



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 05:51 am BST

PDB ID : 6NDC  
Title : RHODOCETIN IN COMPLEX WITH THE INTEGRIN ALPHA2-A DOMAIN WITH CHROMIUM BOUND  
Authors : Stetefeld, J.; McDougall, M.D.; Loewen, P.C.  
Deposited on : 2018-12-13  
Resolution : 3.35 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

<https://www.wwpdb.org/validation/2017/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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

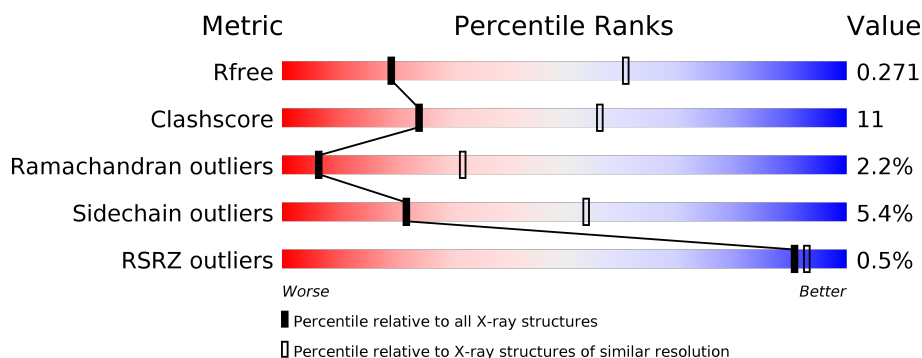
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.35 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	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)
RSRZ outliers	127900	1507 (3.42-3.30)

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 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	A	135	<div> <div>69%</div> <div>26%</div> <div>..</div> </div>
1	D	135	<div> <div>67%</div> <div>25%</div> <div>...</div> </div>
1	G	135	<div> <div>68%</div> <div>27%</div> <div>..</div> </div>
1	J	135	<div> <div>74%</div> <div>21%</div> <div>...</div> </div>
1	M	135	<div> <div>64%</div> <div>29%</div> <div>...</div> </div>
1	P	135	<div> <div>%</div> <div>66%</div> <div>30%</div> <div>...</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	B	124	
2	E	124	
2	H	124	
2	K	124	
2	N	124	
2	Q	124	
3	C	217	
3	F	217	
3	I	217	
3	L	217	
3	O	217	
3	R	217	

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 21590 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 Snaclec rhodocetin subunit gamma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	131	Total	C	N	O	S	0	0	0
			1070	672	189	200	9			
1	D	131	Total	C	N	O	S	0	0	0
			1070	672	189	200	9			
1	G	131	Total	C	N	O	S	0	0	0
			1070	672	189	200	9			
1	J	131	Total	C	N	O	S	0	0	0
			1070	672	189	200	9			
1	M	131	Total	C	N	O	S	0	0	0
			1070	672	189	200	9			
1	P	131	Total	C	N	O	S	0	0	0
			1070	672	189	200	9			

- Molecule 2 is a protein called Snaclec rhodocetin subunit delta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	122	Total	C	N	O	S	0	0	0
			1029	667	174	179	9			
2	E	122	Total	C	N	O	S	0	0	0
			1029	667	174	179	9			
2	H	122	Total	C	N	O	S	0	0	0
			1029	667	174	179	9			
2	K	122	Total	C	N	O	S	0	0	0
			1029	667	174	179	9			
2	N	122	Total	C	N	O	S	0	0	0
			1029	667	174	179	9			
2	Q	122	Total	C	N	O	S	0	0	0
			1029	667	174	179	9			

- Molecule 3 is a protein called Integrin alpha-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	191	Total 1482	C 940	N 250	O 287	S 5	0	0	0
3	F	191	Total 1482	C 940	N 250	O 287	S 5	0	0	0
3	I	191	Total 1482	C 940	N 250	O 287	S 5	0	0	0
3	L	191	Total 1482	C 940	N 250	O 287	S 5	0	0	0
3	O	191	Total 1482	C 940	N 250	O 287	S 5	0	0	0
3	R	191	Total 1482	C 940	N 250	O 287	S 5	0	0	0

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	150	MET	-	expression tag	UNP P17301
C	151	GLY	-	expression tag	UNP P17301
C	152	SER	-	expression tag	UNP P17301
C	153	SER	-	expression tag	UNP P17301
C	154	HIS	-	expression tag	UNP P17301
C	155	HIS	-	expression tag	UNP P17301
C	156	HIS	-	expression tag	UNP P17301
C	157	HIS	-	expression tag	UNP P17301
C	158	HIS	-	expression tag	UNP P17301
C	159	HIS	-	expression tag	UNP P17301
C	160	SER	-	expression tag	UNP P17301
C	161	SER	-	expression tag	UNP P17301
C	162	GLY	-	expression tag	UNP P17301
C	163	LEU	-	expression tag	UNP P17301
C	164	VAL	-	expression tag	UNP P17301
C	165	PRO	-	expression tag	UNP P17301
C	166	ARG	-	expression tag	UNP P17301
C	167	GLY	-	expression tag	UNP P17301
C	168	GLY	-	expression tag	UNP P17301
C	169	SER	-	expression tag	UNP P17301
F	150	MET	-	expression tag	UNP P17301
F	151	GLY	-	expression tag	UNP P17301
F	152	SER	-	expression tag	UNP P17301
F	153	SER	-	expression tag	UNP P17301
F	154	HIS	-	expression tag	UNP P17301
F	155	HIS	-	expression tag	UNP P17301
F	156	HIS	-	expression tag	UNP P17301
F	157	HIS	-	expression tag	UNP P17301

