



# wwPDB X-ray Structure Validation Summary Report

Feb 17, 2024 – 05:33 PM EST

PDB ID : 3T6G  
Title : Structure of the complex between NSP3 (SHEP1) and p130Cas  
Authors : Mace, P.D.; Robinson, H.; Riedl, S.J.  
Deposited on : 2011-07-28  
Resolution : 2.50 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

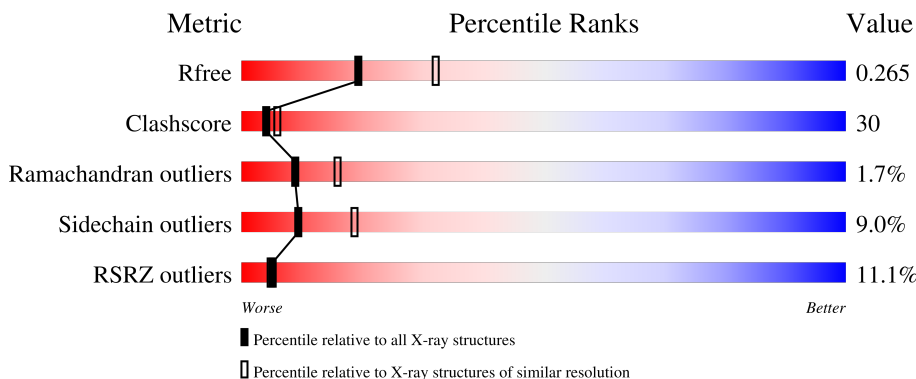
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	331	
1	C	331	
2	B	229	
2	D	229	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6598 atoms, of which 3 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 SH2 domain-containing protein 3C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	299	2348	1495	418	423	12	1	1	0
1	C	273	2139	1361	383	384	11	1	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	381	MET	-	initiating methionine	UNP Q8N5H7
A	497	SER	CYS	engineered mutation	UNP Q8N5H7
A	598	SER	CYS	engineered mutation	UNP Q8N5H7
A	704	LEU	-	expression tag	UNP Q8N5H7
A	705	GLU	-	expression tag	UNP Q8N5H7
A	706	HIS	-	expression tag	UNP Q8N5H7
A	707	HIS	-	expression tag	UNP Q8N5H7
A	708	HIS	-	expression tag	UNP Q8N5H7
A	709	HIS	-	expression tag	UNP Q8N5H7
A	710	HIS	-	expression tag	UNP Q8N5H7
A	711	HIS	-	expression tag	UNP Q8N5H7
C	381	MET	-	initiating methionine	UNP Q8N5H7
C	497	SER	CYS	engineered mutation	UNP Q8N5H7
C	598	SER	CYS	engineered mutation	UNP Q8N5H7
C	704	LEU	-	expression tag	UNP Q8N5H7
C	705	GLU	-	expression tag	UNP Q8N5H7
C	706	HIS	-	expression tag	UNP Q8N5H7
C	707	HIS	-	expression tag	UNP Q8N5H7
C	708	HIS	-	expression tag	UNP Q8N5H7
C	709	HIS	-	expression tag	UNP Q8N5H7
C	710	HIS	-	expression tag	UNP Q8N5H7
C	711	HIS	-	expression tag	UNP Q8N5H7

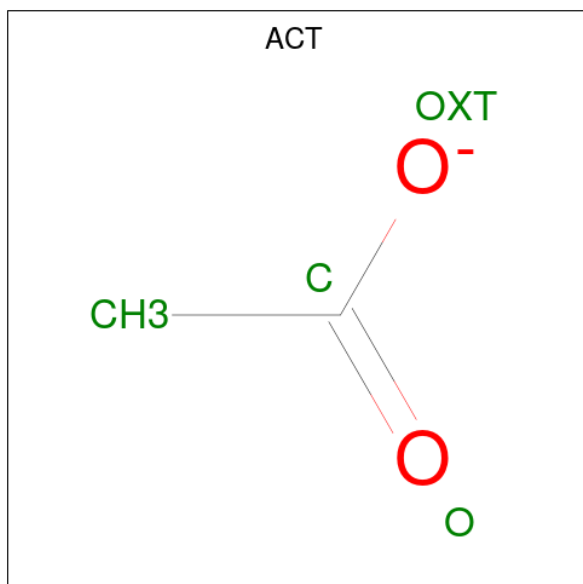
- Molecule 2 is a protein called Breast cancer anti-estrogen resistance protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	134	Total	C	N	O	S	0	0	0
			1035	657	183	192	3			
2	D	131	Total	C	N	O	S	0	0	0
			1010	644	177	186	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	644	MET	-	initiating methionine	UNP P56945
B	871	LEU	-	expression tag	UNP P56945
B	872	GLU	-	expression tag	UNP P56945
D	644	MET	-	initiating methionine	UNP P56945
D	871	LEU	-	expression tag	UNP P56945
D	872	GLU	-	expression tag	UNP P56945

