wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:29 PM GMT

PDB ID	:	1KF9
Title	:	PHAGE DISPLAY DERIVED VARIANT OF HUMAN GROWTH HORMONE COMPLEXED WITH TWO COPIES OF THE EXTRACELLULAR DOMAIN OF ITS RECEPTOR
Authors	:	Schiffer, C.A.; Ultsch, M.; Walsh, S.; Somers, W.; De Vos, A.M.; Kossiakoff, A.A.
Deposited on	:	2001-11-19
Resolution	:	2.60 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

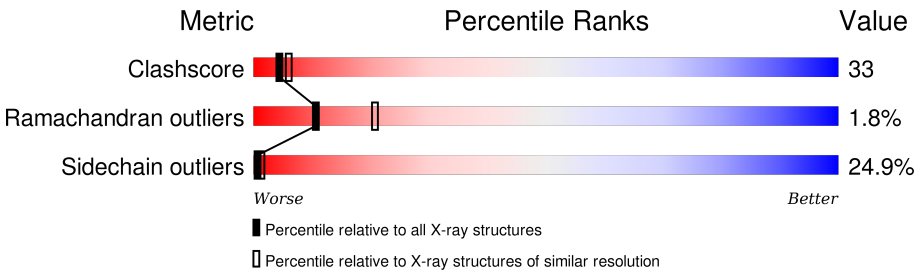
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	191	33% 38% 11% • 16%
1	D	191	37% 36% 9% • 17%
2	B	238	38% 30% 13% 19%
2	C	238	30% 32% 12% 26%
2	E	238	34% 35% 12% 19%
2	F	238	29% 36% 12% 23%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8678 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 PHAGE DISPLAY DERIVED VARIANT HUMAN GROWTH HORMONE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	160	1284	825	211	242	6	0	0	0
1	D	158	1264	811	208	239	6	0	0	0

- Molecule 2 is a protein called EXTRACELLULAR DOMAIN HUMAN GROWTH HORMONE RECEPTOR (1-238).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	193	1577	1011	258	298	10	0	0	0
2	C	175	1419	916	229	265	9	0	0	0
2	E	193	1576	1011	258	298	9	0	0	0
2	F	184	1484	952	239	284	9	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	11	11	11	0	0
3	B	13	13	13	0	0
3	C	28	28	28	0	0
3	D	3	3	3	0	0
3	E	12	12	12	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	F	7	7	7	0	0

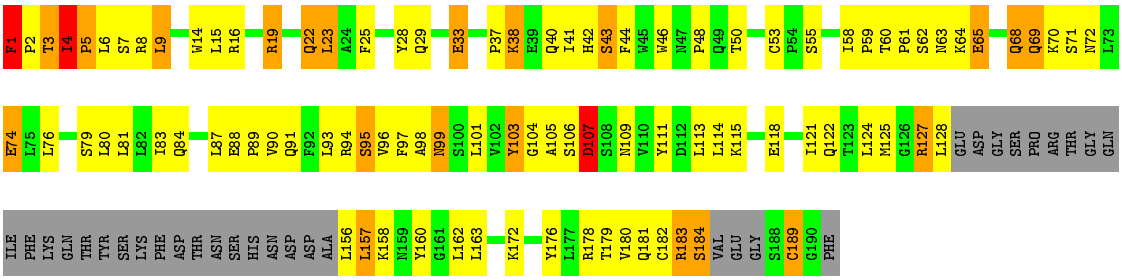
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

Note EDS was not executed.

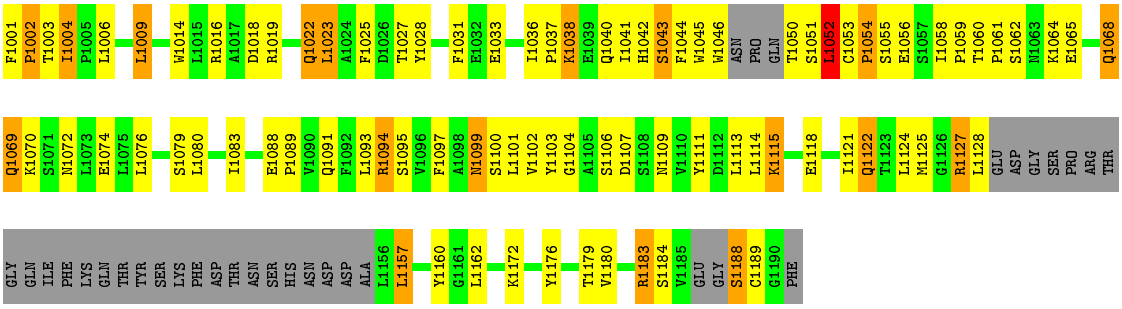
• Molecule 1: PHAGE DISPLAY DERIVED VARIANT HUMAN GROWTH HORMONE