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	158	HIS	-	expression tag	UNP P17301
F	159	HIS	-	expression tag	UNP P17301
F	160	SER	-	expression tag	UNP P17301
F	161	SER	-	expression tag	UNP P17301
F	162	GLY	-	expression tag	UNP P17301
F	163	LEU	-	expression tag	UNP P17301
F	164	VAL	-	expression tag	UNP P17301
F	165	PRO	-	expression tag	UNP P17301
F	166	ARG	-	expression tag	UNP P17301
F	167	GLY	-	expression tag	UNP P17301
F	168	GLY	-	expression tag	UNP P17301
F	169	SER	-	expression tag	UNP P17301
I	150	MET	-	expression tag	UNP P17301
I	151	GLY	-	expression tag	UNP P17301
I	152	SER	-	expression tag	UNP P17301
I	153	SER	-	expression tag	UNP P17301
I	154	HIS	-	expression tag	UNP P17301
I	155	HIS	-	expression tag	UNP P17301
I	156	HIS	-	expression tag	UNP P17301
I	157	HIS	-	expression tag	UNP P17301
I	158	HIS	-	expression tag	UNP P17301
I	159	HIS	-	expression tag	UNP P17301
I	160	SER	-	expression tag	UNP P17301
I	161	SER	-	expression tag	UNP P17301
I	162	GLY	-	expression tag	UNP P17301
I	163	LEU	-	expression tag	UNP P17301
I	164	VAL	-	expression tag	UNP P17301
I	165	PRO	-	expression tag	UNP P17301
I	166	ARG	-	expression tag	UNP P17301
I	167	GLY	-	expression tag	UNP P17301
I	168	GLY	-	expression tag	UNP P17301
I	169	SER	-	expression tag	UNP P17301
L	150	MET	-	expression tag	UNP P17301
L	151	GLY	-	expression tag	UNP P17301
L	152	SER	-	expression tag	UNP P17301
L	153	SER	-	expression tag	UNP P17301
L	154	HIS	-	expression tag	UNP P17301
L	155	HIS	-	expression tag	UNP P17301
L	156	HIS	-	expression tag	UNP P17301
L	157	HIS	-	expression tag	UNP P17301
L	158	HIS	-	expression tag	UNP P17301
L	159	HIS	-	expression tag	UNP P17301

*Continued on next page...*

*Continued from previous page...*

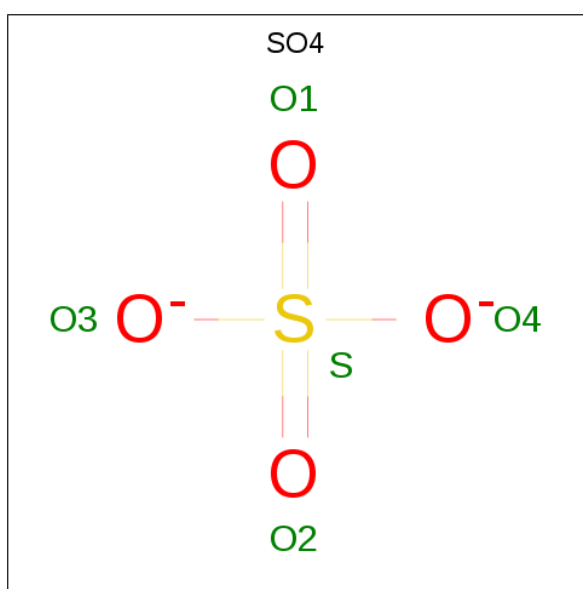
Chain	Residue	Modelled	Actual	Comment	Reference
L	160	SER	-	expression tag	UNP P17301
L	161	SER	-	expression tag	UNP P17301
L	162	GLY	-	expression tag	UNP P17301
L	163	LEU	-	expression tag	UNP P17301
L	164	VAL	-	expression tag	UNP P17301
L	165	PRO	-	expression tag	UNP P17301
L	166	ARG	-	expression tag	UNP P17301
L	167	GLY	-	expression tag	UNP P17301
L	168	GLY	-	expression tag	UNP P17301
L	169	SER	-	expression tag	UNP P17301
O	150	MET	-	expression tag	UNP P17301
O	151	GLY	-	expression tag	UNP P17301
O	152	SER	-	expression tag	UNP P17301
O	153	SER	-	expression tag	UNP P17301
O	154	HIS	-	expression tag	UNP P17301
O	155	HIS	-	expression tag	UNP P17301
O	156	HIS	-	expression tag	UNP P17301
O	157	HIS	-	expression tag	UNP P17301
O	158	HIS	-	expression tag	UNP P17301
O	159	HIS	-	expression tag	UNP P17301
O	160	SER	-	expression tag	UNP P17301
O	161	SER	-	expression tag	UNP P17301
O	162	GLY	-	expression tag	UNP P17301
O	163	LEU	-	expression tag	UNP P17301
O	164	VAL	-	expression tag	UNP P17301
O	165	PRO	-	expression tag	UNP P17301
O	166	ARG	-	expression tag	UNP P17301
O	167	GLY	-	expression tag	UNP P17301
O	168	GLY	-	expression tag	UNP P17301
O	169	SER	-	expression tag	UNP P17301
R	150	MET	-	expression tag	UNP P17301
R	151	GLY	-	expression tag	UNP P17301
R	152	SER	-	expression tag	UNP P17301
R	153	SER	-	expression tag	UNP P17301
R	154	HIS	-	expression tag	UNP P17301
R	155	HIS	-	expression tag	UNP P17301
R	156	HIS	-	expression tag	UNP P17301
R	157	HIS	-	expression tag	UNP P17301
R	158	HIS	-	expression tag	UNP P17301
R	159	HIS	-	expression tag	UNP P17301
R	160	SER	-	expression tag	UNP P17301
R	161	SER	-	expression tag	UNP P17301

