



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:11 PM GMT

PDB ID : 4O3U  
Title : Zymogen HGF-beta/MET with Zymogen Activator Peptide ZAP2.3  
Authors : Eigenbrot, C.; Landgraf, K.E.; Steffek, M.  
Deposited on : 2013-12-18  
Resolution : 3.04 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

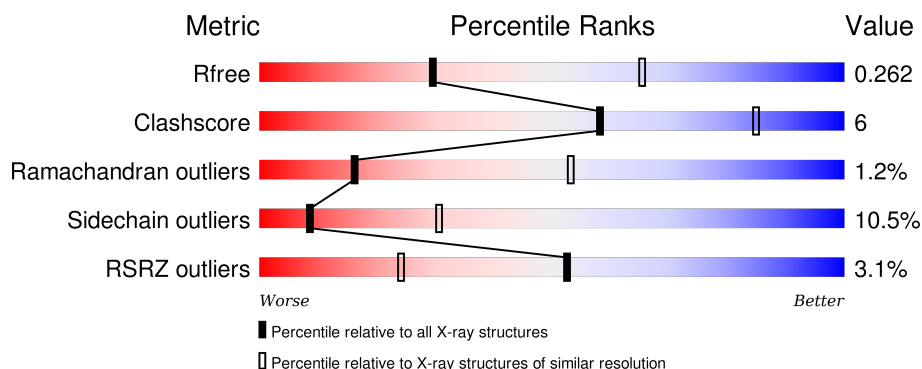
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.04 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))
$R_{free}$	91344	1995 (3.08-3.00)
Clashscore	102246	2351 (3.08-3.00)
Ramachandran outliers	100387	2272 (3.08-3.00)
Sidechain outliers	100360	2275 (3.08-3.00)
RSRZ outliers	91569	2013 (3.08-3.00)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	240	 2% 73% 16% • 9%
2	B	551	 3% 69% 19% • 9%
3	P	15	 73% 20% 7%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5786 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 Hepatocyte growth factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	219	Total	C	N	O	S	0	0	0
			1703	1083	301	305	14			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	495	GLY	VAL	ENGINEERED MUTATION	UNP P14210
A	604	SER	CYS	ENGINEERED MUTATION	UNP P14210
A	729	HIS	-	EXPRESSION TAG	UNP P14210
A	730	HIS	-	EXPRESSION TAG	UNP P14210
A	731	HIS	-	EXPRESSION TAG	UNP P14210
A	732	HIS	-	EXPRESSION TAG	UNP P14210
A	733	HIS	-	EXPRESSION TAG	UNP P14210
A	734	HIS	-	EXPRESSION TAG	UNP P14210

- Molecule 2 is a protein called Hepatocyte growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	500	Total	C	N	O	S	0	0	0
			3969	2522	675	743	29			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	568	HIS	-	EXPRESSION TAG	UNP P08581
B	569	HIS	-	EXPRESSION TAG	UNP P08581
B	570	HIS	-	EXPRESSION TAG	UNP P08581
B	571	HIS	-	EXPRESSION TAG	UNP P08581
B	572	HIS	-	EXPRESSION TAG	UNP P08581
B	573	HIS	-	EXPRESSION TAG	UNP P08581
B	574	HIS	-	EXPRESSION TAG	UNP P08581
B	575	HIS	-	EXPRESSION TAG	UNP P08581

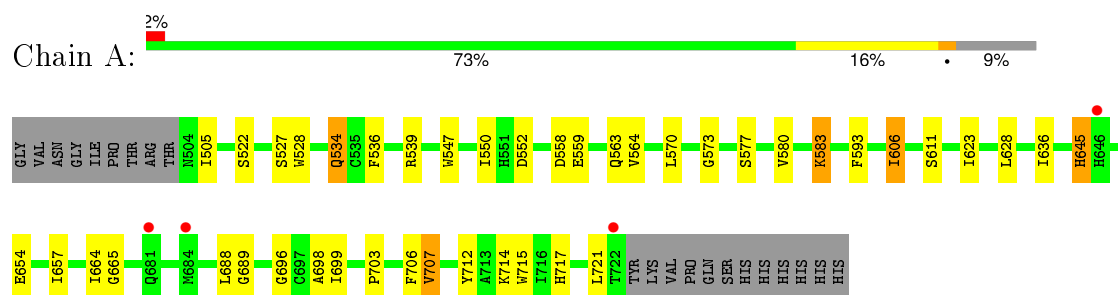
- Molecule 3 is a protein called ZAP 2.3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	14	Total	C	N	O	S	0	0	0
			114	72	18	21	3			

