



# wwPDB X-ray Structure Validation Summary Report

Sep 16, 2014 – 07:04 PM EDT

PDB ID : 4MQW  
Title : Structure of follicle-stimulating hormone in complex with the entire ectodomain of its receptor (P31)  
Authors : Jiang, X.; Liu, H.; Chen, X.; He, X.  
Deposited on : 2013-09-16  
Resolution : 2.90 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

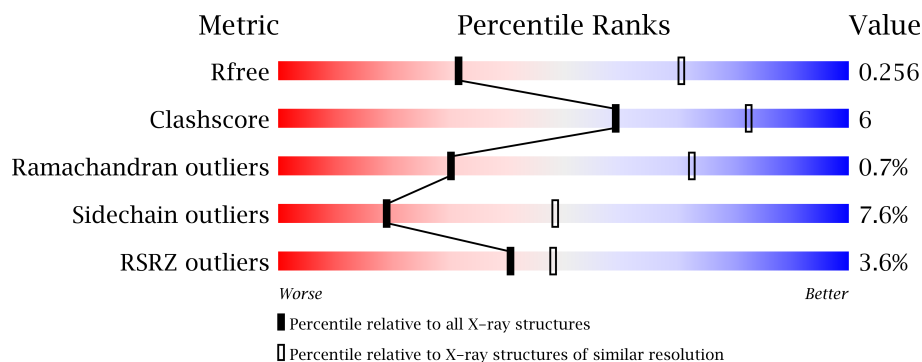
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23489  
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) : stable23489

# 1 Overall quality at a glance

The reported resolution of this entry is 2.90 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	1053 (2.90-2.90)
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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. 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	102	
1	D	102	
1	G	102	
2	B	111	
2	E	111	
2	H	111	
3	X	361	
3	Y	361	
3	Z	361	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	NAG	B	201	-	X
4	NAG	B	202	-	X
4	NAG	E	201	-	X
4	NAG	E	202	-	X
4	NAG	H	201	-	X
4	NAG	H	202	-	X
5	EDO	X	401	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 11798 atoms, of which 0 are hydrogen 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 Glycoprotein hormones, alpha polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	88	Total	C	N	O	S	0	0	0
			675	417	118	127	13			
1	D	89	Total	C	N	O	S	0	0	0
			682	422	119	128	13			
1	G	88	Total	C	N	O	S	0	0	0
			675	417	118	127	13			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	93	ALA	-	EXPRESSION TAG	UNP Q96QJ4
A	94	ALA	-	EXPRESSION TAG	UNP Q96QJ4
A	95	ALA	-	EXPRESSION TAG	UNP Q96QJ4
A	96	HIS	-	EXPRESSION TAG	UNP Q96QJ4
A	97	HIS	-	EXPRESSION TAG	UNP Q96QJ4
A	98	HIS	-	EXPRESSION TAG	UNP Q96QJ4
A	99	HIS	-	EXPRESSION TAG	UNP Q96QJ4
A	100	HIS	-	EXPRESSION TAG	UNP Q96QJ4
A	101	HIS	-	EXPRESSION TAG	UNP Q96QJ4
A	102	HIS	-	EXPRESSION TAG	UNP Q96QJ4
D	93	ALA	-	EXPRESSION TAG	UNP Q96QJ4
D	94	ALA	-	EXPRESSION TAG	UNP Q96QJ4
D	95	ALA	-	EXPRESSION TAG	UNP Q96QJ4
D	96	HIS	-	EXPRESSION TAG	UNP Q96QJ4
D	97	HIS	-	EXPRESSION TAG	UNP Q96QJ4
D	98	HIS	-	EXPRESSION TAG	UNP Q96QJ4
D	99	HIS	-	EXPRESSION TAG	UNP Q96QJ4
D	100	HIS	-	EXPRESSION TAG	UNP Q96QJ4
D	101	HIS	-	EXPRESSION TAG	UNP Q96QJ4
D	102	HIS	-	EXPRESSION TAG	UNP Q96QJ4
G	93	ALA	-	EXPRESSION TAG	UNP Q96QJ4
G	94	ALA	-	EXPRESSION TAG	UNP Q96QJ4
G	95	ALA	-	EXPRESSION TAG	UNP Q96QJ4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	96	HIS	-	EXPRESSION TAG	UNP Q96QJ4
G	97	HIS	-	EXPRESSION TAG	UNP Q96QJ4
G	98	HIS	-	EXPRESSION TAG	UNP Q96QJ4
G	99	HIS	-	EXPRESSION TAG	UNP Q96QJ4
G	100	HIS	-	EXPRESSION TAG	UNP Q96QJ4
G	101	HIS	-	EXPRESSION TAG	UNP Q96QJ4
G	102	HIS	-	EXPRESSION TAG	UNP Q96QJ4

