



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 3, 2017 – 07:19 PM EDT

PDB ID : 1GMO  
Title : CRYSTAL STRUCTURES OF NK1-HEPARIN COMPLEXES REVEAL THE BASIS FOR NK1 ACTIVITY AND ENABLE ENGINEERING OF POTENT AGONISTS OF THE MET RECEPTOR  
Authors : Lietha, D.; Chirgadze, D.Y.; Mulloy, B.; Blundell, T.L.; Gherardi, E.  
Deposited on : unknown  
Resolution : 3.00 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20029824

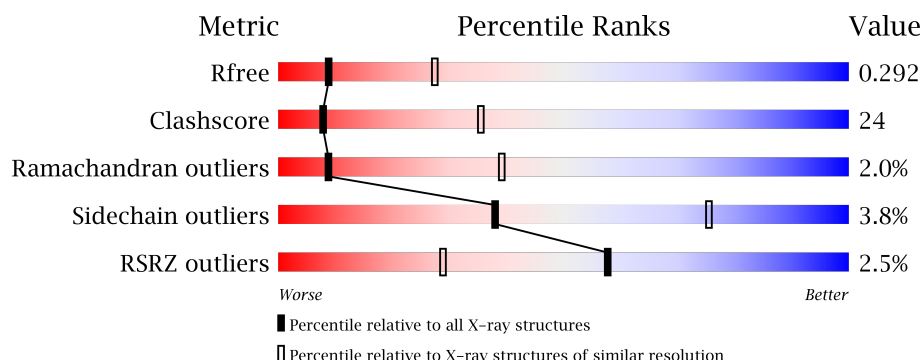
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	100719	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (3.00-3.00)

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	183	<div> <div>3%</div> <div>56%</div> <div>34%</div> <div>5%</div> </div>
1	B	183	<div> <div>%</div> <div>59%</div> <div>34%</div> <div>5%</div> </div>
1	C	183	<div> <div>%</div> <div>55%</div> <div>37%</div> <div>5%</div> </div>
1	D	183	<div> <div>%</div> <div>52%</div> <div>38%</div> <div>5%</div> </div>
1	E	183	<div> <div>%</div> <div>59%</div> <div>32%</div> <div>5%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	183	
1	G	183	
1	H	183	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	IDS	A	305	X	-	-	-
2	IDS	B	315	X	-	-	-
3	SGN	A	302	-	-	-	X
3	SGN	A	304	-	-	-	X
3	SGN	B	302	-	-	-	X
3	SGN	C	304	-	-	-	X
3	SGN	G	305	-	-	-	X
4	SO4	A	307	-	-	-	X
4	SO4	C	301	-	-	X	-
4	SO4	F	307	-	-	X	-
5	EPE	F	309	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11555 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	173	Total	C	N	O	S	0	0	1
			1324	833	229	250	12			
1	B	173	Total	C	N	O	S	0	0	1
			1339	843	232	252	12			
1	C	173	Total	C	N	O	S	0	0	1
			1341	846	232	251	12			
1	D	173	Total	C	N	O	S	0	0	1
			1305	821	224	248	12			
1	E	174	Total	C	N	O	S	0	0	1
			1297	812	223	250	12			
1	F	174	Total	C	N	O	S	0	0	1
			1331	836	232	251	12			
1	G	165	Total	C	N	O	S	0	0	1
			1241	783	213	233	12			
1	H	172	Total	C	N	O	S	0	0	1
			1299	814	225	248	12			

There are 16 discrepancies between the modelled and reference sequences:

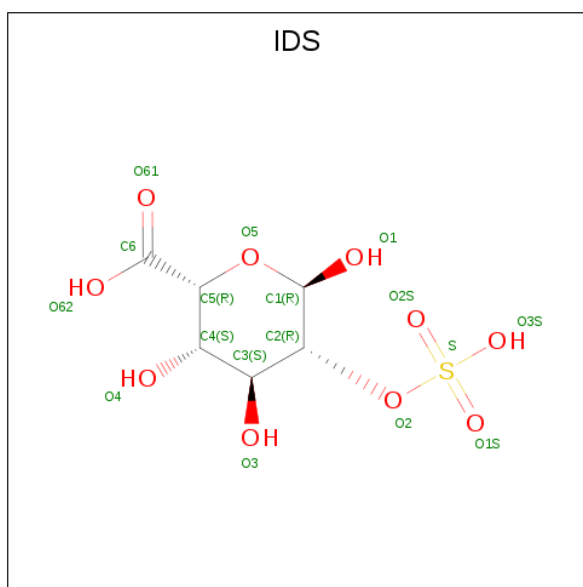
Chain	Residue	Modelled	Actual	Comment	Reference
A	29	VAL	ALA	engineered mutation	UNP P14210
A	72	ASP	ASN	engineered mutation	UNP P14210
B	29	VAL	ALA	engineered mutation	UNP P14210
B	72	ASP	ASN	engineered mutation	UNP P14210
C	29	VAL	ALA	engineered mutation	UNP P14210
C	72	ASP	ASN	engineered mutation	UNP P14210
D	29	VAL	ALA	engineered mutation	UNP P14210
D	72	ASP	ASN	engineered mutation	UNP P14210
E	29	VAL	ALA	engineered mutation	UNP P14210
E	72	ASP	ASN	engineered mutation	UNP P14210
F	29	VAL	ALA	engineered mutation	UNP P14210
F	72	ASP	ASN	engineered mutation	UNP P14210
G	29	VAL	ALA	engineered mutation	UNP P14210

