



wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 02:25 PM GMT

PDB ID : 3IO6
Title : Huntingtin amino-terminal region with 17 Gln residues - crystal C92-a
Authors : Kim, M.W.; Chelliah, Y.; Kim, S.W.; Otwinowski, Z.; Bezprozvanny, I.
Deposited on : 2009-08-13
Resolution : 3.70 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

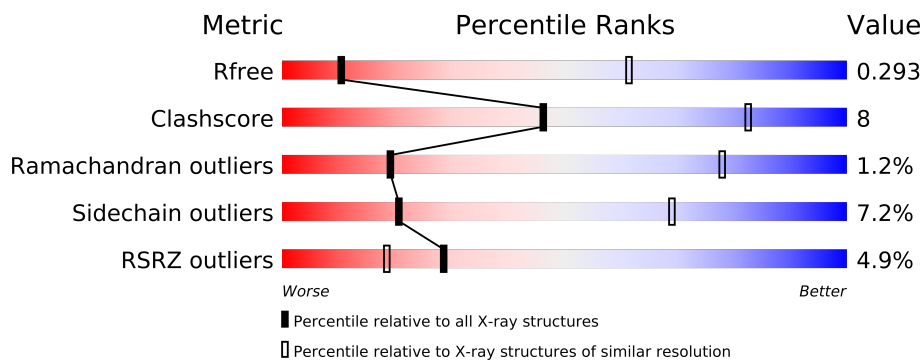
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.70 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}	66092	1098 (4.00-3.40)
Clashscore	79885	1009 (3.94-3.46)
Ramachandran outliers	78287	1016 (3.98-3.42)
Sidechain outliers	78261	1014 (3.98-3.42)
RSRZ outliers	66119	1099 (4.00-3.40)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	449	
1	B	449	
1	C	449	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9265 atoms, of which 0 are hydrogen and 0 are deuterium.

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 periplasmic protein, HUNTINGTIN FUSION PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	400	Total	C	N	O	S	8	0	0
			3059	1969	499	583	8			
1	C	397	Total	C	N	O	S	4	0	0
			3076	1976	504	588	8			
1	B	401	Total	C	N	O	S	20	0	0
			3116	1998	513	597	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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	435	ILE	-	EXPRESSION TAG	UNP P42858
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
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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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
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
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
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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total 2	Zn 2	0	0
2	A	1	Total 1	Zn 1	0	0
2	C	2	Total 2	Zn 2	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total 2	Ca 2	0	0
3	A	3	Total 3	Ca 3	0	0
3	C	4	Total 4	Ca 4	0	0



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	163.40Å 101.24Å 138.19Å 90.00° 92.00° 90.00°	Depositor
Resolution (Å)	38.67 – 3.70 29.73 – 3.70	Depositor EDS
% Data completeness (in resolution range)	92.4 (38.67-3.70) 92.5 (29.73-3.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.02	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.23 (at 3.75Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.245 , 0.293 0.247 , 0.293	Depositor DCC
R_{free} test set	1200 reflections (5.66%)	DCC
Wilson B-factor (Å ²)	98.6	Xtriage
Anisotropy	0.234	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , 94.0	EDS
Estimated twinning fraction	0.012 for -h,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 22387 reflections	Xtriage
F_o, F_c correlation	0.77	EDS
Total number of atoms	9265	wwPDB-VP
Average B, all atoms (Å ²)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.40	1/3128 (0.0%)	0.57	10/4243 (0.2%)
1	B	0.46	2/3187 (0.1%)	0.60	10/4321 (0.2%)
1	C	0.43	1/3147 (0.0%)	0.51	2/4268 (0.0%)
All	All	0.43	4/9462 (0.0%)	0.56	22/12832 (0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	256	LYS	CB-CG	-13.33	1.16	1.52
1	C	144	LYS	CB-CG	-12.71	1.18	1.52
1	A	140	LYS	CB-CG	-10.35	1.24	1.52
1	B	277	LYS	CB-CG	9.03	1.76	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	144	LYS	CA-CB-CG	9.46	134.20	113.40
1	B	34	LYS	CA-CB-CG	-9.24	93.07	113.40
1	B	34	LYS	CB-CG-CD	-8.77	88.79	111.60
1	B	256	LYS	CA-CB-CG	-7.62	96.64	113.40
1	B	277	LYS	CA-CB-CG	-7.27	97.41	113.40

