



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 09:29 AM GMT

PDB ID : 3IO4
Title : Huntingtin amino-terminal region with 17 Gln residues - Crystal C90
Authors : Kim, M.W.; Chelliah, Y.; Kim, S.W.; Otwinowski, Z.; Bezprozvanny, I.
Deposited on : 2009-08-13
Resolution : 3.63 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

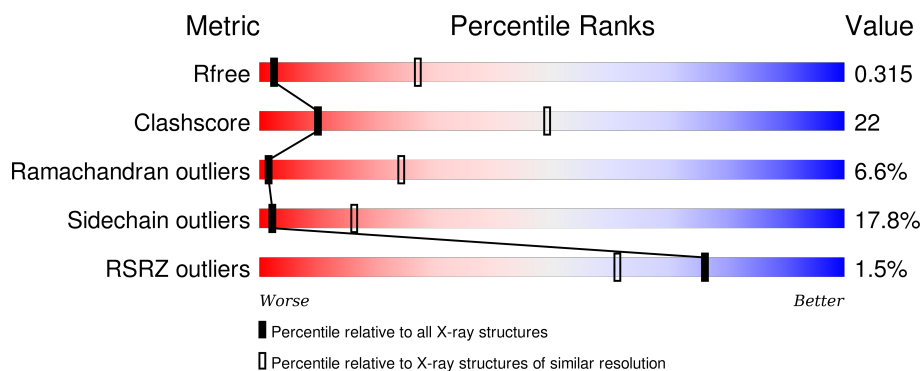
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.63 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	1014 (3.80-3.48)
Clashscore	102246	1130 (3.80-3.48)
Ramachandran outliers	100387	1084 (3.80-3.48)
Sidechain outliers	100360	1083 (3.80-3.48)
RSRZ outliers	91569	1021 (3.80-3.48)

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	449	<div> <div>%</div> <div> <div></div> <div>46%</div> <div>37%</div> <div>9%</div> <div>7%</div> </div> </div>
1	B	449	<div> <div>%</div> <div> <div></div> <div>43%</div> <div>33%</div> <div>10%</div> <div>•</div> <div>12%</div> </div> </div>
1	C	449	<div> <div>%</div> <div> <div></div> <div>51%</div> <div>32%</div> <div>5%</div> <div>•</div> <div>11%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9331 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 Maltose-binding protein, huntingtin fusion protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	417	Total	C	N	O	S	4	0	0
			3176	2046	520	602	8			
1	B	397	Total	C	N	O	S	20	1	0
			3076	1974	506	588	8			
1	C	398	Total	C	N	O	S	8	0	0
			3065	1969	503	585	8			

There are 111 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	359	ALA	-	LINKER	UNP P0AEX9
A	360	ALA	-	LINKER	UNP P0AEX9
A	361	LEU	-	LINKER	UNP P0AEX9
A	362	ALA	-	LINKER	UNP P0AEX9
A	363	ALA	-	LINKER	UNP P0AEX9
A	364	ALA	-	LINKER	UNP P0AEX9
A	365	GLN	-	LINKER	UNP P0AEX9
A	366	THR	-	LINKER	UNP P0AEX9
A	367	ASN	-	LINKER	UNP P0AEX9
A	368	ALA	-	LINKER	UNP P0AEX9
A	369	ALA	-	LINKER	UNP P0AEX9
A	370	ALA	-	LINKER	UNP P0AEX9
A	?	-	GLN	DELETION	UNP P42858
A	?	-	GLN	DELETION	UNP P42858
A	?	-	GLN	DELETION	UNP P42858
A	?	-	GLN	DELETION	UNP P42858
A	?	-	GLN	DELETION	UNP P42858
A	?	-	GLN	DELETION	UNP P42858
A	431	GLN	-	EXPRESSION TAG	UNP P42858
A	432	SER	-	EXPRESSION TAG	UNP P42858
A	433	TYR	-	EXPRESSION TAG	UNP P42858
A	434	GLN	-	EXPRESSION TAG	UNP P42858
A	435	ILE	-	EXPRESSION TAG	UNP P42858