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
R	162	GLY	-	expression tag	UNP P17301
R	163	LEU	-	expression tag	UNP P17301
R	164	VAL	-	expression tag	UNP P17301
R	165	PRO	-	expression tag	UNP P17301
R	166	ARG	-	expression tag	UNP P17301
R	167	GLY	-	expression tag	UNP P17301
R	168	GLY	-	expression tag	UNP P17301
R	169	SER	-	expression tag	UNP P17301

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0
4	G	1	Total O S 5 4 1	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	1	Total O S 5 4 1	0	0
4	J	1	Total O S 5 4 1	0	0
4	K	1	Total O S 5 4 1	0	0
4	N	1	Total O S 5 4 1	0	0
4	O	1	Total O S 5 4 1	0	0
4	Q	1	Total O S 5 4 1	0	0
4	Q	1	Total O S 5 4 1	0	0
4	R	1	Total O S 5 4 1	0	0

- Molecule 5 is CHROMIUM ION (three-letter code: CR) (formula: Cr).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	I	1	Total Cr 1 1	0	0
5	C	1	Total Cr 1 1	0	0
5	O	1	Total Cr 1 1	0	0
5	R	1	Total Cr 1 1	0	0
5	L	1	Total Cr 1 1	0	0
5	F	1	Total Cr 1 1	0	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	I	1	Total Na 1 1	0	0
6	C	1	Total Na 1 1	0	0
6	O	1	Total Na 1 1	0	0

*Continued on next page...*

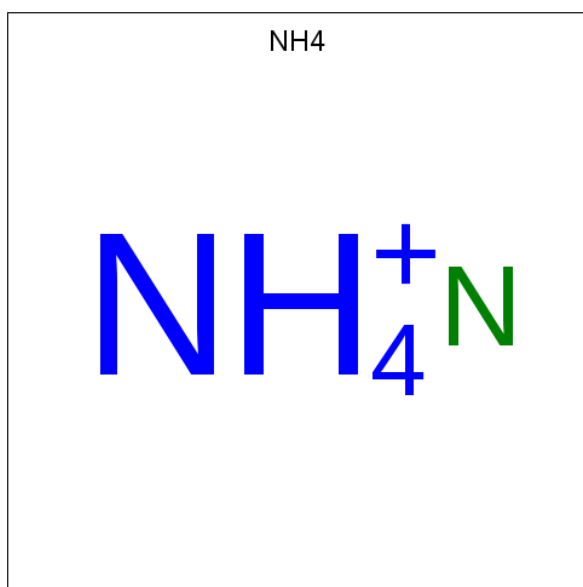
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	R	1	Total 1	Na 1	0	0
6	L	1	Total 1	Na 1	0	0
6	F	1	Total 1	Na 1	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	G	1	Total 1	Cl 1	0	0
7	J	1	Total 1	Cl 1	0	0
7	K	1	Total 1	Cl 1	0	0
7	I	1	Total 1	Cl 1	0	0
7	C	1	Total 1	Cl 1	0	0
7	R	1	Total 1	Cl 1	0	0
7	L	1	Total 1	Cl 1	0	0
7	F	1	Total 1	Cl 1	0	0
7	M	1	Total 1	Cl 1	0	0

- Molecule 8 is AMMONIUM ION (three-letter code: NH4) (formula: H<sub>4</sub>N).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	D	1	Total N 1 1	0	0

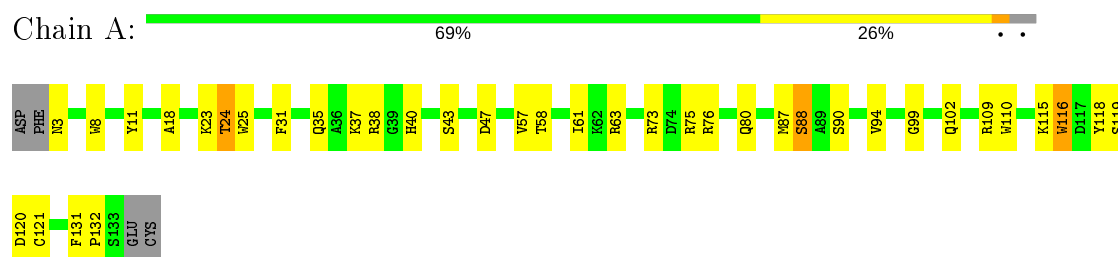
- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	5	Total O 5 5	0	0
9	D	1	Total O 1 1	0	0
9	F	1	Total O 1 1	0	0

