



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 03:46 PM GMT

PDB ID : 4CQM  
Title : Crystal structure of heterotetrameric human ketoacyl reductase complexed with NAD and NADP  
Authors : Venkatesan, R.; SahTeli, S.K.; Awoniyi, L.O.; Jiang, G.; Prus, P.; Kastoniatis, A.J.; Hiltunen, J.K.; Wierenga, R.K.; Chen, Z.  
Deposited on : 2014-02-19  
Resolution : 2.34 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

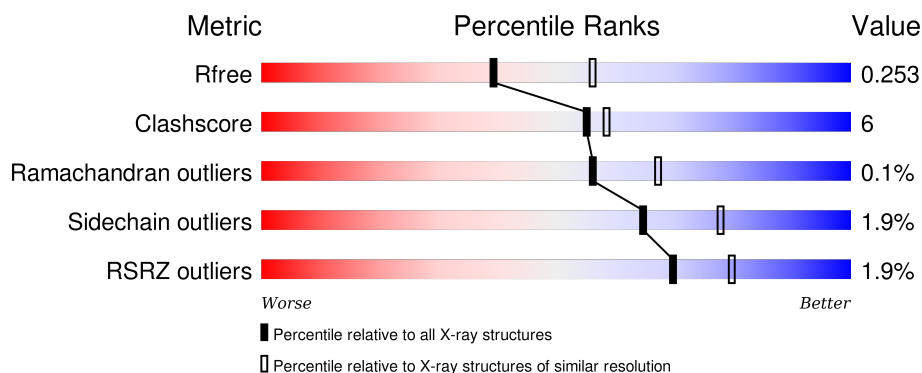
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.34 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1406 (2.36-2.32)
Clashscore	102246	1509 (2.36-2.32)
Ramachandran outliers	100387	1490 (2.36-2.32)
Sidechain outliers	100360	1491 (2.36-2.32)
RSRZ outliers	91569	1412 (2.36-2.32)

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	261	<div> <div></div> <div>84% 11% 5%</div> </div>
1	E	261	<div> <div></div> <div>84% 11% 5%</div> </div>
1	H	261	<div> <div></div> <div>81% 14% 5%</div> </div>
1	I	261	<div> <div></div> <div>83% 12% 5%</div> </div>
1	L	261	<div> <div></div> <div>80% 14% 6%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	261	
1	P	261	
2	B	244	
2	C	244	
2	F	244	
2	G	244	
2	J	244	
2	K	244	
2	N	244	
2	O	244	
3	D	261	

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
4	NAP	N	301	-	-	-	X
5	GOL	E	401	-	-	-	X
5	GOL	M	401	-	-	-	X
6	EDO	D	401	-	-	-	X
6	EDO	M	402	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 29356 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 ESTRADIOL 17-BETA-DEHYDROGENASE 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	248	Total	C	N	O	S	0	0	0
			1783	1106	319	347	11			
1	E	248	Total	C	N	O	S	0	1	0
			1775	1102	315	346	12			
1	H	249	Total	C	N	O	S	0	1	0
			1785	1106	322	346	11			
1	I	249	Total	C	N	O	S	0	1	0
			1802	1121	322	348	11			
1	L	246	Total	C	N	O	S	0	1	0
			1786	1109	322	344	11			
1	M	249	Total	C	N	O	S	0	2	0
			1819	1126	330	351	12			
1	P	248	Total	C	N	O	S	0	1	0
			1785	1110	317	347	11			

- Molecule 2 is a protein called CARBONYL REDUCTASE FAMILY MEMBER 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	232	Total	C	N	O	S	0	0	0
			1680	1053	315	303	9			
2	C	232	Total	C	N	O	S	0	1	0
			1667	1051	300	308	8			
2	F	241	Total	C	N	O	S	0	1	0
			1767	1107	327	323	10			
2	G	224	Total	C	N	O	S	0	0	0
			1628	1022	298	300	8			
2	J	240	Total	C	N	O	S	0	0	0
			1727	1085	318	314	10			
2	K	231	Total	C	N	O	S	0	0	0
			1649	1036	305	300	8			
2	N	232	Total	C	N	O	S	0	3	0
			1727	1083	327	308	9			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	O	235	Total	C	N	O	S	0	0	0
			1640	1030	298	304	8			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	MET	-	EXPRESSION TAG	UNP Q8N4T8
B	-5	HIS	-	EXPRESSION TAG	UNP Q8N4T8
B	-4	HIS	-	EXPRESSION TAG	UNP Q8N4T8
B	-3	HIS	-	EXPRESSION TAG	UNP Q8N4T8
B	-2	HIS	-	EXPRESSION TAG	UNP Q8N4T8
B	-1	HIS	-	EXPRESSION TAG	UNP Q8N4T8
B	0	HIS	-	EXPRESSION TAG	UNP Q8N4T8
C	-6	MET	-	EXPRESSION TAG	UNP Q8N4T8
C	-5	HIS	-	EXPRESSION TAG	UNP Q8N4T8
C	-4	HIS	-	EXPRESSION TAG	UNP Q8N4T8
C	-3	HIS	-	EXPRESSION TAG	UNP Q8N4T8
C	-2	HIS	-	EXPRESSION TAG	UNP Q8N4T8
C	-1	HIS	-	EXPRESSION TAG	UNP Q8N4T8
C	0	HIS	-	EXPRESSION TAG	UNP Q8N4T8
F	-6	MET	-	EXPRESSION TAG	UNP Q8N4T8
F	-5	HIS	-	EXPRESSION TAG	UNP Q8N4T8
F	-4	HIS	-	EXPRESSION TAG	UNP Q8N4T8
F	-3	HIS	-	EXPRESSION TAG	UNP Q8N4T8
F	-2	HIS	-	EXPRESSION TAG	UNP Q8N4T8
F	-1	HIS	-	EXPRESSION TAG	UNP Q8N4T8
F	0	HIS	-	EXPRESSION TAG	UNP Q8N4T8
G	-6	MET	-	EXPRESSION TAG	UNP Q8N4T8
G	-5	HIS	-	EXPRESSION TAG	UNP Q8N4T8
G	-4	HIS	-	EXPRESSION TAG	UNP Q8N4T8
G	-3	HIS	-	EXPRESSION TAG	UNP Q8N4T8
G	-2	HIS	-	EXPRESSION TAG	UNP Q8N4T8
G	-1	HIS	-	EXPRESSION TAG	UNP Q8N4T8
G	0	HIS	-	EXPRESSION TAG	UNP Q8N4T8
J	-6	MET	-	EXPRESSION TAG	UNP Q8N4T8
J	-5	HIS	-	EXPRESSION TAG	UNP Q8N4T8
J	-4	HIS	-	EXPRESSION TAG	UNP Q8N4T8
J	-3	HIS	-	EXPRESSION TAG	UNP Q8N4T8
J	-2	HIS	-	EXPRESSION TAG	UNP Q8N4T8
J	-1	HIS	-	EXPRESSION TAG	UNP Q8N4T8
J	0	HIS	-	EXPRESSION TAG	UNP Q8N4T8
K	-6	MET	-	EXPRESSION TAG	UNP Q8N4T8