- Molecule 2 is a protein called Follitropin subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	109	Total	C	N	O	S	0	0	0
			848	527	142	166	13			
2	E	108	Total	C	N	O	S	0	0	0
			840	522	141	165	12			
2	H	108	Total	C	N	O	S	0	0	0
			840	522	141	165	12			

- Molecule 3 is a protein called Follicle-stimulating hormone receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	X	289	Total	C	N	O	S	0	0	0
			2275	1439	397	427	12			
3	Y	284	Total	C	N	O	S	0	0	0
			2258	1428	392	428	10			
3	Z	290	Total	C	N	O	S	0	0	0
			2312	1463	406	431	12			

There are 33 discrepancies between the modelled and reference sequences:

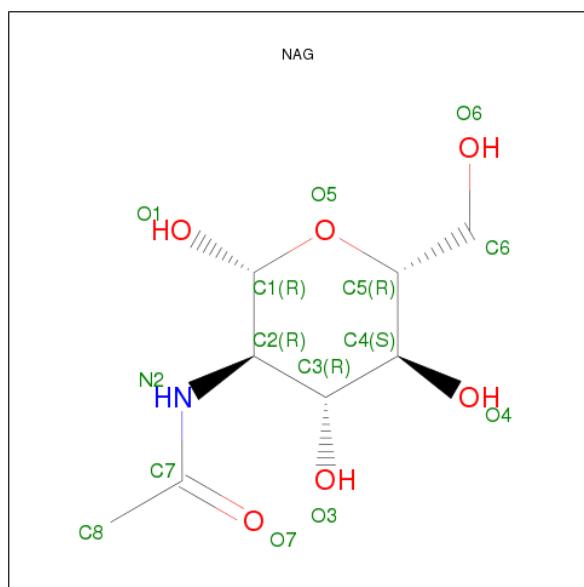
Chain	Residue	Modelled	Actual	Comment	Reference
X	188	SER	CYS	ENGINEERED MUTATION	UNP P23945
X	367	ALA	-	EXPRESSION TAG	UNP P23945
X	368	ALA	-	EXPRESSION TAG	UNP P23945
X	369	ALA	-	EXPRESSION TAG	UNP P23945
X	370	HIS	-	EXPRESSION TAG	UNP P23945
X	371	HIS	-	EXPRESSION TAG	UNP P23945
X	372	HIS	-	EXPRESSION TAG	UNP P23945
X	373	HIS	-	EXPRESSION TAG	UNP P23945
X	374	HIS	-	EXPRESSION TAG	UNP P23945
X	375	HIS	-	EXPRESSION TAG	UNP P23945
X	376	HIS	-	EXPRESSION TAG	UNP P23945

*Continued on next page...*

Continued from previous page...