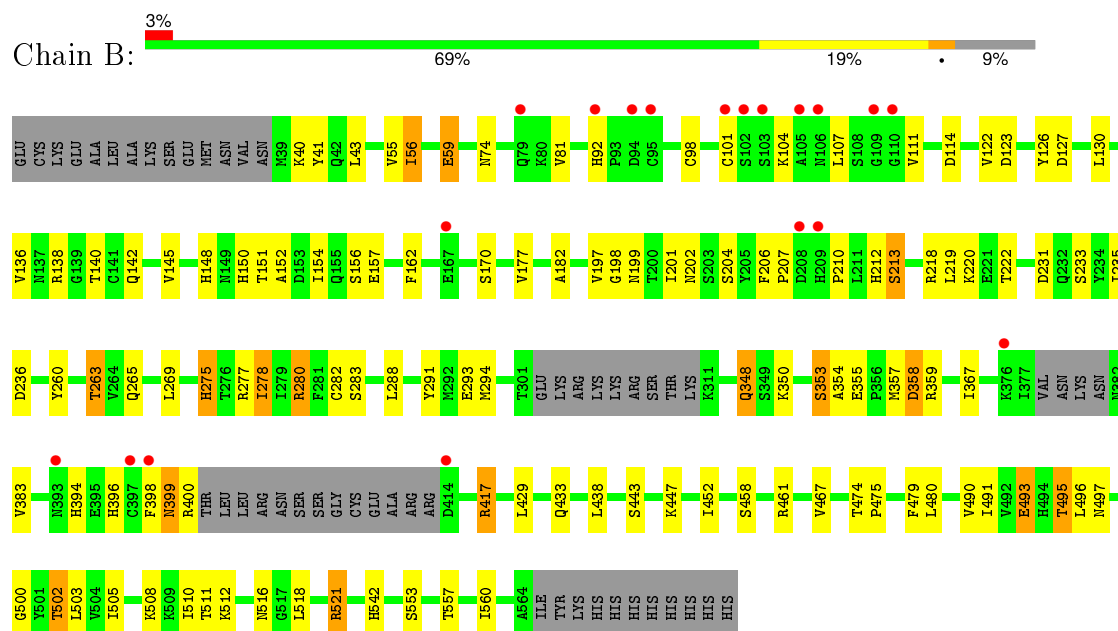
### 3 Residue-property plots

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 ( $\text{RSRZ} > 2$ ). 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.

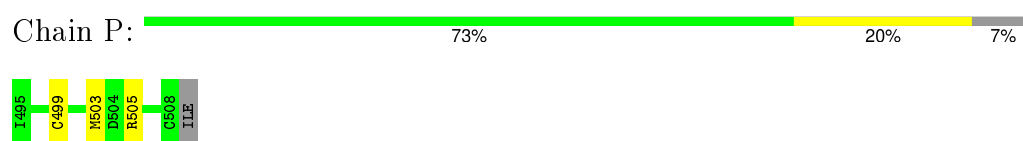
- Molecule 1: Hepatocyte growth factor



- Molecule 2: Hepatocyte growth factor receptor



- Molecule 3: ZAP 2.3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.10Å 139.62Å 66.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.76 – 3.04 33.81 – 3.04	Depositor EDS
% Data completeness (in resolution range)	97.5 (34.76-3.04) 97.5 (33.81-3.04)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.56 (at 3.06Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, $R_{free}$	0.207 , 0.251 0.221 , 0.262	Depositor DCC
$R_{free}$ test set	1007 reflections (4.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	88.2	Xtriage
Anisotropy	0.653	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 67.7	EDS
Estimated twinning fraction	0.016 for k,h,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 24340 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5786	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.52% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.47	0/1744	0.75	0/2362
2	B	0.48	0/4069	0.76	0/5524
3	P	0.50	0/117	0.82	0/157
All	All	0.48	0/5930	0.76	0/8043

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	1703	0	1688	18	0
2	B	3969	0	3816	49	0
3	P	114	0	99	1	0
All	All	5786	0	5603	67	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 6.

The worst 5 of 67 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:534:GLN:HE21	1:A:534:GLN:H	1.39	0.70

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:461:ARG:HG2	2:B:480:LEU:HD23	1.78	0.65
1:A:606:ILE:HD13	1:A:636:ILE:HG13	1.79	0.64
1:A:606:ILE:HD11	1:A:657:ILE:HD11	1.79	0.63
1:A:577:SER:HB3	1:A:712:TYR:CZ	2.34	0.63

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	217/240 (90%)	197 (91%)	19 (9%)	1 (0%)	34	74
2	B	492/551 (89%)	452 (92%)	32 (6%)	8 (2%)	12	45
3	P	12/15 (80%)	11 (92%)	1 (8%)	0	100	100
All	All	721/806 (90%)	660 (92%)	52 (7%)	9 (1%)	16	53

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	204	SER
2	B	148	HIS
2	B	283	SER
2	B	553	SER
1	A	664	ILE

### 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	183/203 (90%)	165 (90%)	18 (10%)	10	34
2	B	451/499 (90%)	403 (89%)	48 (11%)	8	30
3	P	12/13 (92%)	10 (83%)	2 (17%)	3	12
All	All	646/715 (90%)	578 (90%)	68 (10%)	8	31

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	197	VAL
2	B	278	ILE
2	B	516	ASN
2	B	213	SER
2	B	231	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	318	GLN
2	B	348	GLN
2	B	433	GLN
2	B	275	HIS
2	B	396	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 [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](#)

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	A	219/240 (91%)	-0.08	4 (1%) 71 43	54, 81, 126, 155	0
2	B	500/551 (90%)	0.08	19 (3%) 44 19	57, 94, 143, 173	7 (1%)
3	P	14/15 (93%)	-0.28	0 100 100	63, 79, 91, 100	0
All	All	733/806 (90%)	0.03	23 (3%) 52 24	54, 89, 138, 173	7 (0%)

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	110	GLY	7.4
2	B	109	GLY	6.2
2	B	208	ASP	5.0
2	B	105	ALA	4.9
2	B	397	CYS	4.2

### 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 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.