wwPDB X-ray Structure Validation Summary Report (i)

Aug 3, 2023 – 06:03 AM EDT

PDB ID	:	1I7X
Title	:	BETA-CATENIN/E-CADHERIN COMPLEX
Authors	:	Huber, A.H.; Weis, W.I.
Deposited on	:	2001-03-10
Resolution	:	3.00 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

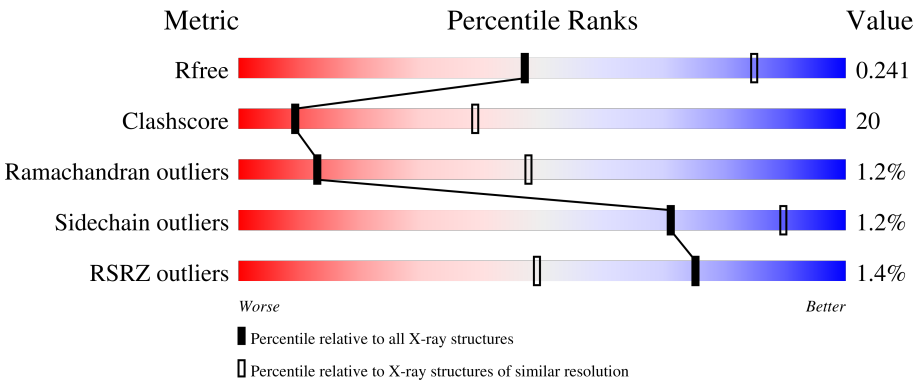
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.34
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 >=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 <=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	538	2% 64% 33%
1	C	538	62% 33%
2	B	151	2% 20% 16% 62%
2	D	151	% 25% 17% 58%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8930 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 BETA-CATENIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	522	3988	2503	723	736	26	0	0	0
1	C	521	3978	2500	721	731	26	0	0	0

- Molecule 2 is a protein called EPITHELIAL-CADHERIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	57	442	277	72	92	1	0	0	0
2	D	63	502	314	77	110	1	0	0	0

- Molecule 3 is water.

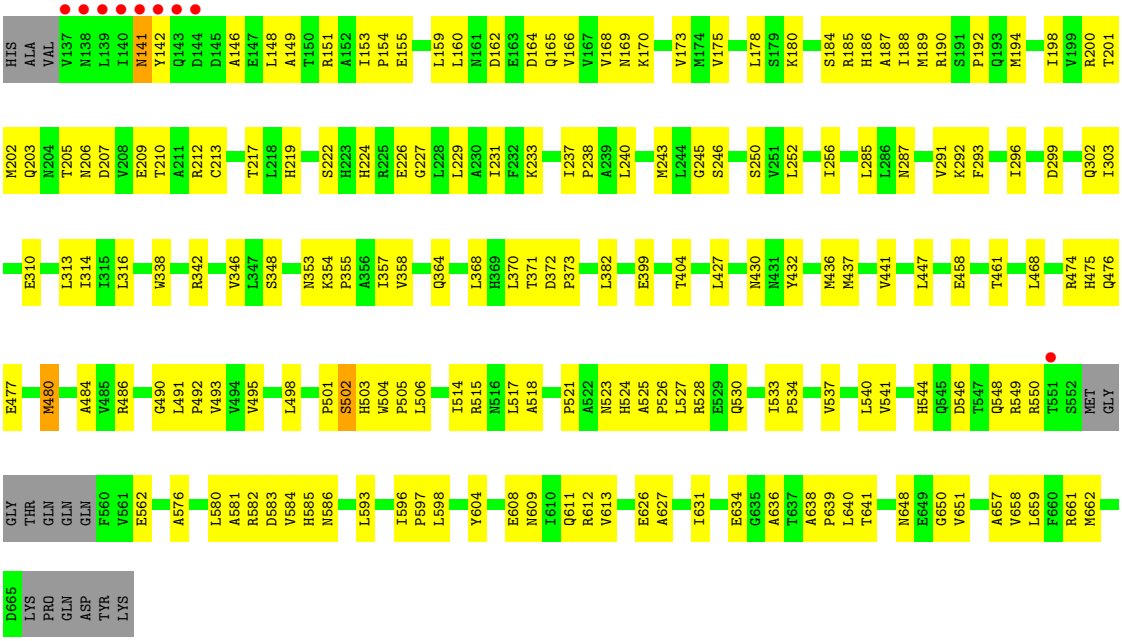
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	11	11	11	0	0
3	C	9	9	9	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

