wwPDB X-ray Structure Validation Summary Report (i)

Sep 26, 2017 – 03:05 AM EDT

PDB ID : 1CN4
Title : ERYTHROPOIETIN COMPLEXED WITH EXTRACELLULAR DOMAINS
OF ERYTHROPOIETIN RECEPTOR
Authors : Stroud, R.M.; Reid, S.W.
Deposited on : unknown
Resolution : 2.80 Å(reported)

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

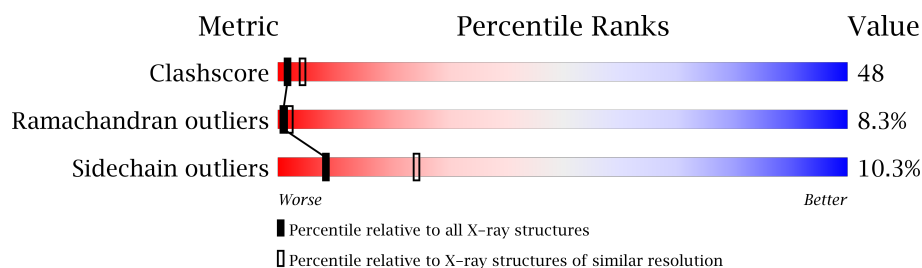|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030345

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 2.80 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		48
Ramachandran outliers		8.3%
Sidechain outliers		10.3%

▮ 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	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)

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	228	31% · 53% · 11% · 5%
1	B	228	30% · 57% · 9% ·
2	C	166	24% · 55% · 12% · 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4604 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 PROTEIN (ERYTHROPOIETIN RECEPTOR).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	217	Total	C	N	O	S	0	0	0
			1688	1071	296	314	7			
1	B	218	Total	C	N	O	S	0	0	0
			1695	1075	297	316	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	52	GLN	ASN	ENGINEERED	UNP P19235
B	52	GLN	ASN	ENGINEERED	UNP P19235
A	164	GLN	ASN	ENGINEERED	UNP P19235
B	164	GLN	ASN	ENGINEERED	UNP P19235
A	211	GLU	ALA	ENGINEERED	UNP P19235
B	211	GLU	ALA	ENGINEERED	UNP P19235

- Molecule 2 is a protein called PROTEIN (ERYTHROPOIETIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	154	Total	C	N	O	S	0	0	0
			1206	767	213	221	5			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	24	LYS	ASN	ENGINEERED	UNP P01588
C	38	LYS	ASN	ENGINEERED	UNP P01588
C	83	LYS	ASN	ENGINEERED	UNP P01588

- Molecule 3 is water.

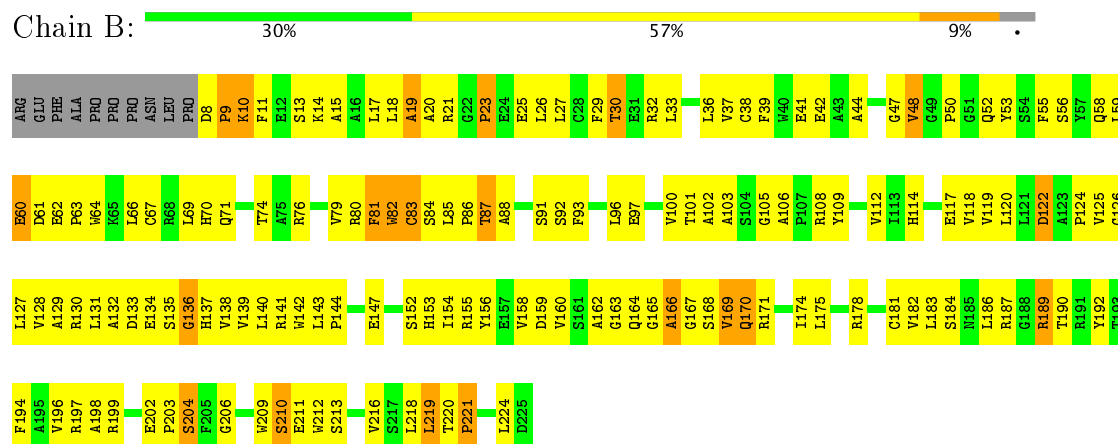| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total 4	O 4	0	0
3	B	5	Total 5	O 5	0	0
3	C	6	Total 6	O 6	0	0

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 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.
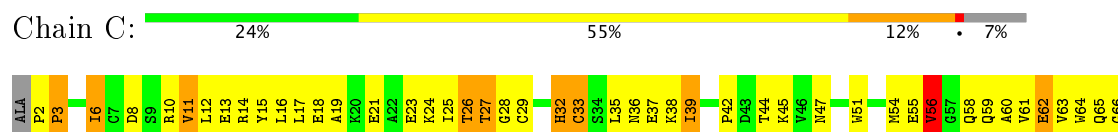
Note EDS was not executed.