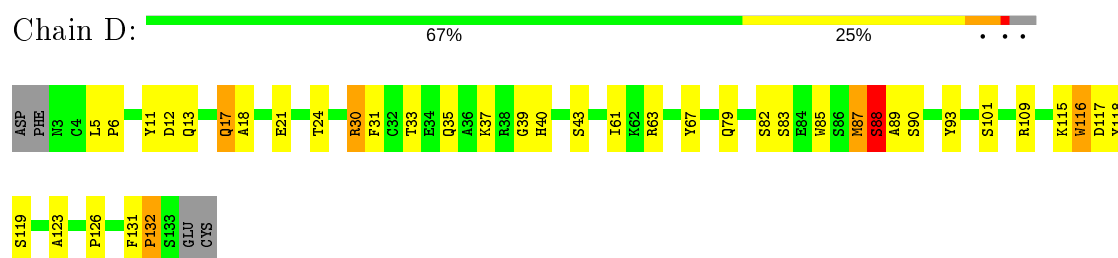
### 3 Residue-property plots

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.

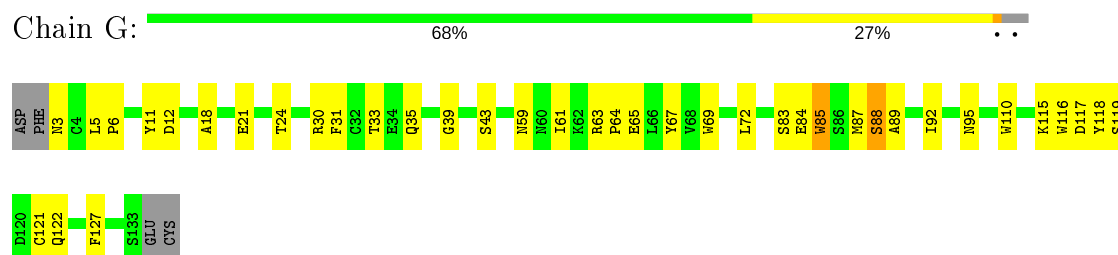
- Molecule 1: Snaclec rhodocetin subunit gamma



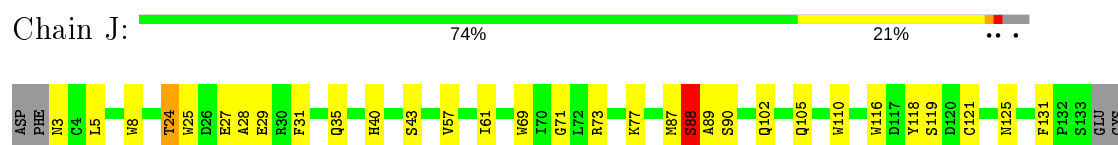
- Molecule 1: Snaclec rhodocetin subunit gamma



- Molecule 1: Snaclec rhodocetin subunit gamma



- Molecule 1: Snaclec rhodocetin subunit gamma



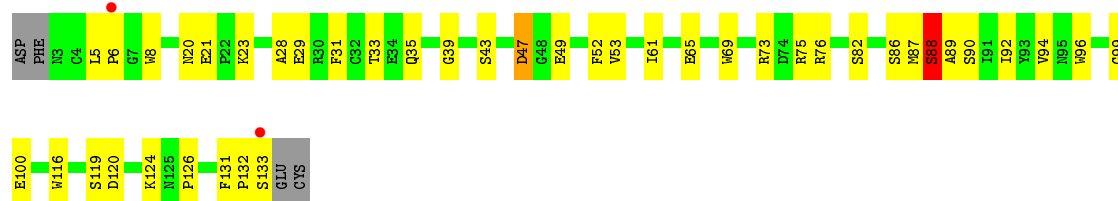
- Molecule 1: Snaclec rhodocetin subunit gamma

Chain M:  64% 29% . . .




• Molecule 1: Snaclec rhodocetin subunit gamma

Chain P:  66% 30% . . .



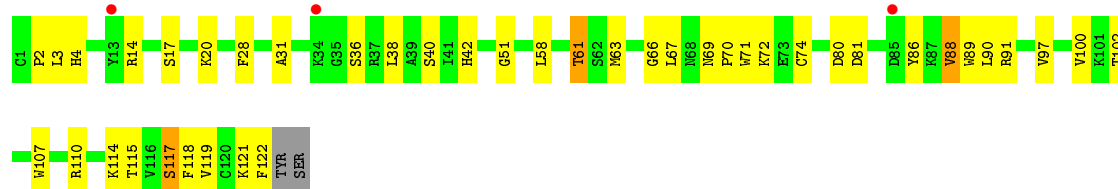
• Molecule 2: Snaclec rhodocetin subunit delta

Chain B:  75% 22% . .



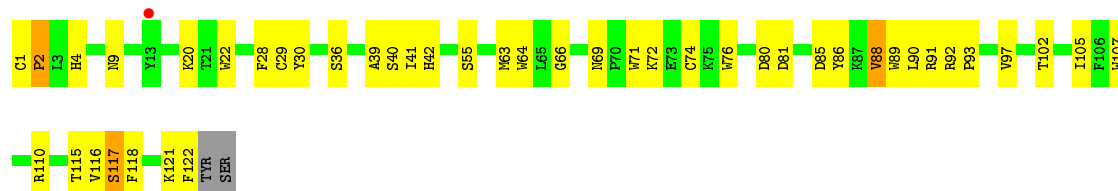
• Molecule 2: Snaclec rhodocetin subunit delta

Chain E:  65% 31% . .