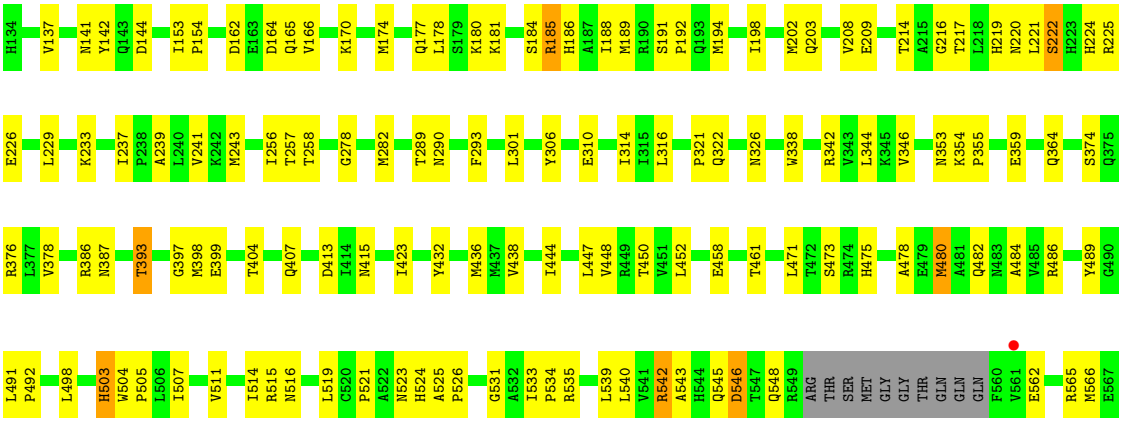
• Molecule 1: BETA-CATENIN

Chain A:



• Molecule 1: BETA-CATENIN

Chain C:



E568 I569 V570 T574 G575 H578 I579 D583 V584 H585 N586 R587 I588 V589 I590 R591 G592 L593 N594 T595 I596 P597 L598 F599 L602 L603 Y604 I607 E608 N609 I610 Q611 R612 V613 A614 A615 G616 V617 L618 L621 D624 A627 I631 E632 A636 T637 A638 P639 L640 T641

E642 L643 L644 H645 T653 Y654 A655 A656 A657 V658 L659 F660 R661 M662 S663 E664 ASP LYS PRO GLN ASP TYR LYS

• Molecule 2: EPITHELIAL-CADHERIN

Chain B: 2% 20% 16% • 62%

ARG ARG ARG THR VAL VAL LYS GLU PRO LEU LEU PRO PRO ASP ASP ASP THR ARG ASP ASN VAL TYR TYR TYR ASP GLU GLU GLY GLY GLY GLU GLU ASP GLN ASP PHE ASP LEU SER GLN LEU LEU HIS ARG GLY LEU ASP ALA ARG PRO GLU V628 T629 R630 A634 P635 T636 L637 M638 S639

V640 P641 Q642 Y643 R644 P645 R646 P647 A648 N649 P650 I653 I657 L661 T669 P672 S675 L676 L677 V678 Y681 E682 G683 S684 GLY SER GLU ALA ALA SER LEU SER SER LEU ASN SER SER GLU SER ASP GLN ASP GLN ASP TYR ASP TYR LEU ASN GLU TRP GLY ASN

ARG PHE LYS LYS LEU ALA ASP MET TYR GLY GLY GLY GLU ASP ASP

• Molecule 2: EPITHELIAL-CADHERIN

Chain D: % 25% 17% 58%

ARG ARG ARG THR VAL VAL LYS GLU PRO LEU LEU PRO PRO ASP ASP ASP THR ARG ASP ASN VAL TYR TYR TYR ASP GLU GLU GLY GLY GLY GLU GLU ASP GLN ASP PHE ASP LEU SER GLN LEU LEU HIS ARG GLY LEU ASP ALA ARG PRO GLU VAL THR R630 N631 D632 V633 A634 P635 THR LEU

