



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 22, 2025 – 10:07 AM EDT

PDB ID : 9PFN / pdb_00009pfn
Title : Structure of POU2F3 POU domains bound to coactivator OCA-T2 and DNA
(2.8 angstrom resolution)
Authors : Ipsaro, J.J.; Alpsoy, A.; Vakoc, C.R.; Joshua-Tor, L.
Deposited on : 2025-07-06
Resolution : 2.79 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

<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 : 4-5-2 with Phenix2.0
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.46

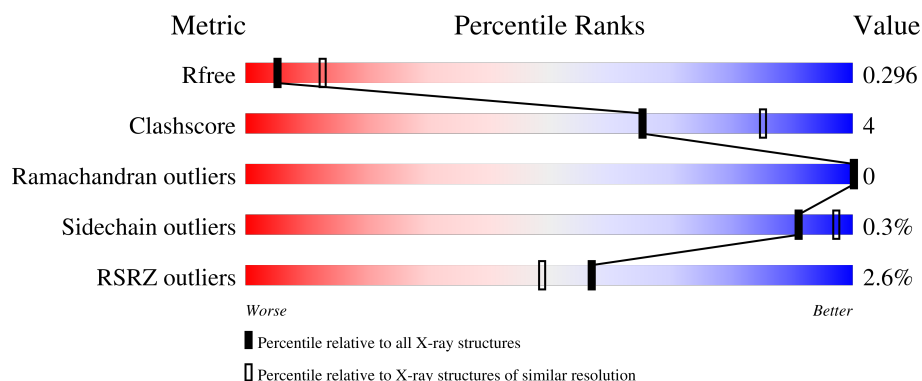
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 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}	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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 $\leq 5\%$. 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	160	<div> <div>6%</div> <div>78% 9% 13%</div> </div>
1	E	160	<div> <div>6%</div> <div>75% 9% 16%</div> </div>
2	B	39	<div> <div>77% 5% 18%</div> </div>
2	F	39	<div> <div>49% 8% 44%</div> </div>
3	C	15	<div> <div>80% 20%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	15	 93% 7%
4	D	15	 40% 60%
4	H	15	 60% 40%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7401 atoms, of which 3482 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 POU domain, class 2, transcription factor 3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	139	Total	C	H	N	O	S	0	0	0
			2311	716	1170	205	213	7			
1	E	135	Total	C	H	N	O	S	0	0	0
			2265	703	1148	201	206	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	182	GLY	-	expression tag	UNP Q9UKI9
E	182	GLY	-	expression tag	UNP Q9UKI9

- Molecule 2 is a protein called POU class 2 homeobox associating factor 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	32	Total	C	H	N	O	0	0	0
			520	156	269	51	44			
2	F	22	Total	C	H	N	O	0	0	0
			393	118	207	39	29			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	GLY	-	expression tag	UNP A8K830
F	0	GLY	-	expression tag	UNP A8K830

- Molecule 3 is a DNA chain called DNA containing POU domain recognition element octamer (sense strand).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	15	Total	C	H	N	O	P	0	0	0
			481	149	171	61	86	14			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	G	15	Total	C	H	N	O	P	0	0	0
			481	149	171	61	86	14			

- Molecule 4 is a DNA chain called DNA containing POU domain recognition element octamer (antisense strand).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	15	Total	C	H	N	O	P	0	0	0
			472	146	173	49	90	14			
4	H	15	Total	C	H	N	O	P	0	0	0
			472	146	173	49	90	14			

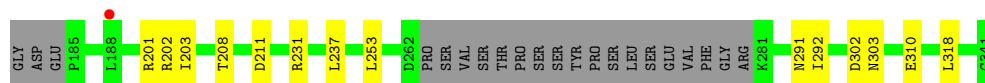
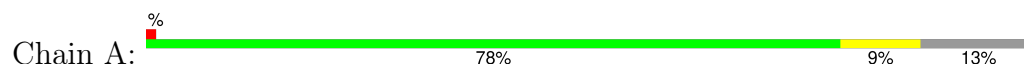
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	O	0	0
			1	1		
5	C	1	Total	O	0	0
			1	1		
5	D	1	Total	O	0	0
			1	1		
5	E	1	Total	O	0	0
			1	1		
5	G	1	Total	O	0	0
			1	1		
5	H	1	Total	O	0	0
			1	1		