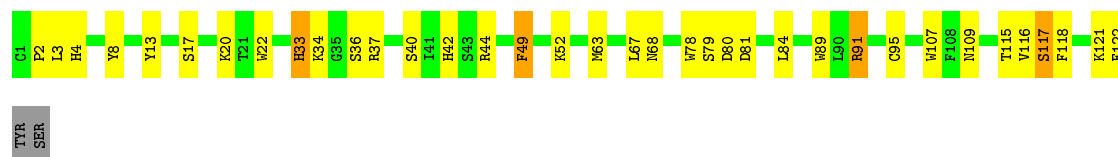
• Molecule 2: Snaclec rhodocetin subunit delta

Chain H:  63% 33% . .

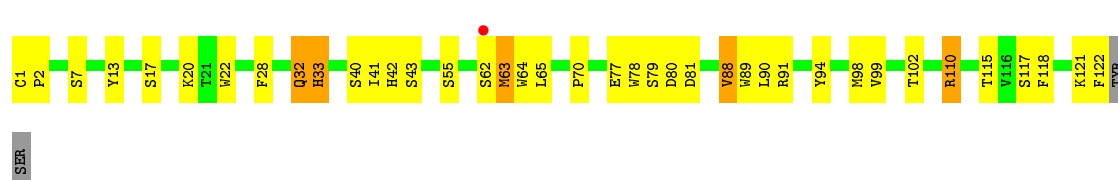


• Molecule 2: Snaclec rhodocetin subunit delta

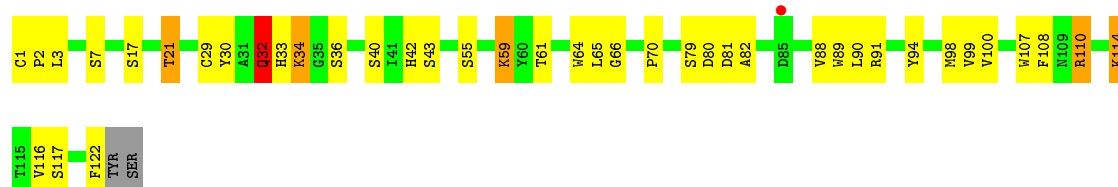
Chain K:  69% 26% . .



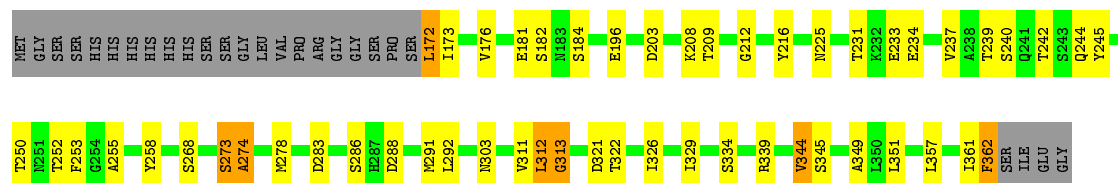
- Molecule 2: Snaclec rhodocetin subunit delta



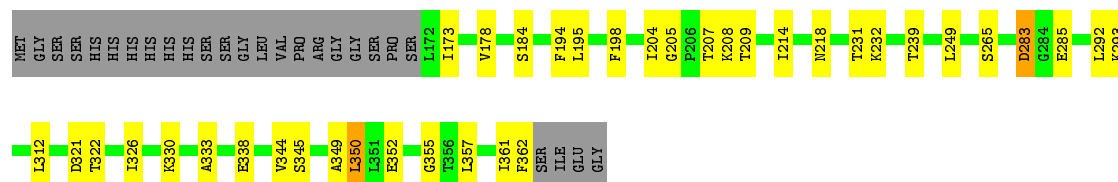
- Molecule 2: Snaclec rhodocetin subunit delta