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	1	Total	C	H	O	0	0
			7	2	3	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	41	Total	O	0	0
			41	41		

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	B	14	Total	O	0	0
			14	14		
4	C	4	Total	O	0	0
			4	4		





## 4 Data and refinement statistics i

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	171.88Å 171.88Å 78.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.55 – 2.50 29.55 – 2.50	Depositor EDS
% Data completeness (in resolution range)	94.1 (29.55-2.50) 99.9 (29.55-2.50)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.38 (at 2.51Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, $R_{free}$	0.197 , 0.266 0.197 , 0.265	Depositor DCC
$R_{free}$ test set	1985 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.5	Xtrriage
Anisotropy	0.626	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 71.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.019 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6598	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	100.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.47	0/2393	0.64	1/3235 (0.0%)
1	C	0.30	0/2176	0.45	0/2936
2	B	0.45	0/1052	0.62	0/1428
2	D	0.32	0/1027	0.47	0/1395
All	All	0.39	0/6648	0.56	1/8994 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	437	VAL	CB-CA-C	-5.20	101.53	111.40

There are no chirality outliers.

There are no planarity outliers.

### 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	2348	0	2427	96	0
1	C	2139	0	2212	194	0
2	B	1035	0	1054	39	0
2	D	1010	0	1032	83	0
3	C	4	3	3	1	0
4	A	41	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	14	0	0	1	0
4	C	4	0	0	2	0
All	All	6595	3	6728	397	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 30.

The worst 5 of 397 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:C:454:MET:HA	1:C:457:LEU:HD12	1.21	1.17
1:A:545:GLN:O	1:A:549:THR:HG22	1.46	1.15
2:D:748:LEU:HD13	2:D:864:LEU:HD23	1.31	1.09
1:A:387:VAL:HB	1:A:388:PRO:HD3	1.09	1.09
1:C:489:MET:HE2	1:C:662:GLN:HG2	1.24	1.08

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	296/331 (89%)	274 (93%)	19 (6%)	3 (1%)	15	28
1	C	269/331 (81%)	238 (88%)	26 (10%)	5 (2%)	8	13
2	B	132/229 (58%)	124 (94%)	5 (4%)	3 (2%)	6	10
2	D	129/229 (56%)	110 (85%)	16 (12%)	3 (2%)	6	10
All	All	826/1120 (74%)	746 (90%)	66 (8%)	14 (2%)	9	16

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	387	VAL
2	B	803	ALA
1	C	670	GLN
2	B	868	ALA
2	D	787	LEU

### 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	255/282 (90%)	230 (90%)	25 (10%)	8	15
1	C	229/282 (81%)	213 (93%)	16 (7%)	15	29
2	B	111/193 (58%)	101 (91%)	10 (9%)	9	19
2	D	108/193 (56%)	96 (89%)	12 (11%)	6	11
All	All	703/950 (74%)	640 (91%)	63 (9%)	9	19

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	819	LEU
2	D	807	ASP
1	C	424	LEU
2	D	802	GLN
2	D	856	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	802	GLN
1	C	662	GLN
1	C	622	HIS
1	C	455	GLN
1	C	631	HIS

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	ACT	C	1	-	3,3,3	0.76	0	3,3,3	1.30	0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1	ACT	1	0

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	299/331 (90%)	0.09	17 (5%) 23 25	48, 69, 99, 148	1 (0%)
1	C	273/331 (82%)	0.90	54 (19%) 1 1	79, 129, 174, 210	1 (0%)
2	B	134/229 (58%)	-0.13	2 (1%) 73 75	52, 69, 103, 139	0
2	D	131/229 (57%)	0.71	20 (15%) 2 1	67, 123, 194, 223	0
All	All	837/1120 (74%)	0.41	93 (11%) 5 5	48, 89, 169, 223	2 (0%)

The worst 5 of 93 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	413	LEU	6.9
2	D	803	ALA	6.7
1	C	450	VAL	6.5
1	C	415	VAL	6.0
1	C	527	SER	5.4

### 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 monosaccharides 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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	B-factors( $\text{\AA}^2$ )	Q<0.9
3	ACT	C	1	4/4	0.86	0.29	105,110,126,126	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.