



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Mar 2, 2017 – 11:59 am GMT

PDB ID : 4UMM
EMDB ID: : EMD-2631
Title : The Cryo-EM structure of the palindromic DNA-bound USP-EcR nuclear receptor reveals an asymmetric organization with allosteric domain positioning
Authors : Maletta, M.; Orlov, I.; Moras, D.; Billas, I.M.L.; Klaholz, B.P.
Deposited on : 2014-05-19
Resolution : 11.60 Å(reported)
Based on PDB ID : 1R1K

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : recalc29047

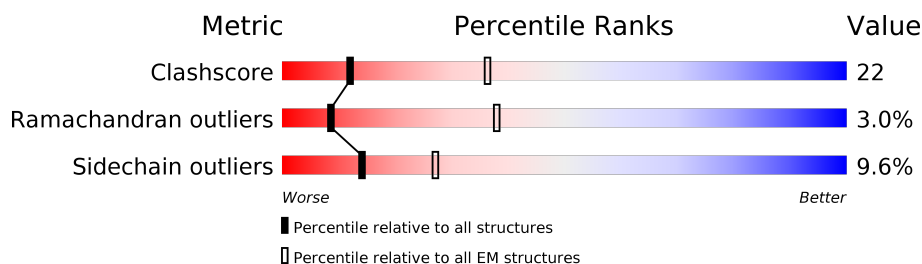
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 11.60 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)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	78	
2	C	20	
3	D	20	
4	E	87	
5	F	264	
6	G	266	

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 6030 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 ECR-USP.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	78	Total	C	N	O	S	0	0
			639	386	129	113	11		

- Molecule 2 is a DNA chain called 5'-D(*CP*AP*AP*GP*GP*GP*TP*TP*CP*AP*AP*TP*GP*CP *AP*CP*TP*TP*GP*TP)-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	20	Total	C	N	O	P	0	0
			407	196	74	118	19		

- Molecule 3 is a DNA chain called 5'-D(*DGP*AP*CP*AP*AP*GP*TP*GP*CP*AP*TP*TP*GP*DAP *AP*CP*CP*CP*TP*T)-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	20	Total	C	N	O	P	0	0
			405	195	75	116	19		

- Molecule 4 is a protein called ECDYSONE RECEPTOR.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	87	Total	C	N	O	S	0	0
			692	424	132	120	16		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	193	MET	ILE	CONFLICT	UNP O18473

- Molecule 5 is a protein called GENE REGULATION PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	241	Total	C	N	O	S	0	0
			1925	1233	335	345	12		

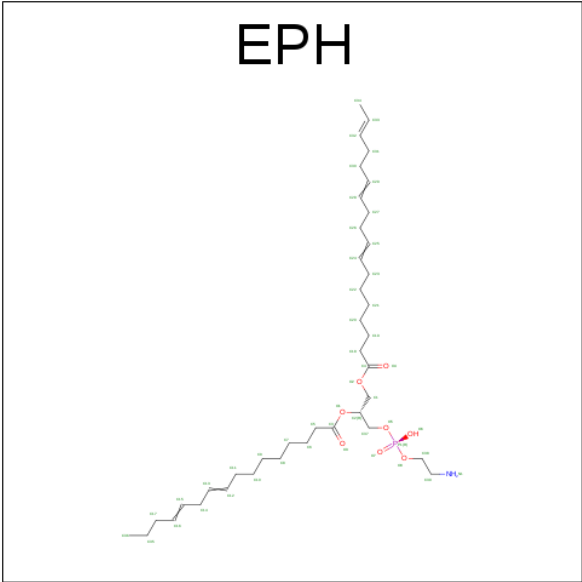
- Molecule 6 is a protein called ECDYSONE RECEPTOR.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	G	236	Total	C	N	O	S	0	0
			1872	1194	316	346	16		

There are 22 discrepancies between the modelled and reference sequences:

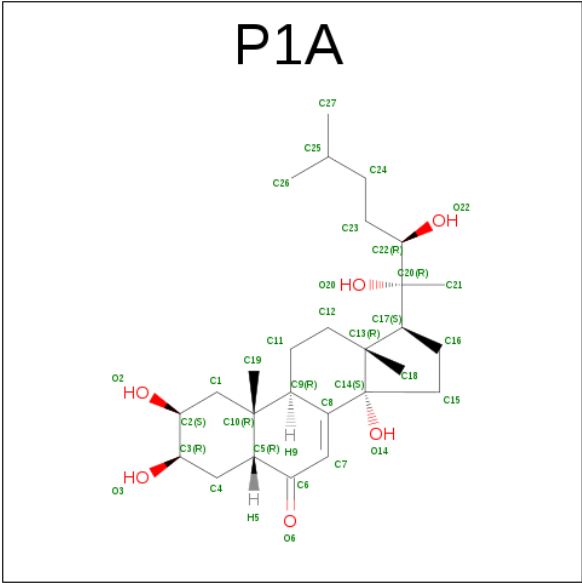
Chain	Residue	Modelled	Actual	Comment	Reference
G	267	GLY	-	EXPRESSION TAG	UNP O18473
G	268	SER	-	EXPRESSION TAG	UNP O18473
G	269	HIS	-	EXPRESSION TAG	UNP O18473
G	270	MET	-	EXPRESSION TAG	UNP O18473
G	271	ALA	-	EXPRESSION TAG	UNP O18473
G	272	SER	-	EXPRESSION TAG	UNP O18473
G	273	MET	-	EXPRESSION TAG	UNP O18473
G	274	THR	-	EXPRESSION TAG	UNP O18473
G	275	GLY	-	EXPRESSION TAG	UNP O18473
G	276	GLY	-	EXPRESSION TAG	UNP O18473
G	277	GLN	-	EXPRESSION TAG	UNP O18473
G	278	GLN	-	EXPRESSION TAG	UNP O18473
G	279	MET	-	EXPRESSION TAG	UNP O18473
G	280	GLY	-	EXPRESSION TAG	UNP O18473
G	281	ARG	-	EXPRESSION TAG	UNP O18473
G	282	ASP	-	EXPRESSION TAG	UNP O18473
G	283	PRO	-	EXPRESSION TAG	UNP O18473
G	322	THR	-	INSERTION	UNP O18473
G	323	TRP	-	INSERTION	UNP O18473
G	324	GLN	-	INSERTION	UNP O18473
G	483	CYS	GLY	CONFLICT	UNP O18473
G	489	LYS	GLU	CONFLICT	UNP O18473

- Molecule 7 is L-ALPHA-PHOSPHATIDYL-BETA-OLEOYL-GAMMA-PALMITOYL-PHOSPHATIDYLETHANOLAMINE (three-letter code: EPH) (formula: C₃₉H₆₈NO₈P).



Mol	Chain	Residues	Atoms					AltConf
7	F	1	Total	C	N	O	P	0
			49	39	1	8	1	

- Molecule 8 is 2,3,14,20,22-PENTAHYDROXYCHOLEST-7-EN-6-ONE (three-letter code: P1A) (formula: $C_{27}H_{44}O_6$).



Mol	Chain	Residues	Atoms			AltConf
8	G	1	Total	C	O	0
			33	27	6	

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		AltConf
9	F	1	Total 5	O 5	0
9	G	2	Total 3	O 3	0
9	F	2	Total 5	O 5	0
9	G	1	Total 3	O 3	0
9	F	2	Total 5	O 5	0

3 Residue-property plots [i](#)

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.

• Molecule 1: ECR-USP

Chain A:  92% 8%

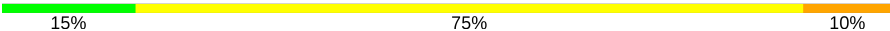


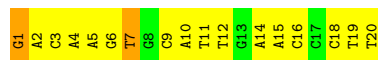
• Molecule 2: 5'-D(*CP*AP*AP*GP*GP*GP*TP*TP*CP*AP*AP*TP*GP*CP *AP*CP*TP*TP*GP*TP)-3'

Chain C:  45% 55%




• Molecule 3: 5'-D(*DGP*AP*CP*AP*AP*GP*TP*GP*CP*AP*TP*TP*GP*DAP *AP*CP*CP*CP*TP*T)-3'