- Molecule 3: Integrin alpha-2

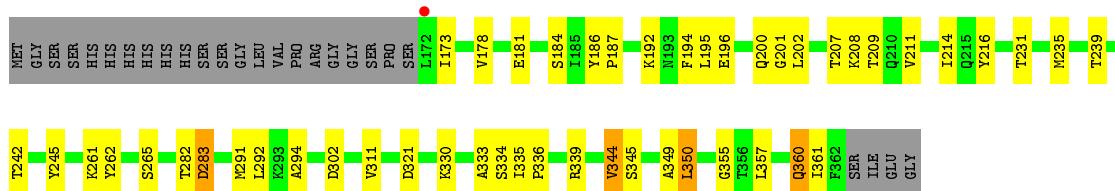


- Molecule 3: Integrin alpha-2



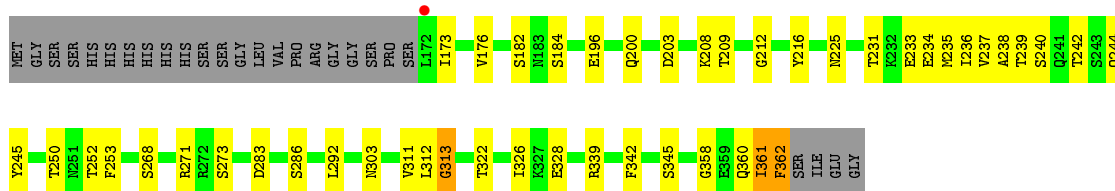
- Molecule 3: Integrin alpha-2





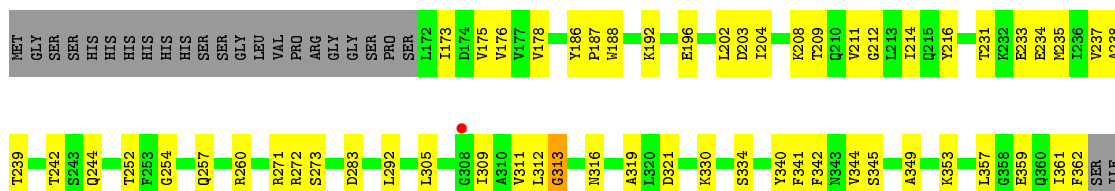
- Molecule 3: Integrin alpha-2

Chain L: 66% 20% 12%



- Molecule 3: Integrin alpha-2

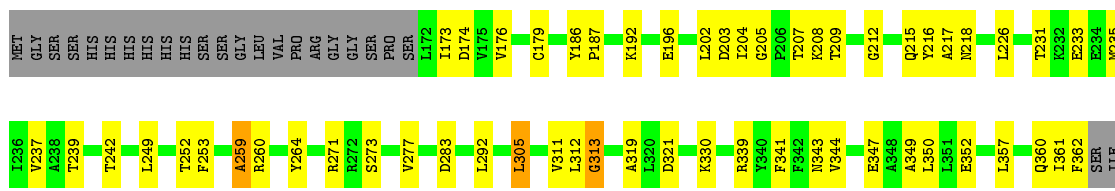
Chain O: 62% 26% 12%



GLU  
GLY

- Molecule 3: Integrin alpha-2

Chain R: 62% 25% 12%



GLU  
GLY

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	131.14Å 131.14Å 251.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	92.90 – 3.35 92.73 – 3.35	Depositor EDS
% Data completeness (in resolution range)	79.8 (92.90-3.35) 79.8 (92.73-3.35)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.16	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.08 (at 3.33Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.215 , 0.276 0.219 , 0.271	Depositor DCC
$R_{free}$ test set	2430 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.6	Xtriage
Anisotropy	0.324	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 15.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.428 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	21590	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.81% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, NH4, SO4, CR, CL

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	0/1101	0.78	0/1491
1	D	0.67	0/1101	0.79	0/1491
1	G	0.68	0/1101	0.81	0/1491
1	J	0.67	0/1101	0.80	0/1491
1	M	0.69	0/1101	0.80	0/1491
1	P	0.69	0/1101	0.82	0/1491
2	B	0.66	0/1064	0.83	0/1439
2	E	0.67	0/1064	0.78	0/1439
2	H	0.66	0/1064	0.78	0/1439
2	K	0.66	0/1064	0.80	0/1439
2	N	0.69	0/1064	0.81	0/1439
2	Q	0.69	0/1064	0.82	0/1439
3	C	0.70	0/1506	0.80	0/2040
3	F	0.69	0/1506	0.81	0/2040
3	I	0.69	0/1506	0.81	0/2040
3	L	0.69	0/1506	0.79	0/2040
3	O	0.70	0/1506	0.79	0/2040
3	R	0.70	0/1506	0.80	0/2040
All	All	0.68	0/22026	0.80	0/29820

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
1	A	0	1
2	B	0	1
2	H	0	1
2	K	0	1
2	Q	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
3	L	0	1
All	All	0	6

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	132	PRO	Peptide
2	B	4	HIS	Peptide
2	H	39	ALA	Peptide
2	K	2	PRO	Peptide
3	L	339	ARG	Peptide

## 5.2 Too-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 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	1070	0	988	31	0
1	D	1070	0	988	27	0
1	G	1070	0	988	28	0
1	J	1070	0	988	23	0
1	M	1070	0	988	44	0
1	P	1070	0	988	34	0
2	B	1029	0	977	26	0
2	E	1029	0	977	26	0
2	H	1029	0	977	32	0
2	K	1029	0	977	26	0
2	N	1029	0	977	28	0
2	Q	1029	0	977	31	0
3	C	1482	0	1478	38	0
3	F	1482	0	1479	27	0
3	I	1482	0	1478	32	0
3	L	1482	0	1478	28	0
3	O	1482	0	1478	34	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	R	1482	0	1478	33	0
4	A	10	0	0	0	0
4	C	5	0	0	0	0
4	D	5	0	0	0	0
4	E	10	0	0	0	0
4	G	5	0	0	0	0
4	H	5	0	0	0	0
4	J	5	0	0	0	0
4	K	5	0	0	1	0
4	N	5	0	0	0	0
4	O	5	0	0	1	0
4	Q	10	0	0	0	0
4	R	5	0	0	0	0
5	C	1	0	0	0	0
5	F	1	0	0	0	0
5	I	1	0	0	0	0
5	L	1	0	0	0	0
5	O	1	0	0	0	0
5	R	1	0	0	0	0
6	C	1	0	0	0	0
6	F	1	0	0	0	0
6	I	1	0	0	0	0
6	L	1	0	0	0	0
6	O	1	0	0	0	0
6	R	1	0	0	0	0
7	C	1	0	0	0	0
7	F	1	0	0	0	0
7	G	1	0	0	0	0
7	I	1	0	0	0	0
7	J	1	0	0	0	0
7	K	1	0	0	0	0
7	L	1	0	0	0	0
7	M	1	0	0	0	0
7	R	1	0	0	0	0
8	D	1	0	0	0	0
9	A	5	0	0	0	0
9	D	1	0	0	0	0
9	F	1	0	0	0	0
All	All	21590	0	20659	463	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 11.

