wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:51 PM GMT

PDB ID	:	1QXP
Title	:	Crystal Structure of a mu-like calpain
Authors	:	Pal, G.P.; Veyra, T.D.; Elce, J.S.; Jia, Z.
Deposited on	:	2003-09-08
Resolution	:	2.80 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:

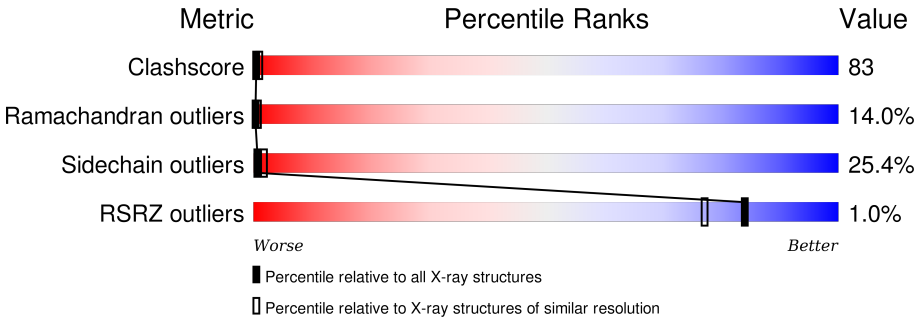
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.80 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))
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	900	% 5% 26% 38% 19% 13%
1	B	900	% 6% 27% 36% 19% 12%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12368 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 mu-like calpain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	783	Total	C	N	O	S	0	0	0
			6053	3846	1037	1143	27			
1	B	788	Total	C	N	O	S	0	0	0
			6003	3830	1015	1129	29			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	105	SER	CYS	ENGINEERED	UNP P97571
B	105	SER	CYS	ENGINEERED	UNP P97571
A	702A	GLY	-	CLONING ARTIFACT	UNP Q07009
A	702B	LYS	-	CLONING ARTIFACT	UNP Q07009
A	702C	LEU	-	CLONING ARTIFACT	UNP Q07009
A	702D	ALA	-	CLONING ARTIFACT	UNP Q07009
A	702E	ALA	-	CLONING ARTIFACT	UNP Q07009
A	702F	ALA	-	CLONING ARTIFACT	UNP Q07009
A	702G	ILE	-	CLONING ARTIFACT	UNP Q07009
A	702H	GLU	-	CLONING ARTIFACT	UNP Q07009
A	702I	HIS	-	EXPRESSION TAG	UNP Q07009
A	702J	HIS	-	EXPRESSION TAG	UNP Q07009
A	702K	HIS	-	EXPRESSION TAG	UNP Q07009
A	702L	HIS	-	EXPRESSION TAG	UNP Q07009
A	702M	HIS	-	EXPRESSION TAG	UNP Q07009
A	702N	HIS	-	EXPRESSION TAG	UNP Q07009
B	702A	GLY	-	CLONING ARTIFACT	UNP Q07009
B	702B	LYS	-	CLONING ARTIFACT	UNP Q07009
B	702C	LEU	-	CLONING ARTIFACT	UNP Q07009
B	702D	ALA	-	CLONING ARTIFACT	UNP Q07009
B	702E	ALA	-	CLONING ARTIFACT	UNP Q07009
B	702F	ALA	-	CLONING ARTIFACT	UNP Q07009
B	702G	ILE	-	CLONING ARTIFACT	UNP Q07009
B	702H	GLU	-	CLONING ARTIFACT	UNP Q07009
B	702I	HIS	-	EXPRESSION TAG	UNP Q07009

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	702J	HIS	-	EXPRESSION TAG	UNP Q07009
B	702K	HIS	-	EXPRESSION TAG	UNP Q07009
B	702L	HIS	-	EXPRESSION TAG	UNP Q07009
B	702M	HIS	-	EXPRESSION TAG	UNP Q07009
B	702N	HIS	-	EXPRESSION TAG	UNP Q07009

- Molecule 2 is water.