*Continued on next page...*

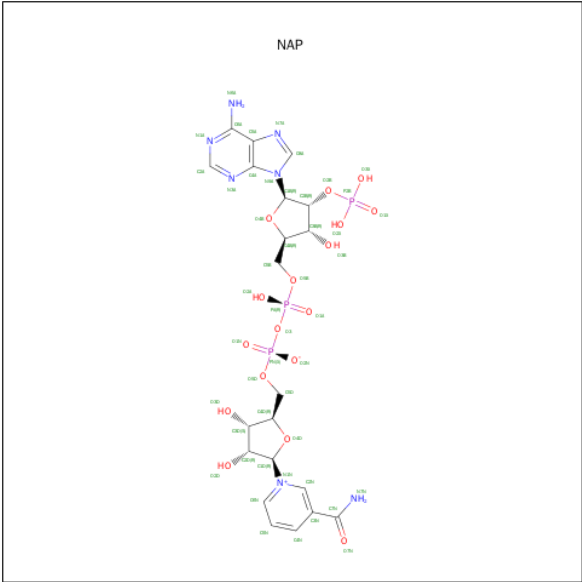
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
K	-5	HIS	-	EXPRESSION TAG	UNP Q8N4T8
K	-4	HIS	-	EXPRESSION TAG	UNP Q8N4T8
K	-3	HIS	-	EXPRESSION TAG	UNP Q8N4T8
K	-2	HIS	-	EXPRESSION TAG	UNP Q8N4T8
K	-1	HIS	-	EXPRESSION TAG	UNP Q8N4T8
K	0	HIS	-	EXPRESSION TAG	UNP Q8N4T8
N	-6	MET	-	EXPRESSION TAG	UNP Q8N4T8
N	-5	HIS	-	EXPRESSION TAG	UNP Q8N4T8
N	-4	HIS	-	EXPRESSION TAG	UNP Q8N4T8
N	-3	HIS	-	EXPRESSION TAG	UNP Q8N4T8
N	-2	HIS	-	EXPRESSION TAG	UNP Q8N4T8
N	-1	HIS	-	EXPRESSION TAG	UNP Q8N4T8
N	0	HIS	-	EXPRESSION TAG	UNP Q8N4T8
O	-6	MET	-	EXPRESSION TAG	UNP Q8N4T8
O	-5	HIS	-	EXPRESSION TAG	UNP Q8N4T8
O	-4	HIS	-	EXPRESSION TAG	UNP Q8N4T8
O	-3	HIS	-	EXPRESSION TAG	UNP Q8N4T8
O	-2	HIS	-	EXPRESSION TAG	UNP Q8N4T8
O	-1	HIS	-	EXPRESSION TAG	UNP Q8N4T8
O	0	HIS	-	EXPRESSION TAG	UNP Q8N4T8

- Molecule 3 is a protein called ESTRADIOL 17-BETA-DEHYDROGENASE 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	248	Total	C	N	O	S	0	1	0
			1772	1101	310	350	11			

- Molecule 4 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: C<sub>21</sub>H<sub>28</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



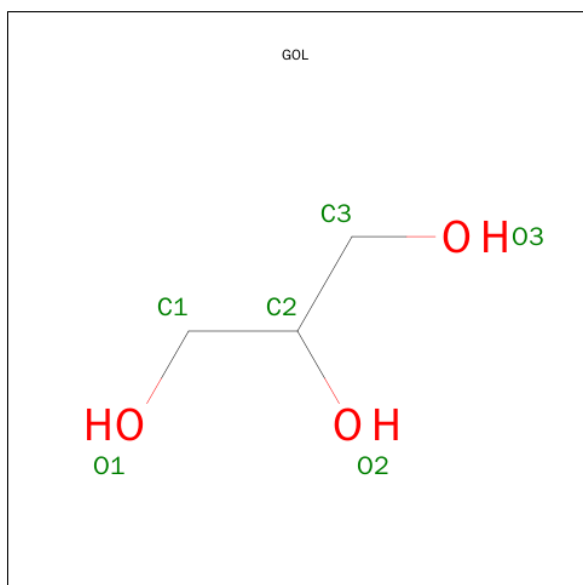
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
4	C	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
4	D	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	E	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	F	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
4	G	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
4	H	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	I	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	J	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
4	K	1	Total	C	N	O	P	0	0
			39	15	5	16	3		
4	L	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	M	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	N	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	O	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
4	P	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	I	1	Total	C	O	0	0
			6	3	3		
5	M	1	Total	C	O	0	0
			6	3	3		

