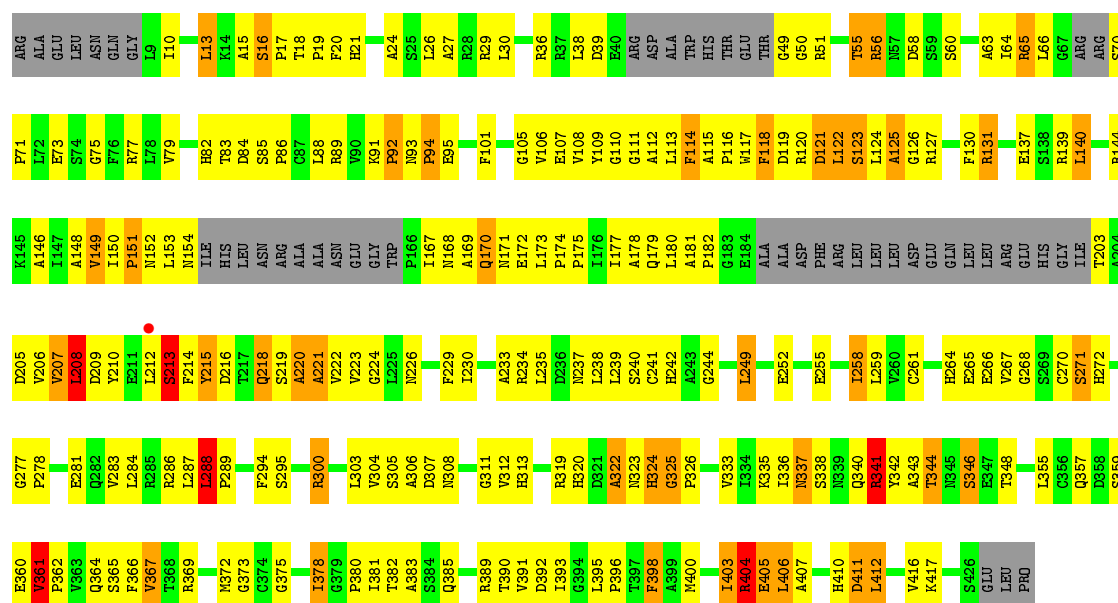
• Molecule 1: Probable M18-family aminopeptidase 2