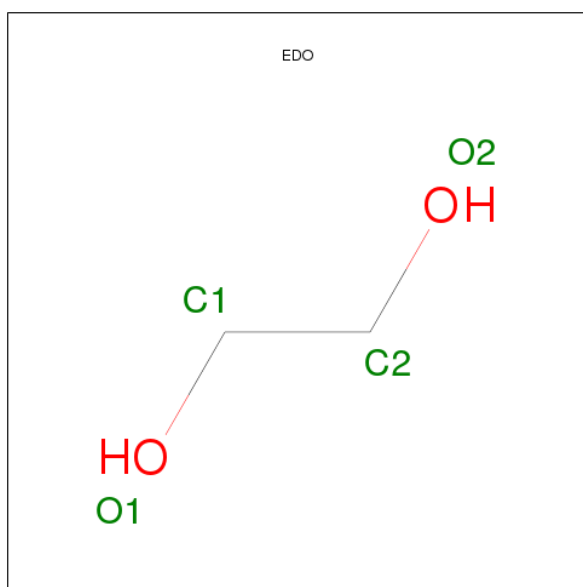
Chain	Residue	Modelled	Actual	Comment	Reference
Y	188	SER	CYS	ENGINEERED MUTATION	UNP P23945
Y	367	ALA	-	EXPRESSION TAG	UNP P23945
Y	368	ALA	-	EXPRESSION TAG	UNP P23945
Y	369	ALA	-	EXPRESSION TAG	UNP P23945
Y	370	HIS	-	EXPRESSION TAG	UNP P23945
Y	371	HIS	-	EXPRESSION TAG	UNP P23945
Y	372	HIS	-	EXPRESSION TAG	UNP P23945
Y	373	HIS	-	EXPRESSION TAG	UNP P23945
Y	374	HIS	-	EXPRESSION TAG	UNP P23945
Y	375	HIS	-	EXPRESSION TAG	UNP P23945
Y	376	HIS	-	EXPRESSION TAG	UNP P23945
Z	188	SER	CYS	ENGINEERED MUTATION	UNP P23945
Z	367	ALA	-	EXPRESSION TAG	UNP P23945
Z	368	ALA	-	EXPRESSION TAG	UNP P23945
Z	369	ALA	-	EXPRESSION TAG	UNP P23945
Z	370	HIS	-	EXPRESSION TAG	UNP P23945
Z	371	HIS	-	EXPRESSION TAG	UNP P23945
Z	372	HIS	-	EXPRESSION TAG	UNP P23945
Z	373	HIS	-	EXPRESSION TAG	UNP P23945
Z	374	HIS	-	EXPRESSION TAG	UNP P23945
Z	375	HIS	-	EXPRESSION TAG	UNP P23945
Z	376	HIS	-	EXPRESSION TAG	UNP P23945

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	H	1	Total	C	N	O	0	0
			14	8	1	5		
4	H	1	Total	C	N	O	0	0
			14	8	1	5		
4	X	1	Total	C	N	O	0	0
			14	8	1	5		
4	Y	1	Total	C	N	O	0	0
			14	8	1	5		
4	Z	1	Total	C	N	O	0	0
			14	8	1	5		

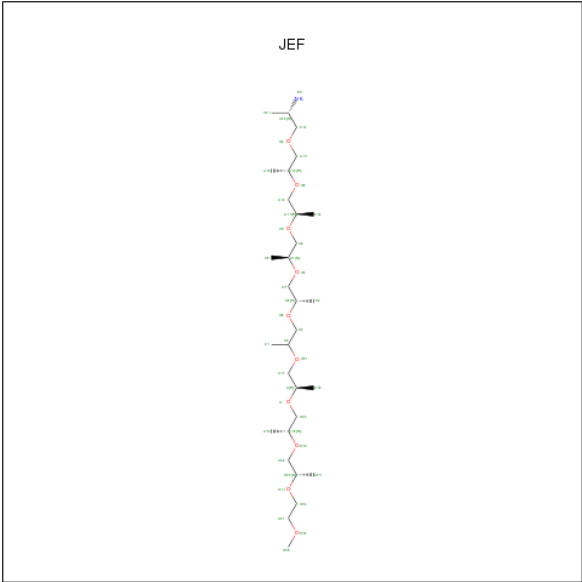
- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	X	1	Total	C	O	0	0
			4	2	2		
5	X	1	Total	C	O	0	0
			4	2	2		
5	Y	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is O-(O-(2-AMINOPROPYL)-O'-(2-METHOXYETHYL)POLYPROPYLEN EGLYCOL 500) (three-letter code: JEF) (formula: C<sub>30</sub>H<sub>63</sub>NO<sub>10</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	Z	1	Total	C	O	0	0
			8	5	3		

- Molecule 7 is water.

