W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2020 – 11:45 am BST

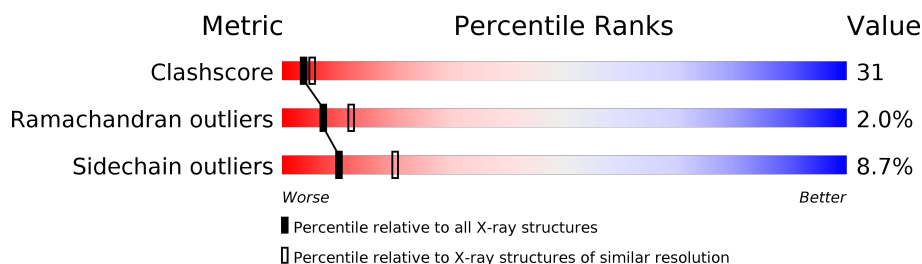|		
PDB ID	:	1JPW
Title	:	Crystal Structure of a Human Tcf-4 / beta-Catenin Complex
Authors	:	Poy, F.; Lepourcelet, M.; Shivdasani, R.A.; Eck, M.J.
Deposited on	:	2001-08-03
Resolution	:	2.50 Å(reported)

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	Percentile Ranks	Value
Clashscore		31
Ramachandran outliers		2.0%
Sidechain outliers		8.7%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (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 on 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	540	56% 35% • 7%
1	B	540	46% 39% 7% • 7%
1	C	540	47% 39% 6% • 7%
2	D	49	29% 16% • 51%
2	E	49	31% 18% 51%
2	F	49	27% 20% • 51%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12767 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
1	A	502	Total 3803	C 2393	N 691	O 693	S 26	0	0	0
1	B	502	Total 3803	C 2393	N 691	O 693	S 26	0	0	0
1	C	502	Total 3803	C 2393	N 691	O 693	S 26	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	132	GLY	LEU	CONFLICT	UNP P35222
A	133	SER	LYS	CONFLICT	UNP P35222
B	132	GLY	LEU	CONFLICT	UNP P35222
B	133	SER	LYS	CONFLICT	UNP P35222
C	132	GLY	LEU	CONFLICT	UNP P35222
C	133	SER	LYS	CONFLICT	UNP P35222