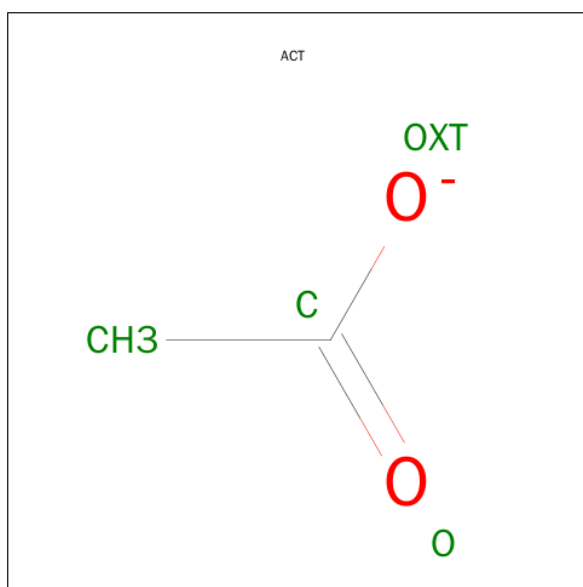
- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total	C	O	0	0
			4	2	2		
6	E	1	Total	C	O	0	0
			4	2	2		
6	F	1	Total	C	O	0	0
			4	2	2		
6	I	1	Total	C	O	0	0
			4	2	2		
6	I	1	Total	C	O	0	0
			4	2	2		
6	L	1	Total	C	O	0	0
			4	2	2		
6	M	1	Total	C	O	0	0
			4	2	2		
6	P	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	H	1	Total	C	O	0	0
			4	2	2		
7	M	1	Total	C	O	0	0
			4	2	2		
7	P	1	Total	C	O	0	0
			4	2	2		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	63	Total	O	0	0
			63	63		
8	B	36	Total	O	0	0
			36	36		
8	C	46	Total	O	0	0
			46	46		
8	D	40	Total	O	0	0
			40	40		
8	E	66	Total	O	0	0
			66	66		
8	F	49	Total	O	0	0
			49	49		
8	G	24	Total	O	0	0
			24	24		
8	H	36	Total	O	0	0
			36	36		
8	I	60	Total	O	0	0
			60	60		

*Continued on next page...*

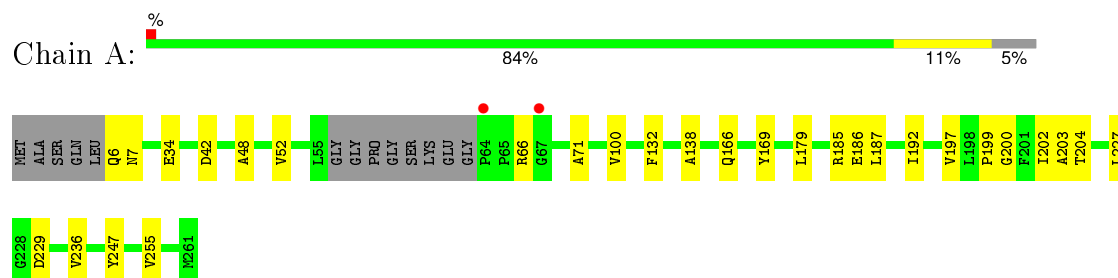
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	J	48	Total 48	O 48	0	0
8	K	47	Total 47	O 47	0	0
8	L	56	Total 56	O 56	0	0
8	M	93	Total 93	O 93	0	0
8	N	52	Total 52	O 52	0	0
8	O	38	Total 38	O 38	0	0
8	P	52	Total 52	O 52	0	0

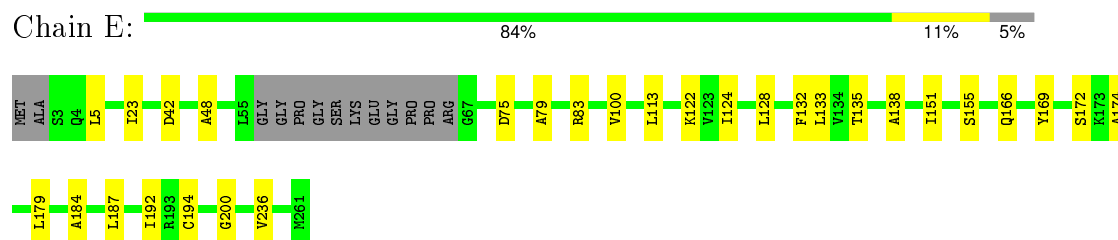
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

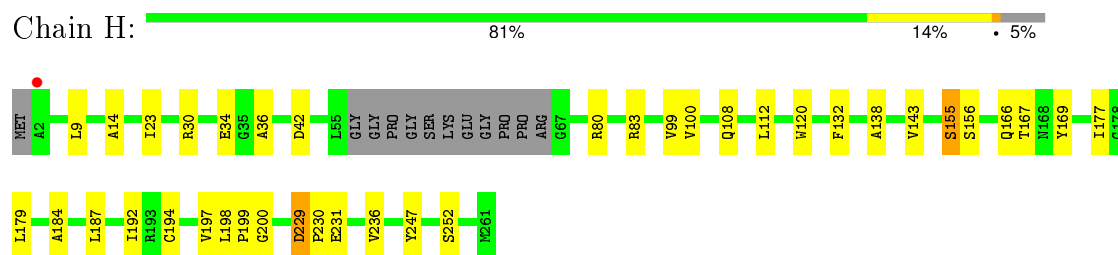
- Molecule 1: ESTRADIOL 17-BETA-DEHYDROGENASE 8



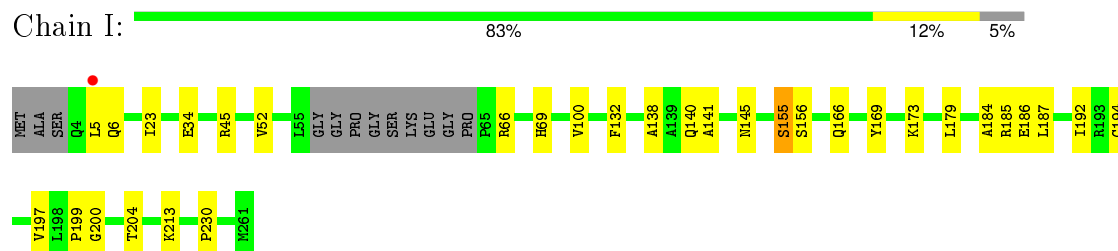
- Molecule 1: ESTRADIOL 17-BETA-DEHYDROGENASE 8