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	72	ASP	ASN	engineered mutation	UNP P14210
H	29	VAL	ALA	engineered mutation	UNP P14210
H	72	ASP	ASN	engineered mutation	UNP P14210

- Molecule 2 is 2-O-sulfo-alpha-L-idopyranuronic acid (three-letter code: IDS) (formula:  $C_6H_{10}O_{10}S$ ).



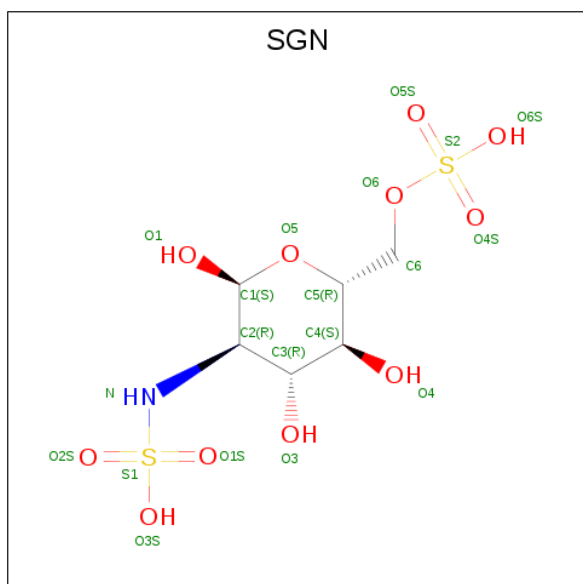
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	0
			17	6	10	1		
2	A	1	Total	C	O	S	0	0
			16	6	9	1		
2	A	1	Total	C	O	S	0	0
			16	6	9	1		
2	B	1	Total	C	O	S	0	0
			17	6	10	1		
2	B	1	Total	C	O	S	0	0
			16	6	9	1		
2	B	1	Total	C	O	S	0	0
			16	6	9	1		
2	B	1	Total	C	O	S	0	0
			15	6	8	1		
2	B	1	Total	C	O	S	0	0
			17	6	10	1		
2	B	1	Total	C	O	S	0	0
			16	6	9	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	O	S	0	0
			16	6	9	1		
2	B	1	Total	C	O	S	0	0
			16	6	9	1		
2	C	1	Total	C	O	S	0	0
			16	6	9	1		
2	C	1	Total	C	O	S	0	0
			16	6	9	1		
2	C	1	Total	C	O	S	0	0
			16	6	9	1		
2	C	1	Total	C	O	S	0	0
			16	6	9	1		
2	F	1	Total	C	O	S	0	0
			16	6	9	1		
2	F	1	Total	C	O	S	0	0
			16	6	9	1		
2	F	1	Total	C	O	S	0	0
			15	6	8	1		
2	G	1	Total	C	O	S	0	0
			16	6	9	1		
2	G	1	Total	C	O	S	0	0
			16	6	9	1		
2	G	1	Total	C	O	S	0	0
			15	6	8	1		

- Molecule 3 is N,O6-DISULFO-GLUCOSAMINE (three-letter code: SGN) (formula:  $C_6H_{13}NO_{11}S_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	A	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	A	1	Total	C	N	O	S	0	0
			18	6	1	9	2		
3	B	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	B	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	B	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	B	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	B	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	B	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	B	1	Total	C	N	O	S	0	0
			18	6	1	9	2		
3	C	1	Total	C	N	O	S	0	0
			20	6	1	11	2		
3	C	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	C	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	C	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	C	1	Total	C	N	O	S	0	0
			18	6	1	9	2		
3	F	1	Total	C	N	O	S	0	0
			20	6	1	11	2		
3	F	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	F	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	G	1	Total	C	N	O	S	0	0
			20	6	1	11	2		
3	G	1	Total	C	N	O	S	0	0
			19	6	1	10	2		
3	G	1	Total	C	N	O	S	0	0
			19	6	1	10	2		

