



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 12:30 PM GMT

PDB ID : 3VH7
Title : Structure of HIV-1 gp41 NHR/fusion inhibitor complex P21
Authors : Yao, X.; Waltersperger, S.; Wang, M.T.; Cui, S.
Deposited on : 2011-08-23
Resolution : 2.02 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

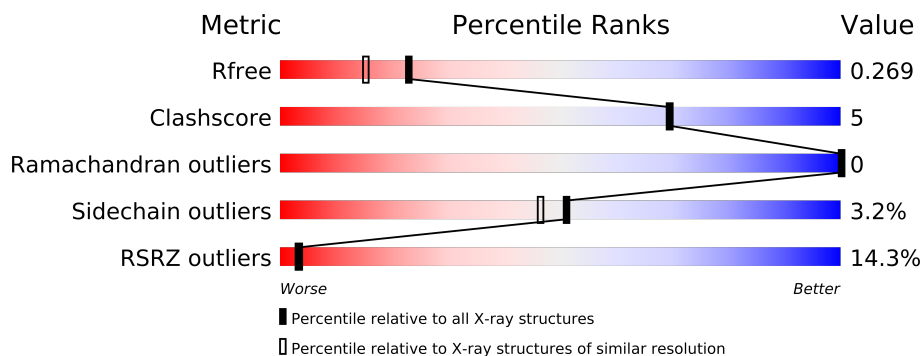
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.02 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}	66092	6003 (2.04-2.00)
Clashscore	79885	7467 (2.04-2.00)
Ramachandran outliers	78287	7370 (2.04-2.00)
Sidechain outliers	78261	7368 (2.04-2.00)
RSRZ outliers	66119	6006 (2.04-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	58	
1	C	58	
1	E	58	
2	B	34	
2	D	34	
2	F	34	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4273 atoms, of which 2092 are hydrogens and 0 are deuterium.

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 Envelope glycoprotein gp160.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	52	Total	C	H	N	O	0	0	0
			840	261	428	76	75			
1	C	51	Total	C	H	N	O	0	0	0
			827	258	421	74	73			
1	E	56	Total	C	H	N	O	0	0	0
			885	274	450	80	80			

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	GLY	-	EXPRESSION TAG	UNP P03375
A	-9	SER	-	EXPRESSION TAG	UNP P03375
A	-8	ALA	-	EXPRESSION TAG	UNP P03375
A	-7	MET	-	EXPRESSION TAG	UNP P03375
A	-6	ALA	-	EXPRESSION TAG	UNP P03375
A	-5	ASP	-	EXPRESSION TAG	UNP P03375
A	-4	ILE	-	EXPRESSION TAG	UNP P03375
A	-3	GLY	-	EXPRESSION TAG	UNP P03375
A	-2	SER	-	EXPRESSION TAG	UNP P03375
A	-1	GLU	-	EXPRESSION TAG	UNP P03375
A	0	PHE	-	EXPRESSION TAG	UNP P03375
A	1000	SER	-	EXPRESSION TAG	UNP P03375
A	1001	GLY	-	EXPRESSION TAG	UNP P03375
A	1002	GLY	-	EXPRESSION TAG	UNP P03375
A	1003	ARG	-	EXPRESSION TAG	UNP P03375
C	-10	GLY	-	EXPRESSION TAG	UNP P03375
C	-9	SER	-	EXPRESSION TAG	UNP P03375
C	-8	ALA	-	EXPRESSION TAG	UNP P03375
C	-7	MET	-	EXPRESSION TAG	UNP P03375
C	-6	ALA	-	EXPRESSION TAG	UNP P03375
C	-5	ASP	-	EXPRESSION TAG	UNP P03375
C	-4	ILE	-	EXPRESSION TAG	UNP P03375
C	-3	GLY	-	EXPRESSION TAG	UNP P03375

