



wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 11:27 am BST

PDB ID : 5XUA
Title : The ligand-free dimer of chemoreceptor MCP2201 ligand binding domain
Authors : Hong, Y.; Li, D.F.; Wang, D.C.
Deposited on : 2017-06-23
Resolution : 2.80 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 ⓘ symbol.

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	:	2.11
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)	:	2.11

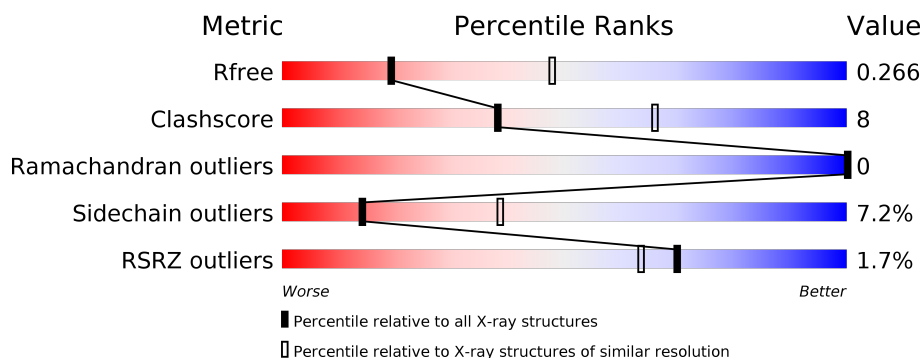
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (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 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 $\leq 5\%$. 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	154	<div> <div>%</div> <div> <div></div> <div>68%</div> <div>20%</div> <div>5%</div> <div>8%</div> </div> </div>
1	B	154	<div> <div>74%</div> <div>14%</div> <div>•</div> <div>11%</div> </div>
1	C	154	<div> <div>68%</div> <div>19%</div> <div>•</div> <div>10%</div> </div>
1	D	154	<div> <div>5%</div> <div>73%</div> <div>13%</div> <div>•</div> <div>11%</div> </div>
1	E	154	<div> <div>%</div> <div>68%</div> <div>18%</div> <div>•</div> <div>11%</div> </div>
1	F	154	<div> <div>2%</div> <div>70%</div> <div>18%</div> <div>12%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	154	<div><div>%</div><div><div></div><div>70%</div><div>18%</div><div>10%</div></div></div>
1	H	154	<div><div>3%</div><div><div></div><div>66%</div><div>19%</div><div>12%</div></div></div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8363 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 Methyl-accepting chemotaxis sensory transducer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	142	Total	C	N	O	S	0	0	0
			1082	675	185	218	4			
1	B	137	Total	C	N	O	S	0	0	0
			1040	649	179	208	4			
1	C	138	Total	C	N	O	S	0	0	0
			1047	654	180	209	4			
1	D	137	Total	C	N	O	S	0	0	0
			1040	649	179	208	4			
1	E	137	Total	C	N	O	S	0	0	0
			1040	649	179	208	4			
1	F	136	Total	C	N	O	S	0	0	0
			1032	644	178	207	3			
1	G	139	Total	C	N	O	S	0	0	0
			1056	659	181	212	4			
1	H	135	Total	C	N	O	S	0	0	0
			1026	641	177	204	4			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	50	MET	-	expression tag	UNP D0IVL9
A	51	HIS	-	expression tag	UNP D0IVL9
A	52	HIS	-	expression tag	UNP D0IVL9
A	53	HIS	-	expression tag	UNP D0IVL9
A	54	HIS	-	expression tag	UNP D0IVL9
A	55	HIS	-	expression tag	UNP D0IVL9
A	56	HIS	-	expression tag	UNP D0IVL9
B	50	MET	-	expression tag	UNP D0IVL9
B	51	HIS	-	expression tag	UNP D0IVL9
B	52	HIS	-	expression tag	UNP D0IVL9
B	53	HIS	-	expression tag	UNP D0IVL9
B	54	HIS	-	expression tag	UNP D0IVL9
B	55	HIS	-	expression tag	UNP D0IVL9

Continued on next page...

Continued from previous page...