MET SER VAL PRO GLN TYR ARG PRO ARG PRO ALA ASN PRO ASP GLU ILE G654 N655 F656 I657 L661 S666 D667 P668 T669 A670 P671 P672 Y673 L676 E682 G683 S684 GLY SER GLU ALA ALA SER LEU SER SER LEU ASN SER SER E698 S699 D700 Q701 D702 Q703 D704 Y705 D706 Y707

L708 N709 R714 L718 M721 Y722 G723 GLY GLY GLU ASP ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	180.80Å 134.20Å 94.20Å 90.00° 94.30° 90.00°	Depositor
Resolution (Å)	30.00 – 3.00 29.71 – 2.99	Depositor EDS
% Data completeness (in resolution range)	94.5 (30.00-3.00) 94.0 (29.71-2.99)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.97 (at 3.00Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.196 , 0.242 0.194 , 0.241	Depositor DCC
R_{free} test set	3449 reflections (7.75%)	wwPDB-VP
Wilson B-factor (Å^2)	50.3	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 63.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	8930	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.34	0/4043	0.59	0/5487
1	C	0.34	0/4034	0.60	0/5476
2	B	0.36	0/454	0.63	0/624
2	D	0.32	0/513	0.52	0/695
All	All	0.34	0/9044	0.59	0/12282

There are no bond length outliers.

There are no bond angle outliers.

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	3988	0	4125	149	0
1	C	3978	0	4117	179	0
2	B	442	0	413	26	0
2	D	502	0	432	23	0
3	A	11	0	0	0	0
3	C	9	0	0	1	0
All	All	8930	0	9087	356	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 20.

The worst 5 of 356 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:638:ALA:HB3	1:C:639:PRO:HD3	1.47	0.95
1:C:473:SER:HB3	1:C:516:ASN:ND2	1.82	0.95
1:C:637:THR:HA	1:C:659:LEU:HD21	1.52	0.91
2:D:632:ASP:OD1	2:D:673:TYR:HB2	1.70	0.91
1:C:583:ASP:HB3	1:C:586:ASN:HD22	1.36	0.90

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	518/538 (96%)	473 (91%)	42 (8%)	3 (1%)	25 64
1	C	517/538 (96%)	459 (89%)	50 (10%)	8 (2%)	10 42
2	B	55/151 (36%)	45 (82%)	7 (13%)	3 (6%)	2 10
2	D	57/151 (38%)	52 (91%)	5 (9%)	0	100 100
All	All	1147/1378 (83%)	1029 (90%)	104 (9%)	14 (1%)	13 48

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	640	VAL
2	B	642	GLN
1	C	643	LEU
1	A	368	LEU
1	A	502	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	434/447 (97%)	431 (99%)	3 (1%)	84	94
1	C	432/447 (97%)	424 (98%)	8 (2%)	57	84
2	B	50/131 (38%)	50 (100%)	0	100	100
2	D	53/131 (40%)	52 (98%)	1 (2%)	57	84
All	All	969/1156 (84%)	957 (99%)	12 (1%)	71	90

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	503	HIS
1	C	542	ARG
2	D	706	ASP
1	C	546	ASP
1	C	144	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	594	ASN
1	C	601	GLN
2	D	709	ASN
1	A	503	HIS
1	A	499	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)

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)

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	522/538 (97%)	-0.73	9 (1%)	70	41	13, 40, 114, 201	0
1	C	521/538 (96%)	-0.70	2 (0%)	92	79	16, 43, 122, 201	0
2	B	57/151 (37%)	0.23	3 (5%)	26	10	34, 78, 165, 195	0
2	D	63/151 (41%)	0.23	2 (3%)	47	20	38, 91, 141, 151	0
All	All	1163/1378 (84%)	-0.62	16 (1%)	75	49	13, 46, 131, 201	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	138	ASN	5.8
1	A	137	VAL	5.2
1	C	645	HIS	4.4
1	A	140	ILE	4.0
1	A	139	LEU	3.4

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.