The worst 5 of 463 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:M:25:TRP:O	1:M:25:TRP:CE3	2.20	0.95
1:M:29:GLU:OE2	2:N:79:SER:OG	1.87	0.92
3:C:203:ASP:OD2	3:C:208:LYS:HD3	1.69	0.91
1:P:29:GLU:CD	2:Q:79:SER:HG	1.74	0.91
1:G:31:PHE:O	1:G:35:GLN:HG2	1.74	0.86

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	129/135 (96%)	117 (91%)	9 (7%)	3 (2%)	6	31
1	D	129/135 (96%)	111 (86%)	14 (11%)	4 (3%)	4	25
1	G	129/135 (96%)	112 (87%)	13 (10%)	4 (3%)	4	25
1	J	129/135 (96%)	111 (86%)	15 (12%)	3 (2%)	6	31
1	M	129/135 (96%)	110 (85%)	16 (12%)	3 (2%)	6	31
1	P	129/135 (96%)	112 (87%)	15 (12%)	2 (2%)	9	38
2	B	120/124 (97%)	109 (91%)	7 (6%)	4 (3%)	4	24
2	E	120/124 (97%)	109 (91%)	7 (6%)	4 (3%)	4	24
2	H	120/124 (97%)	108 (90%)	8 (7%)	4 (3%)	4	24
2	K	120/124 (97%)	106 (88%)	12 (10%)	2 (2%)	9	36
2	N	120/124 (97%)	105 (88%)	11 (9%)	4 (3%)	4	24
2	Q	120/124 (97%)	103 (86%)	11 (9%)	6 (5%)	2	15
3	C	189/217 (87%)	168 (89%)	19 (10%)	2 (1%)	14	46
3	F	189/217 (87%)	169 (89%)	17 (9%)	3 (2%)	9	38

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	I	189/217 (87%)	170 (90%)	15 (8%)	4 (2%)	7	32
3	L	189/217 (87%)	168 (89%)	18 (10%)	3 (2%)	9	38
3	O	189/217 (87%)	174 (92%)	14 (7%)	1 (0%)	29	63
3	R	189/217 (87%)	171 (90%)	15 (8%)	3 (2%)	9	38
All	All	2628/2856 (92%)	2333 (89%)	236 (9%)	59 (2%)	6	32

5 of 59 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	313	GLY
1	D	88	SER
2	E	2	PRO
1	G	85	TRP
1	G	88	SER

### 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	114/118 (97%)	107 (94%)	7 (6%)	18	50
1	D	114/118 (97%)	105 (92%)	9 (8%)	12	40
1	G	114/118 (97%)	109 (96%)	5 (4%)	28	59
1	J	114/118 (97%)	109 (96%)	5 (4%)	28	59
1	M	114/118 (97%)	108 (95%)	6 (5%)	22	54
1	P	114/118 (97%)	109 (96%)	5 (4%)	28	59
2	B	110/112 (98%)	103 (94%)	7 (6%)	17	48
2	E	110/112 (98%)	103 (94%)	7 (6%)	17	48
2	H	110/112 (98%)	105 (96%)	5 (4%)	27	59
2	K	110/112 (98%)	101 (92%)	9 (8%)	11	38
2	N	110/112 (98%)	99 (90%)	11 (10%)	7	29
2	Q	110/112 (98%)	100 (91%)	10 (9%)	9	33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	160/181 (88%)	151 (94%)	9 (6%)	21	53
3	F	160/181 (88%)	152 (95%)	8 (5%)	24	56
3	I	160/181 (88%)	150 (94%)	10 (6%)	18	49
3	L	160/181 (88%)	156 (98%)	4 (2%)	47	73
3	O	160/181 (88%)	158 (99%)	2 (1%)	69	84
3	R	160/181 (88%)	155 (97%)	5 (3%)	40	69
All	All	2304/2466 (93%)	2180 (95%)	124 (5%)	22	54

5 of 124 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	122	PHE
1	J	90	SER
2	Q	59	LYS
3	I	184	SER
3	I	321	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	60	ASN
2	K	4	HIS
3	R	241	GLN
1	J	122	GLN
2	K	42	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

Of 37 ligands modelled in this entry, 1 is modelled with single atom and 21 are monoatomic - leaving 15 for Mogul analysis.

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	SO4	E	201	-	4,4,4	0.31	0	6,6,6	0.07	0
4	SO4	G	202	-	4,4,4	0.38	0	6,6,6	0.10	0
4	SO4	A	201	-	4,4,4	0.37	0	6,6,6	0.07	0
4	SO4	O	403	-	4,4,4	0.37	0	6,6,6	0.08	0
4	SO4	D	201	-	4,4,4	0.36	0	6,6,6	0.07	0
4	SO4	N	201	5,6	4,4,4	0.37	0	6,6,6	0.14	0
4	SO4	Q	202	5,6	4,4,4	0.37	0	6,6,6	0.12	0
4	SO4	A	202	-	4,4,4	0.38	0	6,6,6	0.05	0
4	SO4	K	202	6	4,4,4	0.34	0	6,6,6	0.10	0
4	SO4	Q	201	-	4,4,4	0.33	0	6,6,6	0.11	0
4	SO4	R	402	-	4,4,4	0.38	0	6,6,6	0.07	0
4	SO4	E	202	6	4,4,4	0.31	0	6,6,6	0.21	0
4	SO4	C	402	6	4,4,4	0.34	0	6,6,6	0.12	0
4	SO4	J	201	-	4,4,4	0.34	0	6,6,6	0.09	0
4	SO4	H	201	5,6	4,4,4	0.30	0	6,6,6	0.31	0