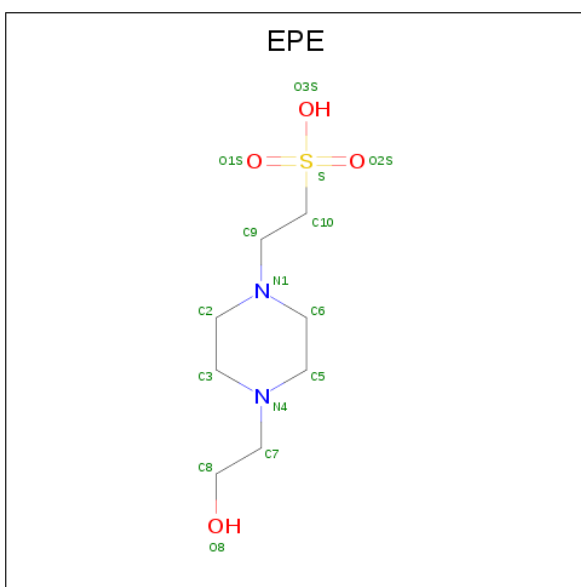
- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 5 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	B	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	C	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	D	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	E	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	F	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	G	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	H	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	21	Total	O	0	0
			21	21		
6	B	34	Total	O	0	0
			34	34		
6	C	27	Total	O	0	0
			27	27		
6	D	22	Total	O	0	0
			22	22		

*Continued on next page...*

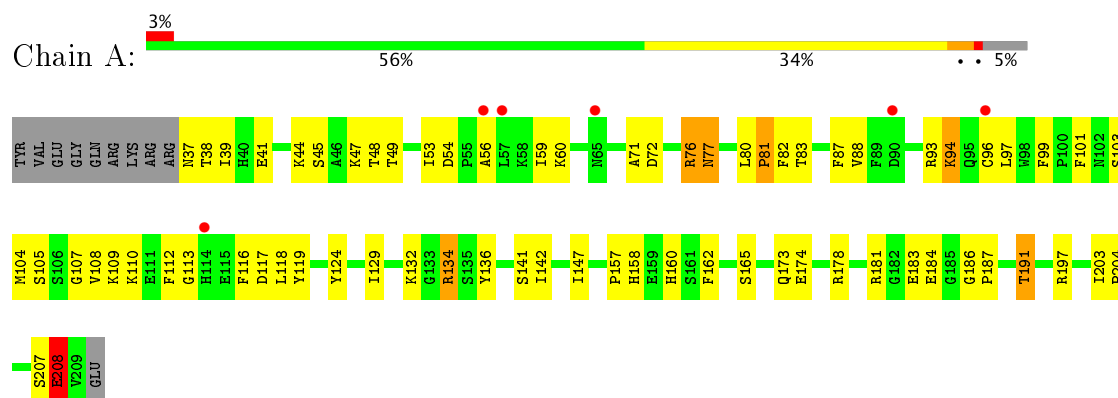
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	E	28	Total 28	O 28	0	0
6	F	20	Total 20	O 20	0	0
6	G	17	Total 17	O 17	0	0
6	H	14	Total 14	O 14	0	0

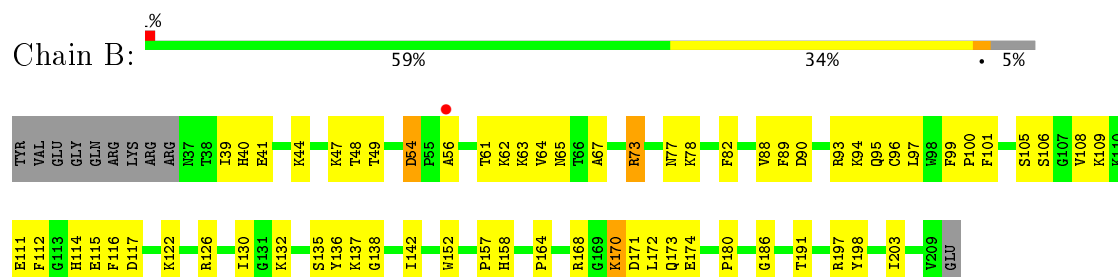
### 3 Residue-property plots [i](#)

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.

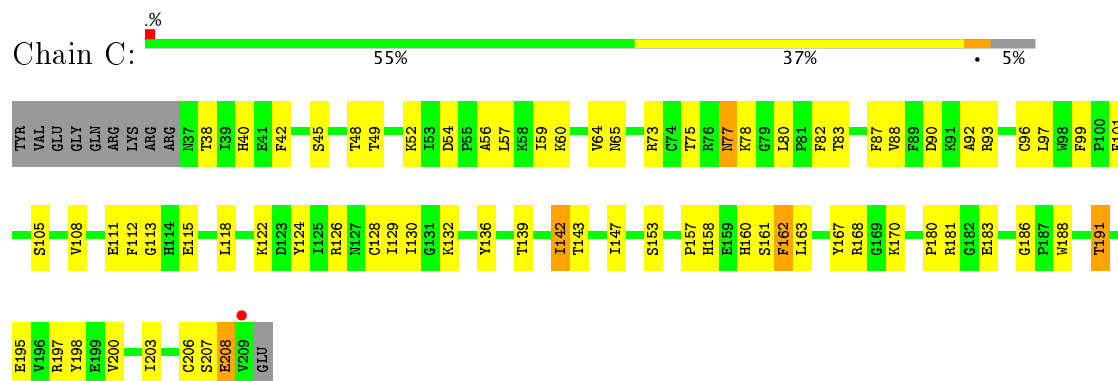
#### • Molecule 1: HEPATOCYTE GROWTH FACTOR



