wwPDB X-ray Structure Validation Summary Report (i)

Dec 13, 2023 – 10:58 am GMT

PDB ID	:	2XF8
Title	:	Structure of the D-Erythrose-4-Phosphate Dehydrogenase from E. coli in complex with a NAD cofactor analog (3-Chloroacetyl adenine pyridine dinucleotide) and sulfate anion
Authors	:	Moniot, S.; Didierjean, C.; Boschi-Muller, S.; Branlant, G.; Corbier, C.
Deposited on	:	2010-05-20
Resolution	:	2.95 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

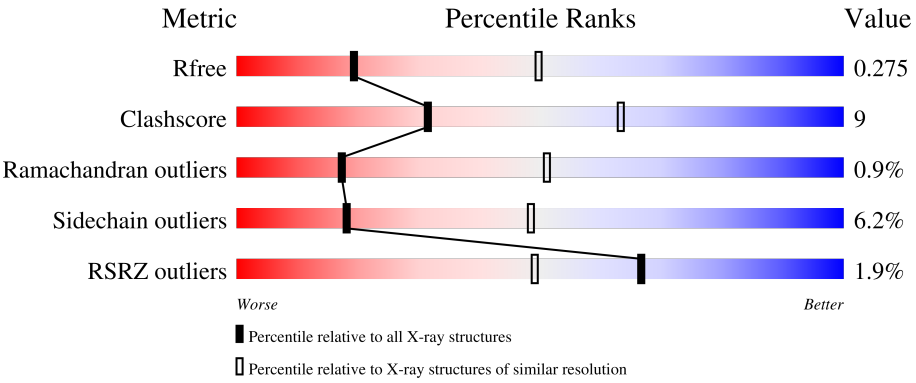
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.95 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.



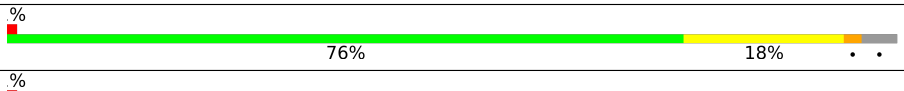
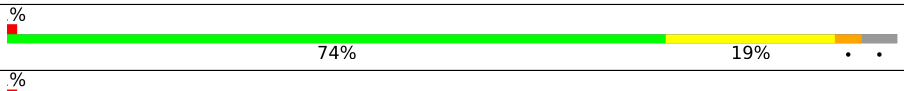
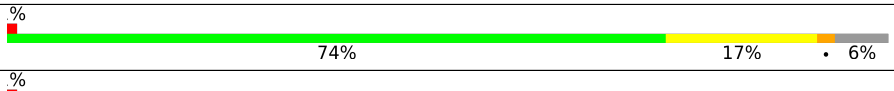
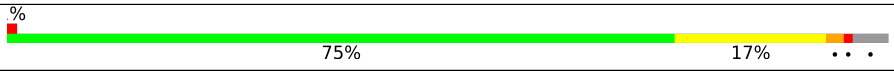
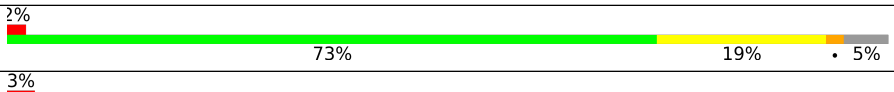
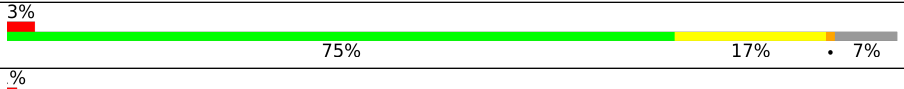
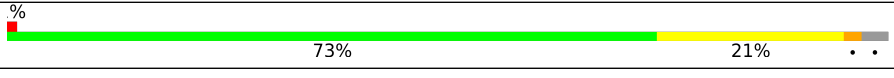
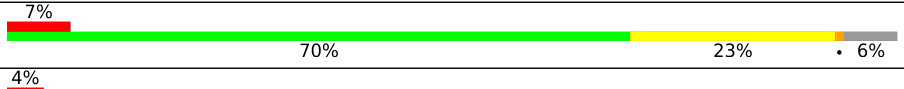
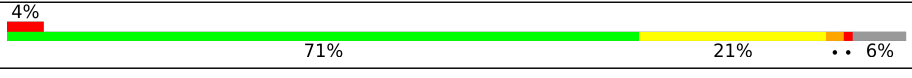
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	338	71% 21%
1	B	338	75% 19%
1	C	338	74% 19%
1	D	338	68% 23% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	338	
1	F	338	
1	G	338	
1	H	338	
1	I	338	
1	J	338	
1	K	338	
1	L	338	
1	M	338	
1	N	338	
1	O	338	
1	P	338	

2 Entry composition (i)

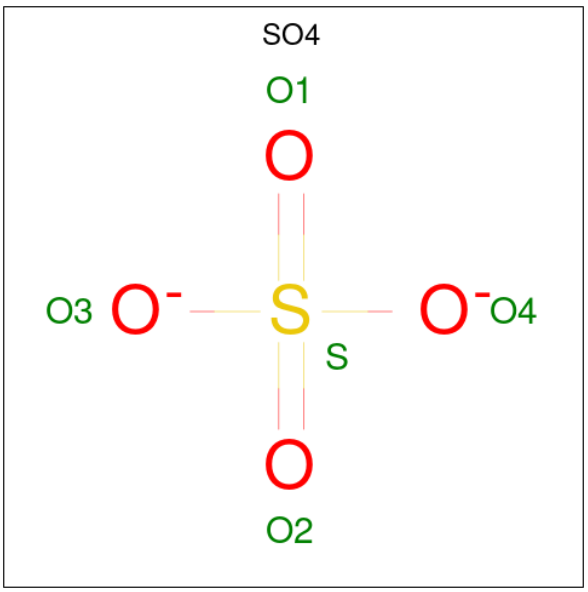
There are 4 unique types of molecules in this entry. The entry contains 40244 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	326	Total	C	N	O	S	0	0	0
			2501	1571	453	470	7			
1	B	325	Total	C	N	O	S	0	0	0
			2503	1574	453	469	7			
1	C	324	Total	C	N	O	S	0	0	0
			2492	1565	452	468	7			
1	D	317	Total	C	N	O	S	0	0	0
			2435	1531	439	458	7			
1	E	325	Total	C	N	O	S	0	0	0
			2496	1568	452	469	7			
1	F	326	Total	C	N	O	S	0	0	1
			2508	1575	457	469	7			
1	G	324	Total	C	N	O	S	0	0	0
			2498	1568	455	468	7			
1	H	317	Total	C	N	O	S	0	0	0
			2425	1526	436	456	7			
1	I	323	Total	C	N	O	S	0	0	0
			2487	1563	451	466	7			
1	J	325	Total	C	N	O	S	0	0	0
			2501	1572	453	469	7			
1	K	320	Total	C	N	O	S	0	0	0
			2465	1550	446	462	7			
1	L	316	Total	C	N	O	S	0	0	0
			2430	1528	438	457	7			
1	M	327	Total	C	N	O	S	0	0	0
			2512	1580	454	471	7			
1	N	325	Total	C	N	O	S	0	0	0
			2507	1575	456	469	7			
1	O	319	Total	C	N	O	S	0	0	0
			2452	1542	442	461	7			
1	P	317	Total	C	N	O	S	0	0	0
			2429	1528	436	458	7			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



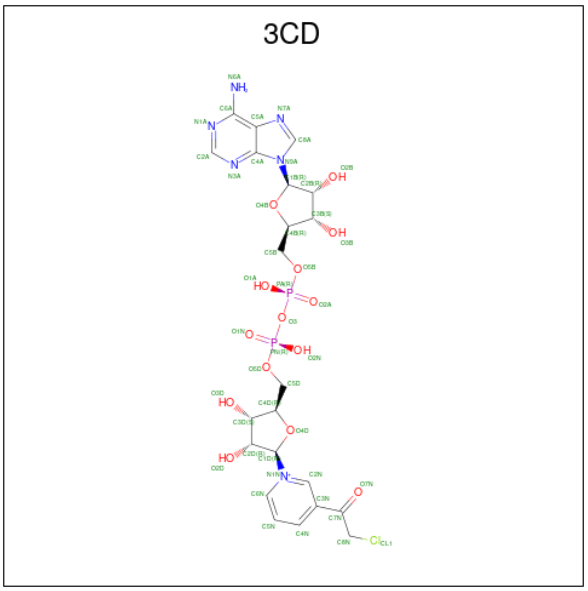
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0
2	E	1	Total O S 5 4 1	0	0
2	F	1	Total O S 5 4 1	0	0
2	G	1	Total O S 5 4 1	0	0
2	H	1	Total O S 5 4 1	0	0
2	I	1	Total O S 5 4 1	0	0
2	J	1	Total O S 5 4 1	0	0
2	K	1	Total O S 5 4 1	0	0
2	L	1	Total O S 5 4 1	0	0
2	M	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	N	1	Total O S 5 4 1	0	0
2	O	1	Total O S 5 4 1	0	0
2	P	1	Total O S 5 4 1	0	0

- Molecule 3 is 3-(CHLOROACETYL) PYRIDINE ADENINE DINUCLEOTIDE (three-letter code: 3CD) (formula: $C_{22}H_{28}ClN_6O_{14}P_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O P 44 22 6 14 2	0	0
3	C	1	Total C N O P 44 22 6 14 2	0	0
3	E	1	Total C N O P 44 22 6 14 2	0	0
3	G	1	Total C N O P 44 22 6 14 2	0	0
3	I	1	Total C N O P 44 22 6 14 2	0	0
3	K	1	Total C N O P 44 22 6 14 2	0	0
3	M	1	Total C N O P 44 22 6 14 2	0	0
3	O	1	Total C N O P 44 22 6 14 2	0	0

- Molecule 4 is water.

