WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Mar 4, 2024 – 09:44 AM EST

PDB ID	:	4NUU
Title	:	Heterotrimer structure of Region II from Plasmodium vivax Duffy Binding Protein (PvDBP) bound to the ectodomain of the Duffy Antigen Receptor for Chemokines (DARC)
Authors	:	Tolia, N.H.
Deposited on	:	2013-12-04
Resolution	:	1.95 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:

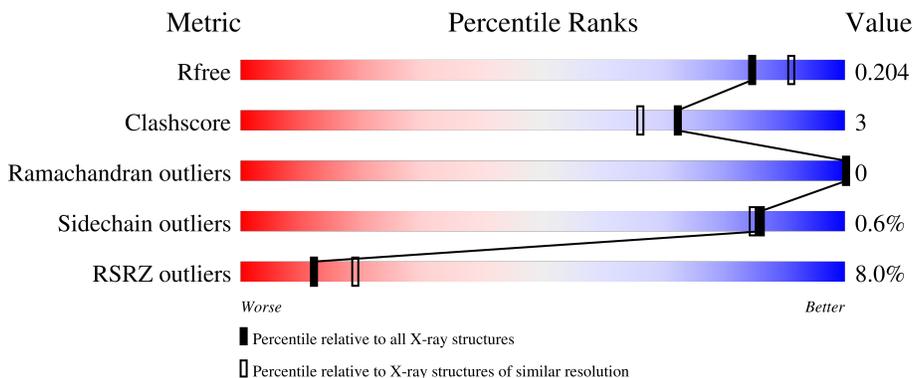|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 1.95 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
Rfree		0.204
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		0.6%
RSRZ outliers		8.0%

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))
R$_{free}$	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	317	8% / 83% 6% 11%
1	B	317	7% / 84% 8% · 8%
2	C	32	6% / 31% 6% 62%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 10257 atoms, of which 4866 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 Duffy receptor.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	282	Total 4724	C 1497	H 2355	N 417	O 437	S 18	0	5	0
1	B	292	Total 4880	C 1555	H 2426	N 435	O 446	S 18	0	3	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	209	ALA	-	expression tag	UNP P22290
A	210	SER	-	expression tag	UNP P22290
B	209	ALA	-	expression tag	UNP P22290
B	210	SER	-	expression tag	UNP P22290

- Molecule 2 is a protein called Duffy antigen/chemokine receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	12	Total 191	C 68	H 85	N 15	O 23	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	12	GLY	-	expression tag	UNP Q16570
C	13	PRO	-	expression tag	UNP Q16570
C	15	GLY	GLU	cloning artifact	UNP Q16570

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	202	Total 202	O 202	0	0

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

| Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	257	Total O 257 257	0	0
3	C	3	Total O 3 3	0	0

3 Residue-property plots ⓘ

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.
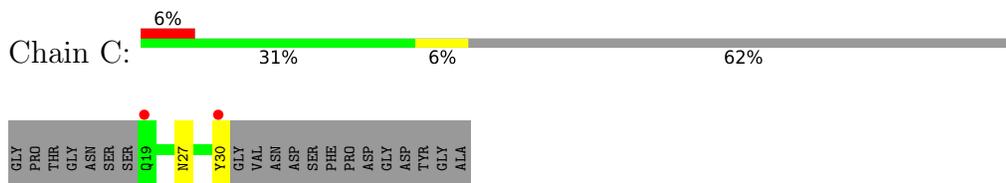
- Molecule 1: Duffy receptor

Chain A:

- Molecule 1: Duffy receptor

Chain B:
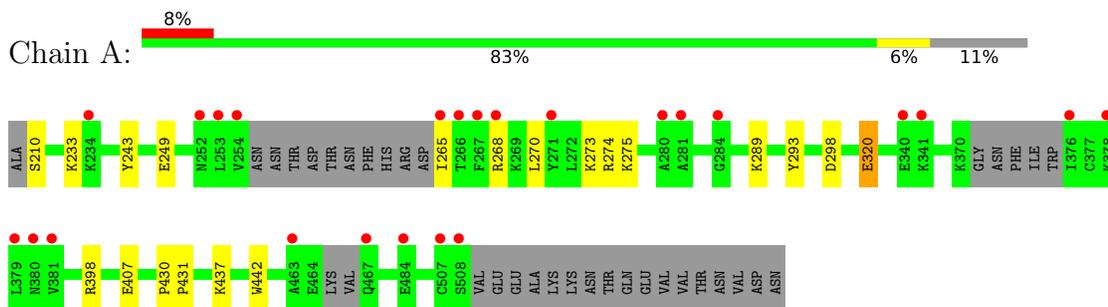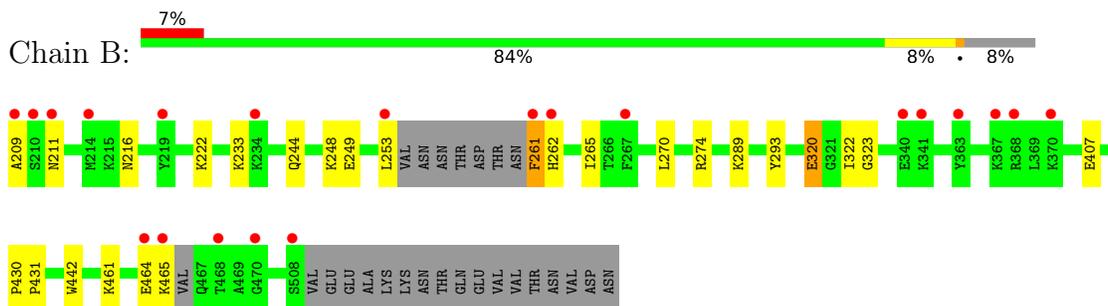
- Molecule 2: Duffy antigen/chemokine receptor

Chain C:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	59.59Å 66.99Å 97.92Å 90.00° 102.11° 90.00°	Depositor
Resolution (Å)	19.86 – 1.95 19.86 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.86-1.95) 99.9 (19.86-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.29 (at 1.94Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.165 , 0.202 0.168 , 0.204	Depositor DCC
R_{free} test set	1870 reflections (3.40%)	wwPDB-VP
Wilson B-factor (Å^2)	27.0	Xtriage
Anisotropy	0.183	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 53.8	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	10257	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 32.97 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.6162e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 ⓘ

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	0.25	0/2416	0.41	0/3245
1	B	0.25	0/2504	0.41	0/3363
2	C	0.24	0/109	0.30	0/148
All	All	0.25	0/5029	0.41	0/6756

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 ⓘ

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	2369	2355	2357	13	0
1	B	2454	2426	2435	21	0
2	C	106	85	85	2	0
3	A	202	0	0	2	0
3	B	257	0	0	1	0
3	C	3	0	0	0	0
All	All	5391	4866	4877	33	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 3.

All (33) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:209:ALA:N	1:B:323:GLY:N	2.32	0.77
1:B:209:ALA:N	1:B:323:GLY:CA	2.58	0.66
1:B:216:ASN:HB2	1:B:244:GLN:O	1.97	0.64
1:B:464:GLU:O	1:B:465:LYS:HB2	1.98	0.63
1:B:248:LYS:NZ	1:B:249:GLU:OE2	2.33	0.61
1:A:274:ARG:NH1	1:B:262:HIS:NE2	2.50	0.60
1:A:265:ILE:N	3:A:786:HOH:O	2.37	0.57
1:A:265:ILE:HG13	1:A:268:ARG:H	1.73	0.53
1:A:298:ASP:OD1	1:A:398:ARG:NH2	2.43	0.52
1:B:253:LEU:HD12	1:B:253:LEU:C	2.32	0.50
1:B:253:LEU:HD13	1:B:261:PHE:N	2.27	0.50
1:B:289:LYS:HG2	1:B:293:TYR:CZ	2.46	0.50
1:B:209:ALA:N	1:B:322:ILE:HG23	2.27	0.49
1:B:222:LYS:HA	1:B:222:LYS:HE2	1.95	0.48
1:A:289:LYS:HG2	1:A:293:TYR:CZ	2.49	0.48
1:B:430:PRO:HB2	1:B:431:PRO:HD3	1.97	0.47
1:A:437:LYS:NZ	3:A:769:HOH:O	2.47	0.47
1:A:233:LYS:HD2	1:A:320:GLU:O	2.15	0.47
1:B:233:LYS:HD2	1:B:320:GLU:O	2.15	0.47
1:B:407:GLU:HG3	1:B:442:TRP:CD2	2.50	0.46
1:B:211:ASN:ND2	1:B:322:ILE:O	2.50	0.45
1:B:461:LYS:O	1:B:464:GLU:O	2.34	0.45
1:A:249:GLU:HB3	1:A:275:LYS:HD2	1.98	0.44
1:B:270:LEU:O	1:B:274:ARG:HG2	2.18	0.43
1:B:464:GLU:O	1:B:465:LYS:CB	2.65	0.43
1:A:407:GLU:HG3	1:A:442:TRP:CD2	2.53	0.43
1:B:209:ALA:N	1:B:323:GLY:H	2.16	0.43
1:B:209:ALA:N	1:B:322:ILE:CG2	2.82	0.43
1:A:273:LYS:NZ	2:C:30:TYR:OH	2.48	0.42
1:B:265:ILE:HD11	2:C:27:ASN:O	2.19	0.42
1:A:430:PRO:HB2	1:A:431:PRO:HD3	2.02	0.41
1:A:270:LEU:CD1	3:B:851:HOH:O	2.69	0.41
1:A:210:SER:HB3	1:A:243:TYR:O	2.21	0.40