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	436	THR	-	EXPRESSION TAG	UNP P42858
A	437	ALA	-	EXPRESSION TAG	UNP P42858
A	438	GLY	-	EXPRESSION TAG	UNP P42858
A	439	LYS	-	EXPRESSION TAG	UNP P42858
A	440	LEU	-	EXPRESSION TAG	UNP P42858
A	441	GLY	-	EXPRESSION TAG	UNP P42858
A	442	THR	-	EXPRESSION TAG	UNP P42858
A	443	GLY	-	EXPRESSION TAG	UNP P42858
A	444	ARG	-	EXPRESSION TAG	UNP P42858
A	445	ARG	-	EXPRESSION TAG	UNP P42858
A	446	PHE	-	EXPRESSION TAG	UNP P42858
A	447	THR	-	EXPRESSION TAG	UNP P42858
A	448	THR	-	EXPRESSION TAG	UNP P42858
A	449	SER	-	EXPRESSION TAG	UNP P42858
B	359	ALA	-	LINKER	UNP P0AEX9
B	360	ALA	-	LINKER	UNP P0AEX9
B	361	LEU	-	LINKER	UNP P0AEX9
B	362	ALA	-	LINKER	UNP P0AEX9
B	363	ALA	-	LINKER	UNP P0AEX9
B	364	ALA	-	LINKER	UNP P0AEX9
B	365	GLN	-	LINKER	UNP P0AEX9
B	366	THR	-	LINKER	UNP P0AEX9
B	367	ASN	-	LINKER	UNP P0AEX9
B	368	ALA	-	LINKER	UNP P0AEX9
B	369	ALA	-	LINKER	UNP P0AEX9
B	370	ALA	-	LINKER	UNP P0AEX9
B	?	-	GLN	DELETION	UNP P42858
B	?	-	GLN	DELETION	UNP P42858
B	?	-	GLN	DELETION	UNP P42858
B	?	-	GLN	DELETION	UNP P42858
B	?	-	GLN	DELETION	UNP P42858
B	?	-	GLN	DELETION	UNP P42858
B	431	GLN	-	EXPRESSION TAG	UNP P42858
B	432	SER	-	EXPRESSION TAG	UNP P42858
B	433	TYR	-	EXPRESSION TAG	UNP P42858
B	434	GLN	-	EXPRESSION TAG	UNP P42858
B	435	ILE	-	EXPRESSION TAG	UNP P42858
B	436	THR	-	EXPRESSION TAG	UNP P42858
B	437	ALA	-	EXPRESSION TAG	UNP P42858
B	438	GLY	-	EXPRESSION TAG	UNP P42858
B	439	LYS	-	EXPRESSION TAG	UNP P42858
B	440	LEU	-	EXPRESSION TAG	UNP P42858

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	441	GLY	-	EXPRESSION TAG	UNP P42858
B	442	THR	-	EXPRESSION TAG	UNP P42858
B	443	GLY	-	EXPRESSION TAG	UNP P42858
B	444	ARG	-	EXPRESSION TAG	UNP P42858
B	445	ARG	-	EXPRESSION TAG	UNP P42858
B	446	PHE	-	EXPRESSION TAG	UNP P42858
B	447	THR	-	EXPRESSION TAG	UNP P42858
B	448	THR	-	EXPRESSION TAG	UNP P42858
B	449	SER	-	EXPRESSION TAG	UNP P42858
C	359	ALA	-	LINKER	UNP P0AEX9
C	360	ALA	-	LINKER	UNP P0AEX9
C	361	LEU	-	LINKER	UNP P0AEX9
C	362	ALA	-	LINKER	UNP P0AEX9
C	363	ALA	-	LINKER	UNP P0AEX9
C	364	ALA	-	LINKER	UNP P0AEX9
C	365	GLN	-	LINKER	UNP P0AEX9
C	366	THR	-	LINKER	UNP P0AEX9
C	367	ASN	-	LINKER	UNP P0AEX9
C	368	ALA	-	LINKER	UNP P0AEX9
C	369	ALA	-	LINKER	UNP P0AEX9
C	370	ALA	-	LINKER	UNP P0AEX9
C	?	-	GLN	DELETION	UNP P42858
C	?	-	GLN	DELETION	UNP P42858
C	?	-	GLN	DELETION	UNP P42858
C	?	-	GLN	DELETION	UNP P42858
C	?	-	GLN	DELETION	UNP P42858
C	?	-	GLN	DELETION	UNP P42858
C	431	GLN	-	EXPRESSION TAG	UNP P42858
C	432	SER	-	EXPRESSION TAG	UNP P42858
C	433	TYR	-	EXPRESSION TAG	UNP P42858
C	434	GLN	-	EXPRESSION TAG	UNP P42858
C	435	ILE	-	EXPRESSION TAG	UNP P42858
C	436	THR	-	EXPRESSION TAG	UNP P42858
C	437	ALA	-	EXPRESSION TAG	UNP P42858
C	438	GLY	-	EXPRESSION TAG	UNP P42858
C	439	LYS	-	EXPRESSION TAG	UNP P42858
C	440	LEU	-	EXPRESSION TAG	UNP P42858
C	441	GLY	-	EXPRESSION TAG	UNP P42858
C	442	THR	-	EXPRESSION TAG	UNP P42858
C	443	GLY	-	EXPRESSION TAG	UNP P42858
C	444	ARG	-	EXPRESSION TAG	UNP P42858
C	445	ARG	-	EXPRESSION TAG	UNP P42858

