



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 26, 2023 – 05:37 PM EDT

PDB ID : 3GCW  
Title : PCSK9:EGFA(H306Y)  
Authors : Kwon, H.J.  
Deposited on : 2009-02-22  
Resolution : 2.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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) : 2.35

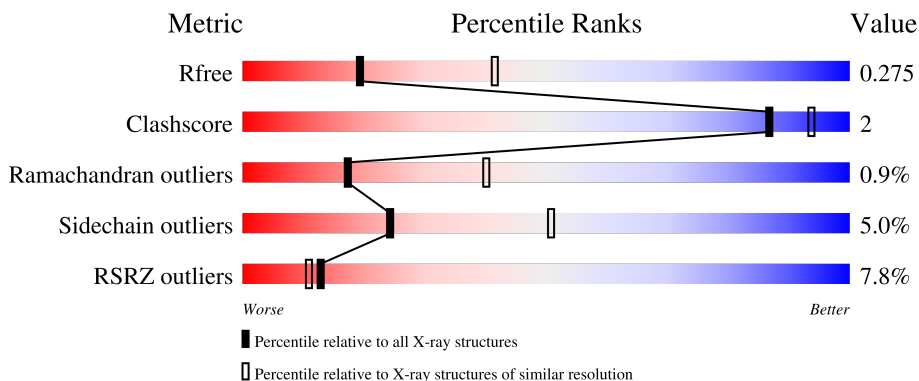
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	P	100	<div style="display: flex; align-items: center;"> <div style="width: 7%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 84%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 8%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 8%; height: 10px; background-color: grey;"></div> </div>
2	A	540	<div style="display: flex; align-items: center;"> <div style="width: 6%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 72%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 8%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 20%; height: 10px; background-color: grey;"></div> </div>
3	E	83	<div style="display: flex; align-items: center;"> <div style="width: 8%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 46%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 44%; height: 10px; background-color: grey;"></div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4279 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 Proprotein convertase subtilisin/kexin type 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	P	92	740	474	133	131	2	0	0	0

- Molecule 2 is a protein called Proprotein convertase subtilisin/kexin type 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	434	3228	2002	588	613	25	0	0	0

- Molecule 3 is a protein called Low-density lipoprotein receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	E	41	310	186	54	63	7	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	290	GLY	-	expression tag	UNP P01130
E	291	ALA	-	expression tag	UNP P01130
E	292	MET	-	expression tag	UNP P01130
E	306	TYR	HIS	engineered mutation	UNP P01130

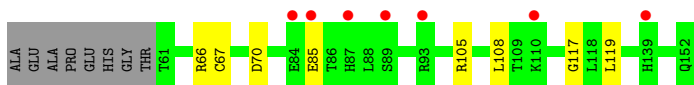
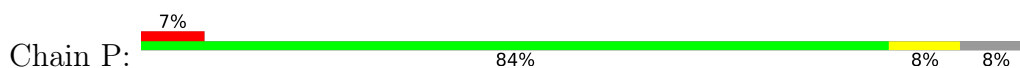
- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	E	1	Total	Ca	0	0
			1	1		

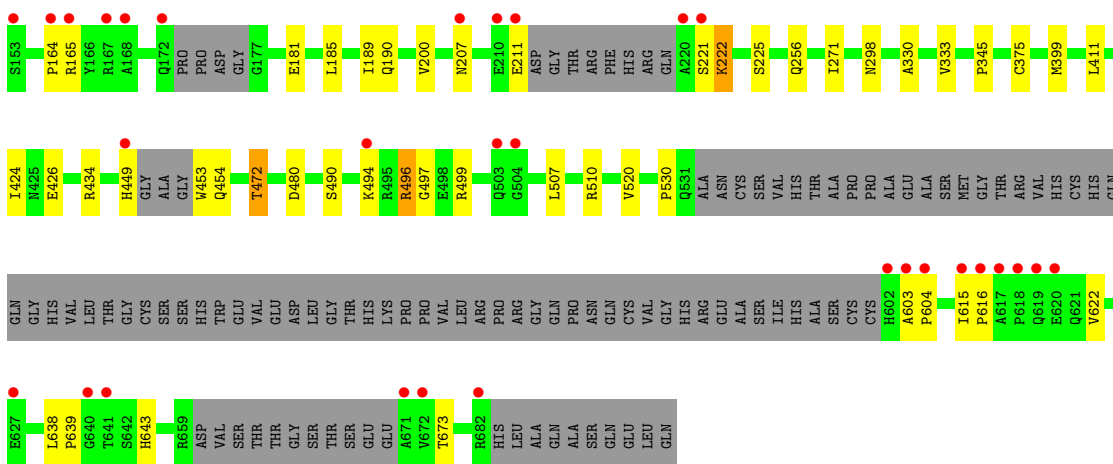
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 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.

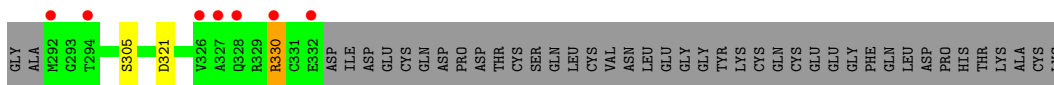
- Molecule 1: Proprotein convertase subtilisin/kexin type 9



- Molecule 2: Proprotein convertase subtilisin/kexin type 9



- Molecule 3: Low-density lipoprotein receptor



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.66Å 115.66Å 133.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.10 – 2.70 37.04 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.3 (39.10-2.70) 99.3 (37.04-2.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.03 (at 2.68Å)	Xtrriage
Refinement program	REFMAC 5.3.0037	Depositor
R, $R_{free}$	0.234 , 0.278 0.233 , 0.275	Depositor DCC
$R_{free}$ test set	1294 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	58.2	Xtrriage
Anisotropy	0.308	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 48.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	4279	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.48% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

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

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	P	0.34	0/757	0.49	0/1023
2	A	0.34	0/3284	0.52	0/4458
3	E	0.34	0/313	0.49	0/420
All	All	0.34	0/4354	0.51	0/5901

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	P	740	0	750	1	0
2	A	3228	0	3181	16	0
3	E	310	0	280	1	0
4	E	1	0	0	0	0
All	All	4279	0	4211	18	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 2.