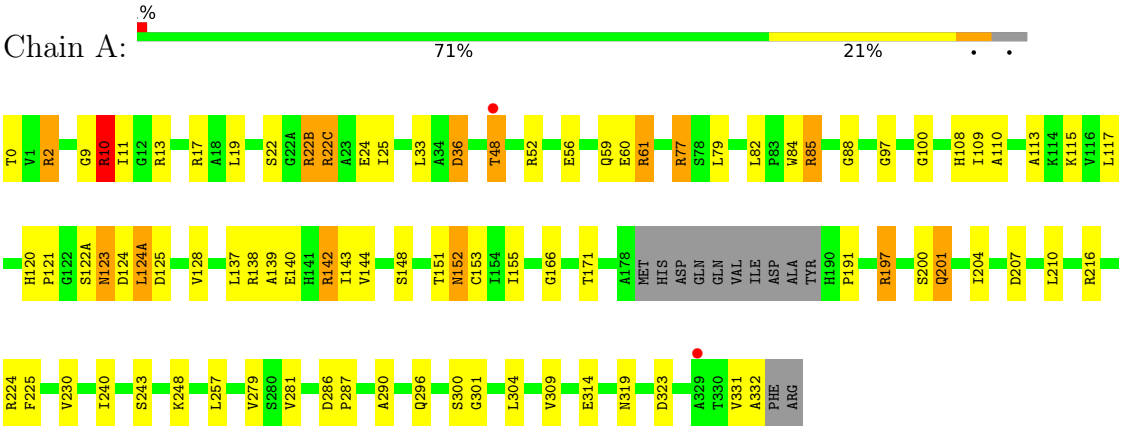
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	15	Total O 15 15	0	0
4	B	12	Total O 12 12	0	0
4	C	11	Total O 11 11	0	0
4	D	8	Total O 8 8	0	0
4	E	11	Total O 11 11	0	0
4	F	8	Total O 8 8	0	0
4	G	9	Total O 9 9	0	0
4	H	8	Total O 8 8	0	0
4	I	8	Total O 8 8	0	0
4	J	12	Total O 12 12	0	0
4	K	13	Total O 13 13	0	0
4	L	15	Total O 15 15	0	0
4	M	7	Total O 7 7	0	0
4	N	16	Total O 16 16	0	0
4	O	8	Total O 8 8	0	0
4	P	10	Total O 10 10	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

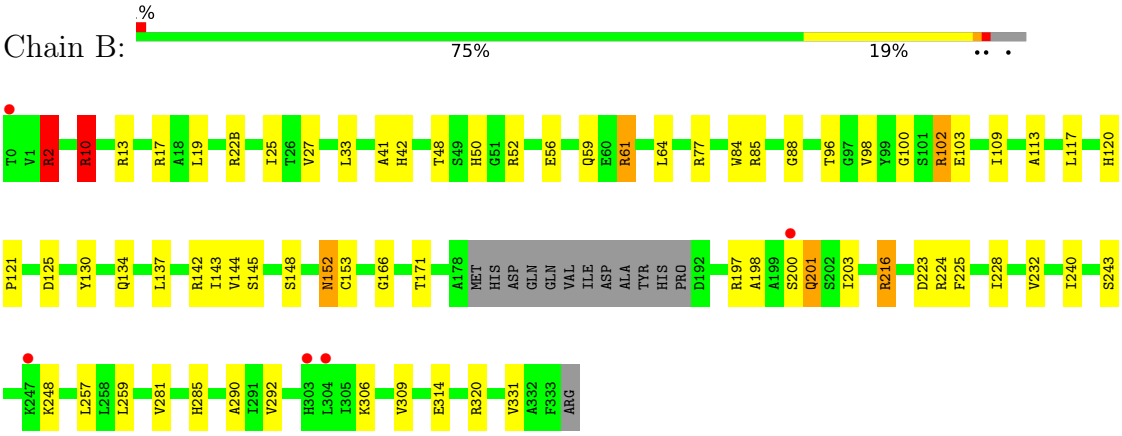
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain A:



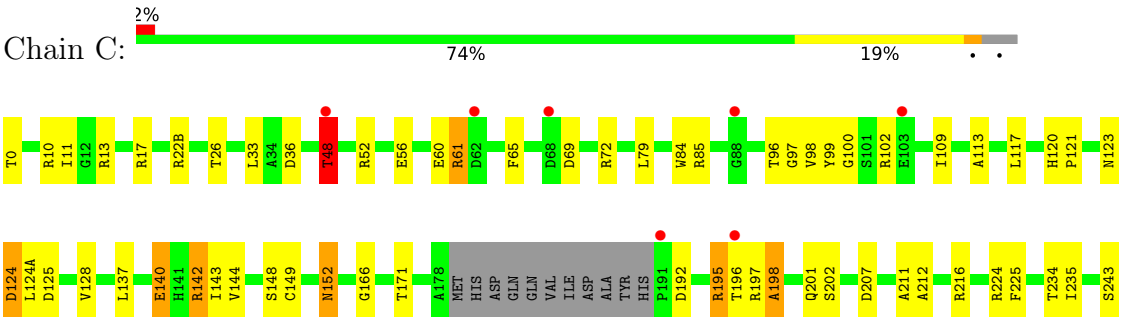
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

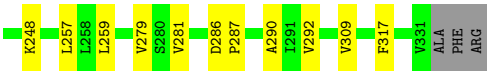
Chain B:



• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

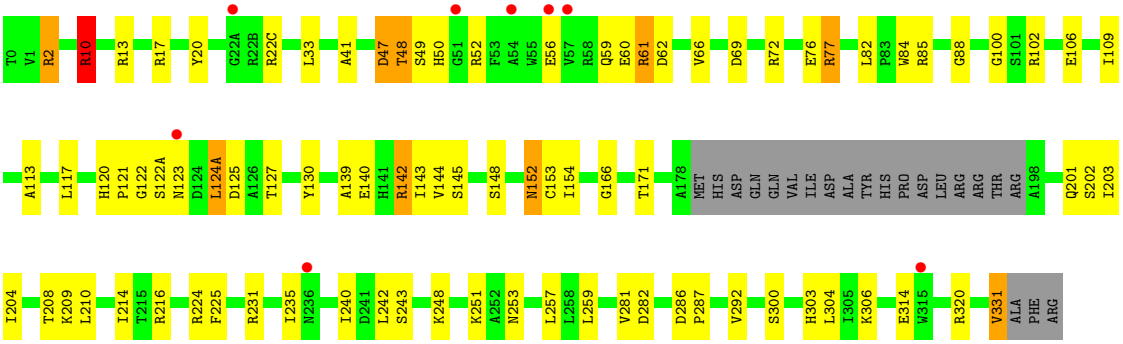
Chain C:





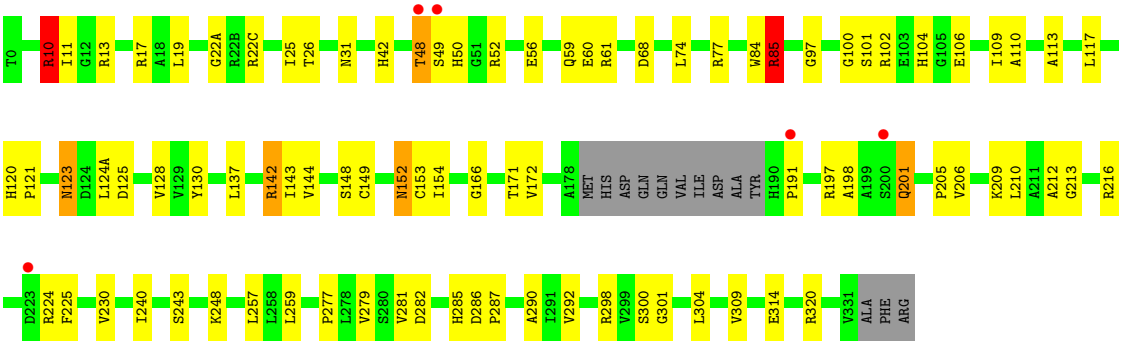
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain D:



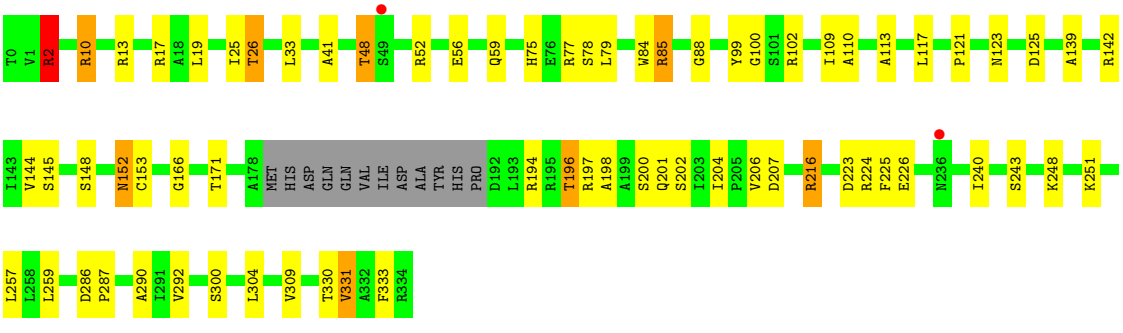
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain E:



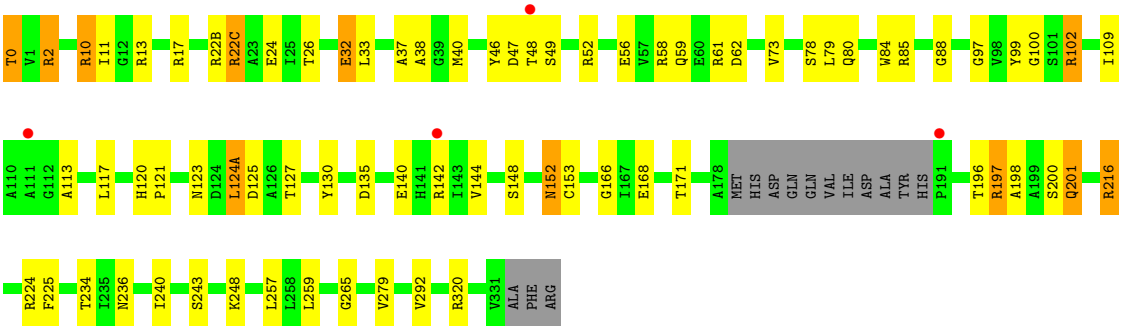
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain F:



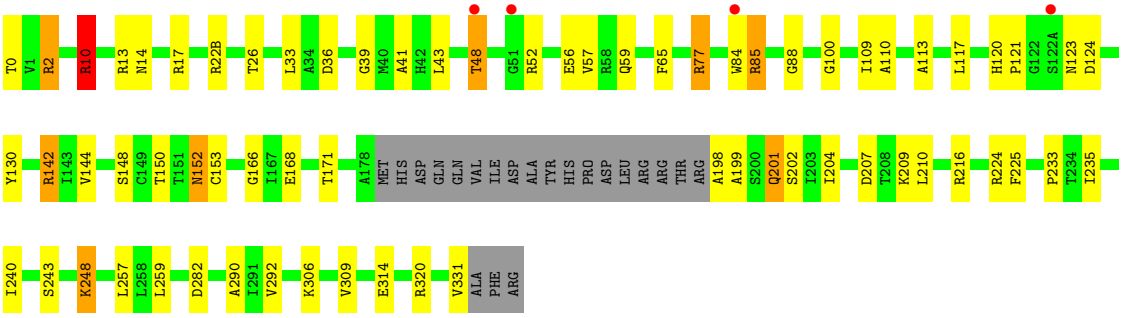
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain G:



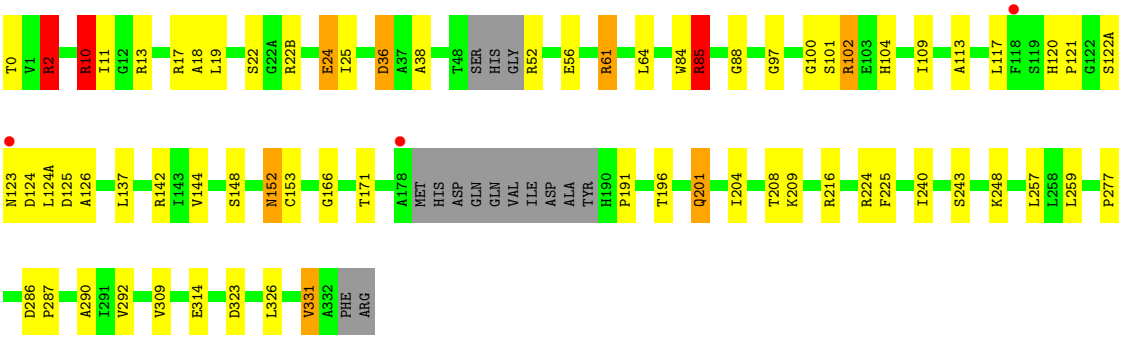
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain H:



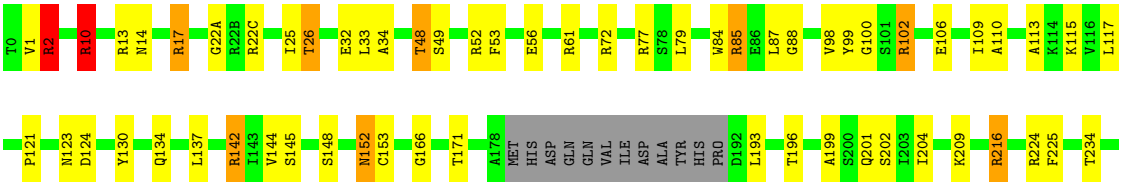
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

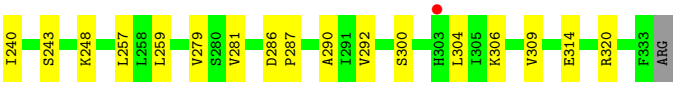
Chain I:



• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

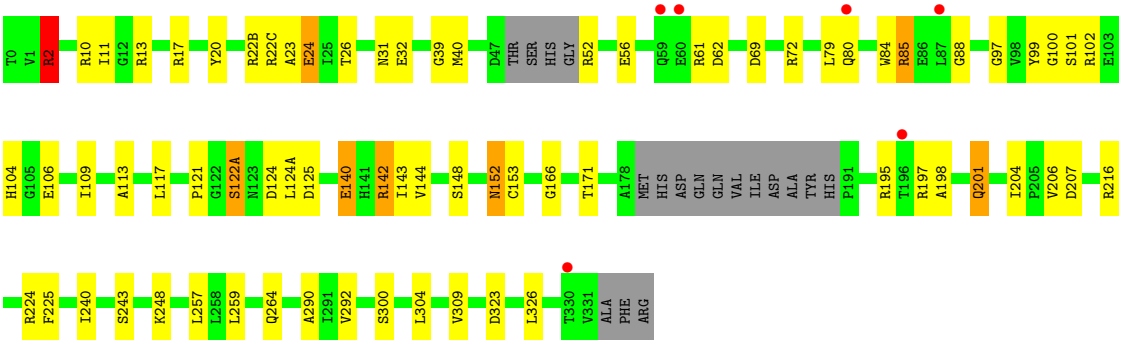
Chain J:





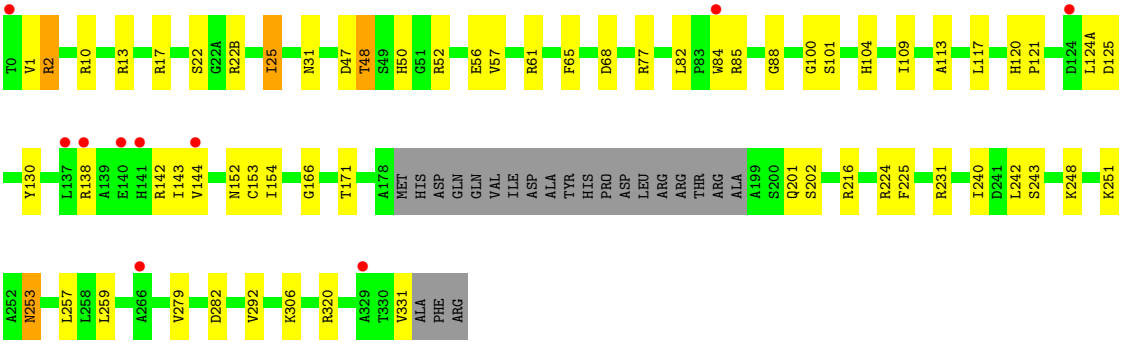
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain K:



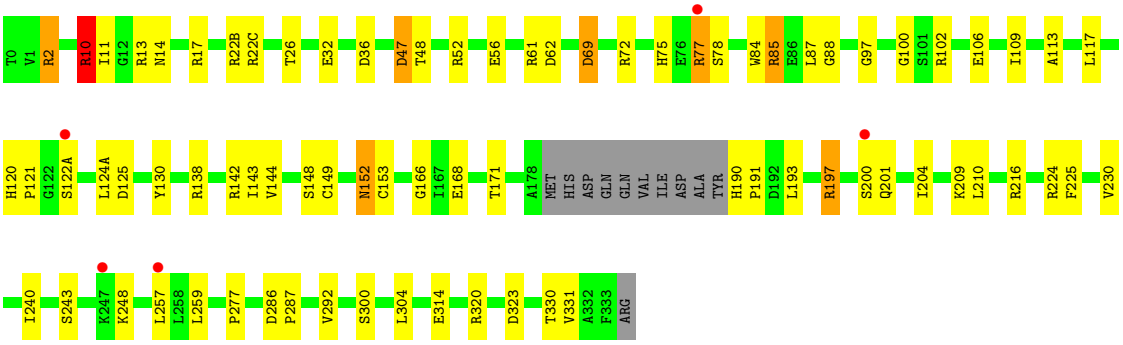
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain L:



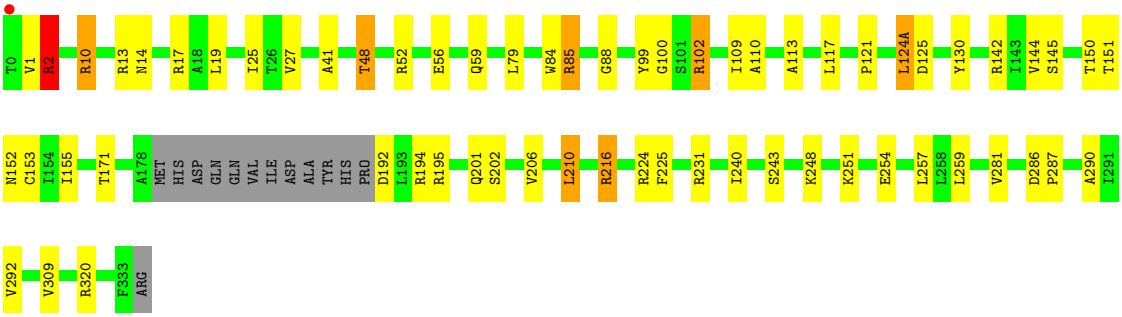
• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

Chain M:

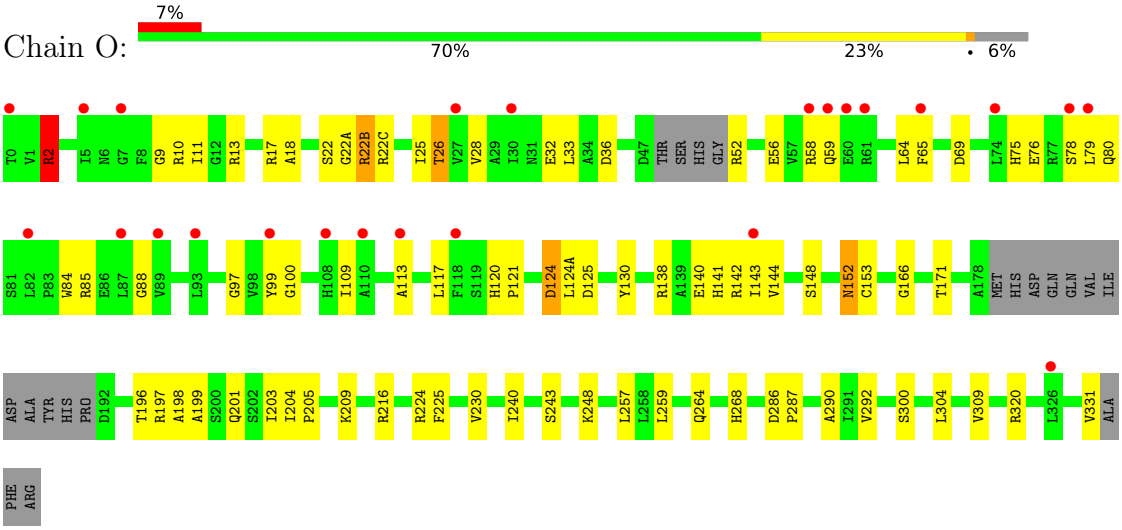


• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE

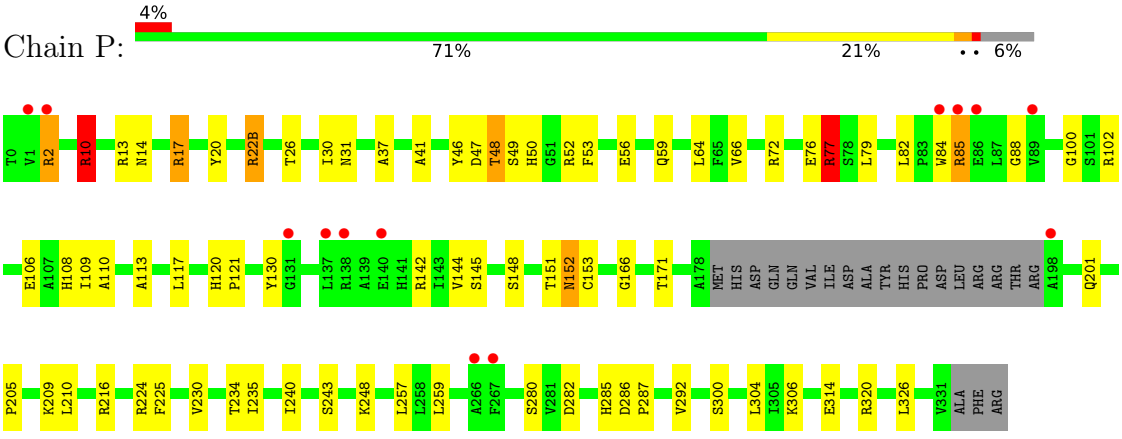
Chain N:



• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE



• Molecule 1: D-ERYTHROSE-4-PHOSPHATE DEHYDROGENASE



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	96.18Å 111.66Å 135.47Å 91.30° 96.11° 88.44°	Depositor
Resolution (Å)	48.72 – 2.95 48.72 – 2.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.72-2.95) 95.4 (48.72-2.95)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.41 (at 2.96Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.237 , 0.280 0.234 , 0.275	Depositor DCC
R_{free} test set	5630 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	50.6	Xtriage
Anisotropy	0.543	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 48.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.043 for -h,k,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	40244	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: 3CD, SO4