#### • Molecule 1: HEPATOCYTE GROWTH FACTOR



#### • Molecule 1: HEPATOCYTE GROWTH FACTOR



#### • Molecule 1: HEPATOCYTE GROWTH FACTOR



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	179.70 Å 174.10 Å 59.90 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 19.97 – 2.97	Depositor EDS
% Data completeness (in resolution range)	95.3 (20.00-3.00) 94.6 (19.97-2.97)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.26 (at 2.98 Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.254 , 0.295 0.252 , 0.292	Depositor DCC
$R_{free}$ test set	1824 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	59.5	Xtriage
Anisotropy	0.123	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 65.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.005 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	11555	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% 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: SO4, EPE, IDS, SGN

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	0.45	0/1358	0.67	0/1837
1	B	0.45	0/1374	0.63	0/1857
1	C	0.51	0/1376	0.67	0/1861
1	D	0.46	0/1338	0.63	0/1815
1	E	0.46	0/1331	0.66	0/1810
1	F	0.46	0/1365	0.66	0/1850
1	G	0.40	0/1273	0.62	0/1726
1	H	0.41	0/1334	0.59	0/1810
All	All	0.45	0/10749	0.64	0/14566

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	1324	0	1221	73	0
1	B	1339	0	1248	62	0
1	C	1341	0	1257	71	0
1	D	1305	0	1193	73	0
1	E	1297	0	1140	72	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1331	0	1223	56	0
1	G	1241	0	1121	55	0
1	H	1299	0	1162	55	0
2	A	49	0	16	0	0
2	B	129	0	41	2	0
2	C	64	0	20	1	0
2	F	47	0	14	0	0
2	G	47	0	14	0	0
3	A	56	0	23	0	0
3	B	132	0	55	3	0
3	C	95	0	41	1	0
3	F	58	0	26	1	0
3	G	58	0	26	0	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0
4	C	5	0	0	2	0
4	D	5	0	0	1	0
4	E	5	0	0	0	0
4	F	10	0	0	2	0
4	H	5	0	0	1	0
5	A	15	0	17	1	0
5	B	15	0	17	2	0
5	C	15	0	17	3	0
5	D	15	0	17	1	0
5	E	15	0	17	1	0
5	F	15	0	17	1	0
5	G	15	0	17	0	0
5	H	15	0	17	0	0
6	A	21	0	0	0	0
6	B	34	0	0	5	0
6	C	27	0	0	6	0
6	D	22	0	0	7	0
6	E	28	0	0	7	0
6	F	20	0	0	3	0
6	G	17	0	0	3	0
6	H	14	0	0	1	0
All	All	11555	0	9977	513	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 24.

The worst 5 of 513 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:E:181:ARG:NH2	1:H:165:SER:HB3	1.63	1.12
1:D:52:LYS:H	1:D:52:LYS:HD3	1.22	1.01
1:A:134:ARG:HD2	1:A:134:ARG:H	1.26	0.98
1:F:64:VAL:HG13	1:F:69:GLN:OE1	1.64	0.96
1:G:132:LYS:H	1:G:132:LYS:HD2	1.31	0.96

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	171/183 (93%)	152 (89%)	15 (9%)	4 (2%)	7	35
1	B	171/183 (93%)	152 (89%)	18 (10%)	1 (1%)	28	70
1	C	171/183 (93%)	147 (86%)	22 (13%)	2 (1%)	15	53
1	D	171/183 (93%)	148 (86%)	19 (11%)	4 (2%)	7	35
1	E	172/183 (94%)	148 (86%)	18 (10%)	6 (4%)	4	23
1	F	172/183 (94%)	149 (87%)	19 (11%)	4 (2%)	7	35
1	G	161/183 (88%)	142 (88%)	14 (9%)	5 (3%)	5	26
1	H	170/183 (93%)	147 (86%)	22 (13%)	1 (1%)	28	70
All	All	1359/1464 (93%)	1185 (87%)	147 (11%)	27 (2%)	9	39

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	38	THR
1	A	208	GLU
1	C	208	GLU
1	E	77	ASN
1	E	209	VAL