All (18) 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)
2:A:449:HIS:HA	2:A:453:TRP:CH2	2.22	0.75
2:A:211:GLU:HB2	2:A:256:GLN:HE21	1.57	0.70
2:A:330:ALA:O	2:A:333:VAL:HG22	1.99	0.61
2:A:185:LEU:HD11	2:A:271:ILE:HD11	1.86	0.57
2:A:490:SER:HB2	2:A:520:VAL:HG12	1.88	0.55
3:E:305:SER:OG	3:E:330:ARG:HA	2.09	0.53
1:P:66:ARG:NH1	1:P:67:CYS:O	2.42	0.49
2:A:426:GLU:HB3	2:A:434:ARG:HG2	1.98	0.46
2:A:496:ARG:HE	2:A:497:GLY:H	1.64	0.44
2:A:615:ILE:HD12	2:A:622:VAL:HG22	2.00	0.43
2:A:345:PRO:HD3	2:A:424:ILE:HG23	2.01	0.43
2:A:603:ALA:HA	2:A:604:PRO:HD3	1.84	0.43
2:A:638:LEU:HB2	2:A:673:THR:HB	2.01	0.42
2:A:472:THR:HG21	2:A:510:ARG:NH2	2.35	0.42
2:A:222:LYS:HE2	2:A:222:LYS:HA	2.02	0.41
2:A:189:ILE:HD13	2:A:200:VAL:HG11	2.03	0.41
2:A:399:MET:HE3	2:A:411:LEU:HG	2.03	0.41
2:A:453:TRP:HD1	2:A:454:GLN:H	1.69	0.40

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	P	90/100 (90%)	88 (98%)	1 (1%)	1 (1%)	14	34
2	A	422/540 (78%)	401 (95%)	17 (4%)	4 (1%)	17	40
3	E	39/83 (47%)	37 (95%)	2 (5%)	0	100	100
All	All	551/723 (76%)	526 (96%)	20 (4%)	5 (1%)	17	40

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	616	PRO
2	A	639	PRO
2	A	164	PRO
1	P	117	GLY
2	A	530	PRO

### 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	P	79/84 (94%)	74 (94%)	5 (6%)	18	40
2	A	345/431 (80%)	329 (95%)	16 (5%)	27	54
3	E	35/71 (49%)	33 (94%)	2 (6%)	20	44
All	All	459/586 (78%)	436 (95%)	23 (5%)	24	51

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	P	70	ASP
1	P	85	GLU
1	P	105	ARG
1	P	108	LEU
1	P	119	LEU
2	A	165	ARG
2	A	181	GLU
2	A	190	GLN
2	A	207	ASN
2	A	221	SER
2	A	222	LYS
2	A	225	SER
2	A	298	ASN
2	A	375	CYS
2	A	472	THR
2	A	480	ASP
2	A	494	LYS
2	A	496	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	A	499	ARG
2	A	507	LEU
2	A	643	HIS
3	E	321	ASP
3	E	330	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	P	101	GLN
2	A	207	ASN
2	A	256	GLN
2	A	298	ASN
2	A	342	GLN
2	A	619	GLN
3	E	300	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

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.

No monomer is involved in short contacts.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	P	92/100 (92%)	0.15	7 (7%) 13 12	45, 54, 64, 67	0
2	A	434/540 (80%)	0.47	30 (6%) 16 15	47, 54, 68, 88	0
3	E	41/83 (49%)	0.78	7 (17%) 1 1	51, 55, 62, 66	0
All	All	567/723 (78%)	0.44	44 (7%) 13 11	45, 55, 66, 88	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	641	THR	6.7
3	E	292	MET	6.4
2	A	220	ALA	5.7
2	A	603	ALA	4.9
2	A	153	SER	4.7
2	A	504	GLY	4.6
2	A	449	HIS	4.1
2	A	494	LYS	4.1
2	A	640	GLY	4.0
2	A	616	PRO	4.0
3	E	328	GLN	4.0
2	A	602	HIS	3.8
2	A	617	ALA	3.7
3	E	330	ARG	3.6
2	A	172	GLN	3.6
2	A	164	PRO	3.4
3	E	327	ALA	3.4
3	E	326	VAL	3.4
2	A	682	ARG	3.1
1	P	85	GLU	3.1
2	A	620	GLU	3.0
3	E	332	GLU	2.9
2	A	672	VAL	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	A	211	GLU	2.8
2	A	619	GLN	2.7
2	A	165	ARG	2.6
1	P	87	HIS	2.6
2	A	615	ILE	2.5
2	A	167	ARG	2.4
2	A	618	PRO	2.4
2	A	210	GLU	2.3
2	A	671	ALA	2.3
1	P	84	GLU	2.2
2	A	503	GLN	2.2
1	P	139	HIS	2.2
3	E	294	THR	2.2
2	A	604	PRO	2.1
2	A	207	ASN	2.1
2	A	627	GLU	2.1
2	A	221	SER	2.1
1	P	110	LYS	2.0
1	P	89	SER	2.0
1	P	93	ARG	2.0
2	A	168	ALA	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	CA	E	1	1/1	0.94	0.07	48,48,48,48	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.