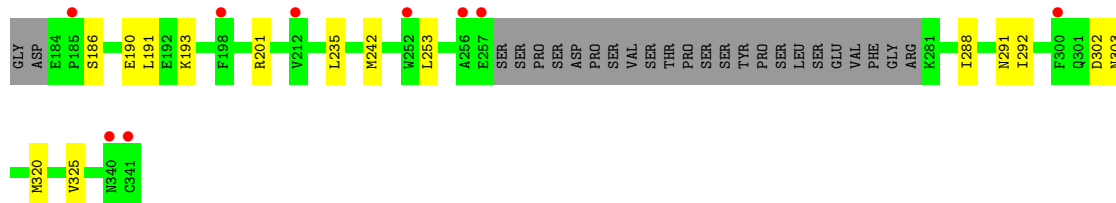
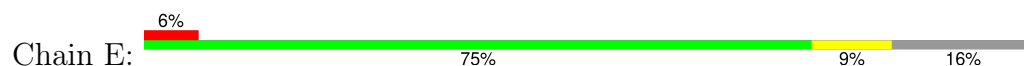
3 Residue-property plots [i](#)

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.

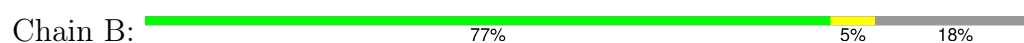
- Molecule 1: POU domain, class 2, transcription factor 3



- Molecule 1: POU domain, class 2, transcription factor 3



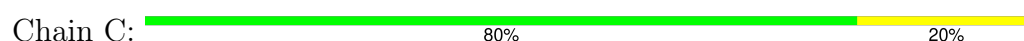
- Molecule 2: POU class 2 homeobox associating factor 3



- Molecule 2: POU class 2 homeobox associating factor 3



- Molecule 3: DNA containing POU domain recognition element octamer (sense strand)



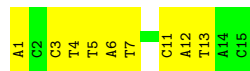
- Molecule 3: DNA containing POU domain recognition element octamer (sense strand)

Chain G:  93% 7%



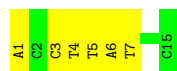
- Molecule 4: DNA containing POU domain recognition element octamer (antisense strand)

Chain D:  40% 60%



- Molecule 4: DNA containing POU domain recognition element octamer (antisense strand)

Chain H:  60% 40%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	81.17Å 36.02Å 88.60Å 90.00° 90.08° 90.00°	Depositor
Resolution (Å)	38.91 – 2.79 38.91 – 2.79	Depositor EDS
% Data completeness (in resolution range)	98.0 (38.91-2.79) 98.0 (38.91-2.79)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.44 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.236 , 0.296 0.233 , 0.296	Depositor DCC
R_{free} test set	664 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	73.3	Xtriage
Anisotropy	0.728	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 66.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.021 for h,-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7401	wwPDB-VP
Average B, all atoms (Å ²)	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.65% 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 [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	0.10	0/1157	0.27	0/1541
1	E	0.12	0/1132	0.31	0/1507
2	B	0.08	0/253	0.21	0/337
2	F	0.07	0/187	0.17	0/248
3	C	0.24	0/349	0.51	0/538
3	G	0.22	0/349	0.50	0/538
4	D	0.24	0/333	0.55	0/511
4	H	0.23	0/333	0.51	0/511
All	All	0.16	0/4093	0.38	0/5731

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	1141	1170	1170	9	0
1	E	1117	1148	1149	8	2
2	B	251	269	269	1	0
2	F	186	207	207	2	0
3	C	310	171	171	1	2
3	G	310	171	171	0	3
4	D	299	173	173	4	5
4	H	299	173	173	3	2

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
5	G	1	0	0	0	0
5	H	1	0	0	0	0
All	All	3919	3482	3483	27	7