- Molecule 1: ESTRADIOL 17-BETA-DEHYDROGENASE 8

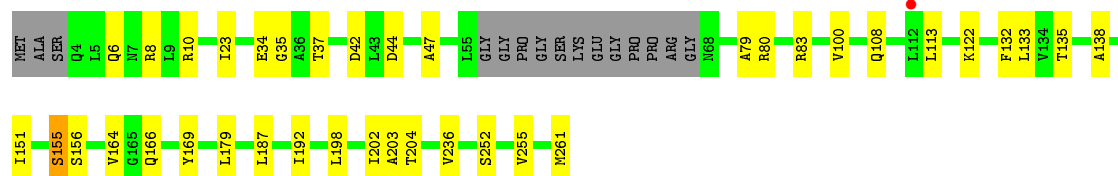


- Molecule 1: ESTRADIOL 17-BETA-DEHYDROGENASE 8




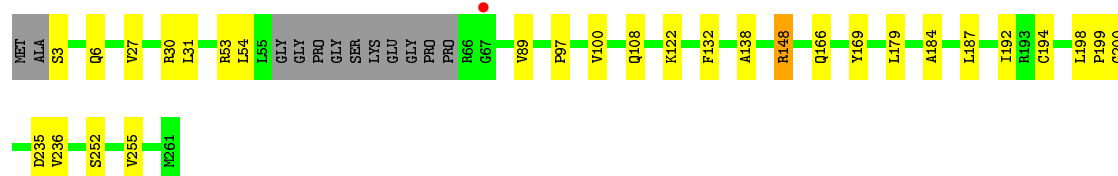
- Molecule 1: ESTRADIOL 17-BETA-DEHYDROGENASE 8

Chain L: 




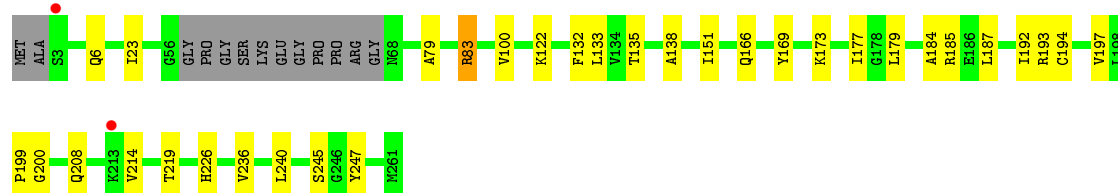
- Molecule 1: ESTRADIOL 17-BETA-DEHYDROGENASE 8

Chain M: 




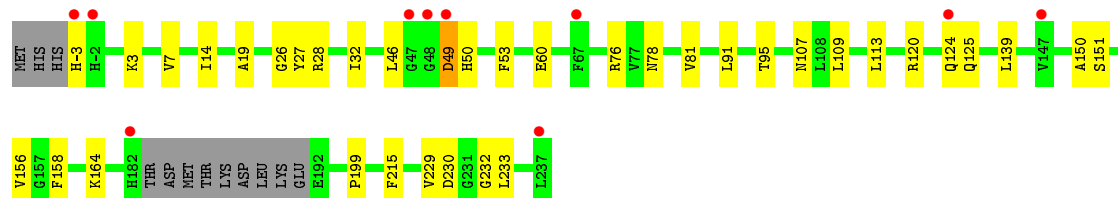
- Molecule 1: ESTRADIOL 17-BETA-DEHYDROGENASE 8

Chain P: 




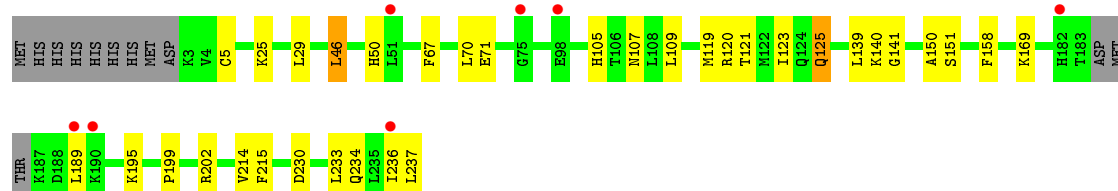
- Molecule 2: CARBONYL REDUCTASE FAMILY MEMBER 4

Chain B: 

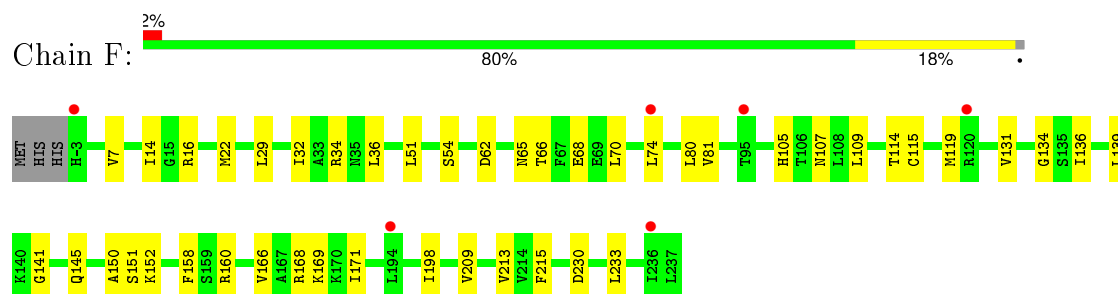


- Molecule 2: CARBONYL REDUCTASE FAMILY MEMBER 4

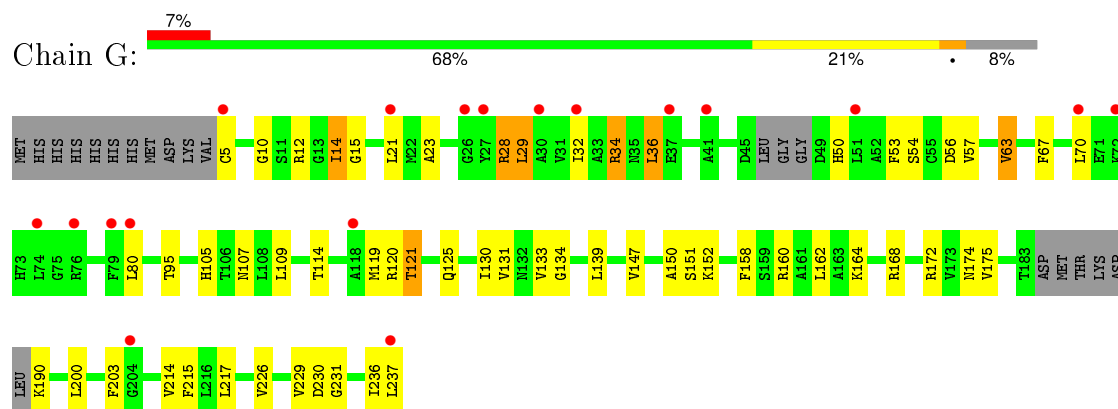
Chain C: 