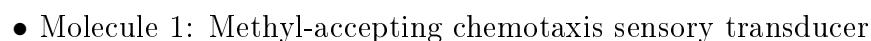
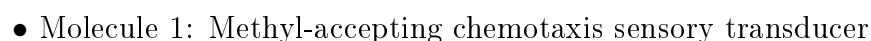
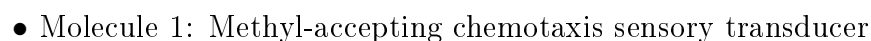
Chain	Residue	Modelled	Actual	Comment	Reference
B	56	HIS	-	expression tag	UNP D0IVL9
C	50	MET	-	expression tag	UNP D0IVL9
C	51	HIS	-	expression tag	UNP D0IVL9
C	52	HIS	-	expression tag	UNP D0IVL9
C	53	HIS	-	expression tag	UNP D0IVL9
C	54	HIS	-	expression tag	UNP D0IVL9
C	55	HIS	-	expression tag	UNP D0IVL9
C	56	HIS	-	expression tag	UNP D0IVL9
D	50	MET	-	expression tag	UNP D0IVL9
D	51	HIS	-	expression tag	UNP D0IVL9
D	52	HIS	-	expression tag	UNP D0IVL9
D	53	HIS	-	expression tag	UNP D0IVL9
D	54	HIS	-	expression tag	UNP D0IVL9
D	55	HIS	-	expression tag	UNP D0IVL9
D	56	HIS	-	expression tag	UNP D0IVL9
E	50	MET	-	expression tag	UNP D0IVL9
E	51	HIS	-	expression tag	UNP D0IVL9
E	52	HIS	-	expression tag	UNP D0IVL9
E	53	HIS	-	expression tag	UNP D0IVL9
E	54	HIS	-	expression tag	UNP D0IVL9
E	55	HIS	-	expression tag	UNP D0IVL9
E	56	HIS	-	expression tag	UNP D0IVL9
F	50	MET	-	expression tag	UNP D0IVL9
F	51	HIS	-	expression tag	UNP D0IVL9
F	52	HIS	-	expression tag	UNP D0IVL9
F	53	HIS	-	expression tag	UNP D0IVL9
F	54	HIS	-	expression tag	UNP D0IVL9
F	55	HIS	-	expression tag	UNP D0IVL9
F	56	HIS	-	expression tag	UNP D0IVL9
G	50	MET	-	expression tag	UNP D0IVL9
G	51	HIS	-	expression tag	UNP D0IVL9
G	52	HIS	-	expression tag	UNP D0IVL9
G	53	HIS	-	expression tag	UNP D0IVL9
G	54	HIS	-	expression tag	UNP D0IVL9
G	55	HIS	-	expression tag	UNP D0IVL9
G	56	HIS	-	expression tag	UNP D0IVL9
H	50	MET	-	expression tag	UNP D0IVL9
H	51	HIS	-	expression tag	UNP D0IVL9
H	52	HIS	-	expression tag	UNP D0IVL9
H	53	HIS	-	expression tag	UNP D0IVL9
H	54	HIS	-	expression tag	UNP D0IVL9
H	55	HIS	-	expression tag	UNP D0IVL9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	56	HIS	-	expression tag	UNP D0IVL9

- Molecule 1: Methyl-accepting chemotaxis sensory transducer



GLN
ILE
GLU
SER
SER
SER
ARG
THR

- Molecule 1: Methyl-accepting chemotaxis sensory transducer



MET HIS HIS HIS HIS HIS HIS HIS M57 M58 Q59 E60 P61 K64 I68 N74 A78 T82 S88 A91 V94 T104 T108 V111 Q114 I115 L118 L119 I127 V134 K149 M153 A154 E155 E156 A157 E158 L161 Y165 V166 Y172

L179 L180 Q183 L187 K190 E193 VAL GLU GLN ILE GLU SER SER ARG THR

- Molecule 1: Methyl-accepting chemotaxis sensory transducer



MET HIS HIS HIS HIS HIS HIS HIS M57 M58 K64 K68 N71 N72 A78 T82 I85 I93 Q94 Q95 S107 V111 L112 K113 Q114 I115 L118 I127 V134 R135 M153 L161 V166 P167 Y172 L175 L176 L179 L180 L187 D188 A189

E193 VAL GLU GLN ILE GLU SER SER ARG THR

- Molecule 1: Methyl-accepting chemotaxis sensory transducer



MET HIS HIS HIS HIS HIS HIS HIS M57 M58 E60 P61 L62 A91 S92 I93 V94 V111 I115 R125 E126 I127 L128 M132 R135 I139 A140 V145 S146 Q147 L148 K149 E155 E158 L161 I162 V166 Y172 L173 K174 E178 L179 A192 E193 V194

E195 GLN ILE GLU SER SER SER ARG THR

- Molecule 1: Methyl-accepting chemotaxis sensory transducer



MET HIS HIS HIS HIS HIS HIS HIS M57 M58 Q59 E60 P61 L62 L67 I68 L93 A78 T82 A86 K87 S88 V111 Q114 I115 L118 I119 E124 R125 D129 Q133 V134 R135 I139 R142 V145 A150 M153 L161 S164 Y165 V166 P167 Y172