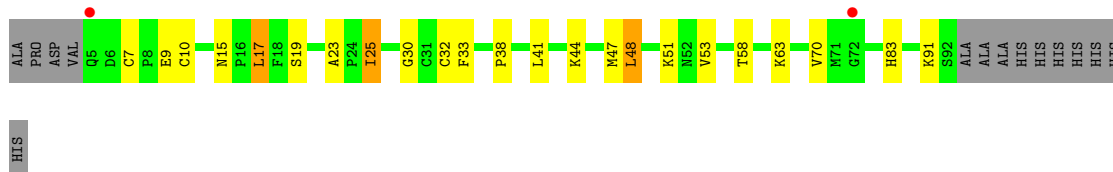
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	13	Total	O	0	0
			13	13		
7	B	16	Total	O	0	0
			16	16		
7	X	22	Total	O	0	0
			22	22		
7	D	10	Total	O	0	0
			10	10		
7	E	13	Total	O	0	0
			13	13		
7	Y	28	Total	O	0	0
			28	28		
7	G	15	Total	O	0	0
			15	15		
7	H	20	Total	O	0	0
			20	20		
7	Z	18	Total	O	0	0
			18	18		

### 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: Glycoprotein hormones, alpha polypeptide

Chain A: 



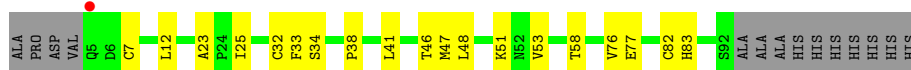
- Molecule 1: Glycoprotein hormones, alpha polypeptide

Chain D: 



- Molecule 1: Glycoprotein hormones, alpha polypeptide

Chain G: 



- Molecule 2: Follitropin subunit beta

Chain B: 



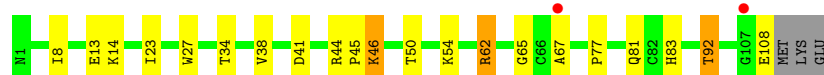
- Molecule 2: Follitropin subunit beta

Chain E: 



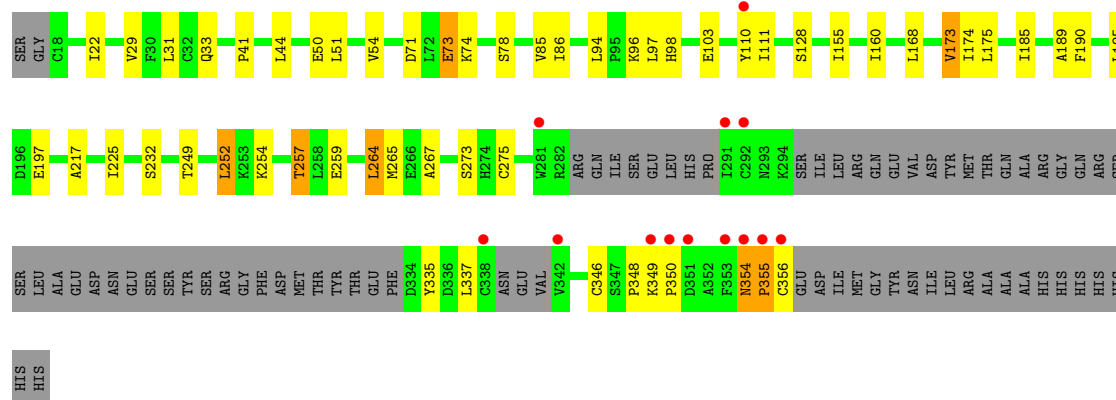
- Molecule 2: Follitropin subunit beta

Chain H: 