Chain D:  15% 75% 10%



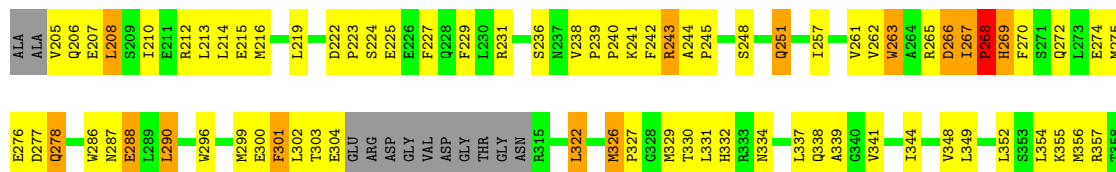
• Molecule 4: ECDYSONE RECEPTOR

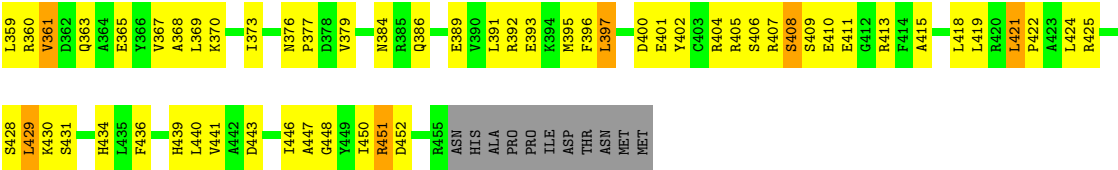
Chain E:  87% 10%



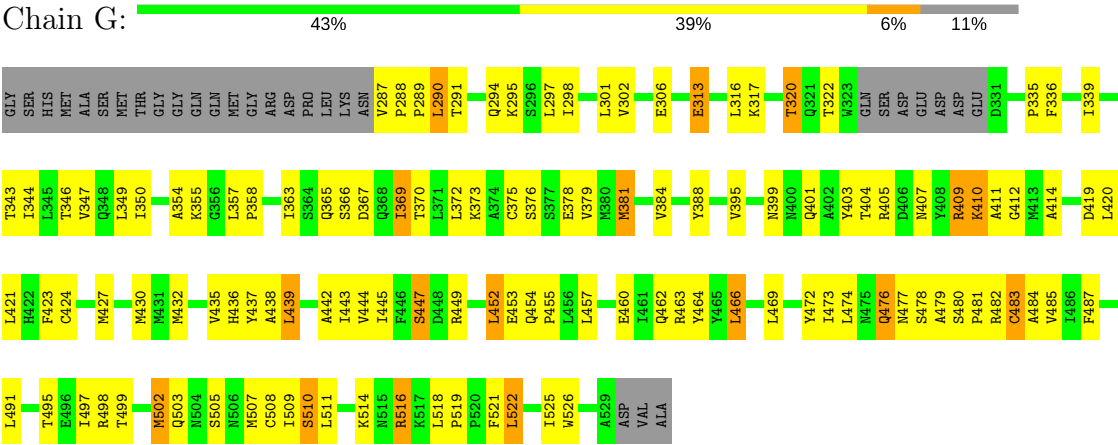
• Molecule 5: GENE REGULATION PROTEIN

Chain F:  41% 43% 7% 9%





● Molecule 6: ECDYSONE RECEPTOR



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	CCD IMAGES 4096X4096	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	100	Depositor
Electron dose ($e^-/\text{\AA}^2$)	20	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	64244	Depositor
Image detector	FEI EAGLE (4K X 4K)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: P1A, EPH

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.61	0/647	0.84	0/854
2	C	1.49	2/456 (0.4%)	2.02	15/703 (2.1%)
3	D	1.47	2/454 (0.4%)	2.13	26/699 (3.7%)
4	E	0.61	0/702	0.85	0/933
5	F	58.15	3/1960 (0.2%)	2.48	7/2646 (0.3%)
6	G	0.55	0/1905	0.72	0/2579
All	All	32.91	7/6124 (0.1%)	1.72	48/8414 (0.6%)

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
5	F	1	2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	268	PRO	N-CA	1657.13	29.64	1.47
5	F	268	PRO	CA-C	1434.96	30.22	1.52
5	F	268	PRO	CA-CB	1349.91	28.53	1.53
3	D	11	DT	C5-C7	6.85	1.54	1.50
3	D	4	DA	C3'-O3'	-6.70	1.35	1.44