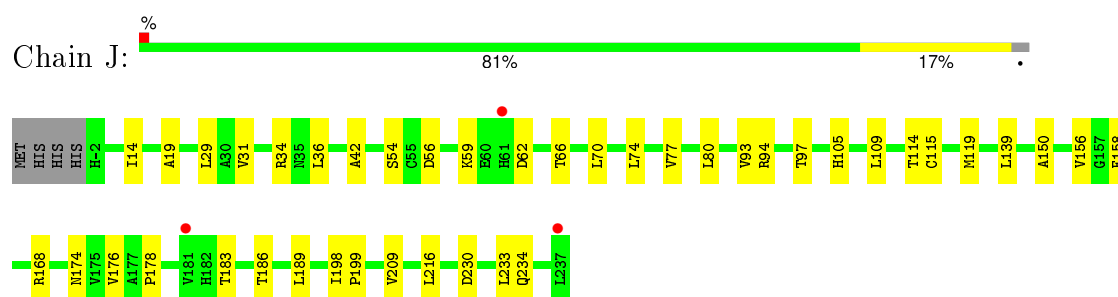
- Molecule 2: CARBONYL REDUCTASE FAMILY MEMBER 4



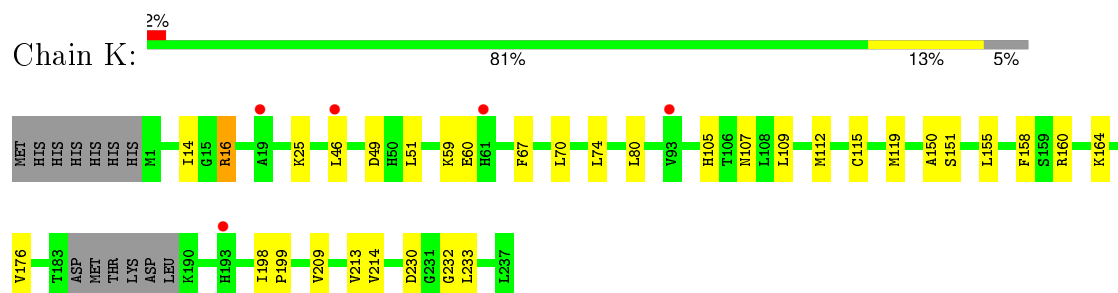
- Molecule 2: CARBONYL REDUCTASE FAMILY MEMBER 4



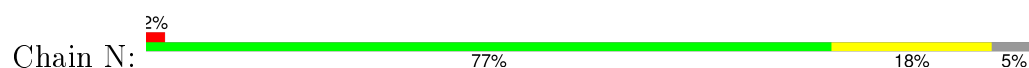
- Molecule 2: CARBONYL REDUCTASE FAMILY MEMBER 4

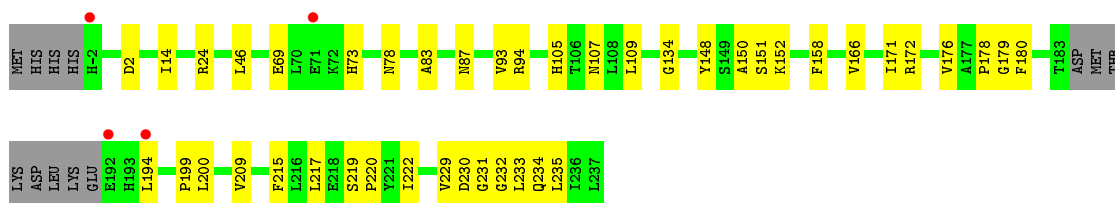


- Molecule 2: CARBONYL REDUCTASE FAMILY MEMBER 4

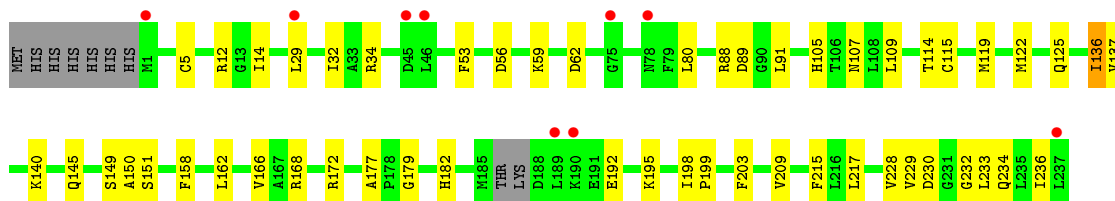
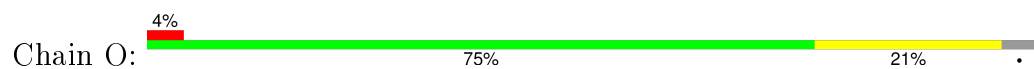


- Molecule 2: CARBONYL REDUCTASE FAMILY MEMBER 4

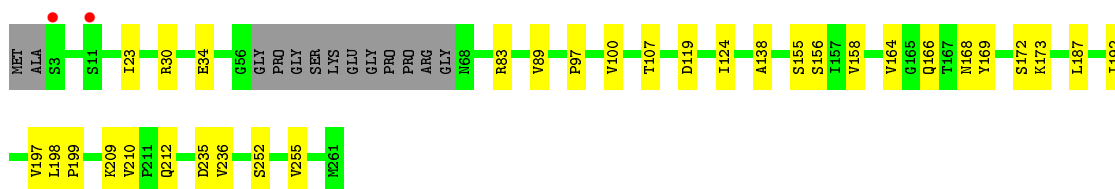
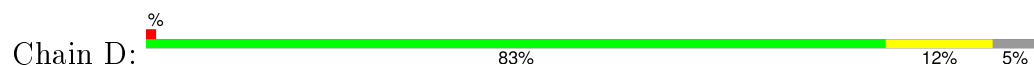




• Molecule 2: CARBONYL REDUCTASE FAMILY MEMBER 4



• Molecule 3: ESTRADIOL 17-BETA-DEHYDROGENASE 8



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.50Å 238.41Å 87.69Å 90.00° 94.18° 90.00°	Depositor
Resolution (Å)	49.25 – 2.34 49.25 – 2.34	Depositor EDS
% Data completeness (in resolution range)	98.6 (49.25-2.34) 90.6 (49.25-2.34)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 2.34Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.203 , 0.235 0.219 , 0.253	Depositor DCC
$R_{free}$ test set	6812 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.0	Xtriage
Anisotropy	0.471	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 38.8	EDS
Estimated twinning fraction	0.120 for L,-K,H 0.116 for l,-k,h	Xtriage
Reported twinning fraction	0.120 for L,-K,H	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	2 of 148308 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	29356	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.08% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NAP, OAS, EDO, ACT