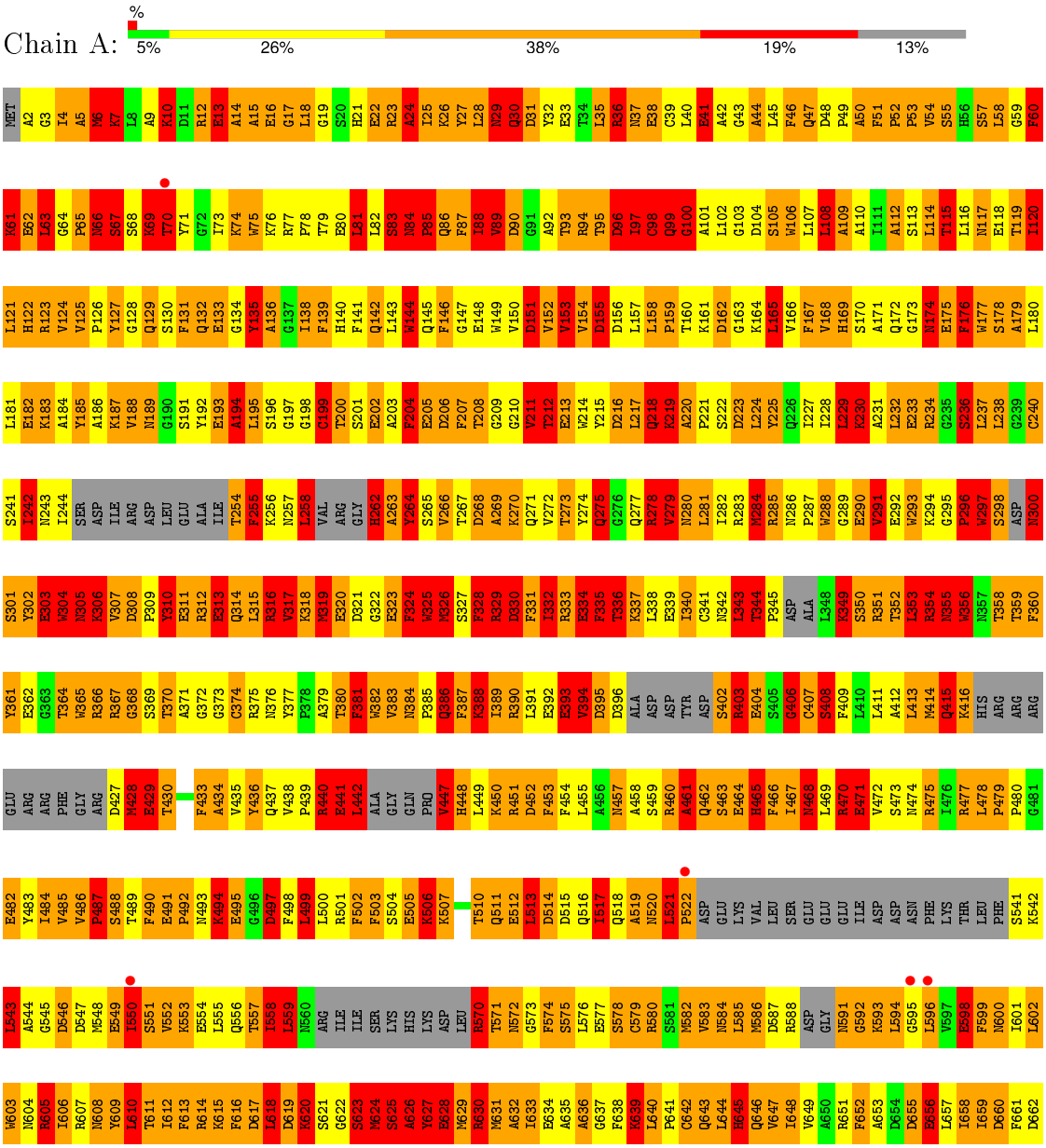
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	172	Total O 172 172	0	0
2	B	140	Total O 140 140	0	0