Chain J: 42% 36% 10% 11%

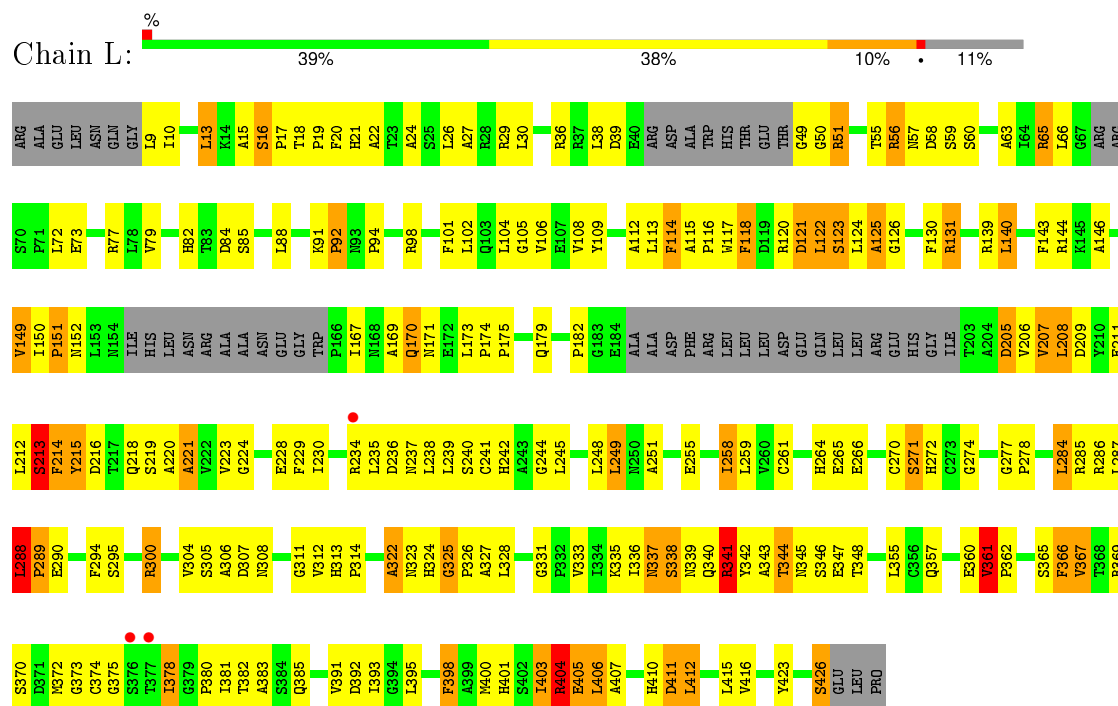


• Molecule 1: Probable M18-family aminopeptidase 2

Chain K: 36% 41% 10% 11%



• Molecule 1: Probable M18-family aminopeptidase 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	134.68Å 134.55Å 134.60Å 60.12° 60.10° 60.16°	Depositor
Resolution (Å)	20.00 – 3.00 20.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	52.9 (20.00-3.00) 65.2 (20.00-2.30)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.13 (at 2.30Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.255 , 0.291 0.330 , 0.360	Depositor DCC
R_{free} test set	6091 reflections (7.35%)	DCC
Wilson B-factor (Å ²)	16.5	Xtriage
Anisotropy	0.180	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 34.5	EDS
Estimated twinning fraction	0.033 for h-l,h,h-k 0.033 for k,k-l,-h+k 0.030 for l,-h+l,-k+l 0.030 for h-k,h-l,h 0.031 for -k+l,l,-h+l 0.031 for k-l,-h+k,k 0.408 for k,l,h 0.408 for l,h,k 0.417 for h-k,-k+l,-k 0.417 for h-l,-l,k-l 0.409 for -k,h-k,-k+l 0.409 for -h+k,-h,-h+l 0.407 for -h+l,-h+k,-h 0.407 for -l,k-l,h-l 0.032 for -h+k,k,k-l 0.033 for -h+l,-k+l,l 0.029 for h,h-k,h-l 0.032 for -h,-l,-k 0.033 for -k,-h,-l 0.033 for -l,-k,-h 0.427 for k-l,h-l,-l 0.427 for -k+l,-k,h-k 0.419 for -h,-h+l,-h+k	Xtriage

Continued on next page...

¹Intensities estimated from amplitudes.

Continued from previous page...

Property	Value	Source
L-test for twinning ²	$\langle L \rangle = 0.50, \langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 329782 reflections	Xtriage
F_o, F_c correlation	0.62	EDS
Total number of atoms	37617	wwPDB-VP
Average B, all atoms (\AA^2)	8.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.08% of the height of the origin peak. No significant pseudotranslation is detected.*

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 [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	0.41	2/2926 (0.1%)	0.57	0/3970
1	B	0.64	1/2926 (0.0%)	0.59	1/3970 (0.0%)
1	C	0.37	2/2926 (0.1%)	0.58	0/3970
1	D	0.35	0/2926	0.63	2/3970 (0.1%)
1	E	0.75	1/2926 (0.0%)	0.59	1/3970 (0.0%)
1	F	0.39	1/2926 (0.0%)	0.57	1/3970 (0.0%)
1	G	0.35	0/2926	0.58	0/3970
1	H	0.42	1/2926 (0.0%)	0.58	1/3970 (0.0%)
1	I	0.58	1/2926 (0.0%)	0.58	0/3970
1	J	0.44	1/2926 (0.0%)	0.59	1/3970 (0.0%)
1	K	0.37	1/2926 (0.0%)	0.57	0/3970
1	L	0.61	1/2926 (0.0%)	0.58	1/3970 (0.0%)
All	All	0.49	12/35112 (0.0%)	0.58	8/47640 (0.0%)

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	C	0	1
1	D	0	2
1	E	0	1
1	G	0	1
1	I	0	1
1	L	0	1
All	All	0	7

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	426	SER	C-O	35.87	1.91	1.23
1	B	426	SER	C-O	29.30	1.79	1.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	426	SER	C-O	27.37	1.75	1.23
1	I	426	SER	C-O	25.28	1.71	1.23
1	J	426	SER	C-O	15.10	1.52	1.23

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	426	SER	CA-C-O	15.91	153.50	120.10
1	B	426	SER	CA-C-O	-8.93	101.35	120.10
1	J	426	SER	CA-C-O	-8.68	101.87	120.10
1	L	426	SER	CA-C-O	-7.23	104.91	120.10
1	E	426	SER	CA-C-O	-6.53	106.38	120.10

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	212	LEU	Peptide
1	D	212	LEU	Peptide
1	D	213	SER	Peptide
1	E	212	LEU	Peptide
1	G	212	LEU	Peptide

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	2870	0	2800	237	0
1	B	2870	0	2800	242	0
1	C	2870	0	2800	225	0
1	D	2870	0	2800	228	0
1	E	2870	0	2800	240	0
1	F	2870	0	2800	225	0
1	G	2870	0	2800	206	0
1	H	2870	0	2800	232	0
1	I	2870	0	2800	259	0
1	J	2870	0	2800	227	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	K	2870	0	2800	236	0
1	L	2870	0	2800	235	0
2	A	262	0	0	132	0
2	B	271	0	0	117	0
2	C	298	0	0	136	0
2	D	247	0	0	105	0
2	E	268	0	0	124	0
2	F	247	0	0	111	0
2	G	255	0	0	93	0
2	H	252	0	0	117	0
2	I	274	0	0	152	0
2	J	258	0	0	129	0
2	K	276	0	0	123	0
2	L	269	0	0	118	0
All	All	37617	0	33600	2700	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 40.