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	446	PHE	-	EXPRESSION TAG	UNP P42858
C	447	THR	-	EXPRESSION TAG	UNP P42858
C	448	THR	-	EXPRESSION TAG	UNP P42858
C	449	SER	-	EXPRESSION TAG	UNP P42858

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	3	Total Zn 3 3	0	0
2	A	3	Total Zn 3 3	0	0
2	C	2	Total Zn 2 2	0	0

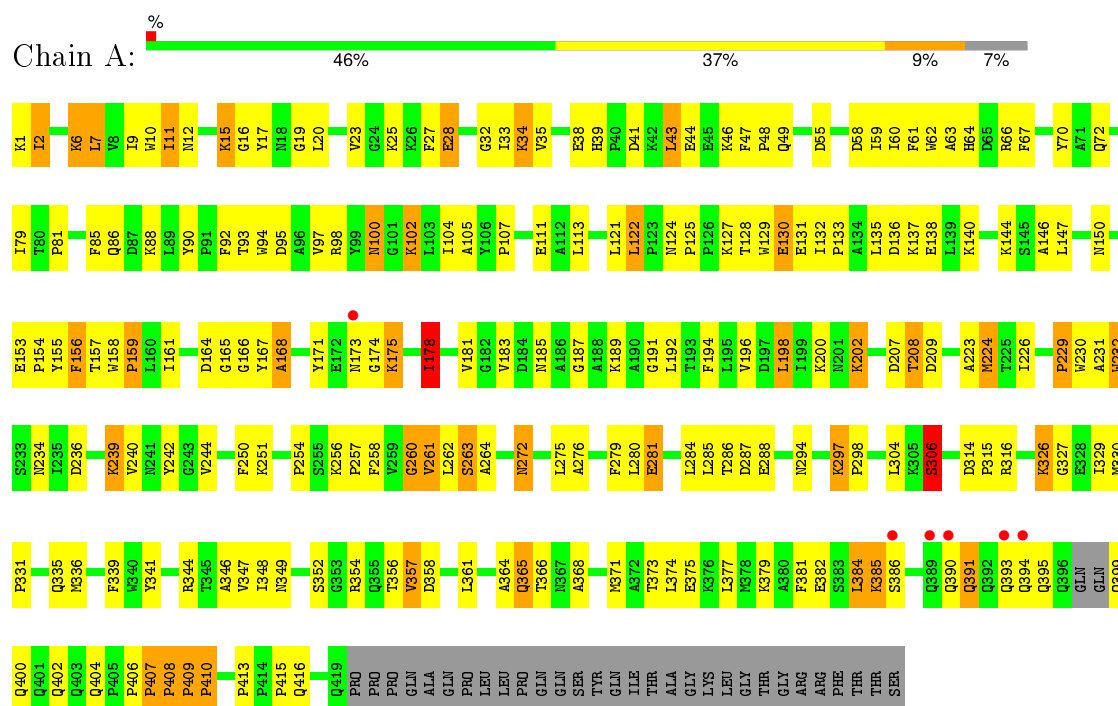
- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

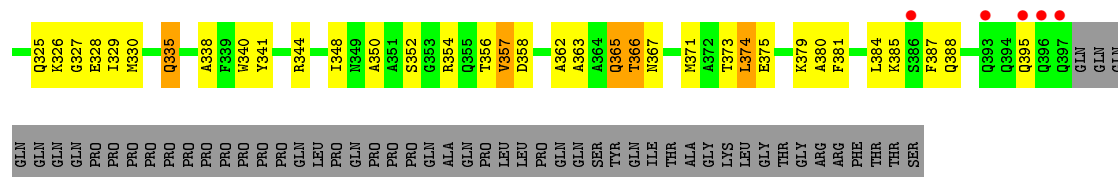
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	3	Total Ca 3 3	0	0
3	C	3	Total Ca 3 3	0	0