### 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	140/165 (85%)	133 (95%)	7 (5%)	28	67
1	B	144/165 (87%)	140 (97%)	4 (3%)	49	82
1	C	145/165 (88%)	141 (97%)	4 (3%)	49	82
1	D	137/165 (83%)	130 (95%)	7 (5%)	28	66
1	E	132/165 (80%)	129 (98%)	3 (2%)	56	85
1	F	141/165 (86%)	135 (96%)	6 (4%)	33	72
1	G	129/165 (78%)	125 (97%)	4 (3%)	45	80
1	H	135/165 (82%)	128 (95%)	7 (5%)	27	65
All	All	1103/1320 (84%)	1061 (96%)	42 (4%)	38	75

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	117	ASP
1	E	77	ASN
1	H	134	ARG
1	D	134	ARG
1	D	191	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	158	HIS
1	E	160	HIS
1	G	65	ASN
1	D	65	ASN
1	G	160	HIS

### 5.3.3 RNA ⓘ

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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

58 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$
2	IDS	A	301	3	14,17,17	1.11	1 (7%)	16,26,26	1.63	3 (18%)
3	SGN	A	302	2	18,19,20	1.33	1 (5%)	21,29,31	2.03	4 (19%)
2	IDS	A	303	3	13,16,17	1.35	1 (7%)	15,24,26	3.15	4 (26%)
3	SGN	A	304	2	18,19,20	1.67	2 (11%)	21,29,31	1.99	5 (23%)
2	IDS	A	305	3	13,16,17	1.29	3 (23%)	15,24,26	4.01	1 (6%)
3	SGN	A	306	2	17,18,20	1.82	2 (11%)	18,27,31	1.35	2 (11%)
4	SO4	A	307	-	4,4,4	1.85	1 (25%)	6,6,6	0.89	0
5	EPE	A	308	-	15,15,15	1.11	1 (6%)	18,20,20	1.01	1 (5%)
2	IDS	B	301	3	14,17,17	1.46	1 (7%)	16,26,26	2.93	4 (25%)
3	SGN	B	302	2	18,19,20	1.46	2 (11%)	21,29,31	1.73	4 (19%)
2	IDS	B	303	3	13,16,17	1.44	2 (15%)	15,24,26	2.17	5 (33%)
3	SGN	B	304	2	18,19,20	1.27	1 (5%)	21,29,31	1.98	5 (23%)
2	IDS	B	305	3	13,16,17	1.43	2 (15%)	15,24,26	1.89	4 (26%)
3	SGN	B	306	2	18,19,20	1.62	3 (16%)	21,29,31	1.87	5 (23%)
2	IDS	B	307	3	12,15,17	1.56	4 (33%)	11,22,26	0.89	0
4	SO4	B	308	-	4,4,4	1.83	1 (25%)	6,6,6	0.84	0
2	IDS	B	309	3	14,17,17	1.56	3 (21%)	16,26,26	2.75	4 (25%)
3	SGN	B	310	2	18,19,20	1.51	2 (11%)	21,29,31	1.65	5 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	IDS	B	311	3	13,16,17	1.32	2 (15%)	15,24,26	1.30	2 (13%)
3	SGN	B	312	2	18,19,20	1.08	1 (5%)	21,29,31	1.62	6 (28%)
2	IDS	B	313	3	13,16,17	1.34	2 (15%)	15,24,26	3.04	4 (26%)
3	SGN	B	314	2	18,19,20	1.60	4 (22%)	21,29,31	2.16	4 (19%)
2	IDS	B	315	3	13,16,17	1.43	3 (23%)	15,24,26	4.13	2 (13%)
3	SGN	B	316	2	17,18,20	1.80	3 (17%)	18,27,31	1.37	2 (11%)
5	EPE	B	317	-	15,15,15	1.37	3 (20%)	18,20,20	1.13	2 (11%)
4	SO4	C	301	-	4,4,4	1.67	1 (25%)	6,6,6	1.07	0
3	SGN	C	302	2	19,20,20	1.64	4 (21%)	24,31,31	2.61	8 (33%)
2	IDS	C	303	3	13,16,17	1.51	4 (30%)	15,24,26	2.33	3 (20%)
3	SGN	C	304	2	18,19,20	1.42	2 (11%)	21,29,31	1.99	6 (28%)
2	IDS	C	305	3	13,16,17	1.18	2 (15%)	15,24,26	1.95	3 (20%)
3	SGN	C	306	2	18,19,20	0.92	0	21,29,31	1.52	5 (23%)
2	IDS	C	307	3	13,16,17	1.31	2 (15%)	15,24,26	1.48	2 (13%)
3	SGN	C	308	2	18,19,20	1.34	2 (11%)	21,29,31	2.23	7 (33%)
2	IDS	C	309	3	13,16,17	1.35	2 (15%)	15,24,26	1.19	1 (6%)
3	SGN	C	310	2	17,18,20	1.51	3 (17%)	18,27,31	1.42	2 (11%)
5	EPE	C	311	-	15,15,15	0.96	1 (6%)	18,20,20	1.10	2 (11%)
4	SO4	D	301	-	4,4,4	1.91	2 (50%)	6,6,6	0.84	0
5	EPE	D	302	-	15,15,15	1.23	2 (13%)	18,20,20	0.95	0
4	SO4	E	301	-	4,4,4	1.88	1 (25%)	6,6,6	0.94	0
5	EPE	E	302	-	15,15,15	1.11	1 (6%)	18,20,20	1.15	2 (11%)
3	SGN	F	301	2	19,20,20	1.72	3 (15%)	24,31,31	1.45	2 (8%)
2	IDS	F	302	3	13,16,17	1.37	2 (15%)	15,24,26	1.85	4 (26%)
3	SGN	F	303	2	18,19,20	1.42	2 (11%)	21,29,31	2.01	5 (23%)
2	IDS	F	304	3	13,16,17	1.34	2 (15%)	15,24,26	2.42	4 (26%)
3	SGN	F	305	2	18,19,20	1.54	2 (11%)	21,29,31	2.06	4 (19%)
2	IDS	F	306	3	12,15,17	1.53	3 (25%)	11,22,26	0.86	0
4	SO4	F	307	-	4,4,4	1.90	1 (25%)	6,6,6	0.88	0
4	SO4	F	308	-	4,4,4	1.89	1 (25%)	6,6,6	0.87	0
5	EPE	F	309	-	15,15,15	1.45	3 (20%)	18,20,20	0.81	0
3	SGN	G	301	2	19,20,20	1.59	3 (15%)	24,31,31	2.65	4 (16%)
2	IDS	G	302	3	13,16,17	1.36	1 (7%)	15,24,26	1.78	2 (13%)
3	SGN	G	303	2	18,19,20	1.55	2 (11%)	21,29,31	2.29	5 (23%)
2	IDS	G	304	3	13,16,17	1.41	2 (15%)	15,24,26	3.12	5 (33%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SGN	G	305	2	18,19,20	1.91	3 (16%)	21,29,31	1.92	4 (19%)
2	IDS	G	306	3	12,15,17	1.49	2 (16%)	11,22,26	0.95	1 (9%)
5	EPE	G	307	-	15,15,15	1.26	1 (6%)	18,20,20	0.92	0
4	SO4	H	301	-	4,4,4	1.86	2 (50%)	6,6,6	0.87	0
5	EPE	H	302	-	15,15,15	1.13	1 (6%)	18,20,20	1.00	1 (5%)