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	0.56	0/2547	1.55	27/3467 (0.8%)
1	B	0.64	1/2548 (0.0%)	1.68	27/3465 (0.8%)
1	C	0.59	0/2537	1.51	24/3451 (0.7%)
1	D	0.64	1/2479 (0.0%)	1.78	29/3373 (0.9%)
1	E	0.65	1/2542 (0.0%)	1.73	25/3460 (0.7%)
1	F	0.63	1/2553 (0.0%)	1.42	27/3472 (0.8%)
1	G	0.61	0/2543	1.53	25/3458 (0.7%)
1	H	0.62	2/2469 (0.1%)	1.76	28/3361 (0.8%)
1	I	0.59	0/2531	1.76	27/3443 (0.8%)
1	J	0.63	3/2546 (0.1%)	1.68	27/3463 (0.8%)
1	K	0.58	0/2508	1.58	26/3410 (0.8%)
1	L	0.68	2/2474 (0.1%)	1.90	27/3366 (0.8%)
1	M	0.63	3/2559 (0.1%)	1.91	28/3483 (0.8%)
1	N	0.62	2/2552 (0.1%)	1.86	27/3470 (0.8%)
1	O	0.61	0/2494	1.60	27/3392 (0.8%)
1	P	0.62	2/2473 (0.1%)	1.77	28/3366 (0.8%)
All	All	0.62	18/40355 (0.0%)	1.69	429/54900 (0.8%)

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
1	D	0	1
1	E	0	3
1	H	0	1
1	I	0	1
1	M	0	2
All	All	0	9

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	253	ASN	CG-ND2	-9.99	1.07	1.32
1	L	253	ASN	CG-OD1	-9.58	1.02	1.24
1	N	14	ASN	CG-ND2	-7.97	1.12	1.32
1	D	303	HIS	CG-CD2	-7.05	1.23	1.35
1	N	14	ASN	CG-OD1	-6.97	1.08	1.24

The worst 5 of 429 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	216	ARG	NE-CZ-NH1	-39.30	100.65	120.30
1	J	216	ARG	NE-CZ-NH1	-38.89	100.86	120.30
1	N	216	ARG	NE-CZ-NH1	-38.41	101.10	120.30
1	P	13	ARG	NE-CZ-NH1	-37.42	101.59	120.30
1	H	13	ARG	NE-CZ-NH1	-36.89	101.86	120.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	22(B)	ARG	Peptide
1	D	122(A)	SER	Peptide
1	E	123	ASN	Peptide
1	E	142	ARG	Sidechain
1	E	85	ARG	Sidechain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2501	0	2485	70	0
1	B	2503	0	2496	50	0
1	C	2492	0	2485	58	0
1	D	2435	0	2428	54	0
1	E	2496	0	2480	54	0
1	F	2508	0	2503	40	0
1	G	2498	0	2496	60	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	H	2425	0	2413	47	0
1	I	2487	0	2480	42	0
1	J	2501	0	2492	51	0
1	K	2465	0	2462	61	0
1	L	2430	0	2423	40	0
1	M	2512	0	2494	44	0
1	N	2507	0	2503	40	0
1	O	2452	0	2443	61	0
1	P	2429	0	2417	52	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
2	G	5	0	0	0	0
2	H	5	0	0	0	0
2	I	5	0	0	0	0
2	J	5	0	0	0	0
2	K	5	0	0	0	0
2	L	5	0	0	0	0
2	M	5	0	0	0	0
2	N	5	0	0	0	0
2	O	5	0	0	0	0
2	P	5	0	0	0	0
3	A	44	0	24	8	0
3	C	44	0	24	6	0
3	E	44	0	24	4	0
3	G	44	0	24	7	0
3	I	44	0	24	7	0
3	K	44	0	24	11	0
3	M	44	0	24	11	0
3	O	44	0	24	13	0
4	A	15	0	0	0	0
4	B	12	0	0	1	0
4	C	11	0	0	2	0
4	D	8	0	0	1	0
4	E	11	0	0	0	0
4	F	8	0	0	0	0
4	G	9	0	0	1	0
4	H	8	0	0	1	0
4	I	8	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	J	12	0	0	1	0
4	K	13	0	0	2	0
4	L	15	0	0	0	0
4	M	7	0	0	0	0
4	N	16	0	0	1	0
4	O	8	0	0	0	0
4	P	10	0	0	1	0
All	All	40244	0	39692	716	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 9.

The worst 5 of 716 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:B:198:ALA:HB3	1:C:48:THR:HG21	1.25	1.14
1:J:48:THR:HG21	1:K:198:ALA:HB3	1.33	1.10
1:N:281:VAL:HG11	1:P:48:THR:HG23	1.33	1.08
1:A:139:ALA:HB2	1:A:332:ALA:HB3	1.38	1.03
1:B:98:VAL:HG11	1:H:77:ARG:HD3	1.42	1.01

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	322/338 (95%)	303 (94%)	14 (4%)	5 (2%)	9 36
1	B	321/338 (95%)	304 (95%)	14 (4%)	3 (1%)	17 51
1	C	320/338 (95%)	297 (93%)	19 (6%)	4 (1%)	12 41
1	D	313/338 (93%)	297 (95%)	12 (4%)	4 (1%)	12 41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	321/338 (95%)	298 (93%)	20 (6%)	3 (1%)	17	51
1	F	322/338 (95%)	306 (95%)	13 (4%)	3 (1%)	17	51
1	G	320/338 (95%)	302 (94%)	16 (5%)	2 (1%)	25	60
1	H	313/338 (93%)	292 (93%)	19 (6%)	2 (1%)	25	60
1	I	317/338 (94%)	302 (95%)	11 (4%)	4 (1%)	12	41
1	J	321/338 (95%)	305 (95%)	14 (4%)	2 (1%)	25	60
1	K	314/338 (93%)	287 (91%)	23 (7%)	4 (1%)	12	41
1	L	312/338 (92%)	298 (96%)	12 (4%)	2 (1%)	25	60
1	M	323/338 (96%)	300 (93%)	20 (6%)	3 (1%)	17	51
1	N	321/338 (95%)	303 (94%)	17 (5%)	1 (0%)	41	73
1	O	313/338 (93%)	292 (93%)	19 (6%)	2 (1%)	25	60
1	P	313/338 (93%)	297 (95%)	14 (4%)	2 (1%)	25	60
All	All	5086/5408 (94%)	4783 (94%)	257 (5%)	46 (1%)	17	51

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	191	PRO
1	A	22(C)	ARG
1	A	191	PRO
1	C	198	ALA
1	K	24	GLU

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	266/279 (95%)	248 (93%)	18 (7%)	16	45
1	B	266/279 (95%)	253 (95%)	13 (5%)	25	58
1	C	266/279 (95%)	249 (94%)	17 (6%)	17	47
1	D	260/279 (93%)	243 (94%)	17 (6%)	17	46

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	266/279 (95%)	249 (94%)	17 (6%)	17	47
1	F	267/279 (96%)	251 (94%)	16 (6%)	19	50
1	G	267/279 (96%)	247 (92%)	20 (8%)	13	40
1	H	258/279 (92%)	242 (94%)	16 (6%)	18	48
1	I	265/279 (95%)	246 (93%)	19 (7%)	14	42
1	J	266/279 (95%)	247 (93%)	19 (7%)	14	43
1	K	263/279 (94%)	249 (95%)	14 (5%)	22	55
1	L	260/279 (93%)	247 (95%)	13 (5%)	24	57
1	M	267/279 (96%)	248 (93%)	19 (7%)	14	43
1	N	267/279 (96%)	253 (95%)	14 (5%)	23	56
1	O	261/279 (94%)	245 (94%)	16 (6%)	18	49
1	P	259/279 (93%)	244 (94%)	15 (6%)	20	51
All	All	4224/4464 (95%)	3961 (94%)	263 (6%)	18	48

5 of 263 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	N	257	LEU
1	O	64	LEU
1	P	225	PHE
1	F	257	LEU
1	F	207	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	236	ASN
1	I	14	ASN
1	I	123	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

24 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	SO4	F	500	-	4,4,4	0.20	0	6,6,6	0.30	0
2	SO4	B	500	-	4,4,4	0.13	0	6,6,6	0.23	0
3	3CD	C	600	1	42,48,49	1.42	3 (7%)	50,73,74	1.71	9 (18%)
3	3CD	I	600	1	42,48,49	1.37	2 (4%)	50,73,74	1.59	11 (22%)
2	SO4	C	500	-	4,4,4	0.07	0	6,6,6	0.56	0
3	3CD	A	600	1	42,48,49	1.26	1 (2%)	50,73,74	1.45	8 (16%)
2	SO4	H	500	-	4,4,4	0.21	0	6,6,6	0.19	0
2	SO4	L	500	-	4,4,4	0.19	0	6,6,6	0.44	0
2	SO4	N	500	-	4,4,4	0.14	0	6,6,6	0.11	0
2	SO4	K	500	-	4,4,4	0.20	0	6,6,6	0.31	0
2	SO4	O	500	-	4,4,4	0.20	0	6,6,6	0.29	0
3	3CD	E	600	1	42,48,49	1.38	6 (14%)	50,73,74	1.72	8 (16%)
2	SO4	J	500	-	4,4,4	0.21	0	6,6,6	0.35	0
2	SO4	D	500	-	4,4,4	0.14	0	6,6,6	0.31	0
2	SO4	E	500	-	4,4,4	0.10	0	6,6,6	0.57	0
2	SO4	G	500	-	4,4,4	0.10	0	6,6,6	0.47	0
3	3CD	O	600	1	42,48,49	1.27	2 (4%)	50,73,74	1.50	6 (12%)
2	SO4	A	500	-	4,4,4	0.17	0	6,6,6	0.27	0
2	SO4	M	500	-	4,4,4	0.14	0	6,6,6	0.24	0
2	SO4	I	500	-	4,4,4	0.17	0	6,6,6	0.22	0
3	3CD	M	600	1	42,48,49	1.22	2 (4%)	50,73,74	1.83	10 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	3CD	K	600	1	42,48,49	1.29	1 (2%)	50,73,74	1.56	10 (20%)
2	SO4	P	500	-	4,4,4	0.18	0	6,6,6	0.27	0
3	3CD	G	600	1	42,48,49	1.44	3 (7%)	50,73,74	1.73	9 (18%)

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	3CD	C	600	1	-	14/26/62/64	0/5/5/5
3	3CD	I	600	1	-	15/26/62/64	0/5/5/5
3	3CD	M	600	1	-	13/26/62/64	0/5/5/5
3	3CD	O	600	1	-	14/26/62/64	0/5/5/5
3	3CD	A	600	1	-	9/26/62/64	0/5/5/5
3	3CD	E	600	1	-	10/26/62/64	0/5/5/5
3	3CD	K	600	1	-	13/26/62/64	0/5/5/5
3	3CD	G	600	1	-	13/26/62/64	0/5/5/5

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	600	3CD	C8N-C7N	-6.83	1.29	1.49
3	C	600	3CD	C8N-C7N	-6.74	1.29	1.49
3	I	600	3CD	C8N-C7N	-6.63	1.30	1.49
3	A	600	3CD	C8N-C7N	-6.39	1.30	1.49
3	K	600	3CD	C8N-C7N	-6.23	1.31	1.49

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	600	3CD	C2N-C3N-C4N	5.53	124.53	118.26
3	O	600	3CD	N3A-C2A-N1A	-5.52	120.04	128.68
3	A	600	3CD	N3A-C2A-N1A	-5.39	120.25	128.68
3	C	600	3CD	N3A-C2A-N1A	-5.18	120.58	128.68
3	G	600	3CD	N3A-C2A-N1A	-5.15	120.63	128.68

There are no chirality outliers.