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	268	PRO	N-CA-CB	-82.73	4.02	103.30
5	F	268	PRO	CA-N-CD	-52.78	37.60	111.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	268	PRO	CB-CA-C	-43.34	3.64	112.00
5	F	268	PRO	N-CA-C	-41.33	4.65	112.10
5	F	268	PRO	CA-C-O	-30.72	46.47	120.20

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	F	268	PRO	CA

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	F	267	ILE	Peptide
5	F	268	PRO	Mainchain

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	639	0	634	3	0
2	C	407	0	225	0	0
3	D	405	0	224	2	0
4	E	692	0	690	7	0
5	F	1925	0	1969	120	0
6	G	1872	0	1880	129	0
7	F	49	0	67	6	0
8	G	33	0	43	3	0
9	F	5	0	0	1	0
9	G	3	0	0	0	0
All	All	6030	0	5732	255	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 22.

The worst 5 of 255 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)
5:F:268:PRO:HB3	5:F:268:PRO:C	1.65	1.15
6:G:480:SER:HB2	6:G:481:PRO:HD2	1.51	0.91
6:G:403:TYR:HA	6:G:407:ASN:HD22	1.35	0.91
6:G:453:GLU:C	6:G:455:PRO:HD3	1.95	0.87
5:F:262:VAL:HG23	5:F:265:ARG:NH2	1.92	0.84

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 PDB entries followed by that with respect to all EM entries.

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	76/78 (97%)	74 (97%)	2 (3%)	0	100	100
4	E	85/87 (98%)	79 (93%)	5 (6%)	1 (1%)	15	57
5	F	237/264 (90%)	196 (83%)	33 (14%)	8 (3%)	4	35
6	G	232/266 (87%)	193 (83%)	29 (12%)	10 (4%)	3	29
All	All	630/695 (91%)	542 (86%)	69 (11%)	19 (3%)	9	37

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	F	268	PRO
5	F	301	PHE
6	G	322	THR
6	G	479	ALA
5	F	408	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 PDB entries followed by that with respect to all EM entries.

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	68/68 (100%)	68 (100%)	0	100	100
4	E	75/75 (100%)	73 (97%)	2 (3%)	50	74
5	F	212/231 (92%)	185 (87%)	27 (13%)	5	25
6	G	205/235 (87%)	180 (88%)	25 (12%)	6	26
All	All	560/609 (92%)	506 (90%)	54 (10%)	14	35

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	F	411	GLU
6	G	306	GLU
6	G	495	THR
5	F	421	LEU
5	F	430	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	338	GLN
5	F	376	ASN
6	G	462	GLN
5	F	343	GLN
5	F	363	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 carbohydrates in this entry.

5.6 Ligand geometry

2 ligands are 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	$\# Z > 2$
7	EPH	F	1456	-	48,48,48	1.62	8 (16%)	50,53,53	1.61	8 (16%)
8	P1A	G	1530	-	36,36,36	3.49	17 (47%)	59,60,60	2.31	18 (30%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	EPH	F	1456	-	-	0/52/52/52	0/0/0/0
8	P1A	G	1530	-	-	0/17/87/87	0/4/4/4

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	G	1530	P1A	O3-C3	-5.95	1.30	1.43
7	F	1456	EPH	P1-O8	-3.41	1.44	1.59
8	G	1530	P1A	O20-C20	-3.33	1.38	1.44
8	G	1530	P1A	O6-C6	-2.51	1.18	1.22
7	F	1456	EPH	P1-O6	-2.21	1.44	1.55

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	F	1456	EPH	O8-P1-O7	-4.78	89.97	109.25
7	F	1456	EPH	O5-P1-O7	-4.76	90.06	109.25
8	G	1530	P1A	C24-C23-C22	-4.76	106.52	112.52
7	F	1456	EPH	O6-P1-O7	-4.20	90.55	112.28
8	G	1530	P1A	O14-C14-C8	-4.11	98.71	106.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	F	1456	EPH	6	0
8	G	1530	P1A	3	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.