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
2	IDS	A	301	3	-	0/5/29/29	0/1/1/1
3	SGN	A	302	2	-	0/10/28/31	0/1/1/1
2	IDS	A	303	3	-	0/5/26/29	0/1/1/1
3	SGN	A	304	2	-	0/10/28/31	0/1/1/1
2	IDS	A	305	3	1/1/6/7	0/5/26/29	0/1/1/1
3	SGN	A	306	2	-	0/10/24/31	0/1/1/1
4	SO4	A	307	-	-	0/0/0/0	0/0/0/0
5	EPE	A	308	-	-	0/9/19/19	0/1/1/1
2	IDS	B	301	3	-	0/5/29/29	0/1/1/1
3	SGN	B	302	2	-	0/10/28/31	0/1/1/1
2	IDS	B	303	3	-	0/5/26/29	0/1/1/1
3	SGN	B	304	2	-	0/10/28/31	0/1/1/1
2	IDS	B	305	3	-	0/5/26/29	0/1/1/1
3	SGN	B	306	2	-	0/10/28/31	0/1/1/1
2	IDS	B	307	3	-	0/5/22/29	1/1/1/1
4	SO4	B	308	-	-	0/0/0/0	0/0/0/0
2	IDS	B	309	3	-	1/5/29/29	0/1/1/1
3	SGN	B	310	2	-	0/10/28/31	0/1/1/1
2	IDS	B	311	3	-	0/5/26/29	0/1/1/1
3	SGN	B	312	2	-	0/10/28/31	0/1/1/1
2	IDS	B	313	3	-	0/5/26/29	0/1/1/1
3	SGN	B	314	2	-	0/10/28/31	0/1/1/1
2	IDS	B	315	3	1/1/6/7	0/5/26/29	0/1/1/1
3	SGN	B	316	2	-	0/10/24/31	0/1/1/1
5	EPE	B	317	-	-	0/9/19/19	0/1/1/1
4	SO4	C	301	-	-	0/0/0/0	0/0/0/0
3	SGN	C	302	2	-	0/10/31/31	0/1/1/1
2	IDS	C	303	3	-	0/5/26/29	0/1/1/1
3	SGN	C	304	2	-	0/10/28/31	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	IDS	C	305	3	-	0/5/26/29	0/1/1/1
3	SGN	C	306	2	-	0/10/28/31	0/1/1/1
2	IDS	C	307	3	-	0/5/26/29	0/1/1/1
3	SGN	C	308	2	-	0/10/28/31	0/1/1/1
2	IDS	C	309	3	-	0/5/26/29	0/1/1/1
3	SGN	C	310	2	-	0/10/24/31	0/1/1/1
5	EPE	C	311	-	-	0/9/19/19	0/1/1/1
4	SO4	D	301	-	-	0/0/0/0	0/0/0/0
5	EPE	D	302	-	-	0/9/19/19	0/1/1/1
4	SO4	E	301	-	-	0/0/0/0	0/0/0/0
5	EPE	E	302	-	-	0/9/19/19	0/1/1/1
3	SGN	F	301	2	-	0/10/31/31	0/1/1/1
2	IDS	F	302	3	-	0/5/26/29	0/1/1/1
3	SGN	F	303	2	-	0/10/28/31	0/1/1/1
2	IDS	F	304	3	-	0/5/26/29	0/1/1/1
3	SGN	F	305	2	-	0/10/28/31	0/1/1/1
2	IDS	F	306	3	-	0/5/22/29	0/1/1/1
4	SO4	F	307	-	-	0/0/0/0	0/0/0/0
4	SO4	F	308	-	-	0/0/0/0	0/0/0/0
5	EPE	F	309	-	-	0/9/19/19	0/1/1/1
3	SGN	G	301	2	-	0/10/31/31	0/1/1/1
2	IDS	G	302	3	-	0/5/26/29	0/1/1/1
3	SGN	G	303	2	-	0/10/28/31	0/1/1/1
2	IDS	G	304	3	-	0/5/26/29	0/1/1/1
3	SGN	G	305	2	-	0/10/28/31	0/1/1/1
2	IDS	G	306	3	-	0/5/22/29	1/1/1/1
5	EPE	G	307	-	-	0/9/19/19	0/1/1/1
4	SO4	H	301	-	-	0/0/0/0	0/0/0/0
5	EPE	H	302	-	-	0/9/19/19	0/1/1/1