5 of 101 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	600	3CD	C5D-O5D-PN-O3
3	A	600	3CD	C3D-C4D-C5D-O5D
3	A	600	3CD	O4D-C4D-C5D-O5D
3	C	600	3CD	C5B-O5B-PA-O2A
3	C	600	3CD	C5B-O5B-PA-O3

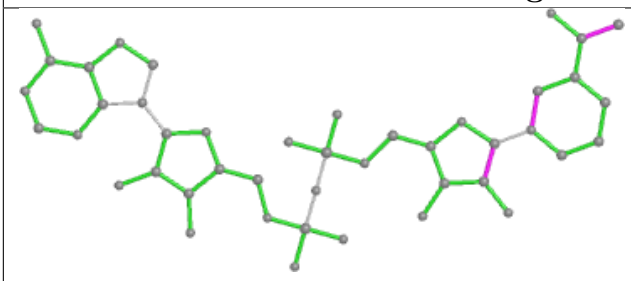
There are no ring outliers.

8 monomers are involved in 67 short contacts:

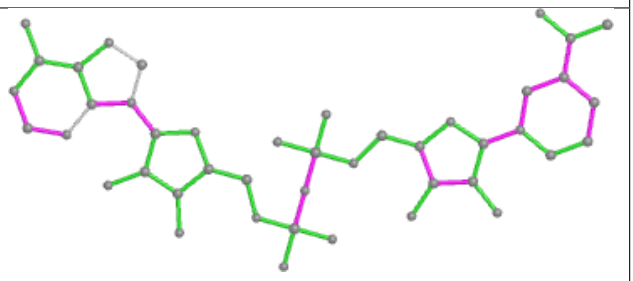
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	600	3CD	6	0
3	I	600	3CD	7	0
3	A	600	3CD	8	0
3	E	600	3CD	4	0
3	O	600	3CD	13	0
3	M	600	3CD	11	0
3	K	600	3CD	11	0
3	G	600	3CD	7	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