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	0.20	0/1803	0.37	0/2446
1	E	0.20	0/1798	0.38	0/2442
1	H	0.20	0/1808	0.37	0/2456
1	I	0.22	0/1826	0.39	0/2476
1	L	0.20	0/1809	0.36	0/2454
1	M	0.21	0/1845	0.38	0/2500
1	P	0.21	0/1808	0.37	0/2454
2	B	0.20	0/1704	0.36	0/2302
2	C	0.21	0/1689	0.37	0/2281
2	F	0.21	0/1794	0.37	0/2421
2	G	0.22	0/1645	0.41	0/2218
2	J	0.20	0/1749	0.37	0/2364
2	K	0.20	0/1668	0.37	0/2251
2	N	0.20	0/1759	0.37	0/2369
2	O	0.22	0/1658	0.37	0/2246
3	D	0.20	0/1784	0.37	0/2421
All	All	0.21	0/28147	0.37	0/38101

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

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	1783	0	1780	16	0
1	E	1775	0	1770	18	0
1	H	1785	0	1771	25	0
1	I	1802	0	1811	20	0
1	L	1786	0	1802	22	0
1	M	1819	0	1839	18	0
1	P	1785	0	1784	24	0
2	B	1680	0	1663	21	0
2	C	1667	0	1666	22	0
2	F	1767	0	1768	27	0
2	G	1628	0	1634	38	0
2	J	1727	0	1725	26	0
2	K	1649	0	1643	20	0
2	N	1727	0	1757	29	0
2	O	1640	0	1591	38	0
3	D	1772	0	1751	22	0
4	A	44	0	25	1	0
4	B	48	0	25	3	0
4	C	32	0	11	0	0
4	D	44	0	25	2	0
4	E	44	0	25	4	0
4	F	48	0	25	2	0
4	G	27	0	11	3	0
4	H	44	0	25	4	0
4	I	44	0	25	3	0
4	J	48	0	25	3	0
4	K	39	0	18	0	0
4	L	44	0	25	2	0
4	M	44	0	25	2	0
4	N	48	0	25	4	0
4	O	48	0	25	4	0
4	P	44	0	25	3	0
5	A	6	0	8	0	0
5	E	6	0	8	2	0
5	I	6	0	8	0	0
5	M	6	0	8	1	0
6	D	4	0	6	1	0
6	E	4	0	6	0	0
6	F	4	0	6	0	0
6	I	8	0	12	0	0
6	L	4	0	6	0	0
6	M	4	0	6	0	0
6	P	4	0	6	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	H	4	0	3	0	0
7	M	4	0	3	0	0
7	P	4	0	3	0	0
8	A	63	0	0	2	0
8	B	36	0	0	2	0
8	C	46	0	0	2	0
8	D	40	0	0	0	0
8	E	66	0	0	0	0
8	F	49	0	0	2	0
8	G	24	0	0	1	0
8	H	36	0	0	0	0
8	I	60	0	0	3	0
8	J	48	0	0	0	0
8	K	47	0	0	2	0
8	L	56	0	0	2	0
8	M	93	0	0	3	0
8	N	52	0	0	2	0
8	O	38	0	0	3	0
8	P	52	0	0	1	0
All	All	29356	0	28209	358	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 358 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:185:ARG:HD3	3:D:164:VAL:HG23	1.67	0.77
2:N:150:ALA:HB2	2:O:158:PHE:HB2	1.67	0.75
1:L:80:ARG:HA	1:L:83:ARG:HE	1.56	0.71
2:G:32:ILE:HG12	2:G:53:PHE:HB2	1.72	0.71
2:K:105:HIS:HA	2:K:109:LEU:HB3	1.73	0.70

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	244/261 (94%)	236 (97%)	7 (3%)	1 (0%)	39	45
1	E	245/261 (94%)	238 (97%)	7 (3%)	0	100	100
1	H	246/261 (94%)	237 (96%)	8 (3%)	1 (0%)	39	45
1	I	246/261 (94%)	238 (97%)	7 (3%)	1 (0%)	39	45
1	L	243/261 (93%)	236 (97%)	5 (2%)	2 (1%)	24	25
1	M	247/261 (95%)	239 (97%)	8 (3%)	0	100	100
1	P	245/261 (94%)	231 (94%)	14 (6%)	0	100	100
2	B	228/244 (93%)	217 (95%)	11 (5%)	0	100	100
2	C	229/244 (94%)	223 (97%)	6 (3%)	0	100	100
2	F	240/244 (98%)	228 (95%)	12 (5%)	0	100	100
2	G	218/244 (89%)	211 (97%)	7 (3%)	0	100	100
2	J	238/244 (98%)	227 (95%)	11 (5%)	0	100	100
2	K	227/244 (93%)	218 (96%)	9 (4%)	0	100	100
2	N	231/244 (95%)	222 (96%)	9 (4%)	0	100	100
2	O	231/244 (95%)	224 (97%)	7 (3%)	0	100	100
3	D	244/261 (94%)	237 (97%)	7 (3%)	0	100	100
All	All	3802/4040 (94%)	3662 (96%)	135 (4%)	5 (0%)	56	67

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	203	ALA
1	L	203	ALA
1	H	155	SER
1	I	155	SER
1	L	155	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	184/198 (93%)	181 (98%)	3 (2%)	70	82
1	E	184/198 (93%)	182 (99%)	2 (1%)	80	89
1	H	182/198 (92%)	179 (98%)	3 (2%)	70	82
1	I	186/198 (94%)	182 (98%)	4 (2%)	60	73
1	L	187/198 (94%)	185 (99%)	2 (1%)	80	89
1	M	192/198 (97%)	191 (100%)	1 (0%)	92	96
1	P	184/198 (93%)	181 (98%)	3 (2%)	70	82
2	B	166/193 (86%)	162 (98%)	4 (2%)	57	69
2	C	165/193 (86%)	160 (97%)	5 (3%)	48	61
2	F	178/193 (92%)	172 (97%)	6 (3%)	44	56
2	G	163/193 (84%)	152 (93%)	11 (7%)	20	22
2	J	171/193 (89%)	170 (99%)	1 (1%)	90	96
2	K	160/193 (83%)	156 (98%)	4 (2%)	55	68
2	N	176/193 (91%)	175 (99%)	1 (1%)	90	96
2	O	154/193 (80%)	153 (99%)	1 (1%)	90	96
3	D	181/197 (92%)	179 (99%)	2 (1%)	80	89
All	All	2813/3127 (90%)	2760 (98%)	53 (2%)	65	78