Chain A: 33% 38% 11% • 16%



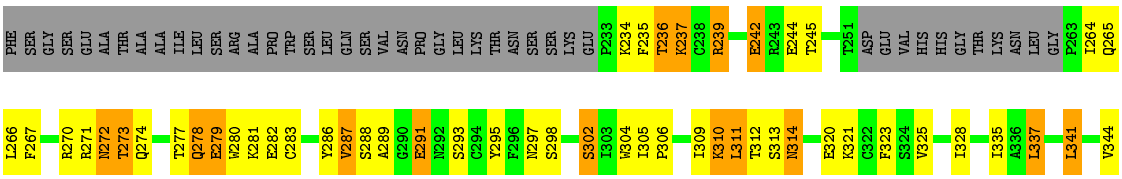
• Molecule 1: PHAGE DISPLAY DERIVED VARIANT HUMAN GROWTH HORMONE

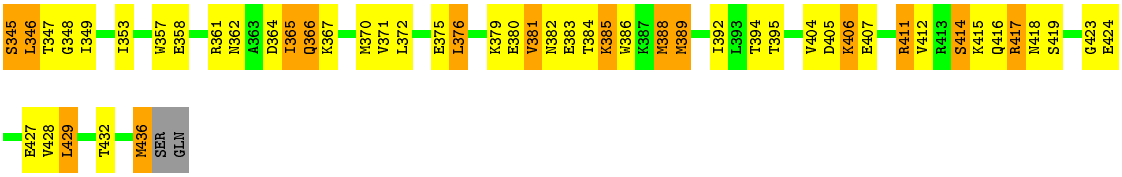
Chain D: 37% 36% 9% • 17%



• Molecule 2: EXTRACELLULAR DOMAIN HUMAN GROWTH HORMONE RECEPTOR (1-238)

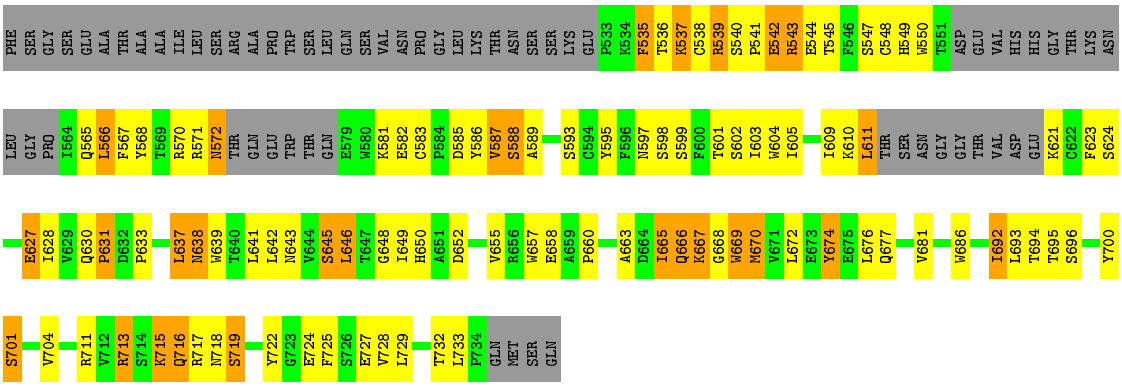
Chain B: 38% 30% 13% 19%





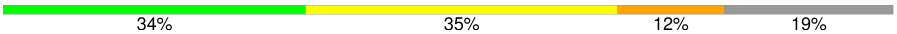
• Molecule 2: EXTRACELLULAR DOMAIN HUMAN GROWTH HORMONE RECEPTOR (1-238)