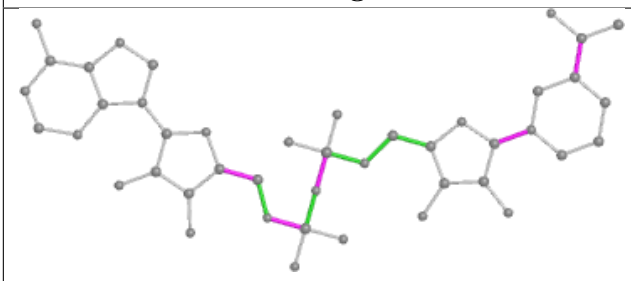
Ligand 3CD C 600



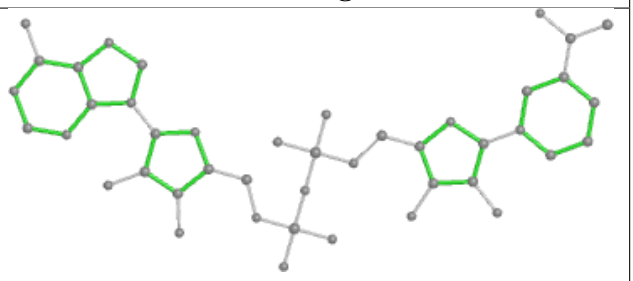
Bond lengths



Bond angles

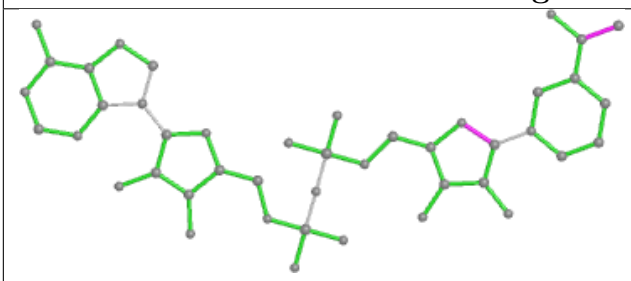


Torsions

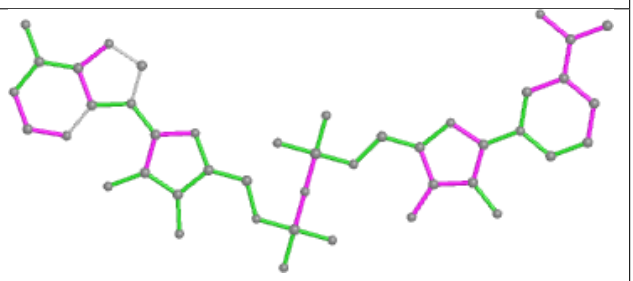


Rings

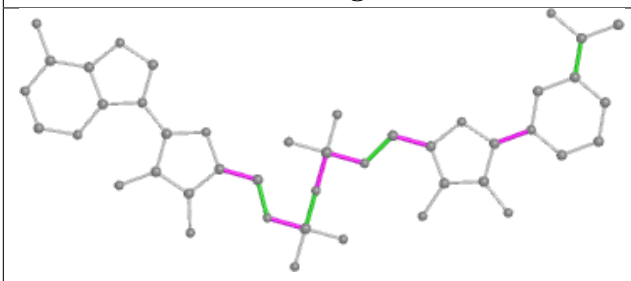
Ligand 3CD I 600



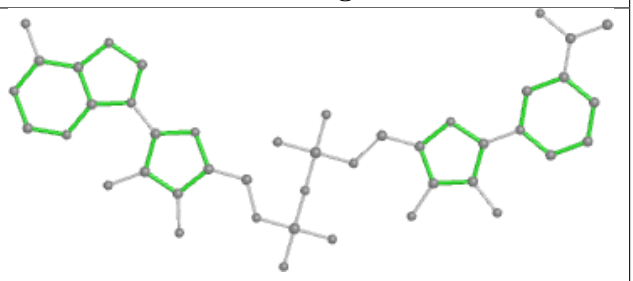
Bond lengths



Bond angles

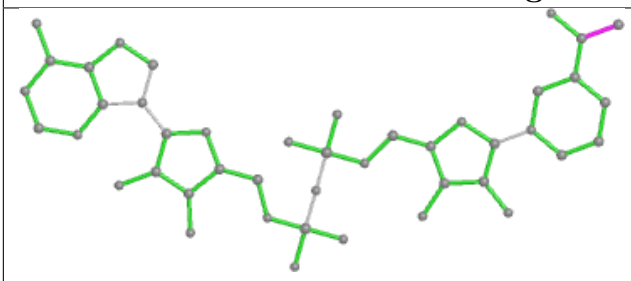


Torsions

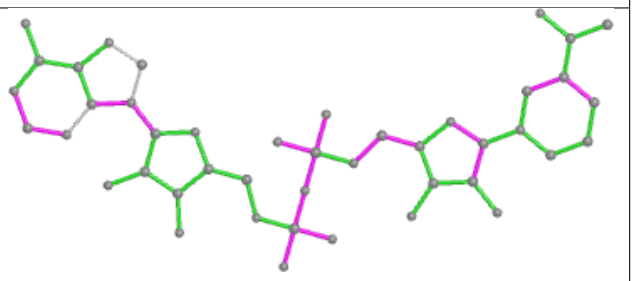


Rings

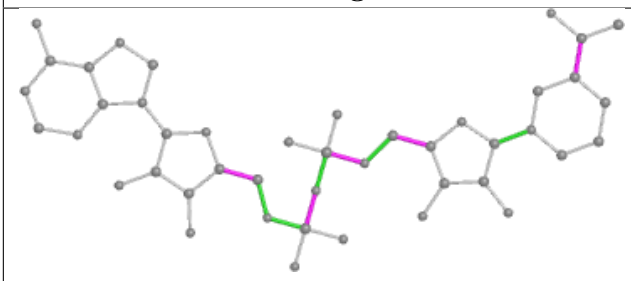
Ligand 3CD A 600



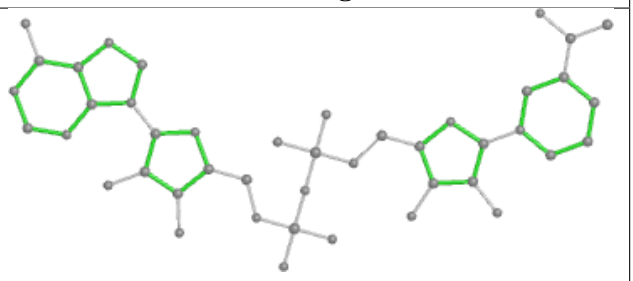
Bond lengths



Bond angles

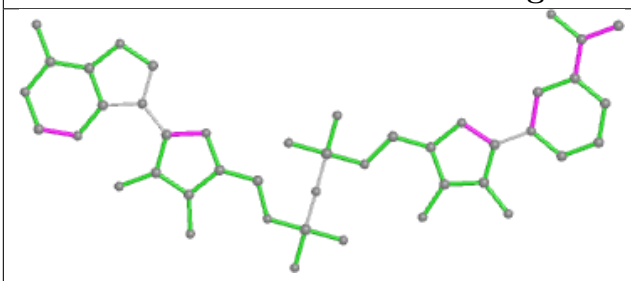


Torsions

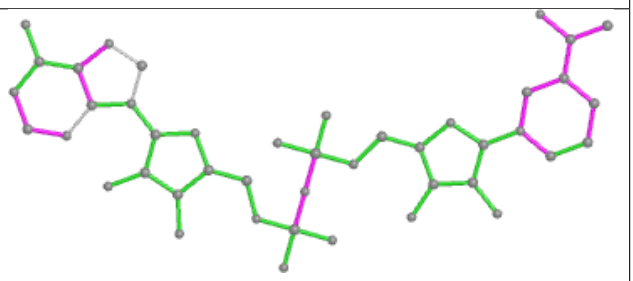


Rings

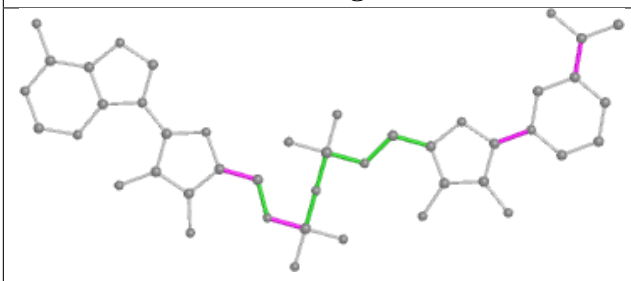
Ligand 3CD E 600



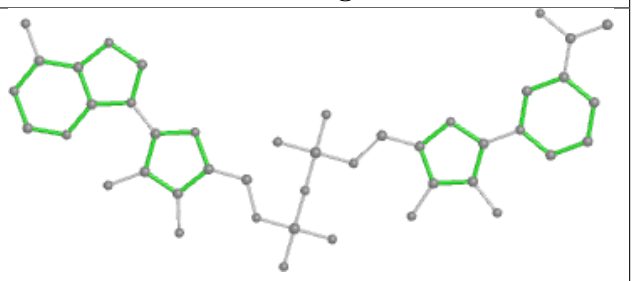
Bond lengths



Bond angles

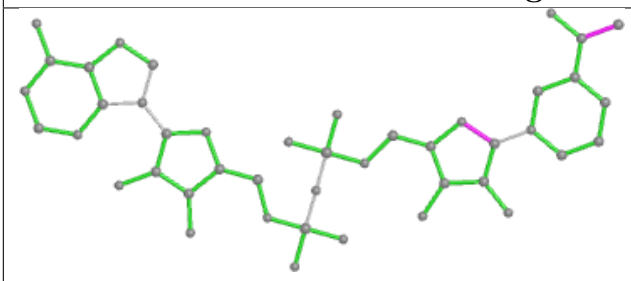


Torsions

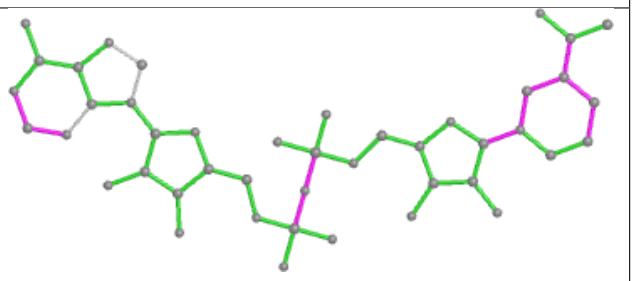


Rings

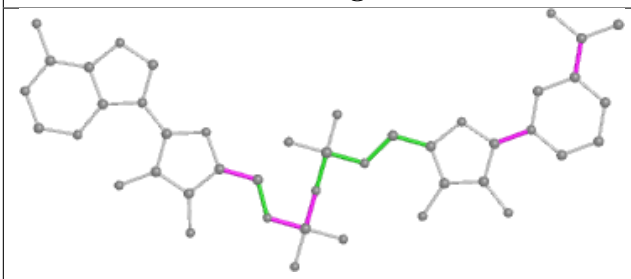
Ligand 3CD O 600



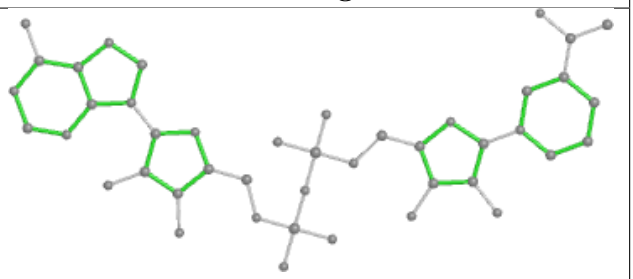
Bond lengths



Bond angles

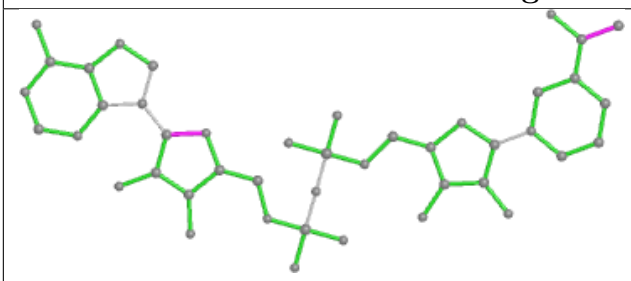


Torsions

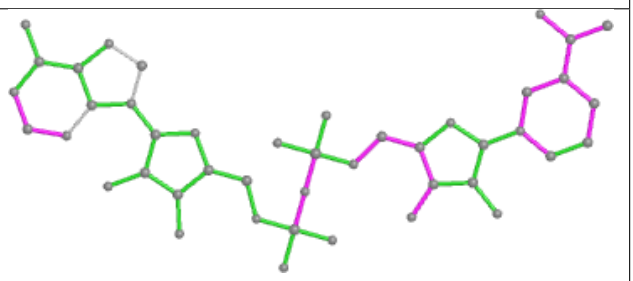


Rings

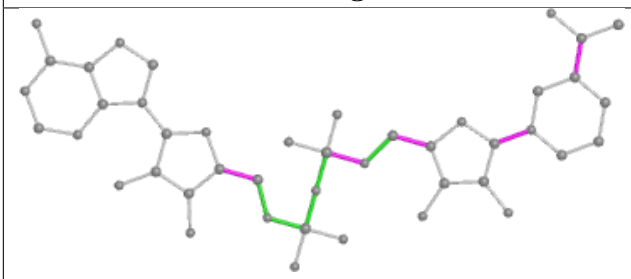
Ligand 3CD M 600



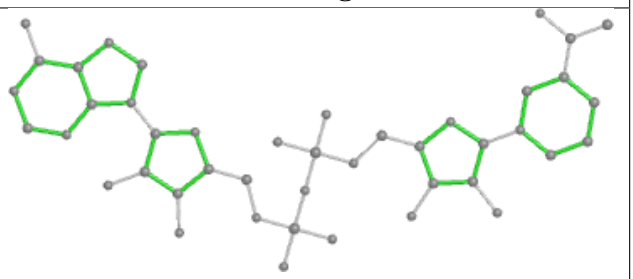
Bond lengths



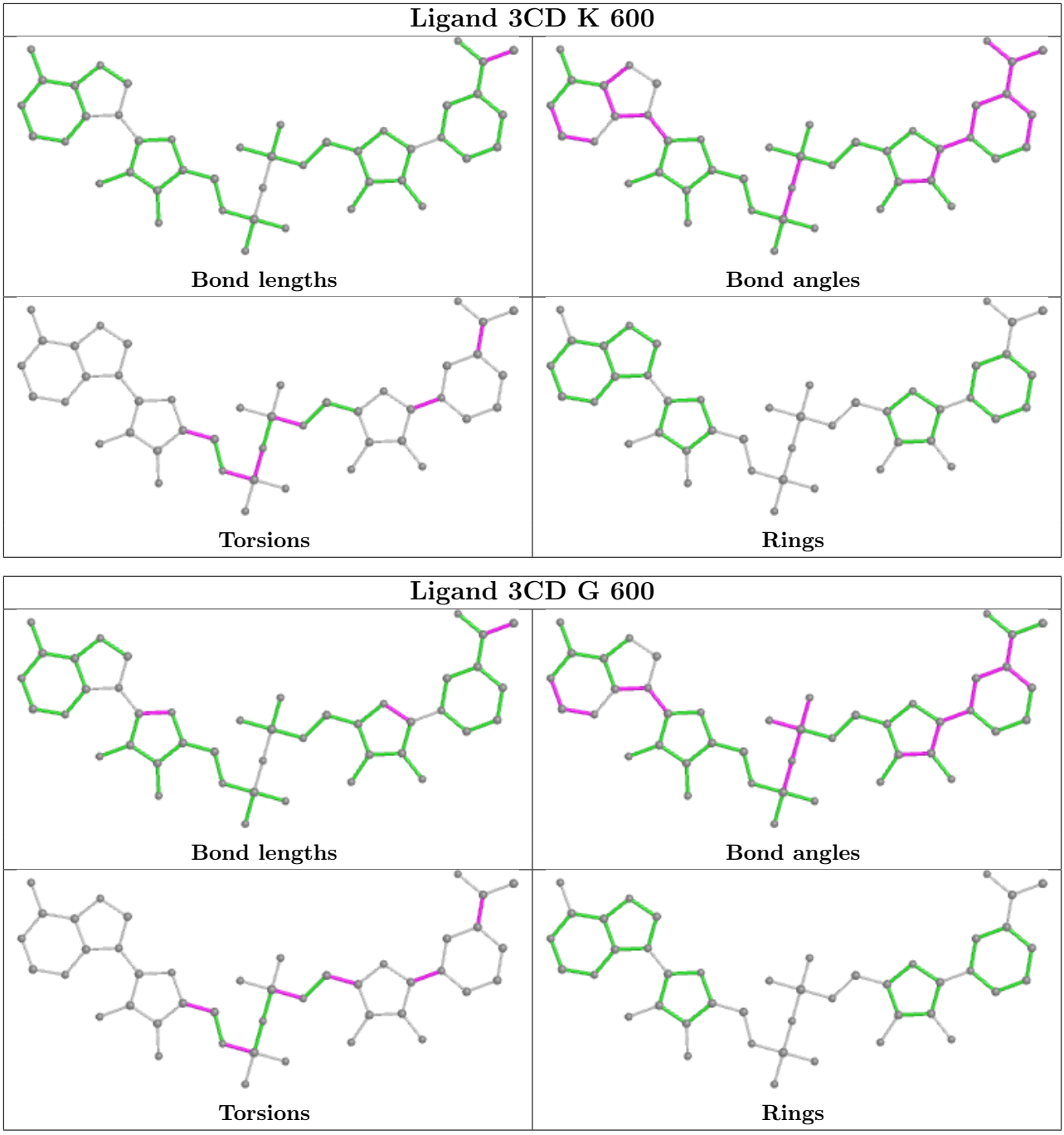
Bond angles



Torsions



Rings



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^{th} 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^2)	Q<0.9
1	A	326/338 (96%)	0.19	2 (0%) 89 78	36, 44, 57, 82	0
1	B	325/338 (96%)	0.26	5 (1%) 73 57	35, 44, 54, 73	0
1	C	324/338 (95%)	0.22	7 (2%) 62 45	33, 44, 63, 79	0
1	D	317/338 (93%)	0.32	8 (2%) 57 40	38, 45, 60, 66	0
1	E	325/338 (96%)	0.28	5 (1%) 73 57	33, 44, 54, 82	0
1	F	326/338 (96%)	0.12	2 (0%) 89 78	29, 44, 54, 66	0
1	G	324/338 (95%)	0.19	4 (1%) 79 63	31, 45, 58, 69	0
1	H	317/338 (93%)	0.23	4 (1%) 77 61	33, 44, 55, 62	0
1	I	323/338 (95%)	0.23	3 (0%) 84 71	32, 44, 54, 72	0
1	J	325/338 (96%)	0.20	1 (0%) 94 87	29, 44, 54, 73	0
1	K	320/338 (94%)	0.27	6 (1%) 66 49	36, 45, 62, 77	0
1	L	316/338 (93%)	0.46	10 (3%) 47 31	32, 45, 61, 68	0
1	M	327/338 (96%)	0.27	5 (1%) 73 57	35, 44, 57, 81	0
1	N	325/338 (96%)	0.19	1 (0%) 94 87	29, 44, 54, 76	0
1	O	319/338 (94%)	0.54	24 (7%) 14 8	36, 45, 61, 75	0
1	P	317/338 (93%)	0.35	13 (4%) 37 24	38, 45, 57, 62	0
All	All	5156/5408 (95%)	0.27	100 (1%) 66 49	29, 44, 58, 82	0

The worst 5 of 100 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	O	59	GLN	5.1
1	O	60	GLU	4.3
1	O	78	SER	4.2
1	C	191	PRO	4.1
1	E	48	THR	4.0

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 monosaccharides 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^{th} 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.