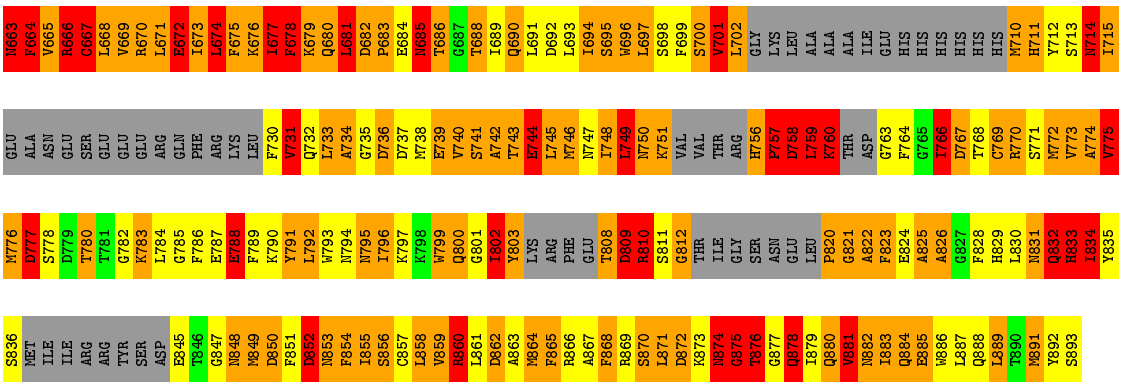
3 Residue-property plots (i)

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: mu-like calpain

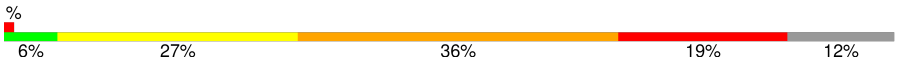
Chain A:

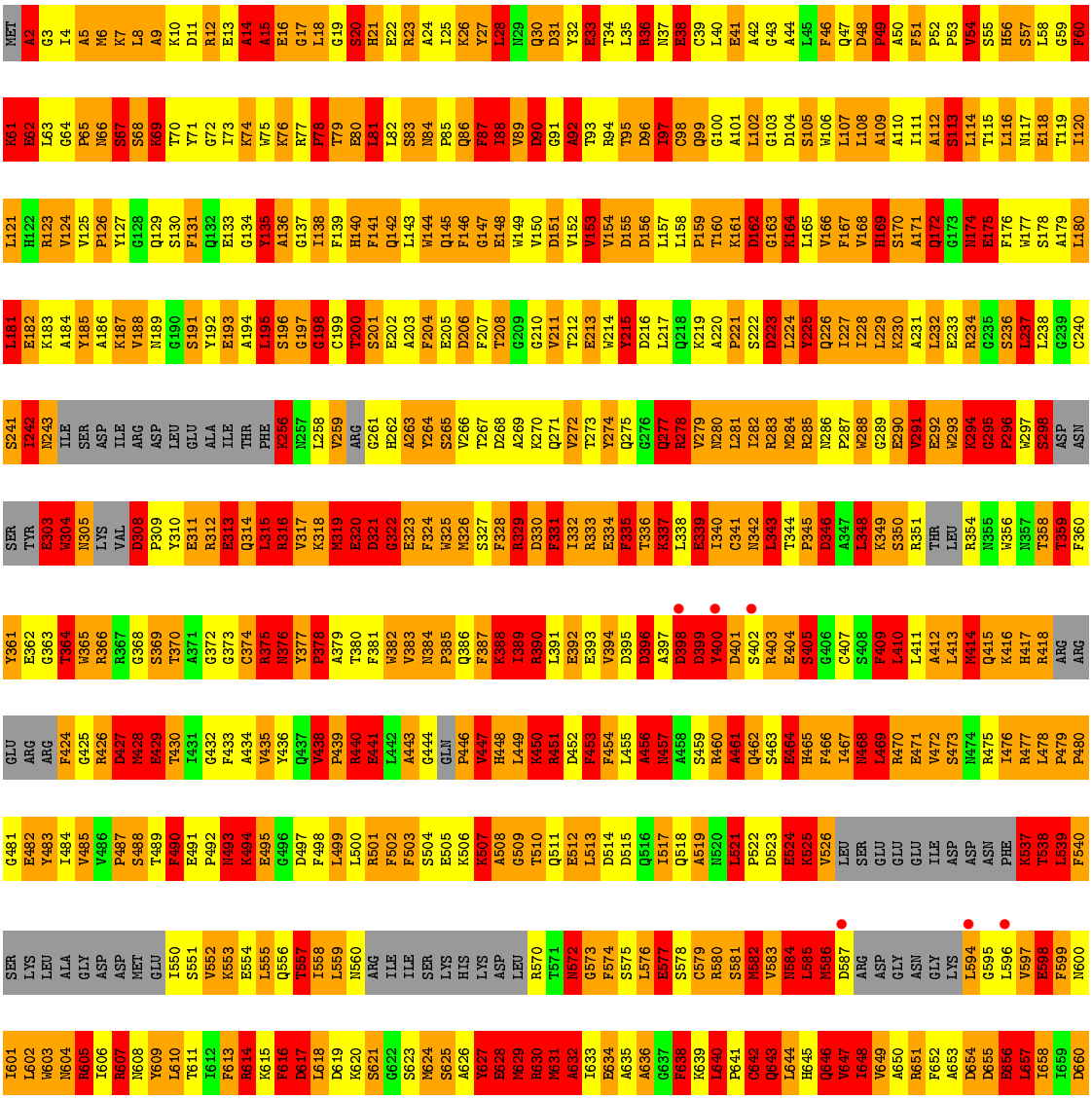


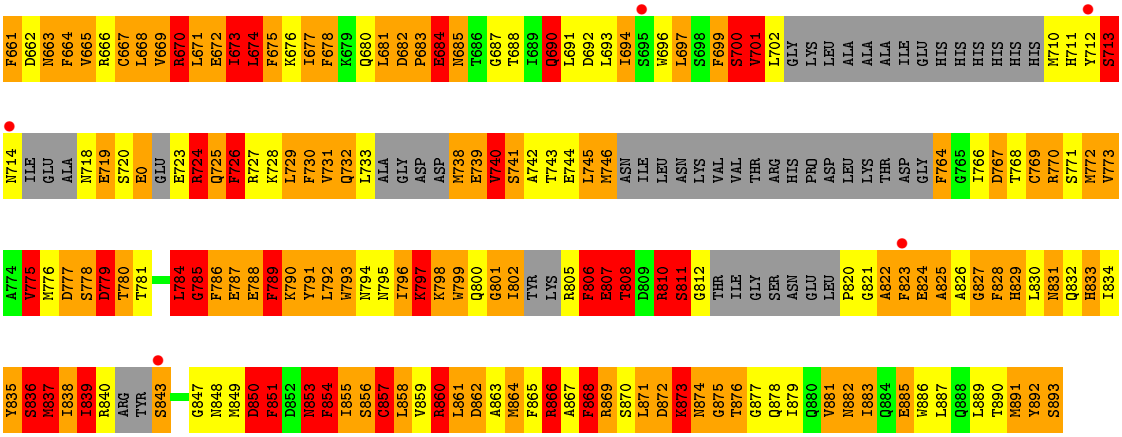


- Molecule 1: mu-like calpain

Chain B:






F661 D662 M663 F664 V665 R666 C667 L668 V669 R670 L671 E672 I673 L674 F675 K676 I677 F678 K679 Q680 L681 D682 P683 E684 N685 T686 G687 T688 I689 Q690 L691 D692 L693 I694 S695 W696 L697 S698 F699 S700 V701 L702 GLY LYS LEU ALA ALA ILE GLU HIS HIS HIS HIS HIS HIS M710 H711 Y712 S713
N714 ILE GLU ALA N718 E719 S720 E0 GLU E723 R724 Q725 F726 R727 K728 L729 F730 V731 Q732 L733 ALA GLY ASP ASP M738 E739 V740 S741 A742 T743 E744 L745 M746 ASN ILE LEU ASN LYS VAL VAL THR ARG HIS PRO ASP LEU LYS THR ASP GLY F764 G765 I766 D767 T768 C769 R770 S771 M772 V773
A774 V775 M776 D777 S778 D779 T780 T781 L784 G785 F786 E787 E788 F789 K790 Y791 L792 W793 N794 N795 I796 K797 K798 W799 Q800 G801 I802 TYR LYS R805 F806 E807 T808 D809 R810 S811 G812 THR ILE GLY SER ASN GLU LEU P820 G821 A822 F823 E824 A825 A826 G827 F828 H829 L830 N831 Q832 H833 I834
Y835 S836 M837 I838 I839 R840 ARG TYR S843 G847 N848 M849 D850 F851 D852 N853 F954 I855 S856 C957 L858 V859 R860 L861 D862 A863 M864 F865 R866 A867 F868 R869 S870 L871 D872 K873 N874 G875 T876 G877 Q878 I879 Q880 V881 N882 I883 Q884 E885 W886 L887 Q888 L889 T890 M891 Y892 S893