L180 L187 K190 A191 ALA GLU VAL GLU GLN ILE GLU SER SER ARG THR

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	64.78Å 105.03Å 91.07Å 90.00° 90.95° 90.00°	Depositor
Resolution (Å)	53.20 – 2.80 55.13 – 2.80	Depositor EDS
% Data completeness (in resolution range)	95.1 (53.20-2.80) 94.8 (55.13-2.80)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.35 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.222 , 0.266 0.222 , 0.266	Depositor DCC
R_{free} test set	1520 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	55.6	Xtriage
Anisotropy	0.320	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 42.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	8363	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.40% 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	0.27	0/1091	0.50	0/1476
1	B	0.26	0/1049	0.47	0/1419
1	C	0.26	0/1056	0.46	0/1429
1	D	0.26	0/1049	0.46	0/1419
1	E	0.27	0/1049	0.47	0/1419
1	F	0.26	0/1041	0.48	0/1409
1	G	0.28	0/1065	0.45	0/1441
1	H	0.26	0/1035	0.47	0/1400
All	All	0.26	0/8435	0.47	0/11412

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	1082	0	1117	24	0
1	B	1040	0	1077	14	0
1	C	1047	0	1086	19	0
1	D	1040	0	1077	15	0
1	E	1040	0	1077	18	0
1	F	1032	0	1068	13	0
1	G	1056	0	1092	18	0
1	H	1026	0	1066	23	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	8363	0	8660	129	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 8.

The worst 5 of 129 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:A:132:MET:SD	1:A:135:ARG:NH2	2.49	0.86
1:C:105:THR:HG21	1:H:150:ALA:HB1	1.66	0.78
1:B:61:PRO:HB2	1:B:187:LEU:HD13	1.72	0.72
1:E:78:ALA:O	1:E:82:THR:OG1	2.13	0.67
1:E:161:LEU:O	1:E:166:VAL:HG23	1.96	0.65

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	140/154 (91%)	137 (98%)	3 (2%)	0	100	100
1	B	135/154 (88%)	129 (96%)	6 (4%)	0	100	100
1	C	136/154 (88%)	135 (99%)	1 (1%)	0	100	100
1	D	135/154 (88%)	129 (96%)	6 (4%)	0	100	100
1	E	135/154 (88%)	131 (97%)	4 (3%)	0	100	100
1	F	134/154 (87%)	130 (97%)	4 (3%)	0	100	100
1	G	137/154 (89%)	133 (97%)	4 (3%)	0	100	100
1	H	133/154 (86%)	128 (96%)	5 (4%)	0	100	100
All	All	1085/1232 (88%)	1052 (97%)	33 (3%)	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	117/129 (91%)	105 (90%)	12 (10%)	7	21
1	B	112/129 (87%)	106 (95%)	6 (5%)	22	53
1	C	113/129 (88%)	106 (94%)	7 (6%)	18	47
1	D	112/129 (87%)	103 (92%)	9 (8%)	12	34
1	E	112/129 (87%)	103 (92%)	9 (8%)	12	34
1	F	111/129 (86%)	105 (95%)	6 (5%)	22	53
1	G	114/129 (88%)	105 (92%)	9 (8%)	12	34
1	H	111/129 (86%)	104 (94%)	7 (6%)	18	46
All	All	902/1032 (87%)	837 (93%)	65 (7%)	14	38

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	104	THR
1	E	94	VAL
1	H	82	THR
1	D	111	VAL
1	D	172	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	95	GLN

5.3.3 RNA ⓘ

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 ⓘ

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(Å ²)	Q<0.9
1	A	142/154 (92%)	-0.08	1 (0%) 87 84	35, 58, 94, 121	0
1	B	137/154 (88%)	-0.13	0 100 100	36, 54, 97, 116	0
1	C	138/154 (89%)	-0.12	0 100 100	41, 56, 91, 102	0
1	D	137/154 (88%)	0.16	7 (5%) 28 19	41, 63, 108, 129	0
1	E	137/154 (88%)	0.04	2 (1%) 73 68	48, 72, 108, 119	0
1	F	136/154 (88%)	0.17	3 (2%) 62 52	41, 61, 102, 133	0
1	G	139/154 (90%)	0.13	2 (1%) 75 70	44, 63, 95, 124	0
1	H	135/154 (87%)	0.24	4 (2%) 50 40	44, 67, 115, 126	0
All	All	1101/1232 (89%)	0.05	19 (1%) 70 63	35, 62, 103, 133	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	119	ILE	4.3
1	H	57	MET	4.0
1	D	102	ALA	3.8
1	G	192	ALA	3.7
1	D	99	ALA	3.2

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 carbohydrates 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.