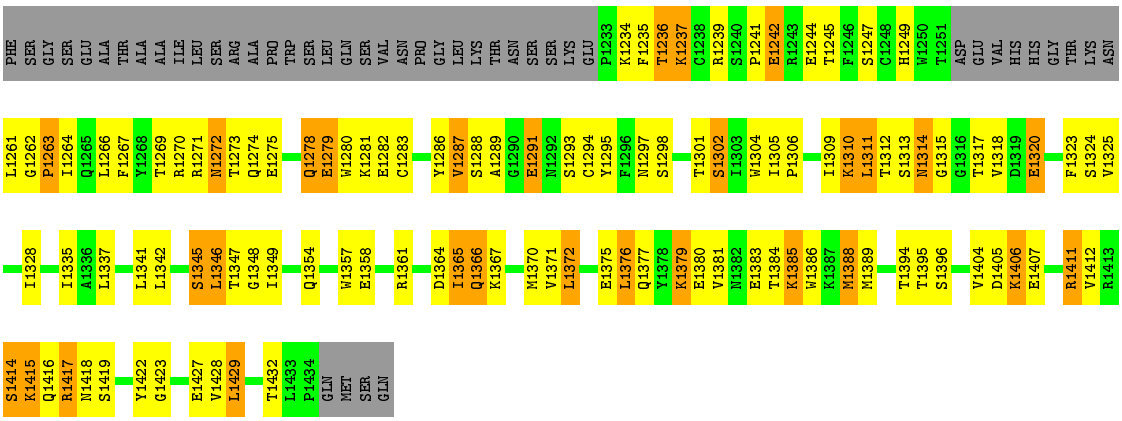
Chain C:



• Molecule 2: EXTRACELLULAR DOMAIN HUMAN GROWTH HORMONE RECEPTOR (1-238)

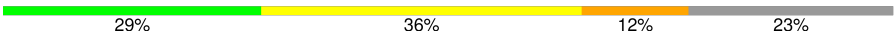
Chain E:

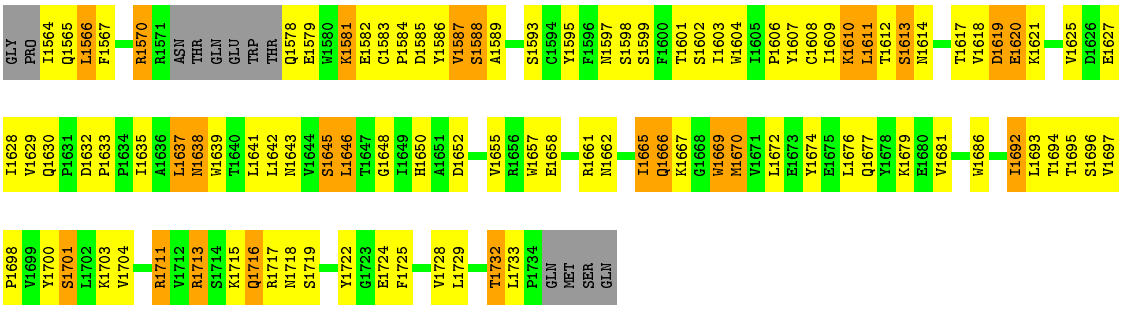




• Molecule 2: EXTRACELLULAR DOMAIN HUMAN GROWTH HORMONE RECEPTOR (1-238)

Chain F:




GLY
PRO
I1564
Q1565
L1566
F1567
R1570
R1571
ASN
THR
GLN
GLU
TRP
THR
Q1578
E1579
W1580
K1581
E1582
C1583
P1584
D1585
Y1586
V1587
S1588
A1589
S1593
C1594
Y1595
F1596
N1597
S1598
S1599
F1600
T1601
S1602
I1603
W1604
I1605
P1606
Y1607
C1608
I1609
K1610
L1611
T1612
S1613
N1614
T1617
V1618
D1619
E1620
K1621
V1625
D1626
E1627
I1628
V1629
Q1630
P1631
D1632
P1633
P1634
I1635
A1636
L1637
N1638
W1639
T1640
L1641
L1642
N1643
V1644
S1645
L1646
T1647
G1648
I1649
H1650
A1651
D1652
V1655
R1656
W1657
E1658
R1661
N1662
I1665
Q1666
K1667
G1668
W1669
M1670
V1671
L1672
E1673
Y1674
E1675
L1676
Q1677
Y1678
K1679
E1680
V1681
W1686
I1692
L1693
T1694
T1695
S1696
V1697
P1698
V1699
Y1700
S1701
L1702
K1703
V1704
R1711
V1712
R1713
S1714
K1715
Q1716
R1717
N1718
S1719
Y1722
G1723
E1724
F1725
V1728
L1729
T1732
L1733
P1734
GLN
MET
SER
GLN