● Molecule 1: PROTEIN (ERYTHROPOIETIN RECEPTOR)
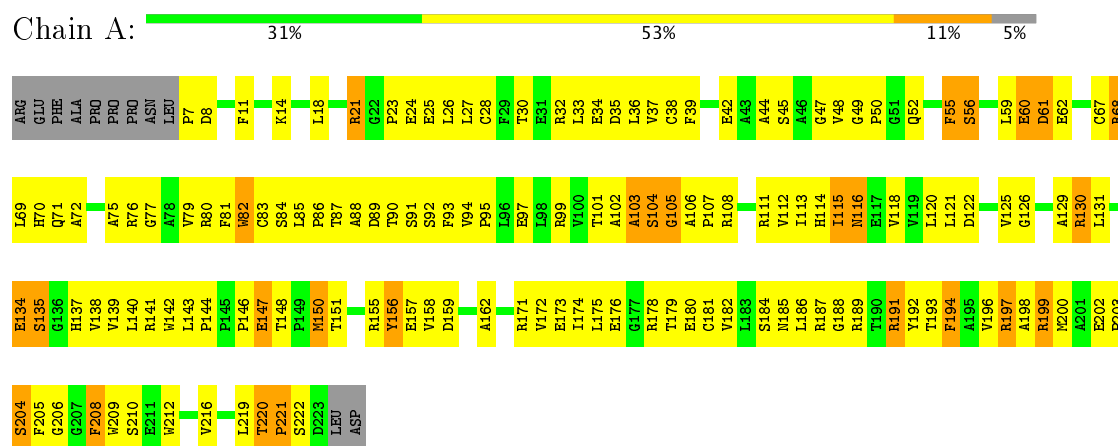
● Molecule 1: PROTEIN (ERYTHROPOIETIN RECEPTOR)

● Molecule 2: PROTEIN (ERYTHROPOIETIN)

L67 A68 L69 L70 S71 E72 A73 V74 L75 R76 A79 L80 L81 V82 K83 S84 S85 Q86 P87 W88 E89 P90 L91 Q92 V95 V99 S100 G101 L102 L105 T106 T107 L108 L109 R110 A111 L112 Q115 K116 E117 A118 I119 S120 P121 P122 D123 ALA ALA SER ALA ALA PRO LEU R131 T132

I133 D136 T137 F138 R139 K140 L141 F142 R143 V144 Y145 S146 N147 F148 L149 R150 G151 K152 L153 K154 L155 Y156 T157 A160 C161 R162 THR GLY ASP ARG

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	73.34Å 80.34Å 134.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.80	Depositor
% Data completeness (in resolution range)	83.4 (20.00-2.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
Refinement program	CNS 0.4	Depositor
R, R_{free}	0.280 , 0.336	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4604	wwPDB-VP
Average B, all atoms (Å^2)	40.0	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	0.41	0/1735	0.69	0/2368
1	B	0.43	0/1741	0.69	1/2375 (0.0%)
2	C	0.49	0/1227	0.71	1/1661 (0.1%)
All	All	0.44	0/4703	0.70	2/6404 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	C	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	133	ILE	N-CA-C	-5.61	95.87	111.00
1	B	23	PRO	N-CA-C	5.11	125.39	112.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	15	TYR	Sidechain

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	1688	0	1635	154	0
1	B	1695	0	1640	161	0
2	C	1206	0	1226	128	0
3	A	4	0	0	0	0
3	B	5	0	0	1	0
3	C	6	0	0	2	0
All	All	4604	0	4501	435	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 48.

The worst 5 of 435 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:115:ILE:H	1:A:115:ILE:HD12	1.09	1.07
1:B:25:GLU:HG2	1:B:26:LEU:H	1.26	1.01
2:C:6:ILE:HG21	2:C:112:LEU:HD11	1.47	0.95
1:A:49:GLY:H	1:A:52:GLN:HG3	1.31	0.95
1:B:55:PHE:HD2	1:B:67:CYS:HB3	1.29	0.94

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	215/228 (94%)	162 (75%)	34 (16%)	19 (9%)	1	2
1	B	216/228 (95%)	159 (74%)	41 (19%)	16 (7%)	1	3
2	C	150/166 (90%)	94 (63%)	43 (29%)	13 (9%)	1	2
All	All	581/622 (93%)	415 (71%)	118 (20%)	48 (8%)	1	2

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	33	LEU
1	A	60	GLU
1	A	104	SER
1	A	147	GLU
1	A	221	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	A	179/190 (94%)	158 (88%)	21 (12%)	6	19
1	B	179/190 (94%)	169 (94%)	10 (6%)	25	57
2	C	129/138 (94%)	110 (85%)	19 (15%)	3	10
All	All	487/518 (94%)	437 (90%)	50 (10%)	8	24

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	81	PHE
1	B	204	SER
2	C	123	ASP
1	B	82	TRP
1	B	120	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	137	HIS
1	B	153	HIS
1	B	170	GLN
1	B	71	GLN
1	B	164	GLN

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 ⓘ

There are no ligands in this entry.

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 (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.