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3059	0	2996	49	0
1	B	3116	0	3087	39	0
1	C	3076	0	3049	84	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
3	A	3	0	0	0	0
3	B	2	0	0	0	0
3	C	4	0	0	0	0
All	All	9265	0	9132	150	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

The worst 5 of 150 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:387:PHE:HB3	1:B:341:TYR:OH	1.30	1.27
1:A:341:TYR:CE2	1:C:387:PHE:CE1	2.27	1.22
1:C:385:LYS:HD2	1:C:385:LYS:O	1.42	1.18
1:A:341:TYR:CZ	1:C:387:PHE:CZ	2.34	1.14
1:A:341:TYR:CE2	1:C:387:PHE:CZ	2.42	1.07

There are no symmetry-related clashes.

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	394/449 (88%)	367 (93%)	19 (5%)	8 (2%)	11 68
1	B	399/449 (89%)	377 (94%)	21 (5%)	1 (0%)	50 92

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	395/449 (88%)	361 (91%)	29 (7%)	5 (1%)	18	76
All	All	1188/1347 (88%)	1105 (93%)	69 (6%)	14 (1%)	19	78

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	405	PRO
1	A	406	PRO
1	A	408	PRO
1	A	413	PRO
1	A	414	PRO

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	307/363 (85%)	288 (94%)	19 (6%)	26	75
1	B	320/363 (88%)	297 (93%)	23 (7%)	21	69
1	C	315/363 (87%)	289 (92%)	26 (8%)	16	62
All	All	942/1089 (86%)	874 (93%)	68 (7%)	21	69

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	160	LEU
1	C	326	LYS
1	B	376	LYS
1	C	170	LYS
1	C	277	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	201	ASN
1	C	218	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	234	ASN
1	C	203	HIS
1	C	234	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 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.

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, 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	400/449 (89%)	-0.03	3 (0%) 83 63	80, 80, 108, 108	27 (6%)
1	B	401/449 (89%)	0.69	31 (7%) 13 10	80, 80, 108, 108	33 (8%)
1	C	397/449 (88%)	0.65	23 (5%) 22 15	80, 80, 80, 108	23 (5%)
All	All	1198/1347 (88%)	0.44	57 (4%) 28 19	80, 80, 108, 108	83 (6%)

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	327	GLY	5.6
1	C	391	GLN	4.7
1	C	260	GLY	4.5
1	B	225	THR	4.1
1	C	109	ALA	3.9

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

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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	ZN	B	451	1/1	0.11	-	107,107,107,107	0
3	CA	B	452	1/1	0.16	-	107,107,107,107	0
3	CA	C	454	1/1	0.18	-	107,107,107,107	0
3	CA	A	453	1/1	0.11	-	107,107,107,107	0
3	CA	B	453	1/1	0.22	-	107,107,107,107	0
3	CA	C	453	1/1	0.15	-	107,107,107,107	0
2	ZN	C	451	1/1	0.38	-	107,107,107,107	0
3	CA	C	452	1/1	0.16	-	107,107,107,107	0
3	CA	A	452	1/1	0.15	-	107,107,107,107	0
2	ZN	B	450	1/1	0.15	-	107,107,107,107	0
2	ZN	A	450	1/1	0.25	-	107,107,107,107	0
2	ZN	C	450	1/1	0.33	-	107,107,107,107	0
3	CA	C	455	1/1	0.21	-	107,107,107,107	0
3	CA	A	451	1/1	0.08	-	107,107,107,107	0

6.5 Other polymers ⓘ

There are no such residues in this entry.