



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 9, 2016 – 10:37 AM EST

PDB ID : 5IOI
Title : X-RAY STRUCTURE OF THE N-TERMINAL DOMAIN OF HUMAN DOUBLECORTIN
Authors : Ruf, A.; Benz, J.; Burger, D.; D'Arcy, B.; Debulpaep, M.; Di Lello, P.; Fry, D.; Huber, W.; Kremer, T.; Laeremans, T.; Matile, H.; Ross, A.; Rudolph, M.G.; Rufer, A.C.; Sharma, A.; Steinmetz, M.O.; Steyaert, J.; Schoch, G.; Stihle, M.; Thoma, R.
Deposited on : 2016-03-08
Resolution : 2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
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)	:	rb-20028442

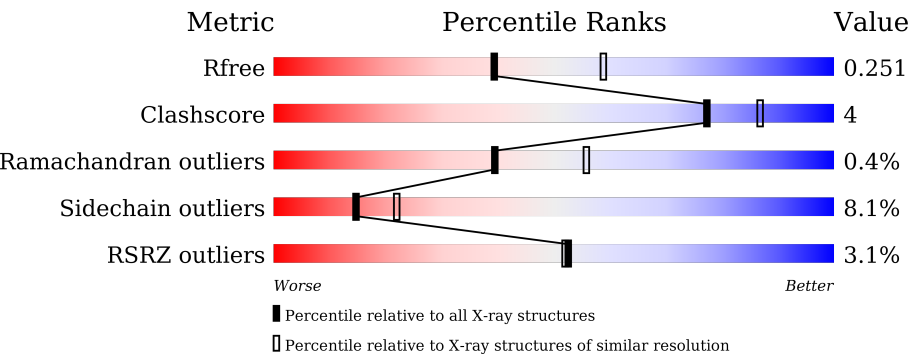
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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. 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	107	<div><div></div><div>84%15%.</div></div>
1	B	107	<div><div>5%</div><div>72%8%18%</div></div>
1	C	107	<div><div>2%</div><div>78%10%10%</div></div>
1	D	107	<div><div>7%</div><div>75%7%17%</div></div>
1	E	107	<div><div>%</div><div>89%11%</div></div>
1	F	107	<div><div>2%</div><div>74%14%11%</div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5085 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 Neuronal migration protein doublecortin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	107	Total	C	N	O	S	0	0	0
			866	544	151	168	3			
1	B	88	Total	C	N	O	S	0	0	0
			716	450	121	143	2			
1	C	96	Total	C	N	O	S	0	0	0
			783	492	134	155	2			
1	D	89	Total	C	N	O	S	0	0	0
			724	455	122	144	3			
1	E	107	Total	C	N	O	S	0	0	0
			866	544	151	168	3			
1	F	95	Total	C	N	O	S	0	0	0
			777	489	133	153	2			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	125	LEU	-	expression tag	UNP O43602
A	126	VAL	-	expression tag	UNP O43602
A	127	PRO	-	expression tag	UNP O43602
A	128	ARG	-	expression tag	UNP O43602
A	129	GLY	-	expression tag	UNP O43602
A	130	SER	-	expression tag	UNP O43602
A	131	HIS	-	expression tag	UNP O43602
A	132	MET	-	expression tag	UNP O43602
A	215	ASP	LYS	engineered mutation	UNP O43602
A	216	ASP	LYS	engineered mutation	UNP O43602
B	125	LEU	-	expression tag	UNP O43602
B	126	VAL	-	expression tag	UNP O43602
B	127	PRO	-	expression tag	UNP O43602
B	128	ARG	-	expression tag	UNP O43602
B	129	GLY	-	expression tag	UNP O43602
B	130	SER	-	expression tag	UNP O43602
B	131	HIS	-	expression tag	UNP O43602