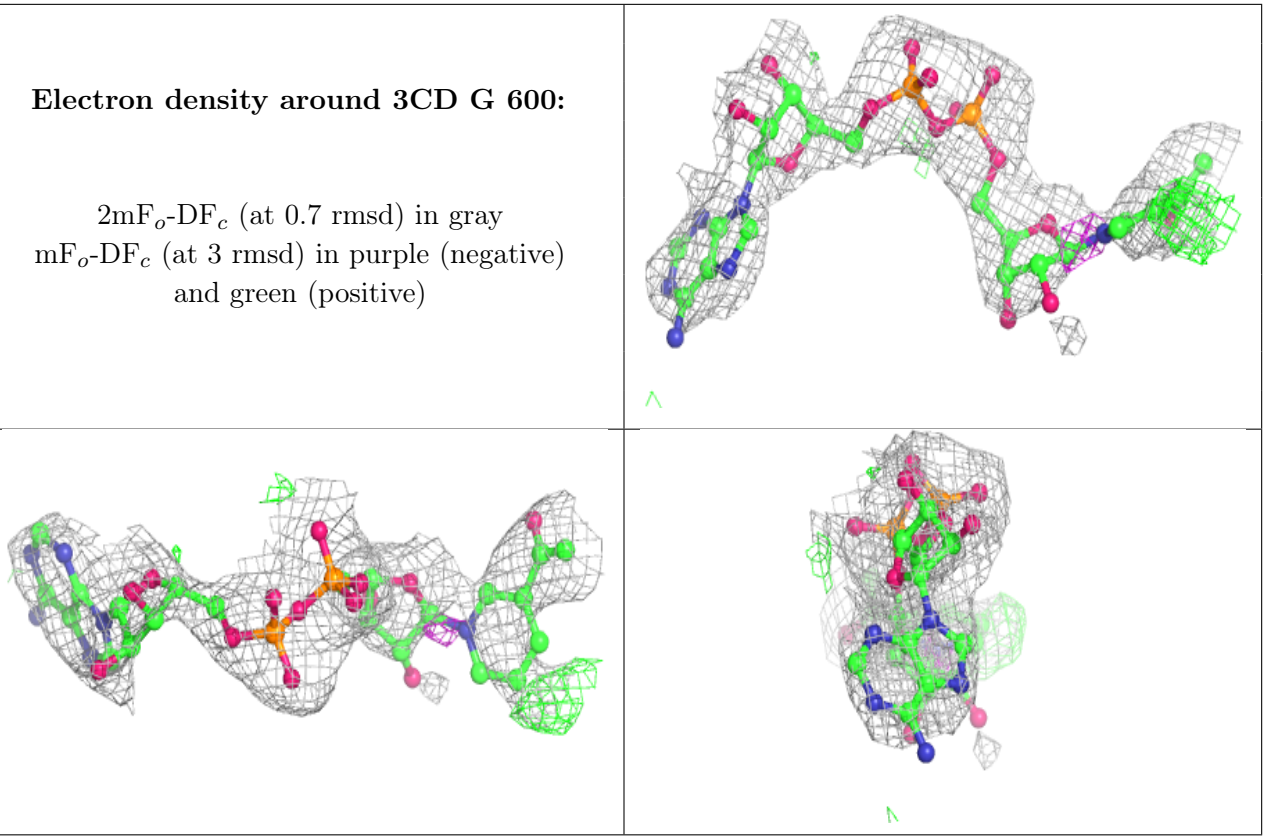
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	N	500	5/5	0.86	0.20	78,79,79,79	0
3	3CD	G	600	44/45	0.86	0.22	57,71,78,80	0
2	SO4	F	500	5/5	0.87	0.24	86,87,87,87	0
2	SO4	B	500	5/5	0.88	0.18	102,102,103,103	0
3	3CD	O	600	44/45	0.88	0.23	54,68,78,79	0
2	SO4	G	500	5/5	0.89	0.21	72,73,74,76	0
2	SO4	I	500	5/5	0.89	0.20	95,95,95,95	0
2	SO4	J	500	5/5	0.89	0.16	82,82,83,83	0
2	SO4	D	500	5/5	0.90	0.21	86,86,87,87	0
3	3CD	M	600	44/45	0.91	0.21	50,57,63,64	0
2	SO4	O	500	5/5	0.91	0.16	79,79,80,80	0
3	3CD	K	600	44/45	0.92	0.25	51,57,60,62	0
3	3CD	E	600	44/45	0.92	0.19	43,48,59,60	0
2	SO4	E	500	5/5	0.92	0.19	74,76,76,76	0
2	SO4	C	500	5/5	0.93	0.17	55,56,58,58	0
2	SO4	H	500	5/5	0.93	0.18	81,81,82,82	0
3	3CD	I	600	44/45	0.93	0.19	37,47,55,57	0
2	SO4	P	500	5/5	0.93	0.13	86,86,87,87	0
3	3CD	A	600	44/45	0.93	0.18	48,52,64,66	0
3	3CD	C	600	44/45	0.93	0.18	44,53,57,58	0
2	SO4	M	500	5/5	0.95	0.12	67,67,68,68	0
2	SO4	A	500	5/5	0.96	0.15	71,71,71,72	0
2	SO4	L	500	5/5	0.97	0.10	58,59,60,60	0
2	SO4	K	500	5/5	0.98	0.11	49,49,50,51	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different

orientation to approximate a three-dimensional view.

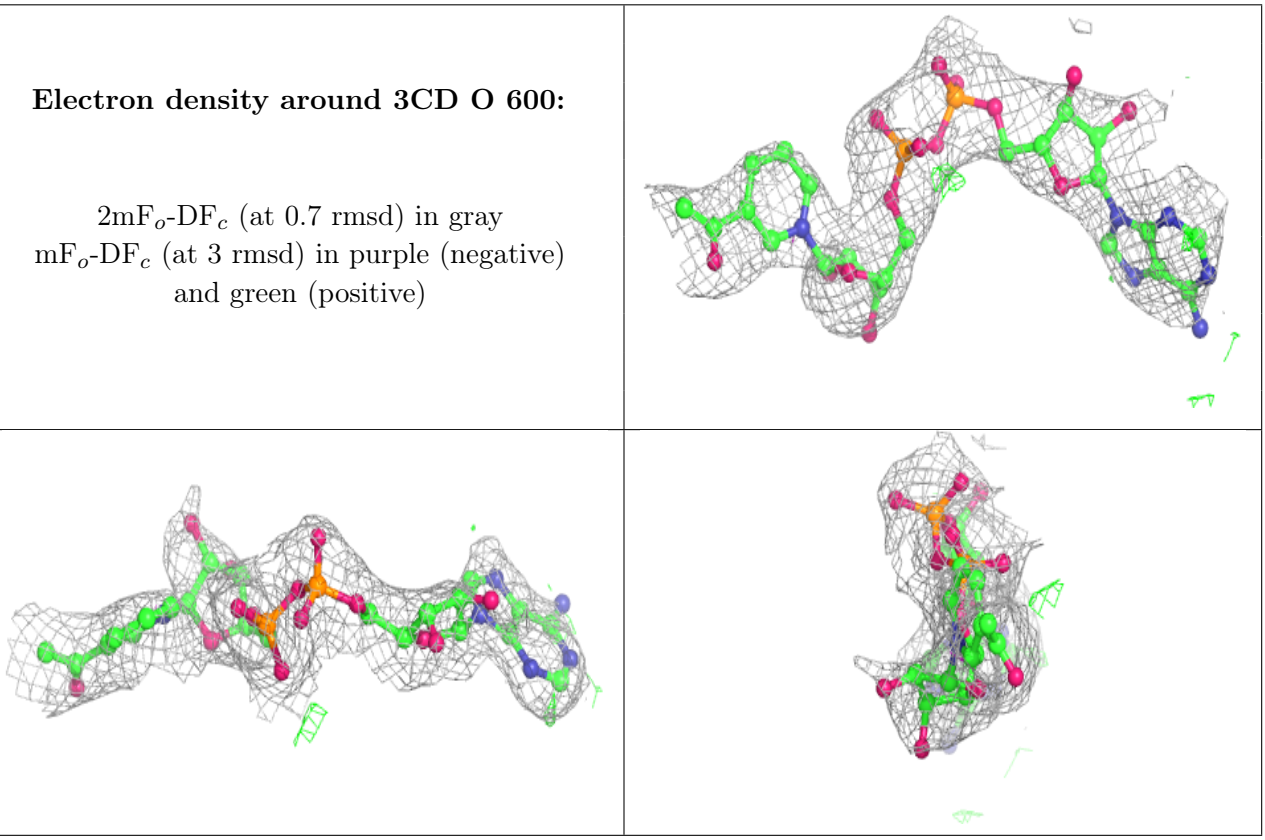
Electron density around 3CD G 600:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



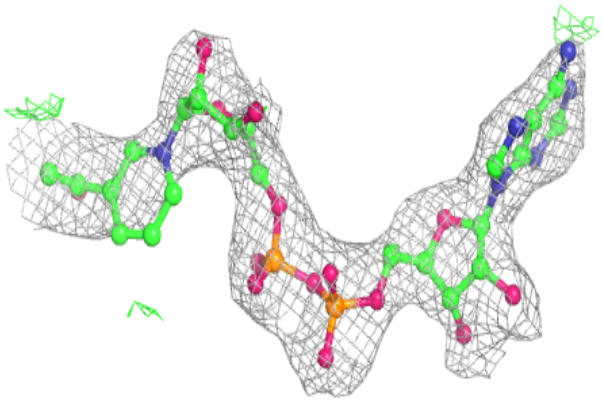
Electron density around 3CD O 600:

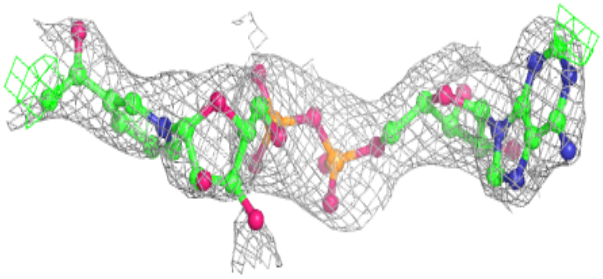
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

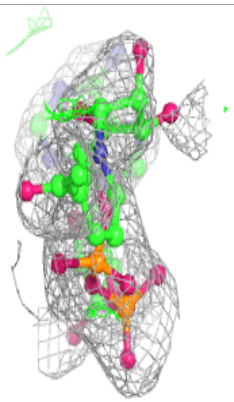


Electron density around 3CD M 600:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