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 (27) 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:E:186:SER:OG	1:E:190:GLU:OE1	2.06	0.73
2:B:15:ILE:HG21	2:B:20:LEU:HD21	1.85	0.59
1:A:291:ASN:OD1	1:A:292:ILE:N	2.37	0.57
1:E:235:LEU:HD22	1:E:242:MET:HE1	1.88	0.56
1:A:201:ARG:HB2	1:A:253:LEU:HD21	1.87	0.56
1:A:292:ILE:HG23	1:A:318:LEU:HD13	1.87	0.56
1:E:291:ASN:OD1	1:E:292:ILE:N	2.42	0.53
4:H:3:DC:H2'	4:H:4:DT:C6	2.46	0.51
2:F:16:THR:HG23	2:F:19:GLU:H	1.76	0.50
1:A:208:THR:HG23	1:A:211:ASP:H	1.77	0.49
1:A:302:ASP:OD1	1:A:303:ASN:N	2.46	0.49
1:E:302:ASP:OD1	1:E:303:ASN:N	2.49	0.46
4:D:3:DC:H2'	4:D:4:DT:C6	2.51	0.46
4:H:4:DT:H2'	4:H:5:DT:C6	2.51	0.46
1:E:201:ARG:HB3	1:E:253:LEU:HD11	1.98	0.46
1:A:303:ASN:ND2	1:A:310:GLU:OE1	2.48	0.45
4:D:11:DC:H2'	4:D:12:DA:C8	2.52	0.45
1:E:320:MET:HE3	1:E:325:VAL:HG22	1.98	0.45
4:D:4:DT:H2'	4:D:5:DT:C6	2.53	0.44
1:E:191:LEU:HD13	1:E:242:MET:HE2	1.99	0.43
4:D:6:DA:H2''	4:D:7:DT:H5''	2.00	0.43
4:H:6:DA:H2''	4:H:7:DT:H5''	2.01	0.43
1:E:242:MET:SD	2:F:17:VAL:HG21	2.59	0.42
1:A:231:ARG:HB3	1:A:237:LEU:HG	2.00	0.42
1:A:202:ARG:HG3	1:A:203:ILE:N	2.36	0.41
3:C:14:DG:H2'	3:C:15:DG:C8	2.56	0.41
1:A:208:THR:HG22	1:A:211:ASP:CG	2.47	0.40

All (7) 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)
4:D:13:DT:OP2	1:E:193:LYS:HZ1[2_555]	1.18	0.42
4:D:13:DT:OP2	1:E:193:LYS:NZ[2_555]	1.88	0.32
3:C:1:DT:H3	4:H:1:DA:N1[1_556]	1.31	0.29
4:D:1:DA:H61	3:G:1:DT:O4[1_565]	1.35	0.25
4:D:1:DA:N1	3:G:1:DT:H3[1_565]	1.45	0.15
4:D:1:DA:N1	3:G:1:DT:N3[1_565]	2.09	0.11
3:C:1:DT:N3	4:H:1:DA:N1[1_556]	2.15	0.05

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	135/160 (84%)	135 (100%)	0	0	100	100
1	E	131/160 (82%)	130 (99%)	1 (1%)	0	100	100
2	B	30/39 (77%)	30 (100%)	0	0	100	100
2	F	20/39 (51%)	20 (100%)	0	0	100	100
All	All	316/398 (79%)	315 (100%)	1 (0%)	0	100	100

There are no Ramachandran outliers to report.

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	129/148 (87%)	129 (100%)	0	100	100
1	E	125/148 (84%)	124 (99%)	1 (1%)	79	93
2	B	26/30 (87%)	26 (100%)	0	100	100
2	F	20/30 (67%)	20 (100%)	0	100	100
All	All	300/356 (84%)	299 (100%)	1 (0%)	91	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	288	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	317	GLN
2	F	9	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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 [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	139/160 (86%)	0.23	1 (0%) 84 79	53, 90, 137, 149	0
1	E	135/160 (84%)	0.49	9 (6%) 25 19	69, 109, 161, 177	0
2	B	32/39 (82%)	0.42	0 100 100	63, 95, 148, 162	0
2	F	22/39 (56%)	0.79	0 100 100	90, 122, 144, 154	0
3	C	15/15 (100%)	-0.39	0 100 100	50, 72, 115, 128	0
3	G	15/15 (100%)	-0.39	0 100 100	64, 81, 121, 135	0
4	D	15/15 (100%)	-0.40	0 100 100	59, 77, 115, 123	0
4	H	15/15 (100%)	-0.32	0 100 100	63, 70, 122, 136	0
All	All	388/458 (84%)	0.27	10 (2%) 57 49	50, 99, 152, 177	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	198	PHE	3.4
1	E	252	TRP	2.7
1	E	212	VAL	2.7
1	E	300	PHE	2.5
1	E	341	CYS	2.4
1	E	256	ALA	2.4
1	A	188	LEU	2.3
1	E	185	PRO	2.1
1	E	340	ASN	2.1
1	E	257	GLU	2.1

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 oligosaccharides 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.