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	132	MET	-	expression tag	UNP O43602
B	215	ASP	LYS	engineered mutation	UNP O43602
B	216	ASP	LYS	engineered mutation	UNP O43602
C	125	LEU	-	expression tag	UNP O43602
C	126	VAL	-	expression tag	UNP O43602
C	127	PRO	-	expression tag	UNP O43602
C	128	ARG	-	expression tag	UNP O43602
C	129	GLY	-	expression tag	UNP O43602
C	130	SER	-	expression tag	UNP O43602
C	131	HIS	-	expression tag	UNP O43602
C	132	MET	-	expression tag	UNP O43602
C	215	ASP	LYS	engineered mutation	UNP O43602
C	216	ASP	LYS	engineered mutation	UNP O43602
D	125	LEU	-	expression tag	UNP O43602
D	126	VAL	-	expression tag	UNP O43602
D	127	PRO	-	expression tag	UNP O43602
D	128	ARG	-	expression tag	UNP O43602
D	129	GLY	-	expression tag	UNP O43602
D	130	SER	-	expression tag	UNP O43602
D	131	HIS	-	expression tag	UNP O43602
D	132	MET	-	expression tag	UNP O43602
D	215	ASP	LYS	engineered mutation	UNP O43602
D	216	ASP	LYS	engineered mutation	UNP O43602
E	125	LEU	-	expression tag	UNP O43602
E	126	VAL	-	expression tag	UNP O43602
E	127	PRO	-	expression tag	UNP O43602
E	128	ARG	-	expression tag	UNP O43602
E	129	GLY	-	expression tag	UNP O43602
E	130	SER	-	expression tag	UNP O43602
E	131	HIS	-	expression tag	UNP O43602
E	132	MET	-	expression tag	UNP O43602
E	215	ASP	LYS	engineered mutation	UNP O43602
E	216	ASP	LYS	engineered mutation	UNP O43602
F	125	LEU	-	expression tag	UNP O43602
F	126	VAL	-	expression tag	UNP O43602
F	127	PRO	-	expression tag	UNP O43602
F	128	ARG	-	expression tag	UNP O43602
F	129	GLY	-	expression tag	UNP O43602
F	130	SER	-	expression tag	UNP O43602
F	131	HIS	-	expression tag	UNP O43602
F	132	MET	-	expression tag	UNP O43602
F	215	ASP	LYS	engineered mutation	UNP O43602

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	216	ASP	LYS	engineered mutation	UNP O43602


- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	95	Total O 95 95	0	0
2	B	50	Total O 50 50	0	0
2	C	58	Total O 58 58	0	0
2	D	26	Total O 26 26	0	0
2	E	69	Total O 69 69	0	0
2	F	55	Total O 55 55	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.

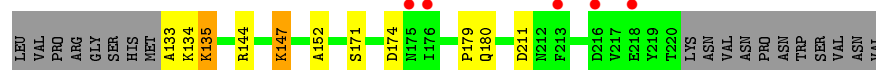
- Molecule 1: Neuronal migration protein doublecortin

Chain A: 




- Molecule 1: Neuronal migration protein doublecortin

Chain B: 




- Molecule 1: Neuronal migration protein doublecortin

Chain C: 




- Molecule 1: Neuronal migration protein doublecortin

Chain D: 



- Molecule 1: Neuronal migration protein doublecortin

Chain E: 



- Molecule 1: Neuronal migration protein doublecortin

Chain F: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	97.72Å 97.72Å 377.09Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.14 – 2.40 35.10 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (35.14-2.40) 100.0 (35.10-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.10 (at 2.39Å)	Xtriage
Refinement program	BUSTER 2.9.2	Depositor
R, R_{free}	0.196 , 0.234 0.211 , 0.251	Depositor DCC
R_{free} test set	2200 reflections (5.41%)	DCC
Wilson B-factor (Å ²)	42.0	Xtriage
Anisotropy	0.060	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 56.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5085	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.55	0/884	0.79	0/1193
1	B	0.49	0/729	0.72	0/980
1	C	0.48	0/799	0.72	0/1077
1	D	0.97	1/737 (0.1%)	1.63	3/990 (0.3%)
1	E	0.53	0/884	0.78	0/1193
1	F	0.52	0/793	0.77	0/1069
All	All	0.61	1/4826 (0.0%)	0.95	3/6502 (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	D	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	132	MET	C-N	-22.87	0.81	1.34