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	O	403	SO4	1	0
4	K	202	SO4	1	0

## 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	A	131/135 (97%)	0.32	0 100 100	31, 48, 75, 86	0
1	D	131/135 (97%)	0.22	0 100 100	25, 46, 70, 100	0
1	G	131/135 (97%)	0.19	0 100 100	22, 45, 72, 92	0
1	J	131/135 (97%)	0.29	0 100 100	29, 50, 73, 84	0
1	M	131/135 (97%)	0.29	0 100 100	33, 61, 101, 117	0
1	P	131/135 (97%)	0.26	2 (1%) 73 76	31, 59, 105, 127	0
2	B	122/124 (98%)	0.35	1 (0%) 86 89	34, 58, 87, 104	0
2	E	122/124 (98%)	0.45	3 (2%) 57 59	29, 55, 80, 113	0
2	H	122/124 (98%)	0.41	1 (0%) 86 89	25, 53, 85, 98	0
2	K	122/124 (98%)	0.32	0 100 100	35, 58, 84, 106	0
2	N	122/124 (98%)	0.25	1 (0%) 86 89	31, 53, 79, 113	0
2	Q	122/124 (98%)	0.26	1 (0%) 86 89	30, 55, 78, 110	0
3	C	191/217 (88%)	0.22	0 100 100	30, 49, 75, 105	0
3	F	191/217 (88%)	0.14	0 100 100	26, 45, 74, 106	0
3	I	191/217 (88%)	0.15	1 (0%) 91 93	24, 45, 75, 97	0
3	L	191/217 (88%)	0.23	1 (0%) 91 93	24, 51, 76, 96	0
3	O	191/217 (88%)	0.15	1 (0%) 91 93	25, 47, 73, 95	0
3	R	191/217 (88%)	0.16	0 100 100	23, 46, 77, 105	0
All	All	2664/2856 (93%)	0.24	12 (0%) 91 93	22, 50, 81, 127	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	13	TYR	2.9
2	H	13	TYR	2.6
1	P	133	SER	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	5	TRP	2.4
3	O	308	GLY	2.4

## 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](#)

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(Å <sup>2</sup> )	Q<0.9
8	NH4	D	202	1/1	0.80	0.39	6,6,6,6	0
4	SO4	R	402	5/5	0.81	0.18	88,90,93,93	5
4	SO4	A	202	5/5	0.85	0.28	143,155,166,174	0
7	CL	I	401	1/1	0.85	0.16	37,37,37,37	0
4	SO4	H	201	5/5	0.86	0.17	57,61,70,81	0
7	CL	L	403	1/1	0.87	0.16	38,38,38,38	0
4	SO4	K	202	5/5	0.87	0.21	83,84,107,119	0
4	SO4	E	202	5/5	0.88	0.20	62,70,81,89	0
7	CL	F	401	1/1	0.88	0.21	37,37,37,37	0
4	SO4	Q	202	5/5	0.89	0.15	90,96,108,123	0
4	SO4	O	403	5/5	0.90	0.14	102,104,111,116	0
6	NA	C	403	1/1	0.90	0.17	31,31,31,31	0
7	CL	G	201	1/1	0.91	0.13	38,38,38,38	0
7	CL	J	202	1/1	0.91	0.11	57,57,57,57	0
4	SO4	C	402	5/5	0.91	0.21	72,78,96,117	0
7	CL	M	201	1/1	0.91	0.23	48,48,48,48	0
6	NA	R	404	1/1	0.92	0.20	35,35,35,35	0
4	SO4	N	201	5/5	0.93	0.16	84,89,105,108	0
6	NA	O	402	1/1	0.94	0.26	32,32,32,32	0
5	CR	R	401	1/1	0.94	0.10	76,76,76,76	0
6	NA	F	403	1/1	0.95	0.14	5,5,5,5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	J	201	5/5	0.96	0.16	67,70,73,80	0
4	SO4	D	201	5/5	0.96	0.15	43,48,50,54	0
4	SO4	A	201	5/5	0.96	0.17	73,76,79,82	0
4	SO4	E	201	5/5	0.97	0.14	58,66,71,76	0
4	SO4	Q	201	5/5	0.97	0.15	61,64,69,73	0
7	CL	K	201	1/1	0.97	0.06	49,49,49,49	0
7	CL	C	404	1/1	0.97	0.09	31,31,31,31	0
5	CR	I	402	1/1	0.97	0.13	44,44,44,44	0
6	NA	I	403	1/1	0.98	0.15	8,8,8,8	0
7	CL	R	403	1/1	0.98	0.07	35,35,35,35	0
5	CR	C	401	1/1	0.98	0.12	52,52,52,52	0
5	CR	O	401	1/1	0.98	0.08	76,76,76,76	0
4	SO4	G	202	5/5	0.98	0.10	46,47,49,50	0
5	CR	F	402	1/1	0.98	0.11	47,47,47,47	0
6	NA	L	402	1/1	0.99	0.13	38,38,38,38	0
5	CR	L	401	1/1	0.99	0.10	53,53,53,53	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.