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	72.74Å 184.60Å 86.37Å 90.00° 100.74° 90.00°	Depositor
Resolution (Å)	91.29 – 2.80 49.81 – 2.69	Depositor EDS
% Data completeness (in resolution range)	91.6 (91.29-2.80) 87.3 (49.81-2.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.04 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.229 , 0.311 0.232 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	55.6	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 98.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Outliers	0 of 55542 reflections	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	12368	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.85% 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	3.70	964/6177 (15.6%)	2.67	470/8354 (5.6%)
1	B	3.71	932/6128 (15.2%)	2.67	489/8288 (5.9%)
All	All	3.70	1896/12305 (15.4%)	2.67	959/16642 (5.8%)

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	45
1	B	0	39
All	All	0	84

The worst 5 of 1896 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	67	SER	CA-CB	28.73	1.96	1.52
1	A	339	GLU	CD-OE1	23.66	1.51	1.25
1	A	320	GLU	CD-OE2	20.06	1.47	1.25
1	B	429	GLU	CD-OE1	19.54	1.47	1.25
1	A	811	SER	CA-CB	-18.37	1.25	1.52

The worst 5 of 959 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	460	ARG	NE-CZ-NH1	-29.77	105.41	120.30
1	A	329	ARG	NE-CZ-NH1	26.14	133.37	120.30
1	A	329	ARG	NE-CZ-NH2	-24.17	108.22	120.30
1	B	514	ASP	CB-CG-OD2	-18.56	101.59	118.30
1	A	285	ARG	NE-CZ-NH2	-15.96	112.32	120.30

There are no chirality outliers.

5 of 84 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	120	ILE	Mainchain
1	A	152	VAL	Mainchain
1	A	24	ALA	Mainchain
1	A	29	ASN	Mainchain
1	A	99	GLN	Mainchain,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	6053	0	5615	959	3
1	B	6003	0	5477	969	1
2	A	172	0	0	53	1
2	B	140	0	0	45	1
All	All	12368	0	11092	1928	3

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 83.

The worst 5 of 1928 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:583:VAL:CB	1:A:583:VAL:CA	1.74	1.65
1:B:0:GLU:CA	1:B:0:GLU:CB	1.75	1.64
1:B:786:PHE:CA	1:B:786:PHE:CB	1.74	1.64
1:A:550:ILE:CB	1:A:550:ILE:CA	1.75	1.64
1:B:8:LEU:CD2	1:B:8:LEU:CG	1.76	1.63

All (3) 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:457:ASN:OD1	2:B:1033:HOH:O[1_454]	1.56	0.64
1:A:442:LEU:CD2	2:A:922:HOH:O[1_455]	1.95	0.25

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:300:ASN:OD1	1:B:303:GLU:N[2_656]	2.08	0.12

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	747/900 (83%)	500 (67%)	143 (19%)	104 (14%)	0 1
1	B	748/900 (83%)	497 (66%)	146 (20%)	105 (14%)	0 1
All	All	1495/1800 (83%)	997 (67%)	289 (19%)	209 (14%)	0 1

5 of 209 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	54	VAL
1	A	60	PHE
1	A	63	LEU
1	A	67	SER
1	A	83	SER

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	606/782 (78%)	459 (76%)	147 (24%)	1 2
1	B	584/782 (75%)	429 (74%)	155 (26%)	0 1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1190/1564 (76%)	888 (75%)	302 (25%)	1 2

5 of 302 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	853	ASN
1	B	111	ILE
1	B	701	VAL
1	A	874	ASN
1	B	41	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	853	ASN
1	B	47	GLN
1	B	646	GLN
1	A	878	GLN
1	A	888	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

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	783/900 (87%)	-0.35	5 (0%)	90	86	16, 49, 81, 98	0
1	B	788/900 (87%)	-0.32	11 (1%)	78	69	17, 51, 84, 102	0
All	All	1571/1800 (87%)	-0.33	16 (1%)	84	77	16, 50, 83, 102	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	587	ASP	3.8
1	B	596	LEU	3.8
1	A	596	LEU	3.4
1	B	712	TYR	3.3
1	B	402	SER	3.3

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.