- Molecule 2 is a protein called transcription factor 7-like 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	24	Total 170	C 104	N 26	O 40	0	0	0
2	E	24	Total 170	C 104	N 26	O 40	0	0	0
2	F	24	Total 170	C 104	N 26	O 40	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	7	SER	GLY	CONFLICT	UNP Q9NQB0
E	7	SER	GLY	CONFLICT	UNP Q9NQB0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	7	SER	GLY	CONFLICT	UNP Q9NQB0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	290	Total 290	O 290	0	0
3	D	19	Total 19	O 19	0	0
3	B	281	Total 281	O 281	0	0
3	E	12	Total 12	O 12	0	0
3	C	236	Total 236	O 236	0	0
3	F	10	Total 10	O 10	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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.
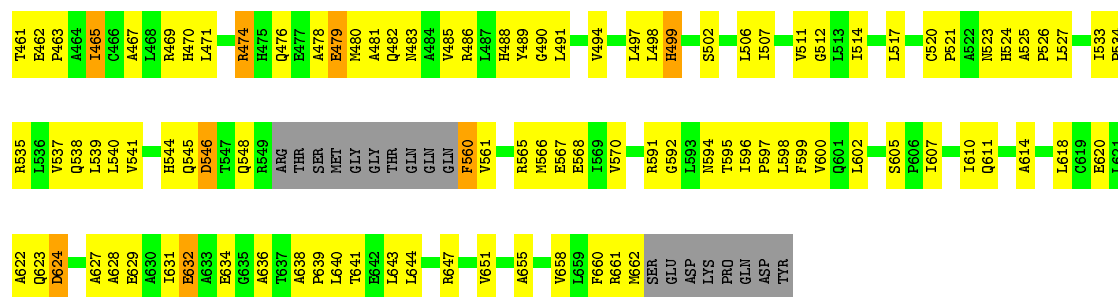
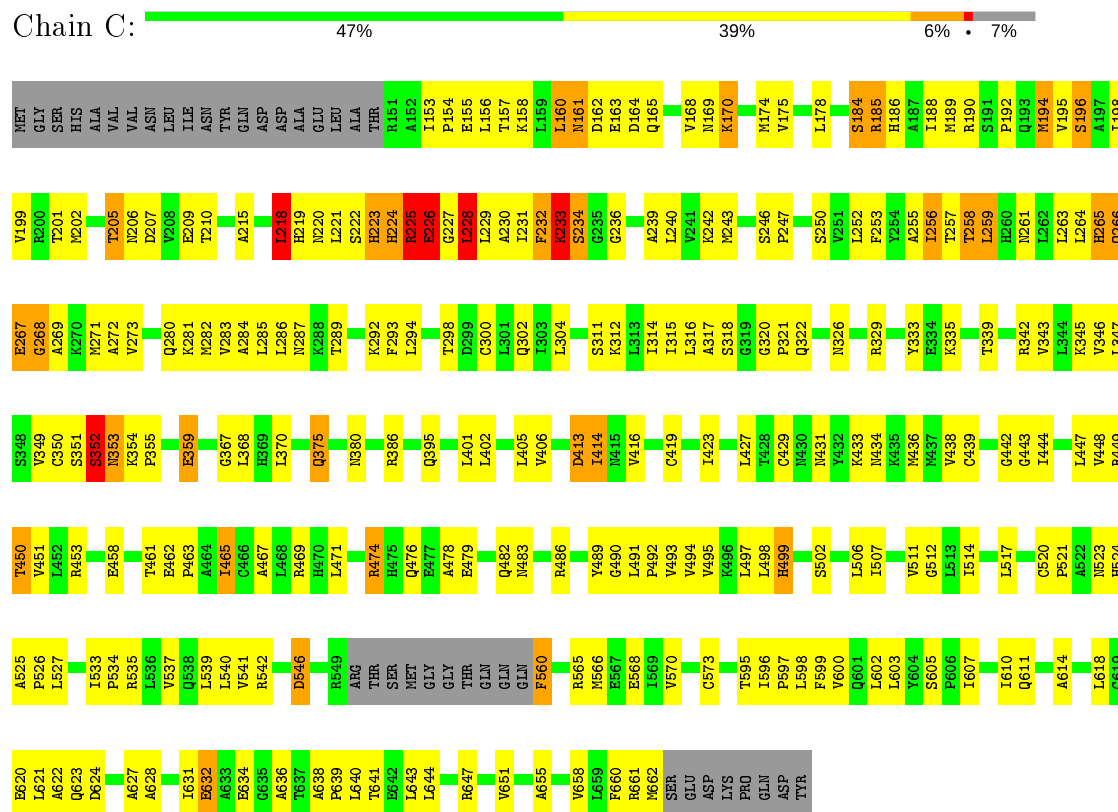Note EDS was not executed.