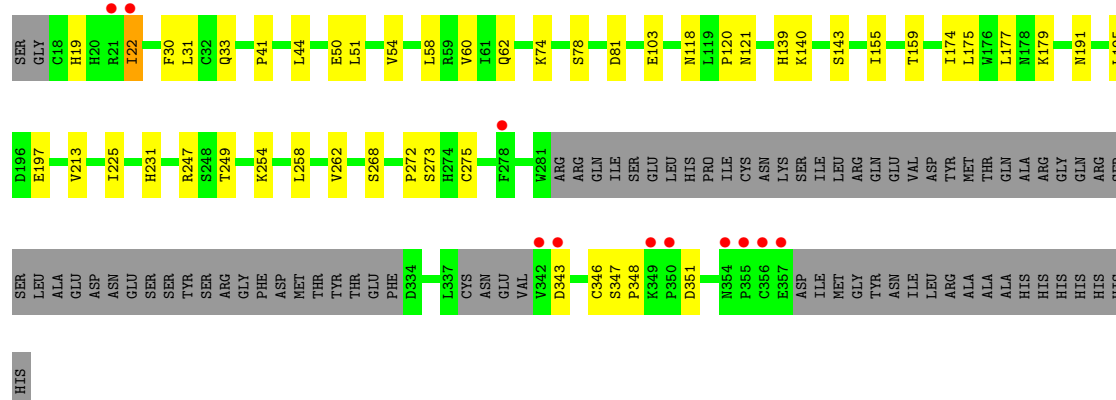
• Molecule 3: Follicle-stimulating hormone receptor

Chain X:



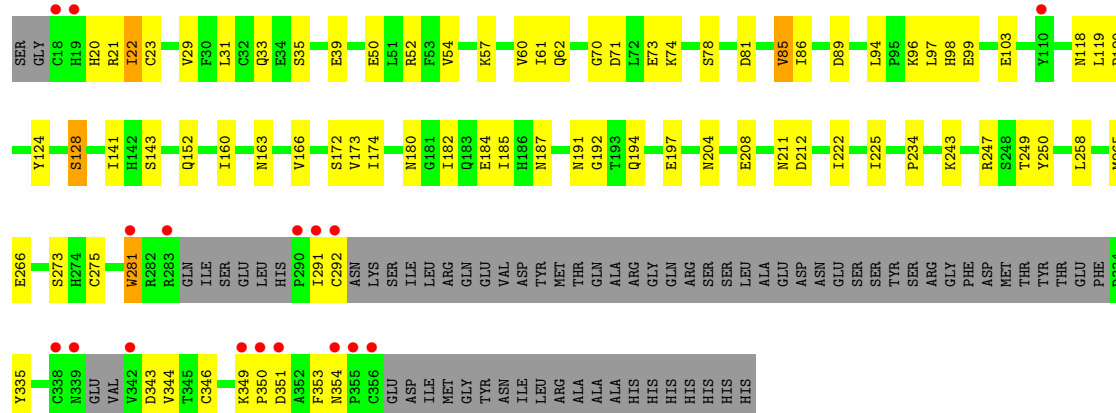
• Molecule 3: Follicle-stimulating hormone receptor

Chain Y:



• Molecule 3: Follicle-stimulating hormone receptor

Chain Z:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.90Å 95.90Å 204.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.93 – 2.90 24.93 – 2.90	Depositor EDS
% Data completeness (in resolution range)	95.6 (24.93-2.90) 95.5 (24.93-2.90)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.71 (at 2.89Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, $R_{free}$	0.174 , 0.237 0.192 , 0.256	Depositor DCC
$R_{free}$ test set	2226 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	57.3	Xtriage
Anisotropy	0.278	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 49.3	EDS
Estimated twinning fraction	0.000 for -h,-k,l 0.000 for h,-h-k,-l 0.000 for -k,-h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 44547 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11798	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.94 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.4222e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities 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: EDO, JEF, NAG, TYS

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	0.46	0/690	0.65	0/933
1	D	0.44	0/697	0.67	0/943
1	G	0.44	0/690	0.67	0/933
2	B	0.44	0/867	0.74	0/1177
2	E	0.44	0/859	0.72	0/1167
2	H	0.46	0/859	0.72	0/1167
3	X	0.44	0/2299	0.73	0/3115
3	Y	0.46	0/2283	0.73	0/3094
3	Z	0.45	0/2339	0.73	0/3168
All	All	0.45	0/11583	0.72	0/15697