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	G	21	LEU
2	G	70	LEU
2	N	235	LEU
2	G	28	ARG
2	G	34	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	G	50	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	OAS	D	156	3	7,8,9	1.17	1 (14%)	6,9,11	1.52	2 (33%)

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	OAS	D	156	3	-	0/5/7/9	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	156	OAS	OG-C1A	2.42	1.46	1.33

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	156	OAS	O-C-CA	-2.47	119.05	125.49
3	D	156	OAS	CB-OG-C1A	2.41	123.25	117.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	156	OAS	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

31 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	NAP	A	301	-	38,48,52	0.83	1 (2%)	47,73,80	1.55	5 (10%)
5	GOL	A	401	-	5,5,5	0.22	0	5,5,5	0.23	0
4	NAP	B	301	-	42,52,52	0.83	1 (2%)	54,80,80	1.44	3 (5%)
4	NAP	C	301	-	27,34,52	0.93	1 (3%)	35,53,80	1.70	3 (8%)
4	NAP	D	301	-	38,48,52	0.81	1 (2%)	47,73,80	1.54	5 (10%)
6	EDO	D	401	-	3,3,3	0.47	0	2,2,2	0.44	0
4	NAP	E	301	-	38,48,52	0.82	1 (2%)	47,73,80	1.55	4 (8%)
5	GOL	E	401	-	5,5,5	0.23	0	5,5,5	0.21	0
6	EDO	E	402	-	3,3,3	0.48	0	2,2,2	0.40	0
4	NAP	F	301	-	42,52,52	0.81	1 (2%)	54,80,80	1.47	3 (5%)
6	EDO	F	401	-	3,3,3	0.47	0	2,2,2	0.42	0
4	NAP	G	301	-	24,29,52	1.01	1 (4%)	29,45,80	1.78	3 (10%)
4	NAP	H	301	-	38,48,52	0.82	1 (2%)	47,73,80	1.52	4 (8%)
7	ACT	H	501	-	1,3,3	1.19	0	0,3,3	0.00	-
4	NAP	I	301	-	38,48,52	0.83	1 (2%)	47,73,80	1.54	4 (8%)
5	GOL	I	401	-	5,5,5	0.22	0	5,5,5	0.23	0
6	EDO	I	402	-	3,3,3	0.47	0	2,2,2	0.42	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	EDO	I	403	-	3,3,3	0.47	0	2,2,2	0.42	0
4	NAP	J	301	-	42,52,52	0.82	1 (2%)	54,80,80	1.51	3 (5%)
4	NAP	K	301	-	35,42,52	0.85	1 (2%)	46,65,80	1.61	4 (8%)
4	NAP	L	301	-	38,48,52	0.83	1 (2%)	47,73,80	1.54	4 (8%)
6	EDO	L	401	-	3,3,3	0.49	0	2,2,2	0.38	0
4	NAP	M	301	-	38,48,52	0.83	1 (2%)	47,73,80	1.55	4 (8%)
5	GOL	M	401	-	5,5,5	0.24	0	5,5,5	0.28	0
6	EDO	M	402	-	3,3,3	0.47	0	2,2,2	0.41	0
7	ACT	M	501	-	1,3,3	1.09	0	0,3,3	0.00	-
4	NAP	N	301	-	42,52,52	0.82	1 (2%)	54,80,80	1.45	3 (5%)
4	NAP	O	301	-	42,52,52	0.82	1 (2%)	54,80,80	1.47	3 (5%)
4	NAP	P	301	-	38,48,52	0.83	1 (2%)	47,73,80	1.52	4 (8%)
6	EDO	P	401	-	3,3,3	0.47	0	2,2,2	0.42	0
7	ACT	P	501	-	1,3,3	1.20	0	0,3,3	0.00	-

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	NAP	A	301	-	-	0/22/62/67	0/5/5/5
5	GOL	A	401	-	-	0/4/4/4	0/0/0/0
4	NAP	B	301	-	-	0/27/67/67	0/5/5/5
4	NAP	C	301	-	-	0/20/40/67	0/3/3/5
4	NAP	D	301	-	-	0/22/62/67	0/5/5/5
6	EDO	D	401	-	-	0/1/1/1	0/0/0/0
4	NAP	E	301	-	-	0/22/62/67	0/5/5/5
5	GOL	E	401	-	-	0/4/4/4	0/0/0/0
6	EDO	E	402	-	-	0/1/1/1	0/0/0/0
4	NAP	F	301	-	-	0/27/67/67	0/5/5/5
6	EDO	F	401	-	-	0/1/1/1	0/0/0/0
4	NAP	G	301	-	-	0/11/31/67	0/3/3/5
4	NAP	H	301	-	-	0/22/62/67	0/5/5/5
7	ACT	H	501	-	-	0/0/0/0	0/0/0/0
4	NAP	I	301	-	-	0/22/62/67	0/5/5/5
5	GOL	I	401	-	-	0/4/4/4	0/0/0/0
6	EDO	I	402	-	-	0/1/1/1	0/0/0/0
6	EDO	I	403	-	-	0/1/1/1	0/0/0/0
4	NAP	J	301	-	-	0/27/67/67	0/5/5/5
4	NAP	K	301	-	-	0/23/56/67	0/4/4/5

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAP	L	301	-	-	0/22/62/67	0/5/5/5
6	EDO	L	401	-	-	0/1/1/1	0/0/0/0
4	NAP	M	301	-	-	0/22/62/67	0/5/5/5
5	GOL	M	401	-	-	0/4/4/4	0/0/0/0
6	EDO	M	402	-	-	0/1/1/1	0/0/0/0
7	ACT	M	501	-	-	0/0/0/0	0/0/0/0
4	NAP	N	301	-	-	0/27/67/67	0/5/5/5
4	NAP	O	301	-	-	0/27/67/67	0/5/5/5
4	NAP	P	301	-	-	0/22/62/67	0/5/5/5
6	EDO	P	401	-	-	0/1/1/1	0/0/0/0
7	ACT	P	501	-	-	0/0/0/0	0/0/0/0