There are no symmetry-related clashes.

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	278/317 (88%)	273 (98%)	5 (2%)	0	100	100
1	B	288/317 (91%)	284 (99%)	4 (1%)	0	100	100
2	C	10/32 (31%)	10 (100%)	0	0	100	100
All	All	576/666 (86%)	567 (98%)	9 (2%)	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	263/292 (90%)	262 (100%)	1 (0%)	91	90
1	B	270/292 (92%)	268 (99%)	2 (1%)	84	82
2	C	12/26 (46%)	12 (100%)	0	100	100
All	All	545/610 (89%)	542 (99%)	3 (1%)	86	85

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	320	GLU
1	B	261	PHE
1	B	320	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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 monosaccharides 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 ⓘ

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(Å²)	Q<0.9
1	A	282/317 (88%)	0.38	24 (8%)	10	17	17, 33, 71, 111	0
1	B	292/317 (92%)	0.23	21 (7%)	15	23	18, 31, 68, 119	0
2	C	12/32 (37%)	0.96	2 (16%)	1	2	34, 46, 75, 98	0
All	All	586/666 (87%)	0.32	47 (8%)	12	19	17, 33, 71, 119	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	209	ALA	10.3
1	A	254	VAL	9.7
1	A	265	ILE	9.1
1	A	267	PHE	6.4
1	B	261	PHE	6.0
1	A	376	ILE	5.5
1	B	253	LEU	5.0
1	A	252	ASN	4.6
1	A	271	TYR	4.4
1	B	340	GLU	4.2
1	A	340	GLU	4.2
1	A	234	LYS	4.1
1	A	507	CYS	4.1
1	A	266	THR	4.0
2	C	30	TYR	3.9
1	A	463	ALA	3.9
1	B	464	GLU	3.8
1	A	253	LEU	3.7
1	B	210	SER	3.5
2	C	19	GLN	3.5
1	B	470	GLY	3.3
1	B	368	ARG	3.1
1	B	465	LYS	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	211	ASN	3.0
1	A	379	LEU	2.9
1	B	219	TYR	2.9
1	B	363	TYR	2.9
1	B	234	LYS	2.9
1	B	508	SER	2.8
1	A	467	GLN	2.8
1	B	367	LYS	2.7
1	A	484	GLU	2.7
1	B	468	THR	2.7
1	B	370	LYS	2.7
1	A	508	SER	2.6
1	A	341	LYS	2.5
1	B	341	LYS	2.5
1	A	380	ASN	2.3
1	A	268	ARG	2.2
1	A	280	ALA	2.1
1	B	214	MET	2.1
1	B	267	PHE	2.1
1	B	262	HIS	2.0
1	A	281	ALA	2.0
1	A	284	GLY	2.0
1	A	378	LYS	2.0
1	A	381	VAL	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.