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	SER	-	EXPRESSION TAG	UNP P03375
C	-1	GLU	-	EXPRESSION TAG	UNP P03375
C	0	PHE	-	EXPRESSION TAG	UNP P03375
C	1000	SER	-	EXPRESSION TAG	UNP P03375
C	1001	GLY	-	EXPRESSION TAG	UNP P03375
C	1002	GLY	-	EXPRESSION TAG	UNP P03375
C	1003	ARG	-	EXPRESSION TAG	UNP P03375
E	-10	GLY	-	EXPRESSION TAG	UNP P03375
E	-9	SER	-	EXPRESSION TAG	UNP P03375
E	-8	ALA	-	EXPRESSION TAG	UNP P03375
E	-7	MET	-	EXPRESSION TAG	UNP P03375
E	-6	ALA	-	EXPRESSION TAG	UNP P03375
E	-5	ASP	-	EXPRESSION TAG	UNP P03375
E	-4	ILE	-	EXPRESSION TAG	UNP P03375
E	-3	GLY	-	EXPRESSION TAG	UNP P03375
E	-2	SER	-	EXPRESSION TAG	UNP P03375
E	-1	GLU	-	EXPRESSION TAG	UNP P03375
E	0	PHE	-	EXPRESSION TAG	UNP P03375
E	1000	SER	-	EXPRESSION TAG	UNP P03375
E	1001	GLY	-	EXPRESSION TAG	UNP P03375
E	1002	GLY	-	EXPRESSION TAG	UNP P03375
E	1003	ARG	-	EXPRESSION TAG	UNP P03375

- Molecule 2 is a protein called CP32M.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	30	Total	C	H	N	O	S	2	0	0
			529	178	252	41	56	2			
2	D	31	Total	C	H	N	O	S	0	0	0
			545	183	261	42	57	2			
2	F	31	Total	C	H	N	O	S	0	2	0
			578	192	280	45	59	2			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	12	Total 12	O 12	0	0
4	B	13	Total 13	O 13	0	0
4	C	8	Total 8	O 8	0	0
4	D	7	Total 7	O 7	0	0
4	E	16	Total 16	O 16	0	0
4	F	12	Total 12	O 12	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 errors 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.

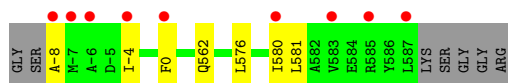
- Molecule 1: Envelope glycoprotein gp160

Chain A: 



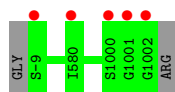
- Molecule 1: Envelope glycoprotein gp160

Chain C: 



- Molecule 1: Envelope glycoprotein gp160

Chain E: 



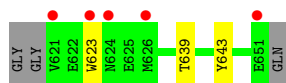
- Molecule 2: CP32M

Chain B: 



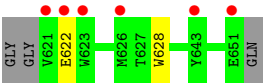
- Molecule 2: CP32M

Chain D: 



- Molecule 2: CP32M

Chain F: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	50.51Å 45.50Å 55.51Å 90.00° 107.65° 90.00°	Depositor
Resolution (Å)	42.61 – 2.02 42.61 – 2.02	Depositor EDS
% Data completeness (in resolution range)	93.4 (42.61-2.02) 93.6 (42.61-2.02)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.57 (at 2.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.222 , 0.274 0.217 , 0.269	Depositor DCC
R_{free} test set	748 reflections (5.01%)	DCC
Wilson B-factor (Å ²)	25.6	Xtriage
Anisotropy	0.712	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.42 , 43.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 14960 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4273	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.54% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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

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	0.22	0/416	0.36	0/560
1	C	0.22	0/410	0.36	0/553
1	E	0.23	0/439	0.36	0/590
2	B	0.44	0/284	0.49	0/383
2	D	0.22	0/291	0.35	0/393
2	F	0.21	0/311	0.38	0/419
All	All	0.26	0/2151	0.38	0/2898

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	412	428	2	6	0
1	C	406	421	2	4	0
1	E	435	450	0	0	0
2	B	277	252	0	5	0
2	D	284	261	0	1	0
2	F	298	280	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	1	0	0	0	0
4	A	12	0	0	0	0
4	B	13	0	0	0	0
4	C	8	0	0	0	0
4	D	7	0	0	0	0
4	E	16	0	0	0	0
4	F	12	0	0	0	0
All	All	2181	2092	4	13	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