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	132	MET	O-C-N	-35.17	66.43	122.70
1	D	132	MET	CA-C-N	24.65	171.43	117.20
1	D	132	MET	C-N-CA	17.91	166.47	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	132	MET	Mainchain

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	866	0	829	14	0
1	B	716	0	679	7	0
1	C	783	0	741	4	0
1	D	724	0	687	4	0
1	E	866	0	829	8	0
1	F	777	0	736	10	0
2	A	95	0	0	0	0
2	B	50	0	0	1	0
2	C	58	0	0	2	0
2	D	26	0	0	0	0
2	E	69	0	0	0	0
2	F	55	0	0	1	0
All	All	5085	0	4501	34	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 4.

All (34) 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:230:ASN:HD22	1:E:150:VAL:H	1.19	0.90
1:A:224:ASN:HD22	1:A:224:ASN:H	1.27	0.81
1:B:147:LYS:HE3	1:B:147:LYS:H	1.49	0.77
1:A:128:ARG:NH2	1:B:171:SER:OG	2.26	0.68
1:A:150:VAL:H	1:E:230:ASN:HD22	1.46	0.60
1:E:159:ARG:HH12	1:F:171:SER:HB3	1.69	0.58
1:A:224:ASN:N	1:A:224:ASN:HD22	1.98	0.57
1:F:211:ASP:C	1:F:213:PHE:H	2.07	0.56
1:D:202:GLU:CD	1:E:170:ARG:HE	2.10	0.56
1:A:227:TRP:HB2	1:E:135:LYS:HE3	1.89	0.53
1:A:137:ARG:NH1	1:A:147:LYS:O	2.41	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:152:ALA:H	1:F:170:ARG:NH2	2.08	0.51
1:C:180:GLN:HE21	1:C:183:ARG:NH2	2.09	0.50
1:A:230:ASN:ND2	1:E:150:VAL:H	1.99	0.50
1:D:132:MET:HG3	1:D:132:MET:O	2.12	0.50
1:F:218:GLU:HB3	1:F:221:LYS:HB2	1.95	0.49
1:D:152:ALA:H	1:F:170:ARG:HH22	1.59	0.48
1:F:183:ARG:HD3	1:F:210:SER:O	2.14	0.48
1:A:150:VAL:H	1:E:230:ASN:ND2	2.12	0.48
1:A:226:ASN:HB2	1:E:227:TRP:CE2	2.49	0.47
1:A:163:ALA:HA	1:B:152:ALA:HB2	1.96	0.47
1:F:170:ARG:HD2	2:F:307:HOH:O	2.15	0.46
1:A:167:ASP:HA	1:B:135:LYS:HD2	1.97	0.46
1:B:147:LYS:N	1:B:147:LYS:HE3	2.27	0.45
1:B:133:ALA:N	2:B:302:HOH:O	2.50	0.45
1:F:180:GLN:HB2	1:F:210:SER:HB3	2.00	0.44
1:C:218:GLU:HB3	1:C:221:LYS:HG2	2.01	0.43
1:F:211:ASP:C	1:F:213:PHE:N	2.71	0.43
1:B:179:PRO:HG2	1:B:180:GLN:NE2	2.35	0.42
1:A:160:SER:HB2	2:C:336:HOH:O	2.20	0.42
1:C:211:ASP:HB2	2:C:315:HOH:O	2.20	0.42
1:F:211:ASP:OD1	1:F:213:PHE:HB2	2.19	0.41
1:C:221:LYS:HD3	1:C:221:LYS:HA	1.89	0.41
1:A:137:ARG:HH22	1:A:224:ASN:HD21	1.69	0.41