3 Residue-property plots

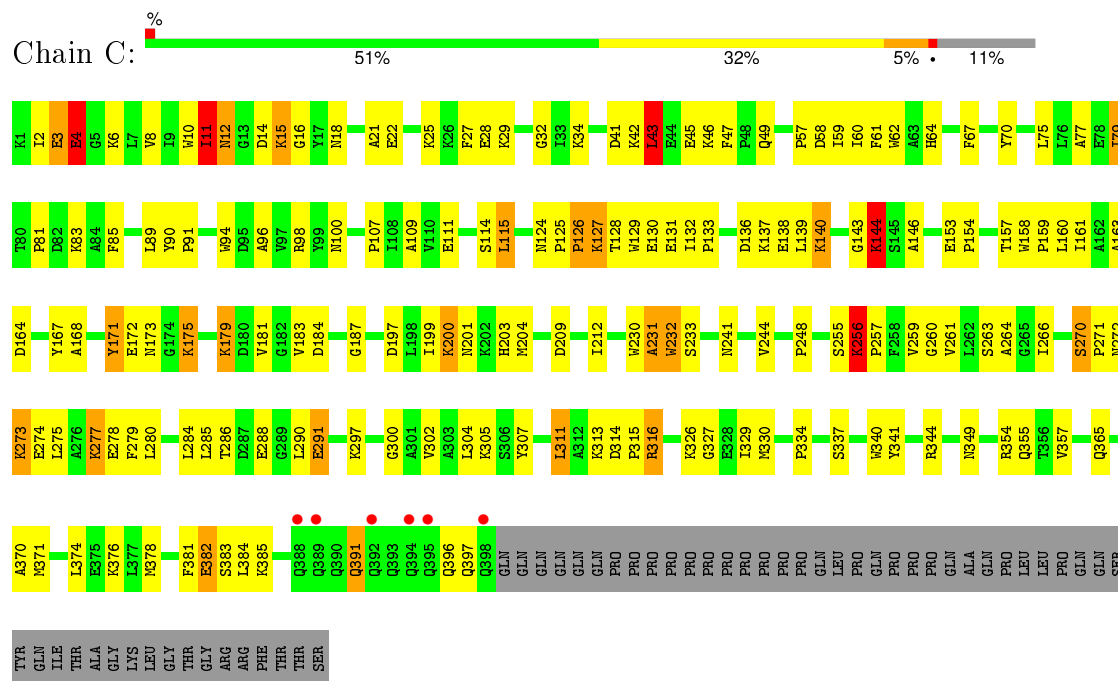
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: Maltose-binding protein, huntingtin fusion protein





• Molecule 1: Maltose-binding protein, huntingtin fusion protein



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	162.95Å 101.21Å 142.05Å 90.00° 90.42° 90.00°	Depositor
Resolution (Å)	37.42 – 3.63 35.51 – 3.63	Depositor EDS
% Data completeness (in resolution range)	64.4 (37.42-3.63) 64.4 (35.51-3.63)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.02	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.43 (at 3.66Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.251 , 0.304 0.261 , 0.315	Depositor DCC
R_{free} test set	842 reflections (5.24%)	DCC
Wilson B-factor (Å ²)	135.3	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.21 , 24.1	EDS
Estimated twinning fraction	0.041 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 16917 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9331	wwPDB-VP
Average B, all atoms (Å ²)	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.22% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.50	1/3258 (0.0%)	0.60	1/4436 (0.0%)
1	B	0.81	7/3147 (0.2%)	0.73	8/4270 (0.2%)
1	C	0.63	1/3136 (0.0%)	0.59	2/4256 (0.0%)
All	All	0.66	9/9541 (0.1%)	0.64	11/12962 (0.1%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	15	LYS	CB-CG	-25.50	0.83	1.52
1	B	140	LYS	CB-CG	20.19	2.07	1.52
1	B	15	LYS	CB-CG	-19.85	0.98	1.52
1	B	111	GLU	CB-CG	-18.43	1.17	1.52
1	A	202	LYS	CB-CG	-12.70	1.18	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	46	LYS	CA-CB-CG	16.52	149.75	113.40
1	B	46	LYS	CB-CG-CD	-13.13	77.47	111.60
1	B	111	GLU	CA-CB-CG	9.89	135.16	113.40
1	B	140	LYS	CA-CB-CG	-9.55	92.39	113.40
1	B	15	LYS	CB-CG-CD	-8.72	88.94	111.60