The worst 5 of 2700 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:L:230:ILE:CD1	2:L:582:HOH:O	1.71	1.33
1:L:230:ILE:HD13	2:L:582:HOH:O	1.23	1.33
1:F:13:LEU:HD22	2:F:655:HOH:O	1.29	1.32
1:I:361:VAL:HG23	2:I:581:HOH:O	1.26	1.32
1:K:181:ALA:HB3	2:K:550:HOH:O	1.13	1.30

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	369/428 (86%)	246 (67%)	82 (22%)	41 (11%)	0	2
1	B	369/428 (86%)	241 (65%)	84 (23%)	44 (12%)	0	2
1	C	369/428 (86%)	245 (66%)	81 (22%)	43 (12%)	0	2
1	D	369/428 (86%)	240 (65%)	86 (23%)	43 (12%)	0	2
1	E	369/428 (86%)	243 (66%)	84 (23%)	42 (11%)	0	2
1	F	369/428 (86%)	243 (66%)	85 (23%)	41 (11%)	0	2
1	G	369/428 (86%)	242 (66%)	90 (24%)	37 (10%)	1	3
1	H	369/428 (86%)	245 (66%)	86 (23%)	38 (10%)	1	3
1	I	369/428 (86%)	240 (65%)	88 (24%)	41 (11%)	0	2
1	J	369/428 (86%)	241 (65%)	85 (23%)	43 (12%)	0	2
1	K	369/428 (86%)	244 (66%)	83 (22%)	42 (11%)	0	2
1	L	369/428 (86%)	244 (66%)	83 (22%)	42 (11%)	0	2
All	All	4428/5136 (86%)	2914 (66%)	1017 (23%)	497 (11%)	0	2

5 of 497 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	116	PRO
1	A	125	ALA
1	A	131	ARG
1	A	207	VAL
1	A	214	PHE

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	303/344 (88%)	268 (88%)	35 (12%)	7	27
1	B	303/344 (88%)	266 (88%)	37 (12%)	6	25
1	C	303/344 (88%)	266 (88%)	37 (12%)	6	25
1	D	303/344 (88%)	269 (89%)	34 (11%)	7	29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	303/344 (88%)	266 (88%)	37 (12%)	6	25
1	F	303/344 (88%)	268 (88%)	35 (12%)	7	27
1	G	303/344 (88%)	267 (88%)	36 (12%)	6	26
1	H	303/344 (88%)	268 (88%)	35 (12%)	7	27
1	I	303/344 (88%)	264 (87%)	39 (13%)	5	23
1	J	303/344 (88%)	269 (89%)	34 (11%)	7	29
1	K	303/344 (88%)	266 (88%)	37 (12%)	6	25
1	L	303/344 (88%)	268 (88%)	35 (12%)	7	27
All	All	3636/4128 (88%)	3205 (88%)	431 (12%)	6	26

5 of 431 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	284	LEU
1	G	366	PHE
1	L	65	ARG
1	F	324	HIS
1	G	65	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 141 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	339	ASN
1	G	315	ASN
1	L	152	ASN
1	E	385	GLN
1	F	299	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 [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	379/428 (88%)	-0.41	0	100	100	2, 4, 30, 51	0
1	B	379/428 (88%)	-0.40	0	100	100	2, 3, 29, 49	0
1	C	379/428 (88%)	-0.38	0	100	100	2, 3, 30, 56	0
1	D	379/428 (88%)	-0.42	2 (0%)	91	76	2, 3, 30, 46	0
1	E	379/428 (88%)	-0.40	0	100	100	2, 4, 29, 45	0
1	F	379/428 (88%)	-0.41	1 (0%)	94	84	2, 4, 28, 63	0
1	G	379/428 (88%)	-0.41	1 (0%)	94	84	2, 4, 29, 42	0
1	H	379/428 (88%)	-0.42	0	100	100	2, 3, 27, 51	0
1	I	379/428 (88%)	-0.38	0	100	100	2, 4, 32, 57	0
1	J	379/428 (88%)	-0.40	1 (0%)	94	84	2, 3, 25, 50	0
1	K	379/428 (88%)	-0.41	1 (0%)	94	84	2, 3, 27, 65	0
1	L	379/428 (88%)	-0.40	3 (0%)	87	67	2, 3, 29, 55	0
All	All	4548/5136 (88%)	-0.40	9 (0%)	95	87	2, 4, 29, 65	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	376	SER	2.6
1	G	376	SER	2.5
1	J	374	CYS	2.5
1	K	212	LEU	2.4
1	L	377	THR	2.2

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.