wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 02:39 am BST

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)

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

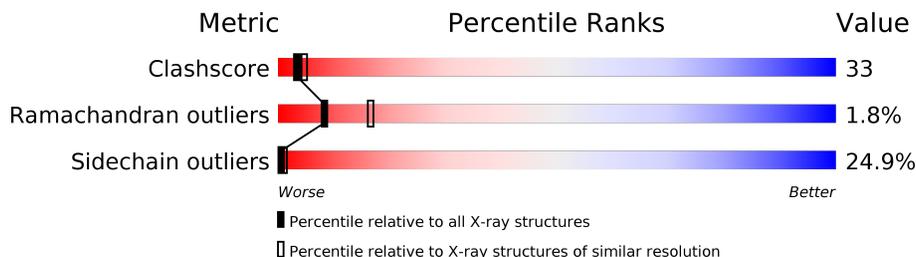|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	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 2.60 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	Percentile Ranks	Value
Clashscore		33
Ramachandran outliers		1.8%
Sidechain outliers		24.9%

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))
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (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 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%

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
1	A	160	Total 1284	C 825	N 211	O 242	S 6	0	0	0
1	D	158	Total 1264	C 811	N 208	O 239	S 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
2	B	193	Total 1577	C 1011	N 258	O 298	S 10	0	0	0
2	C	175	Total 1419	C 916	N 229	O 265	S 9	0	0	0
2	E	193	Total 1576	C 1011	N 258	O 298	S 9	0	0	0
2	F	184	Total 1484	C 952	N 239	O 284	S 9	0	0	0

- Molecule 3 is water.

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

Continued on next page...

Continued from previous page...

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

Note EDS was not executed.

• Molecule 1: PHAGE DISPLAY DERIVED VARIANT HUMAN GROWTH HORMONE

• Molecule 1: PHAGE DISPLAY DERIVED VARIANT HUMAN GROWTH HORMONE

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

S345 L346 T347 G348 I349 I353 W357 E358 R361 N362 A363 D364 I365 Q366 K367 M370 V371 L372 E375 L376 K379 E380 V381 N382 E383 T384 K385 W386 K387 M388 M389 I392 L393 T394 T395 V404 D405 K406 E407 R411 V412 R413 S414 K415 Q416 R417 N418 S419 G423 E424

E427 V428 L429 T432 M436 SER GLN

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

Chain C: 30% 32% 12% 26%

PHE SER GLY SER GLU ALA THR ALA ALA ILE LEU SER ARG ALA PRO TRP SER LEU GLN SER VAL ASN PRO GLY LEU LYS THR ASN SER SER LYS GLU P533 K534 F535 T536 K537 C538 R539 S540 P541 E542 R543 E544 T545 F546 S547 C548 H549 W550 T551 ASP GLU VAL HIS HIS GLY THR LYS ASN

LEU GLY PRO I564 Q565 L566 F567 Y568 T569 R570 R571 N572 THR GLN GLU TRP THR GLN E579 W580 K581 E582 C583 P584 D585 Y586 V587 S588 A589 S593 C594 Y595 F596 N597 S598 S599 F600 T601 S602 I603 W604 I605 I609 K610 L611 THR SER ASN GLY THR VAL ASP GLU K621 G622 F623 S624

E627 I628 V629 Q630 P631 D632 P633 L637 N638 W639 T640 L641 L642 N643 V644 S645 L646 T647 G648 I649 H650 A651 D652 V655 R656 W657 E658 A659 P660 A663 D664 I665 Q666 K667 G668 W669 M670 V671 L672 E673 Y674 E675 L676 Q677 V681 W686 I692 L693 T694 T695 S696 Y700

S701 V704 R711 V712 R713 S714 K715 Q716 R717 N718 S719 Y722 G723 E724 F725 S726 E727 V728 L729 T732 L733 P734 GLN MET SER GLN

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

Chain E: 34% 35% 12% 19%

PHE SER GLY SER GLU ALA THR ALA ALA ILE LEU SER ARG ALA PRO TRP SER LEU GLN SER VAL ASN PRO GLY LEU LYS THR ASN SER SER LYS GLU P1233 K1234 F1235 T1236 K1237 C1238 R1239 S1240 P1241 E1242 R1243 E1244 T1245 F1246 S1247 C1248 H1249 W1250 T1251 ASP GLU VAL HIS HIS GLY THR LYS ASN

L1261 G1262 P1263 I1264 Q1265 L1266 F1267 Y1268 T1269 R1270 R1271 N1272 T1273 Q1274 E1275 Q1278 E1279 W1280 K1281 E1282 C1283 Y1286 V1287 S1288 A1289 G1290 E1291 N1292 S1293 C1294 Y1295 F1296 N1297 S1298 T1301 I1302 I1303 W1304 I1305 P1306 I1309 K1310 L1311 T1312 S1313 N1314 G1315 G1316 T1317 V1318 D1319 E1320 F1323 S1324 V1325

I1328 I1335 A1336 L1337 L1341 L1342 S1345 L1346 T1347 G1348 I1349 Q1354 W1357 E1358 R1361 D1364 I1365 Q1366 K1367 M1370 V1371 L1372 E1375 L1376 Q1377 Y1378 K1379 E1380 V1381 N1382 E1383 T1384 K1385 W1386 K1387 M1388 M1389 T1394 T1395 S1396 V1404 D1405 K1406 E1407 R1411 V1412 R1413

S1414 K1415 Q1416 R1417 N1418 S1419 Y1422 G1423 E1427 V1428 L1429 T1432 L1433 P1434 GLN MET SER GLN

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

Chain F: 29% 36% 12% 23%

PHE SER GLY SER GLU ALA THR ALA ALA ILE LEU SER ARG ALA PRO TRP SER LEU GLN SER VAL ASN PRO GLY LEU LYS THR ASN SER SER LYS GLU P1533 K1534 F1535 T1536 K1537 C1538 R1539 S1540 P1541 E1542 R1543 E1544 S1547 C1548 H1549 W1550 T1551 ASP GLU VAL HIS HIS GLY THR LYS ASN LEU

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore 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%)	3	4
1	D	150/191 (78%)	129 (86%)	17 (11%)	4 (3%)	5	8

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%)	29	52
2	C	167/238 (70%)	144 (86%)	21 (13%)	2 (1%)	13	27
2	E	189/238 (79%)	171 (90%)	16 (8%)	2 (1%)	14	30
2	F	178/238 (75%)	155 (87%)	20 (11%)	3 (2%)	9	18
All	All	1027/1334 (77%)	896 (87%)	113 (11%)	18 (2%)	8	16

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	1
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%)	0	1
2	F	165/218 (76%)	121 (73%)	44 (27%)	0	1
All	All	961/1222 (79%)	722 (75%)	239 (25%)	0	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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.