The worst 5 of 116 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	302	IDS	O2-C2	-3.50	1.41	1.47
2	B	313	IDS	O2-C2	-3.47	1.41	1.47
2	A	303	IDS	O2-C2	-3.39	1.41	1.47
2	G	304	IDS	O2-C2	-3.31	1.41	1.47
2	F	304	IDS	O2-C2	-3.26	1.42	1.47

The worst 5 of 160 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	301	SGN	O4-C4-C3	-11.77	84.74	110.36
3	C	302	SGN	O4-C4-C3	-10.61	87.27	110.36
2	G	304	IDS	O4-C4-C3	-6.35	96.53	110.36
2	A	303	IDS	O4-C4-C3	-6.31	96.62	110.36
3	B	304	SGN	O4-C4-C3	-6.10	97.07	110.36

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	305	IDS	C4
2	B	315	IDS	C4

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	309	IDS	S-O2-C2-C1

All (2) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	306	IDS	C1-C2-C3-C4-C5-O5
2	B	307	IDS	C1-C2-C3-C4-C5-O5

17 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	308	EPE	1	0
2	B	301	IDS	1	0
3	B	314	SGN	2	0
2	B	315	IDS	1	0
3	B	316	SGN	1	0
5	B	317	EPE	2	0
4	C	301	SO4	2	0
3	C	304	SGN	1	0
2	C	305	IDS	1	0
5	C	311	EPE	3	0
4	D	301	SO4	1	0
5	D	302	EPE	1	0
5	E	302	EPE	1	0
3	F	305	SGN	1	0
4	F	307	SO4	2	0
5	F	309	EPE	1	0
4	H	301	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(Å²)	Q<0.9
1	A	173/183 (94%)	0.02	6 (3%)	44	19	22, 50, 91, 102	0
1	B	173/183 (94%)	-0.23	1 (0%)	89	71	18, 47, 84, 94	0
1	C	173/183 (94%)	-0.36	1 (0%)	89	71	10, 35, 57, 72	0
1	D	173/183 (94%)	-0.16	1 (0%)	89	71	17, 48, 87, 90	0
1	E	174/183 (95%)	-0.03	1 (0%)	89	71	22, 54, 88, 98	0
1	F	174/183 (95%)	-0.12	3 (1%)	70	42	18, 50, 92, 109	0
1	G	165/183 (90%)	0.30	8 (4%)	31	12	43, 69, 95, 103	0
1	H	172/183 (93%)	0.25	13 (7%)	15	6	24, 65, 114, 129	0
All	All	1377/1464 (94%)	-0.04	34 (2%)	58	29	10, 52, 94, 129	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	209	VAL	4.1
1	H	66	THR	3.9
1	F	209	VAL	3.5
1	H	65	ASN	3.1
1	A	57	LEU	3.0