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	675	0	647	14	0
1	D	682	0	656	12	0
1	G	675	0	647	14	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	848	0	798	4	0
2	E	840	0	789	12	0
2	H	840	0	789	14	0
3	X	2275	0	2240	33	0
3	Y	2258	0	2235	19	0
3	Z	2312	0	2293	36	0
4	A	28	0	26	1	0
4	B	28	0	26	0	0
4	D	28	0	26	0	0
4	E	28	0	26	0	0
4	G	28	0	26	1	0
4	H	28	0	26	0	0
4	X	14	0	13	0	0
4	Y	14	0	13	0	0
4	Z	14	0	13	0	0
5	B	8	0	12	0	0
5	X	8	0	12	0	0
5	Y	4	0	6	0	0
6	Z	8	0	8	0	0
7	A	13	0	0	0	0
7	B	16	0	0	0	0
7	D	10	0	0	0	0
7	E	13	0	0	0	0
7	G	15	0	0	0	0
7	H	20	0	0	0	0
7	X	22	0	0	0	0
7	Y	28	0	0	0	0
7	Z	18	0	0	1	0
All	All	11798	0	11327	138	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 6.

The worst 5 of 138 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:Y:275:CYS:HG	3:Y:346:CYS:HG	1.05	0.84
1:G:32:CYS:HB3	1:G:58:THR:HG23	1.62	0.81
3:Z:275:CYS:HG	3:Z:346:CYS:HG	0.86	0.81
1:G:47:MET:HE3	1:G:51:LYS:H	1.48	0.78
2:E:41:ASP:O	2:E:44:ARG:HD2	1.85	0.76

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	86/102 (84%)	80 (93%)	6 (7%)	0	100	100
1	D	87/102 (85%)	80 (92%)	7 (8%)	0	100	100
1	G	86/102 (84%)	79 (92%)	7 (8%)	0	100	100
2	B	107/111 (96%)	103 (96%)	4 (4%)	0	100	100
2	E	106/111 (96%)	103 (97%)	2 (2%)	1 (1%)	25	66
2	H	106/111 (96%)	101 (95%)	3 (3%)	2 (2%)	12	42
3	X	280/361 (78%)	244 (87%)	32 (11%)	4 (1%)	16	52
3	Y	277/361 (77%)	246 (89%)	30 (11%)	1 (0%)	43	82
3	Z	281/361 (78%)	241 (86%)	38 (14%)	2 (1%)	30	72
All	All	1416/1722 (82%)	1277 (90%)	129 (9%)	10 (1%)	30	72

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	X	354	ASN
2	H	67	ALA
3	Z	70	GLY
3	X	252	LEU
3	X	337	LEU

### 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	80/90 (89%)	72 (90%)	8 (10%)	11	32

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	81/90 (90%)	77 (95%)	4 (5%)	35	73
1	G	80/90 (89%)	76 (95%)	4 (5%)	34	73
2	B	97/99 (98%)	84 (87%)	13 (13%)	6	15
2	E	96/99 (97%)	86 (90%)	10 (10%)	10	29
2	H	96/99 (97%)	88 (92%)	8 (8%)	16	43
3	X	256/328 (78%)	245 (96%)	11 (4%)	40	78
3	Y	257/328 (78%)	238 (93%)	19 (7%)	20	50
3	Z	263/328 (80%)	241 (92%)	22 (8%)	16	42
All	All	1306/1551 (84%)	1207 (92%)	99 (8%)	19	48

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	83	HIS
3	Y	179	LYS
3	Z	273	SER
3	Y	19	HIS
3	Y	62	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
3	Y	251	ASN
3	Y	354	ASN
3	Z	118	ASN
3	Y	240	ASN
3	Z	93	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

3 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and



the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	TYS	X	335	3	16,16,17	5.52	7 (43%)	20,22,24	1.95	8 (40%)
3	TYS	Y	335	3	16,16,17	5.80	10 (62%)	20,22,24	2.22	7 (35%)
3	TYS	Z	335	3	16,16,17	5.77	11 (68%)	20,22,24	1.52	3 (15%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	TYS	X	335	3	-	0/9/11/13	0/1/1/1
3	TYS	Y	335	3	-	0/9/11/13	0/1/1/1
3	TYS	Z	335	3	-	0/9/11/13	0/1/1/1