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	P	301	NAP	C5A-C4A	3.04	1.47	1.40
4	D	301	NAP	C5A-C4A	3.05	1.47	1.40
4	F	301	NAP	C5A-C4A	3.06	1.47	1.40
4	M	301	NAP	C5A-C4A	3.06	1.47	1.40
4	G	301	NAP	C5A-C4A	3.06	1.47	1.40

The worst 5 of 59 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	301	NAP	N3A-C2A-N1A	-7.01	123.53	128.89
4	O	301	NAP	N3A-C2A-N1A	-6.90	123.61	128.89
4	H	301	NAP	N3A-C2A-N1A	-6.87	123.64	128.89
4	A	301	NAP	N3A-C2A-N1A	-6.87	123.64	128.89
4	K	301	NAP	N3A-C2A-N1A	-6.85	123.65	128.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

18 monomers are involved in 46 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	301	NAP	1	0
4	B	301	NAP	3	0
4	D	301	NAP	2	0
6	D	401	EDO	1	0
4	E	301	NAP	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	401	GOL	2	0
4	F	301	NAP	2	0
4	G	301	NAP	3	0
4	H	301	NAP	4	0
4	I	301	NAP	3	0
4	J	301	NAP	3	0
4	L	301	NAP	2	0
4	M	301	NAP	2	0
5	M	401	GOL	1	0
4	N	301	NAP	4	0
4	O	301	NAP	4	0
4	P	301	NAP	3	0
6	P	401	EDO	2	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	248/261 (95%)	-0.06	2 (0%) 87 92	28, 44, 71, 124	0
1	E	248/261 (95%)	-0.18	0 100 100	28, 41, 64, 90	0
1	H	249/261 (95%)	-0.04	1 (0%) 93 97	33, 52, 78, 90	0
1	I	249/261 (95%)	-0.11	1 (0%) 93 97	27, 44, 72, 111	0
1	L	246/261 (94%)	-0.18	1 (0%) 93 97	29, 44, 69, 92	0
1	M	249/261 (95%)	-0.17	1 (0%) 93 97	26, 40, 64, 103	0
1	P	248/261 (95%)	-0.03	2 (0%) 87 92	31, 47, 72, 110	0
2	B	232/244 (95%)	0.04	10 (4%) 39 51	30, 53, 87, 119	0
2	C	232/244 (95%)	0.19	7 (3%) 54 64	34, 54, 90, 130	0
2	F	241/244 (98%)	0.10	6 (2%) 61 72	30, 50, 88, 135	0
2	G	224/244 (91%)	0.60	18 (8%) 15 23	45, 73, 105, 123	0
2	J	240/244 (98%)	0.05	3 (1%) 79 86	27, 51, 85, 118	0
2	K	231/244 (94%)	0.05	5 (2%) 65 76	29, 49, 84, 127	0
2	N	232/244 (95%)	-0.03	4 (1%) 73 82	31, 45, 77, 102	0
2	O	235/244 (96%)	0.32	9 (3%) 44 56	37, 62, 102, 136	0
3	D	247/261 (94%)	-0.09	2 (0%) 87 92	31, 48, 74, 114	0
All	All	3851/4040 (95%)	0.02	72 (1%) 70 80	26, 49, 86, 136	0

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	67	GLY	15.3
2	O	1	MET	9.5
2	G	51	LEU	6.2
2	F	194	LEU	5.4
2	B	47	GLY	5.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [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. 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
3	OAS	D	156	9/10	0.85	0.17	-	31,43,55,60	0

## 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. 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	GOL	M	401	6/6	0.93	0.24	4.38	47,51,57,70	0
5	GOL	E	401	6/6	0.75	0.20	4.34	62,66,71,82	0
6	EDO	M	402	4/4	0.92	0.19	3.34	47,51,55,55	0
4	NAP	N	301	48/48	0.90	0.23	2.89	37,78,135,205	0
6	EDO	D	401	4/4	0.78	0.21	2.72	60,60,62,63	0
7	ACT	P	501	4/4	0.93	0.20	1.60	58,58,60,62	0
6	EDO	P	401	4/4	0.87	0.22	1.28	56,59,60,63	0
6	EDO	F	401	4/4	0.93	0.17	1.19	47,54,58,61	0
4	NAP	K	301	39/48	0.93	0.16	1.15	35,49,160,161	0
4	NAP	O	301	48/48	0.88	0.19	1.15	40,83,158,213	0
5	GOL	A	401	6/6	0.92	0.16	1.04	59,60,66,74	0
4	NAP	B	301	48/48	0.89	0.16	0.98	50,78,105,128	0
7	ACT	H	501	4/4	0.90	0.15	0.77	63,64,65,66	0
4	NAP	L	301	44/48	0.94	0.14	0.60	40,50,61,209	0
5	GOL	I	401	6/6	0.94	0.13	0.60	52,59,66,67	0
6	EDO	I	402	4/4	0.97	0.15	0.59	34,40,45,47	0
4	NAP	C	301	32/48	0.92	0.14	0.57	47,58,147,147	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
6	EDO	E	402	4/4	0.93	0.17	0.46	47,49,51,52	0
7	ACT	M	501	4/4	0.94	0.13	0.37	37,39,41,43	0
6	EDO	I	403	4/4	0.93	0.13	0.21	54,56,58,60	0
4	NAP	J	301	48/48	0.92	0.14	0.19	41,57,138,141	0
4	NAP	M	301	44/48	0.96	0.13	0.06	31,37,47,52	0
4	NAP	F	301	48/48	0.95	0.13	0.03	33,48,80,127	0
4	NAP	E	301	44/48	0.96	0.12	-0.11	24,39,48,54	0
4	NAP	A	301	44/48	0.96	0.12	-0.13	34,43,48,55	0
4	NAP	I	301	44/48	0.97	0.11	-0.38	31,38,44,49	0
4	NAP	D	301	44/48	0.94	0.12	-0.39	35,52,61,82	0
4	NAP	P	301	44/48	0.96	0.12	-0.41	29,44,54,109	0
4	NAP	H	301	44/48	0.97	0.10	-0.84	34,44,60,95	0
4	NAP	G	301	27/48	0.88	0.14	-0.93	56,59,105,142	0
6	EDO	L	401	4/4	0.73	0.16	-	62,64,65,67	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.