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

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

### 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	EPE	F	309	15/15	0.78	0.41	4.88	102,103,104,105	0
3	SGN	B	302	19/20	0.69	0.42	3.81	138,141,143,143	0
3	SGN	C	304	19/20	0.81	0.30	3.52	94,101,106,106	0
4	SO4	A	307	5/5	0.88	0.27	2.37	126,126,127,127	0
3	SGN	A	302	19/20	0.69	0.34	2.05	120,121,124,124	0
3	SGN	G	305	19/20	0.64	0.44	1.56	125,126,129,129	0
5	EPE	C	311	15/15	0.93	0.26	1.46	46,50,53,53	0
3	SGN	A	304	19/20	0.74	0.46	1.44	123,125,127,127	0
5	EPE	E	302	15/15	0.93	0.25	1.29	64,69,70,70	0
4	SO4	F	308	5/5	0.91	0.26	1.09	133,133,133,134	0
4	SO4	C	301	5/5	0.90	0.23	0.74	108,108,109,110	0
5	EPE	A	308	15/15	0.95	0.20	0.61	35,37,43,44	0
2	IDS	G	304	16/17	0.77	0.31	0.55	123,124,126,127	0
5	EPE	G	307	15/15	0.90	0.28	0.50	73,80,86,86	0
3	SGN	B	310	19/20	0.88	0.21	0.36	95,103,105,105	0
2	IDS	F	306	15/17	0.78	0.30	0.31	127,130,132,132	0
5	EPE	D	302	15/15	0.93	0.20	0.17	62,67,68,68	0
4	SO4	D	301	5/5	0.94	0.21	0.04	102,102,102,103	0
2	IDS	A	303	16/17	0.71	0.29	-0.05	118,121,122,122	0
5	EPE	B	317	15/15	0.95	0.18	-0.34	35,46,53,53	0
4	SO4	E	301	5/5	0.94	0.17	-0.43	94,94,95,95	0
4	SO4	B	308	5/5	0.95	0.15	-0.57	101,101,101,101	0
4	SO4	F	307	5/5	0.93	0.17	-0.63	111,112,112,112	0
4	SO4	H	301	5/5	0.95	0.14	-0.70	82,83,83,83	0
5	EPE	H	302	15/15	0.94	0.16	-0.73	54,55,57,58	0
3	SGN	C	308	19/20	0.92	0.18	-0.75	75,79,81,89	0
3	SGN	B	312	19/20	0.90	0.22	-0.82	70,79,82,89	0
2	IDS	B	313	16/17	0.88	0.18	-1.10	93,97,102,102	0
2	IDS	B	305	16/17	0.76	0.45	-	139,140,141,144	0
2	IDS	C	309	16/17	0.91	0.17	-	90,95,101,114	0
2	IDS	C	307	16/17	0.94	0.14	-	76,78,82,82	0
2	IDS	B	301	17/17	0.61	0.47	-	142,145,146,146	0
3	SGN	B	316	18/20	0.77	0.34	-	124,127,129,129	0
2	IDS	F	302	16/17	0.62	0.45	-	134,140,143,143	0
3	SGN	B	314	19/20	0.84	0.20	-	97,103,107,109	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	IDS	G	302	16/17	0.72	0.39	-	122,129,131,131	0
2	IDS	A	301	17/17	0.74	0.49	-	121,122,123,123	0
3	SGN	C	310	18/20	0.67	0.38	-	118,126,130,130	0
2	IDS	G	306	15/17	0.75	0.44	-	125,125,126,127	0
2	IDS	A	305	16/17	0.82	0.36	-	125,127,130,131	0
2	IDS	C	305	16/17	0.83	0.19	-	71,87,89,91	0
3	SGN	B	306	19/20	0.69	0.42	-	143,145,146,147	0
3	SGN	C	302	20/20	0.71	0.22	-	105,112,122,122	0
3	SGN	B	304	19/20	0.79	0.25	-	134,137,138,138	0
2	IDS	F	304	16/17	0.85	0.26	-	134,135,136,137	0
3	SGN	A	306	18/20	0.62	0.43	-	132,134,142,142	0
3	SGN	G	303	19/20	0.84	0.18	-	119,120,121,122	0
2	IDS	B	307	15/17	0.77	0.48	-	144,144,145,145	0
2	IDS	C	303	16/17	0.80	0.27	-	103,106,106,106	0
2	IDS	B	303	16/17	0.57	0.42	-	138,140,142,142	0
3	SGN	F	301	20/20	0.56	0.50	-	143,150,155,155	0
2	IDS	B	315	16/17	0.83	0.28	-	106,112,115,123	0
3	SGN	F	303	19/20	0.81	0.24	-	127,131,132,133	0
2	IDS	B	311	16/17	0.89	0.26	-	82,90,92,93	0
3	SGN	G	301	20/20	0.65	0.50	-	131,137,140,140	0
2	IDS	B	309	17/17	0.75	0.34	-	106,114,118,118	0
3	SGN	F	305	19/20	0.76	0.32	-	132,135,137,137	0
3	SGN	C	306	19/20	0.89	0.20	-	55,60,67,74	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.