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	105/107 (98%)	104 (99%)	1 (1%)	0	100	100
1	B	86/107 (80%)	84 (98%)	2 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	94/107 (88%)	90 (96%)	3 (3%)	1 (1%)	17	25
1	D	87/107 (81%)	85 (98%)	2 (2%)	0	100	100
1	E	105/107 (98%)	98 (93%)	7 (7%)	0	100	100
1	F	93/107 (87%)	91 (98%)	1 (1%)	1 (1%)	17	25
All	All	570/642 (89%)	552 (97%)	16 (3%)	2 (0%)	39	56

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	227	TRP
1	F	173	SER

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	96/96 (100%)	89 (93%)	7 (7%)	17	27
1	B	78/96 (81%)	72 (92%)	6 (8%)	16	24
1	C	86/96 (90%)	77 (90%)	9 (10%)	8	12
1	D	79/96 (82%)	72 (91%)	7 (9%)	12	18
1	E	96/96 (100%)	90 (94%)	6 (6%)	22	35
1	F	85/96 (88%)	78 (92%)	7 (8%)	14	21
All	All	520/576 (90%)	478 (92%)	42 (8%)	15	22

All (42) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	125	LEU
1	A	132	MET
1	A	134	LYS
1	A	210	SER
1	A	212	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	220	THR
1	A	224	ASN
1	B	134	LYS
1	B	135	LYS
1	B	144	ARG
1	B	147	LYS
1	B	174	ASP
1	B	211	ASP
1	C	134	LYS
1	C	144	ARG
1	C	157	ARG
1	C	159	ARG
1	C	176	ILE
1	C	180	GLN
1	C	211	ASP
1	C	226	ASN
1	C	228	SER
1	D	134	LYS
1	D	135	LYS
1	D	174	ASP
1	D	177	ASN
1	D	202	GLU
1	D	211	ASP
1	D	212	ASN
1	E	132	MET
1	E	174	ASP
1	E	176	ILE
1	E	204	GLU
1	E	211	ASP
1	E	212	ASN
1	F	134	LYS
1	F	157	ARG
1	F	159	ARG
1	F	211	ASP
1	F	212	ASN
1	F	217	VAL
1	F	220	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	224	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	230	ASN
1	B	180	GLN
1	E	230	ASN
1	F	180	GLN
1	F	226	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](#)

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	132:MET	C	133:ALA	N	0.81

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	107/107 (100%)	-0.46	0 100 100	24, 34, 53, 72	0
1	B	88/107 (82%)	0.12	5 (5%) 27 27	30, 49, 86, 89	0
1	C	96/107 (89%)	-0.19	2 (2%) 67 66	30, 47, 78, 82	0
1	D	89/107 (83%)	0.37	8 (8%) 12 11	33, 63, 94, 105	0
1	E	107/107 (100%)	-0.26	1 (0%) 85 85	27, 41, 67, 79	0
1	F	95/107 (88%)	-0.26	2 (2%) 67 66	31, 45, 74, 83	0
All	All	582/642 (90%)	-0.13	18 (3%) 52 52	24, 45, 81, 105	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	213	PHE	4.8
1	B	213	PHE	4.3
1	E	213	PHE	3.9
1	B	176	ILE	3.8
1	F	212	ASN	3.5
1	B	218	GLU	3.3
1	D	216	ASP	3.2
1	F	213	PHE	2.9
1	D	215	ASP	2.9
1	C	213	PHE	2.8
1	D	218	GLU	2.8
1	D	175	ASN	2.7
1	D	180	GLN	2.6
1	D	132	MET	2.5
1	B	175	ASN	2.5
1	D	133	ALA	2.4
1	B	216	ASP	2.4
1	C	212	ASN	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.