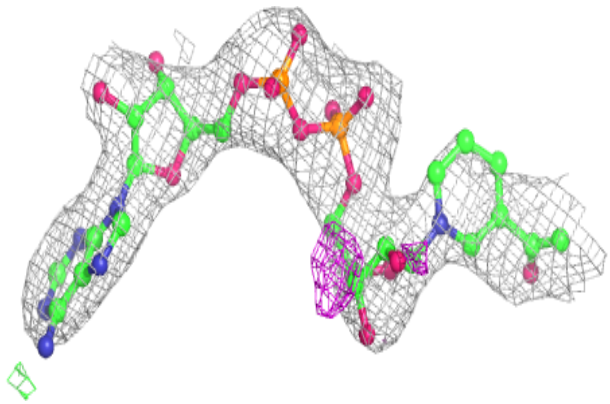


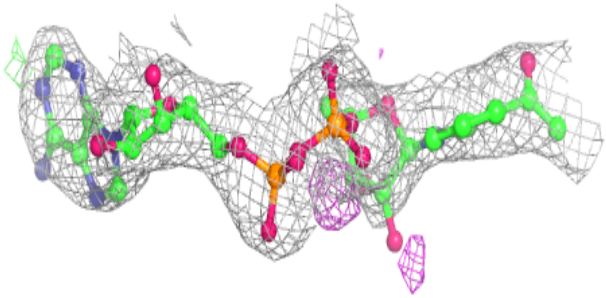


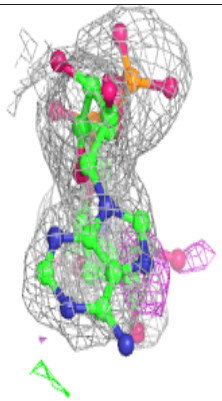


Electron density around 3CD K 600:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

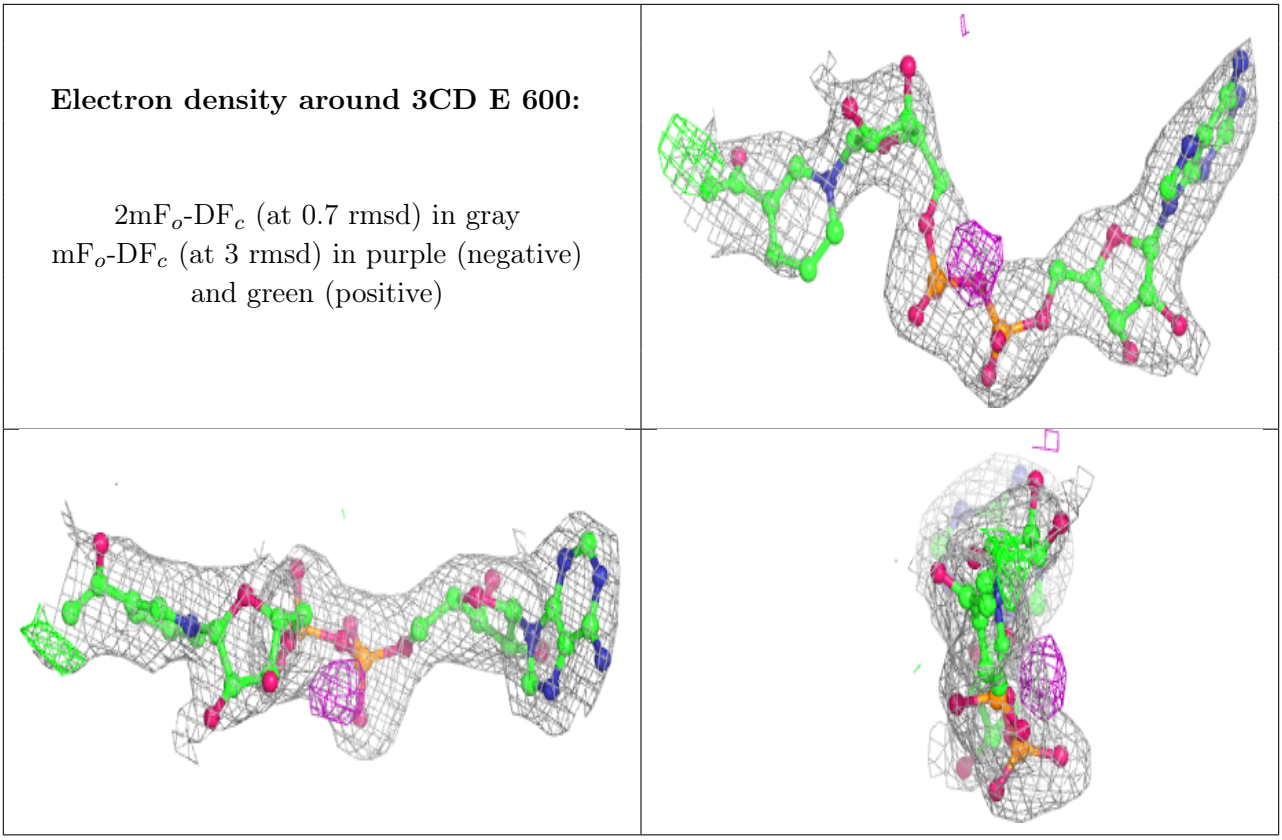






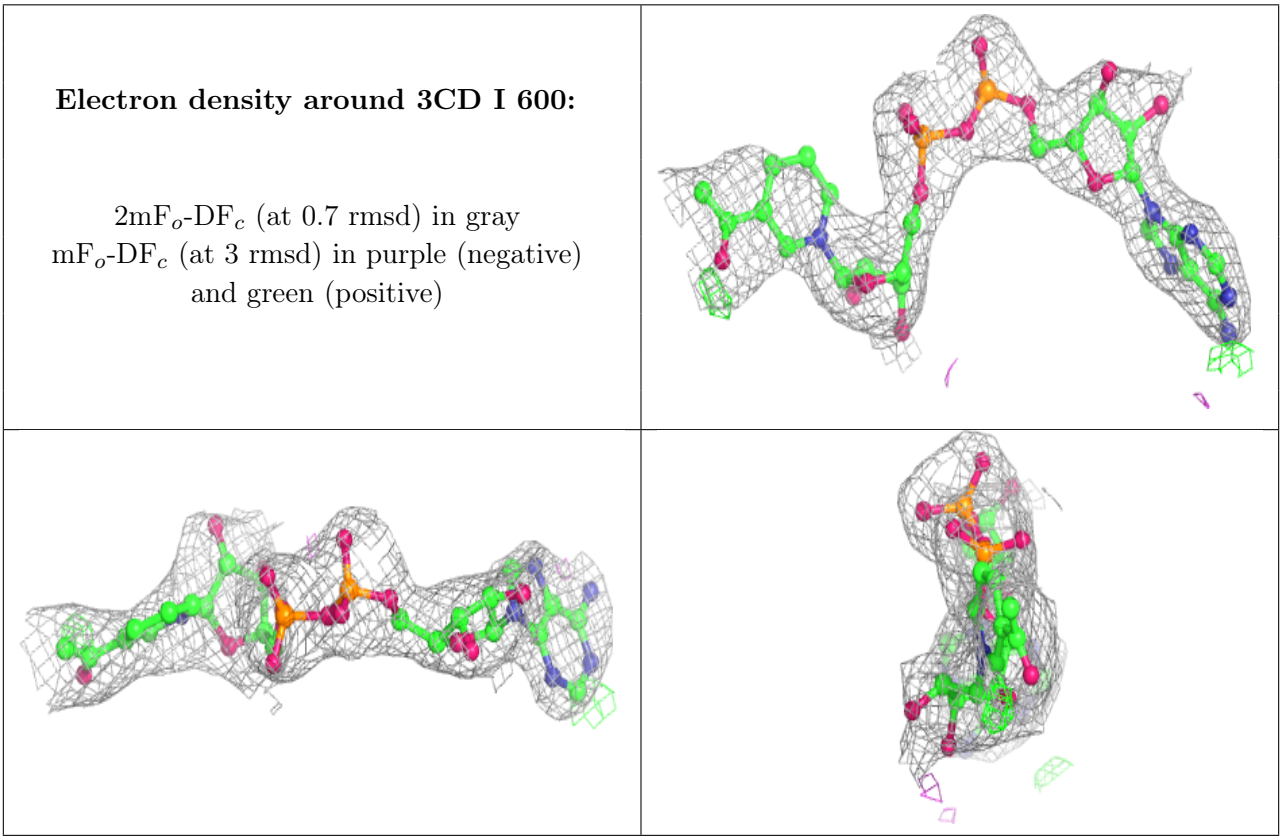
Electron density around 3CD E 600:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



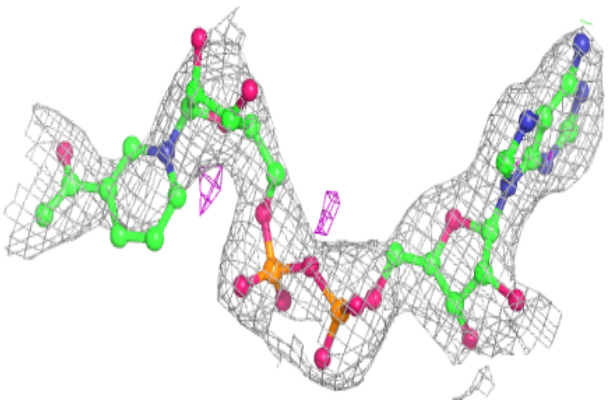
Electron density around 3CD I 600:

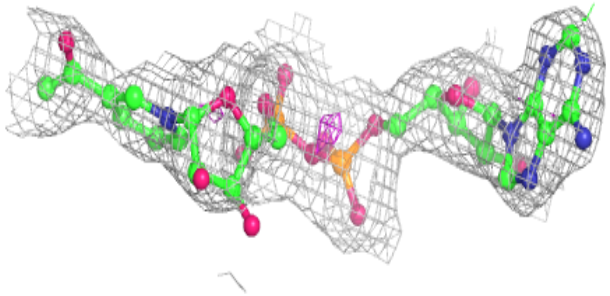
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

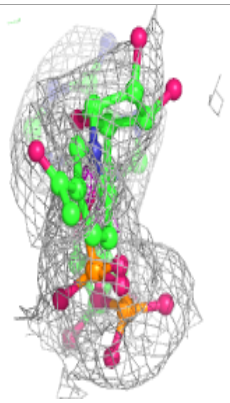


Electron density around 3CD A 600:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

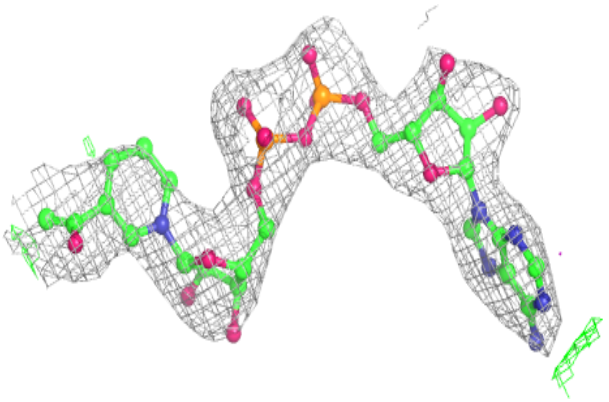


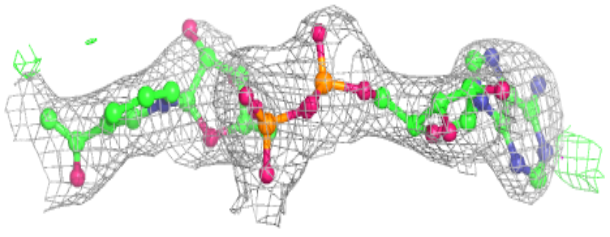


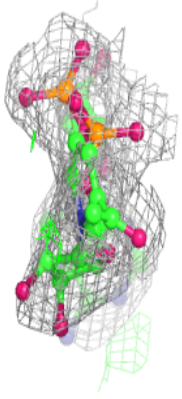


Electron density around 3CD C 600:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.