All (13) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:567:GLN:NE2	2:B:643:TYR:OH	2.32	0.62
1:A:0:PHE:O	1:A:549:VAL:HG22	2.08	0.53
1:A:576:LEU:O	1:A:580:ILE:HG12	2.10	0.51
1:C:576:LEU:O	1:C:580:ILE:HG12	2.11	0.51
2:B:631:TRP:O	2:B:635:ILE:HG12	2.16	0.46
2:B:639:THR:HG22	2:B:643:TYR:CE2	2.52	0.45
1:A:553:ASN:O	1:A:557:ARG:HG3	2.19	0.43
2:D:639:THR:HG22	2:D:643:TYR:CE2	2.53	0.42
2:B:624:ASN:N	2:B:624:ASN:OD1	2.53	0.42
1:C:-8:ALA:O	1:C:-4:ILE:HG12	2.19	0.41
1:A:580:ILE:CD1	1:C:580:ILE:HD12	2.50	0.41
1:C:580:ILE:HD13	1:C:580:ILE:N	2.36	0.40
1:A:574:LYS:HE2	2:B:632:GLU:OE1	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	50/58 (86%)	48 (96%)	2 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	49/58 (84%)	49 (100%)	0	0	100	100
1	E	54/58 (93%)	53 (98%)	1 (2%)	0	100	100
2	B	28/34 (82%)	28 (100%)	0	0	100	100
2	D	29/34 (85%)	29 (100%)	0	0	100	100
2	F	31/34 (91%)	30 (97%)	1 (3%)	0	100	100
All	All	241/276 (87%)	237 (98%)	4 (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	43/46 (94%)	43 (100%)	0	100	100
1	C	42/46 (91%)	39 (93%)	3 (7%)	21	13
1	E	45/46 (98%)	45 (100%)	0	100	100
2	B	30/32 (94%)	29 (97%)	1 (3%)	50	45
2	D	31/32 (97%)	30 (97%)	1 (3%)	51	46
2	F	33/32 (103%)	31 (94%)	2 (6%)	26	18
All	All	224/234 (96%)	217 (97%)	7 (3%)	51	48

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	623	TRP
1	C	0	PHE
1	C	562	GLN
1	C	581	LEU
2	D	623	TRP
2	F	622	GLU
2	F	628	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	567	GLN
2	B	650	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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(Å ²)	Q<0.9
1	A	52/58 (89%)	0.81	9 (17%)	2	2	18, 35, 80, 96	0
1	C	51/58 (87%)	0.66	9 (17%)	2	2	18, 33, 104, 125	0
1	E	56/58 (96%)	0.54	5 (8%)	10	9	17, 36, 77, 101	0
2	B	30/34 (88%)	0.49	2 (6%)	17	16	21, 33, 82, 86	1 (3%)
2	D	31/34 (91%)	0.82	5 (16%)	2	2	24, 37, 92, 104	0
2	F	31/34 (91%)	0.70	6 (19%)	2	2	23, 35, 77, 88	0
All	All	251/276 (90%)	0.67	36 (14%)	3	3	17, 35, 87, 125	1 (0%)

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	623	TRP	6.0
2	D	621	VAL	5.5
1	E	1001	GLY	4.6
2	F	623	TRP	4.4
2	F	621	VAL	4.2
1	C	-4	ILE	3.5
1	A	1001	GLY	3.4
1	C	583	VAL	3.3
1	E	580	ILE	3.3
1	A	-1	GLU	3.3
1	C	580	ILE	3.2
1	A	-4	ILE	3.2
2	D	626	MET	3.1
2	F	626	MET	3.1
1	C	-7	MET	2.9
2	D	623	TRP	2.9
1	C	-6	ALA	2.9
1	E	1000	SER	2.9
2	F	622	GLU	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	580	ILE	2.8
1	C	0	PHE	2.7
1	A	583	VAL	2.7
2	F	651	GLU	2.6
1	A	546	SER	2.6
2	D	651	GLU	2.6
1	C	-8	ALA	2.5
2	F	643	TYR	2.4
1	A	-5	ASP	2.4
1	C	587	LEU	2.4
1	E	1002	GLY	2.3
2	D	624	ASN	2.3
1	C	585	ARG	2.3
1	E	-9	SER	2.2
2	B	624	ASN	2.2
1	A	549	VAL	2.2
1	A	0	PHE	2.1

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 ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å ²)	Q<0.9
3	MG	E	1101	1/1	0.09	-0.81	44,44,44,44	0

6.5 Other polymers ⓘ

There are no such residues in this entry.