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	67.29Å 111.94Å 95.29Å 90.00° 90.06° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60	Depositor
% Data completeness (in resolution range)	90.5 (20.00-2.60)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.843	Depositor
R, R_{free}	0.234 , 0.326	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8678	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

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	0.67	0/1312	0.84	2/1781 (0.1%)
1	D	0.71	0/1289	0.84	2/1747 (0.1%)
2	B	0.70	0/1622	0.87	0/2208
2	C	0.67	0/1459	0.86	1/1987 (0.1%)
2	E	0.69	0/1621	0.86	1/2208 (0.0%)
2	F	0.65	0/1525	0.84	0/2078
All	All	0.68	0/8828	0.85	6/12009 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	3	THR	N-CA-C	-9.03	86.61	111.00
1	D	1107	ASP	CB-CG-OD1	6.76	124.39	118.30
2	C	668	GLY	N-CA-C	-6.00	98.09	113.10
2	E	1324	SER	N-CA-C	-5.23	96.87	111.00
1	A	1	PHE	C-N-CD	5.20	139.32	128.40

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	1284	0	1257	102	0
1	D	1264	0	1241	88	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1577	0	1506	91	0
2	C	1419	0	1338	99	0
2	E	1576	0	1508	104	0
2	F	1484	0	1395	115	0
3	A	11	0	0	2	0
3	B	13	0	0	1	0
3	C	28	0	0	2	0
3	D	3	0	0	0	0
3	E	12	0	0	1	0
3	F	7	0	0	0	0
All	All	8678	0	8245	550	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 33.

The worst 5 of 550 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:104:GLY:CA	2:C:666:GLN:HB3	1.65	1.26
2:E:1346:LEU:HB2	2:F:1701:SER:HB3	1.24	1.17
1:A:104:GLY:HA2	2:C:666:GLN:HB3	1.25	1.14
2:F:1581:LYS:HZ3	2:F:1581:LYS:HB2	1.08	1.09
1:A:104:GLY:HA3	2:C:666:GLN:HB3	1.41	1.02

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 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	154/191 (81%)	129 (84%)	19 (12%)	6 (4%)	4 5
1	D	150/191 (78%)	129 (86%)	17 (11%)	4 (3%)	6 10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	189/238 (79%)	168 (89%)	20 (11%)	1 (0%)	34	60
2	C	167/238 (70%)	144 (86%)	21 (13%)	2 (1%)	16	33
2	E	189/238 (79%)	171 (90%)	16 (8%)	2 (1%)	17	36
2	F	178/238 (75%)	155 (87%)	20 (11%)	3 (2%)	11	22
All	All	1027/1334 (77%)	896 (87%)	113 (11%)	18 (2%)	11	21

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	106	SER
1	A	107	ASP
2	B	274	GLN
1	D	1002	PRO
2	E	1274	GLN

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 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	143/175 (82%)	112 (78%)	31 (22%)	1	2
1	D	142/175 (81%)	112 (79%)	30 (21%)	1	2
2	B	177/218 (81%)	130 (73%)	47 (27%)	0	1
2	C	157/218 (72%)	113 (72%)	44 (28%)	0	1
2	E	177/218 (81%)	134 (76%)	43 (24%)	1	1
2	F	165/218 (76%)	121 (73%)	44 (27%)	0	1
All	All	961/1222 (79%)	722 (75%)	239 (25%)	1	1

5 of 239 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	672	LEU
1	D	1055	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	1658	GLU
2	C	692	ILE
2	C	729	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	716	GLN
1	D	1122	GLN
2	F	1643	ASN
1	D	1068	GLN
1	D	1099	ASN

5.3.3 RNA (i)

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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.