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	Z	335	TYS	O-C	18.79	1.24	1.11
3	Y	335	TYS	O-C	18.27	1.24	1.11
3	X	335	TYS	O-C	18.12	1.23	1.11
3	Y	335	TYS	OH-S	-10.21	1.44	1.63
3	X	335	TYS	OH-S	-9.75	1.45	1.63

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	Y	335	TYS	OH-CZ-CE1	5.15	128.45	118.69
3	X	335	TYS	O2-S-O1	4.19	130.23	112.68
3	Y	335	TYS	OH-CZ-CE2	-4.05	111.03	118.69
3	Y	335	TYS	C-CA-N	3.73	117.56	113.83
3	Z	335	TYS	O2-S-O1	3.62	127.81	112.68

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

21 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	NAG	A	201	1	12,14,15	0.31	0	15,19,21	0.85	1 (6%)
4	NAG	A	202	1	12,14,15	0.36	0	15,19,21	0.44	0
4	NAG	B	201	2	12,14,15	0.37	0	15,19,21	0.70	1 (6%)
4	NAG	B	202	2	12,14,15	0.41	0	15,19,21	0.62	0
5	EDO	B	203	-	3,3,3	0.83	0	2,2,2	0.11	0
5	EDO	B	204	-	3,3,3	0.73	0	2,2,2	0.22	0
4	NAG	D	201	1	12,14,15	0.35	0	15,19,21	0.65	1 (6%)
4	NAG	D	202	1	12,14,15	0.36	0	15,19,21	0.48	0
4	NAG	E	201	2	12,14,15	0.35	0	15,19,21	0.46	0
4	NAG	E	202	2	12,14,15	0.36	0	15,19,21	0.46	0
4	NAG	G	201	1	12,14,15	0.33	0	15,19,21	0.76	1 (6%)
4	NAG	G	202	1	12,14,15	0.33	0	15,19,21	0.33	0
4	NAG	H	201	2	12,14,15	0.36	0	15,19,21	0.43	0
4	NAG	H	202	2	12,14,15	0.33	0	15,19,21	0.35	0
5	EDO	X	401	-	3,3,3	0.82	0	2,2,2	0.07	0
5	EDO	X	402	-	3,3,3	0.71	0	2,2,2	0.30	0
4	NAG	X	403	3	12,14,15	0.37	0	15,19,21	0.35	0
5	EDO	Y	401	-	3,3,3	0.73	0	2,2,2	0.35	0
4	NAG	Y	402	3	12,14,15	0.33	0	15,19,21	0.34	0
6	JEF	Z	401	-	5,7,40	1.79	1 (20%)	5,7,48	1.51	2 (40%)
4	NAG	Z	402	3	12,14,15	0.34	0	15,19,21	0.33	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	201	1	-	0/6/23/26	0/1/1/1
4	NAG	A	202	1	-	0/6/23/26	0/1/1/1
4	NAG	B	201	2	-	0/6/23/26	0/1/1/1
4	NAG	B	202	2	-	0/6/23/26	0/1/1/1
5	EDO	B	203	-	-	0/1/1/1	0/0/0/0
5	EDO	B	204	-	-	0/1/1/1	0/0/0/0
4	NAG	D	201	1	-	0/6/23/26	0/1/1/1
4	NAG	D	202	1	-	0/6/23/26	0/1/1/1
4	NAG	E	201	2	-	0/6/23/26	0/1/1/1
4	NAG	E	202	2	-	0/6/23/26	0/1/1/1
4	NAG	G	201	1	-	0/6/23/26	0/1/1/1
4	NAG	G	202	1	-	0/6/23/26	0/1/1/1
4	NAG	H	201	2	-	0/6/23/26	0/1/1/1
4	NAG	H	202	2	-	0/6/23/26	0/1/1/1
5	EDO	X	401	-	-	0/1/1/1	0/0/0/0
5	EDO	X	402	-	-	0/1/1/1	0/0/0/0
4	NAG	X	403	3	-	0/6/23/26	0/1/1/1
5	EDO	Y	401	-	-	0/1/1/1	0/0/0/0
4	NAG	Y	402	3	-	0/6/23/26	0/1/1/1
6	JEF	Z	401	-	-	0/5/6/46	0/0/0/0
4	NAG	Z	402	3	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	Z	401	JEF	O4-C11	3.53	1.47	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	201	NAG	C2-N2-C7	2.54	126.47	123.39
4	A	201	NAG	C2-N2-C7	2.44	126.34	123.39
6	Z	401	JEF	O5-C10-C11	-2.36	109.99	113.35
4	B	201	NAG	C2-N2-C7	2.34	126.22	123.39
4	D	201	NAG	C2-N2-C7	2.09	125.92	123.39