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	3176	0	3103	172	0
1	B	3076	0	3033	141	0
1	C	3065	0	3018	105	1
2	A	3	0	0	0	1
2	B	3	0	0	0	0
2	C	2	0	0	0	0
3	B	3	0	0	0	0
3	C	3	0	0	0	0
All	All	9331	0	9154	404	1

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 22.

The worst 5 of 404 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:B:142:LYS:NZ	1:B:142:LYS:CE	1.70	1.51
1:A:175:LYS:CE	1:A:175:LYS:O	1.79	1.28
1:A:175:LYS:HE3	1:A:175:LYS:O	1.13	1.25
1:C:381:PHE:CD1	1:C:385:LYS:HE3	1.87	1.09
1:A:239:LYS:HA	1:A:239:LYS:HE2	1.20	1.09

All (1) 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:C:138:GLU:OE2	2:A:452:ZN:ZN[3_445]	1.54	0.66

5.3 Torsion angles

5.3.1 Protein backbone

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	413/449 (92%)	332 (80%)	56 (14%)	25 (6%)	2	26
1	B	396/449 (88%)	308 (78%)	57 (14%)	31 (8%)	1	18
1	C	396/449 (88%)	325 (82%)	47 (12%)	24 (6%)	2	26
All	All	1205/1347 (90%)	965 (80%)	160 (13%)	80 (7%)	1	24

5 of 80 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	ILE
1	A	168	ALA
1	A	224	MET
1	A	229	PRO
1	A	230	TRP

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	320/363 (88%)	266 (83%)	54 (17%)	2	19
1	B	313/363 (86%)	240 (77%)	73 (23%)	1	7
1	C	310/363 (85%)	268 (86%)	42 (14%)	5	30
All	All	943/1089 (87%)	774 (82%)	169 (18%)	2	16

5 of 169 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	119	LYS
1	B	211	SER
1	C	277	LYS
1	B	132	ILE
1	B	164	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	86	GLN
1	B	294	ASN
1	C	218	ASN
1	B	201	ASN
1	A	173	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 14 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	417/449 (92%)	-0.50	6 (1%)	78 63	103, 103, 135, 135	49 (11%)
1	B	397/449 (88%)	-0.47	6 (1%)	76 61	103, 103, 135, 135	42 (10%)
1	C	398/449 (88%)	-0.56	6 (1%)	76 61	103, 103, 135, 136	36 (9%)
All	All	1212/1347 (89%)	-0.51	18 (1%)	76 61	103, 103, 135, 136	127 (10%)

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	396	GLN	4.6
1	A	394	GLN	4.6
1	C	394	GLN	4.5
1	B	395	GLN	3.7
1	B	397	GLN	3.2

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](#)

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(\AA^2)	Q<0.9
3	CA	B	454	1/1	0.87	0.68	-	153,153,153,153	0
2	ZN	A	450	1/1	0.93	0.13	-	151,151,151,151	0
2	ZN	B	450	1/1	0.93	0.25	-	128,128,128,128	0
3	CA	B	453	1/1	0.80	0.14	-	137,137,137,137	0
2	ZN	B	451	1/1	0.96	0.25	-	109,109,109,109	0
3	CA	C	452	1/1	0.89	0.15	-	132,132,132,132	0
2	ZN	B	452	1/1	0.92	0.08	-	203,203,203,203	0
2	ZN	A	452	1/1	0.96	0.17	-	149,149,149,149	0
3	CA	B	455	1/1	0.82	0.07	-	147,147,147,147	0
2	ZN	A	451	1/1	0.96	0.23	-	179,179,179,179	0
3	CA	C	454	1/1	0.82	0.21	-	133,133,133,133	0
2	ZN	C	451	1/1	0.88	0.07	-	185,185,185,185	0
3	CA	C	453	1/1	0.56	0.27	-	141,141,141,141	0
2	ZN	C	450	1/1	0.95	0.28	-	170,170,170,170	0

6.5 Other polymers

There are no such residues in this entry.