- Molecule 1: BETA-CATENIN
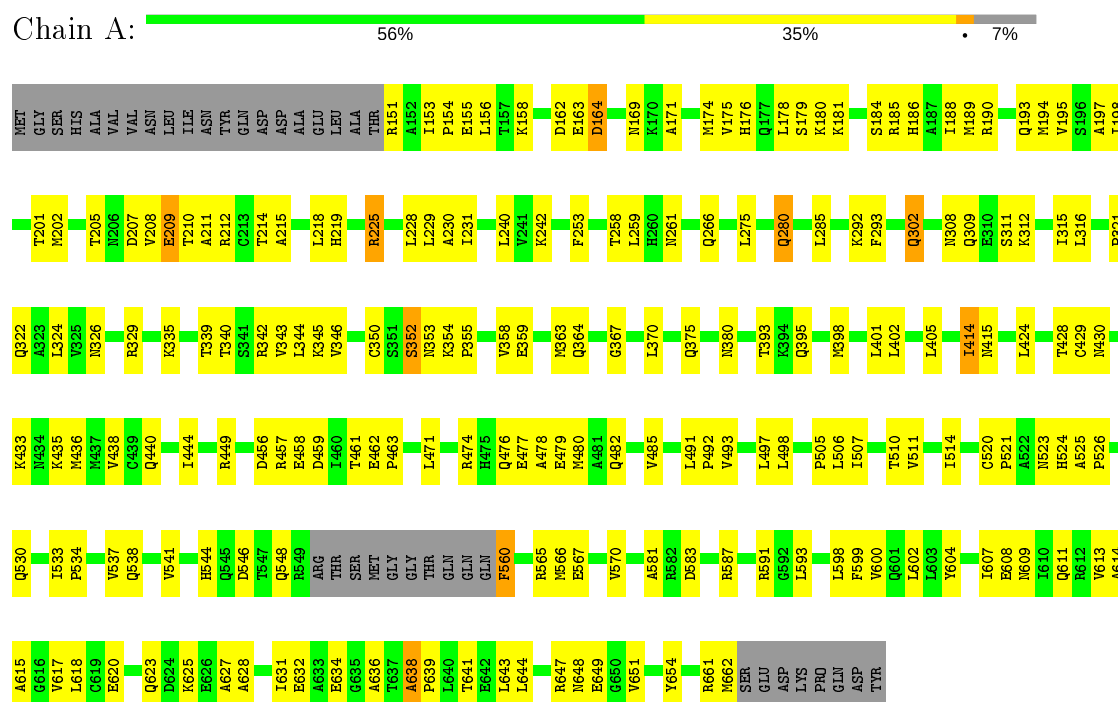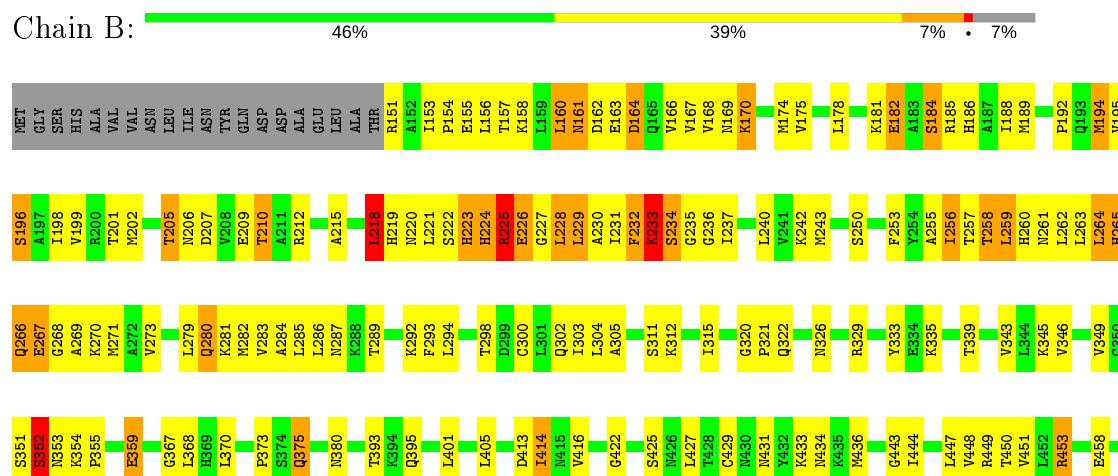
- Molecule 1: BETA-CATENIN

• Molecule 1: BETA-CATENIN

Chain C: 47% 39% 6% • 7%

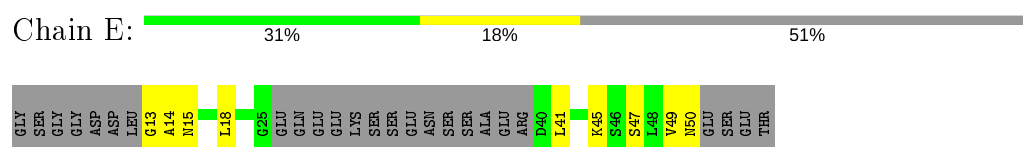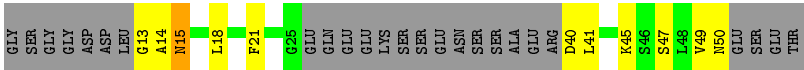• Molecule 2: transcription factor 7-like 2

Chain D: 29% 16% • 51%

• Molecule 2: transcription factor 7-like 2

Chain E: 31% 18% 51%

• Molecule 2: transcription factor 7-like 2

Chain F:

27%   20%   •   51%
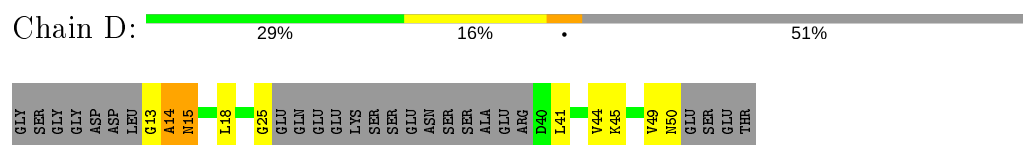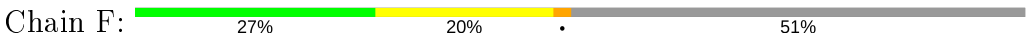
4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	376.53Å 92.65Å 49.25Å 90.00° 95.98° 90.00°	Depositor
Resolution (Å)	25.00 – 2.50	Depositor
% Data completeness (in resolution range)	86.0 (25.00-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.304 , 0.390	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12767	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

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	0.43	0/3857	0.68	0/5238
1	B	0.36	0/3857	0.67	6/5238 (0.1%)
1	C	0.32	0/3857	0.63	4/5238 (0.1%)
2	D	0.41	0/169	0.57	0/226
2	E	0.38	0/169	0.56	0/226
2	F	0.41	0/169	0.57	0/226
All	All	0.37	0/12078	0.65	10/16392 (0.1%)

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	B	0	2
1	C	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	225	ARG	CG-CD-NE	-10.66	89.42	111.80
1	B	225	ARG	NE-CZ-NH1	7.67	124.14	120.30
1	B	225	ARG	NE-CZ-NH2	-6.35	117.12	120.30
1	C	225	ARG	N-CA-C	6.19	127.70	111.00
1	C	218	LEU	CA-CB-CG	6.03	129.16	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	225	ARG	Sidechain
1	B	560	PHE	Sidechain
1	C	560	PHE	Sidechain

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	3803	0	3929	192	0
1	B	3803	0	3929	266	0
1	C	3803	0	3929	273	1
2	D	170	0	151	20	0
2	E	170	0	151	10	0
2	F	170	0	151	12	0
3	A	290	0	0	36	0
3	B	281	0	0	63	0
3	C	236	0	0	46	0
3	D	19	0	0	2	0
3	E	12	0	0	2	0
3	F	10	0	0	1	0
All	All	12767	0	12240	746	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 31.

The worst 5 of 746 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:225:ARG:HH21	1:C:225:ARG:CB	1.22	1.52
1:C:225:ARG:NH2	1:C:225:ARG:HB3	1.34	1.42
1:A:280:GLN:H	1:A:280:GLN:NE2	1.32	1.25
1:C:414:ILE:HD12	1:C:414:ILE:H	1.06	1.19
1:B:280:GLN:N	1:B:280:GLN:HE21	1.39	1.19

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:223:HIS:O	1:C:225:ARG:NH2[2_655]	2.02	0.18

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	498/540 (92%)	469 (94%)	22 (4%)	7 (1%)	11	20
1	B	498/540 (92%)	449 (90%)	39 (8%)	10 (2%)	7	12
1	C	498/540 (92%)	449 (90%)	39 (8%)	10 (2%)	7	12
2	D	20/49 (41%)	17 (85%)	1 (5%)	2 (10%)	0	0
2	E	20/49 (41%)	17 (85%)	2 (10%)	1 (5%)	2	2
2	F	20/49 (41%)	17 (85%)	2 (10%)	1 (5%)	2	2
All	All	1554/1767 (88%)	1418 (91%)	105 (7%)	31 (2%)	7	12

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	15	ASN
1	B	225	ARG
1	B	232	PHE
1	B	233	LYS
1	B	266	GLN

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	409/448 (91%)	393 (96%)	16 (4%)	32	57
1	B	409/448 (91%)	361 (88%)	48 (12%)	5	10
1	C	409/448 (91%)	361 (88%)	48 (12%)	5	10
2	D	18/41 (44%)	18 (100%)	0	100	100
2	E	18/41 (44%)	18 (100%)	0	100	100
2	F	18/41 (44%)	18 (100%)	0	100	100
All	All	1281/1467 (87%)	1169 (91%)	112 (9%)	10	20

5 of 112 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	414	ILE
1	B	624	ASP
1	C	465	ILE
1	B	450	THR
1	B	479	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	302	GLN
1	B	380	ASN
1	C	434	ASN
1	B	322	GLN
1	B	326	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)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.