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, 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	88/102 (86%)	-0.51	2 (2%) 57 66	40, 67, 114, 147	0
1	D	89/102 (87%)	-0.48	3 (3%) 43 51	41, 67, 121, 159	0
1	G	88/102 (86%)	-0.50	1 (1%) 77 84	48, 71, 114, 145	0
2	B	109/111 (98%)	-0.43	2 (1%) 65 74	43, 72, 126, 162	0
2	E	108/111 (97%)	-0.43	1 (0%) 81 88	48, 75, 116, 158	0
2	H	108/111 (97%)	-0.31	2 (1%) 64 72	49, 76, 126, 152	0
3	X	289/361 (80%)	-0.20	13 (4%) 32 39	41, 86, 140, 210	0
3	Y	284/361 (78%)	-0.24	11 (3%) 37 45	44, 86, 144, 187	0
3	Z	290/361 (80%)	-0.10	17 (5%) 22 25	42, 92, 143, 195	0
All	All	1453/1722 (84%)	-0.28	52 (3%) 41 48	40, 82, 139, 210	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	4	VAL	6.7
3	Z	356	CYS	6.4
3	X	351	ASP	5.6
3	Z	290	PRO	5.2
3	Z	351	ASP	5.1

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

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, 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	TYS	Z	335	16/17	0.28	2.05	105,114,138,139	0
3	TYS	X	335	16/17	0.26	1.00	107,121,139,141	0
3	TYS	Y	335	16/17	0.24	0.62	114,125,145,147	0

### 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, 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	B	202	14/15	0.37	18.38	162,171,176,179	0
4	NAG	H	202	14/15	0.34	13.74	135,143,151,154	0
5	EDO	X	401	4/4	0.32	13.26	66,67,68,68	0
4	NAG	E	202	14/15	0.44	12.92	144,150,159,164	0
4	NAG	E	201	14/15	0.28	4.91	145,152,158,163	0
4	NAG	B	201	14/15	0.36	4.13	126,136,146,146	0
4	NAG	H	201	14/15	0.35	4.05	140,145,153,154	0
4	NAG	Y	402	14/15	0.28	1.85	106,118,125,127	0
4	NAG	Z	402	14/15	0.24	1.83	103,118,123,124	0
4	NAG	X	403	14/15	0.28	1.50	109,117,127,127	0
6	JEF	Z	401	8/41	0.21	0.91	108,114,116,117	0
5	EDO	B	203	4/4	0.23	0.79	76,77,78,79	0
5	EDO	Y	401	4/4	0.17	-0.01	74,76,77,77	0
4	NAG	D	202	14/15	0.15	-0.09	75,78,86,87	0
4	NAG	G	202	14/15	0.15	-0.14	85,88,95,98	0
4	NAG	D	201	14/15	0.13	-0.33	71,76,80,83	0
4	NAG	A	201	14/15	0.12	-0.60	70,77,86,89	0
4	NAG	A	202	14/15	0.11	-0.80	72,78,86,92	0
4	NAG	G	201	14/15	0.12	-0.85	69,79,88,92	0
5	EDO	X	402	4/4	0.11	-1.16	97,98,99,99	0
5	EDO	B	204	4/4	0.41	-	109,109,110,110	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.