wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:13 PM GMT

PDB ID	:	1JC1
Title	:	CRYSTAL STRUCTURE ANALYSIS OF A REDOX-SENSITIVE GREEN FLUORESCENT PROTEIN VARIANT IN A OXIDIZED FORM
Authors	:	Hanson, G.T.; Aggeler, R.; Oglesbee, D.; Cannon, M.; Capaldi, R.A.; Tsien, R.Y.; Remington, S.J.
Deposited on	:	2001-06-07
Resolution	:	1.90 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

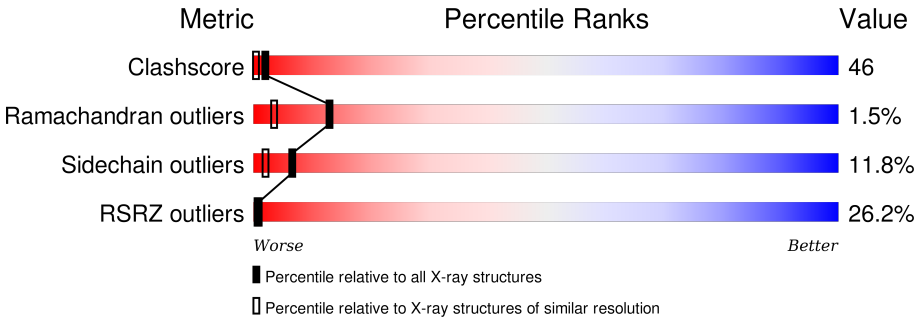
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.90 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))
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	236	13% 40% 43% 11%
1	B	236	15% 40% 40% 13%
1	C	236	47% 22% 53% 20%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5394 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 GREEN FLUORESCENT PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	227	Total	C	N	O	S	0	0	0
			1782	1133	298	343	8			
1	B	226	Total	C	N	O	S	0	0	0
			1760	1119	293	340	8			
1	C	226	Total	C	N	O	S	0	0	0
			1678	1067	276	328	7			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	48	SER	CYS	ENGINEERED	UNP P42212
A	64	LEU	PHE	ENGINEERED	UNP P42212
A	66	CRO	SER	CHROMOPHORE	UNP P42212
A	66	CRO	TYR	CHROMOPHORE	UNP P42212
A	66	CRO	GLY	CHROMOPHORE	UNP P42212
A	80	ARG	GLN	ENGINEERED	UNP P42212
A	147	CYS	SER	ENGINEERED	UNP P42212
A	204	CYS	GLN	ENGINEERED	UNP P42212
B	48	SER	CYS	ENGINEERED	UNP P42212
B	64	LEU	PHE	ENGINEERED	UNP P42212
B	66	CRO	SER	CHROMOPHORE	UNP P42212
B	66	CRO	TYR	CHROMOPHORE	UNP P42212
B	66	CRO	GLY	CHROMOPHORE	UNP P42212
B	80	ARG	GLN	ENGINEERED	UNP P42212
B	147	CYS	SER	ENGINEERED	UNP P42212
B	204	CYS	GLN	ENGINEERED	UNP P42212
C	48	SER	CYS	ENGINEERED	UNP P42212
C	64	LEU	PHE	ENGINEERED	UNP P42212
C	66	CRO	SER	CHROMOPHORE	UNP P42212
C	66	CRO	TYR	CHROMOPHORE	UNP P42212
C	66	CRO	GLY	CHROMOPHORE	UNP P42212
C	80	ARG	GLN	ENGINEERED	UNP P42212
C	147	CYS	SER	ENGINEERED	UNP P42212

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	204	CYS	GLN	ENGINEERED	UNP P42212

- Molecule 2 is water.

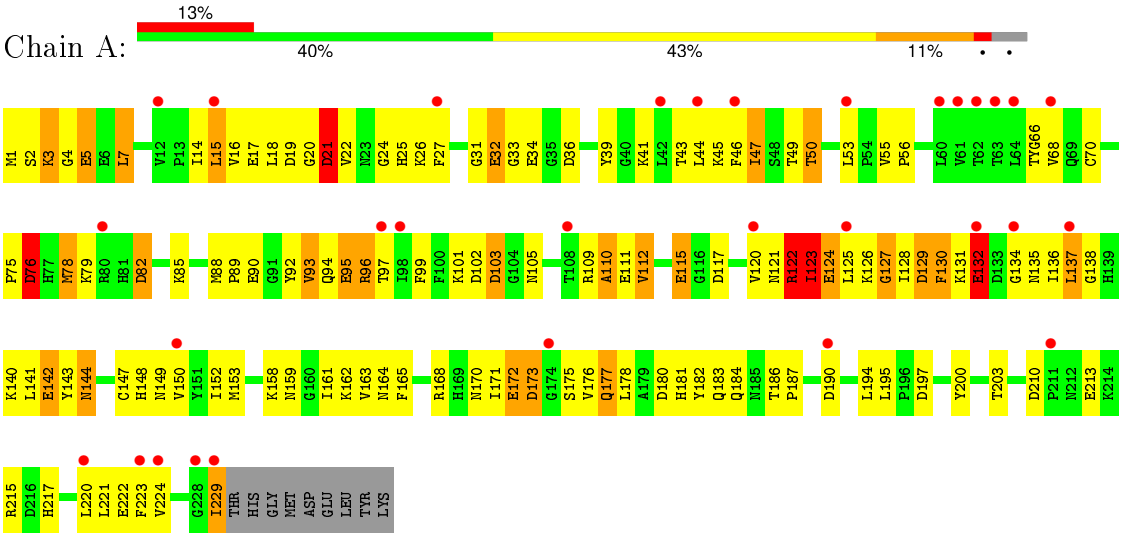
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	93	Total O 93 93	0	0
2	B	58	Total O 58 58	0	0
2	C	23	Total O 23 23	0	0

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.

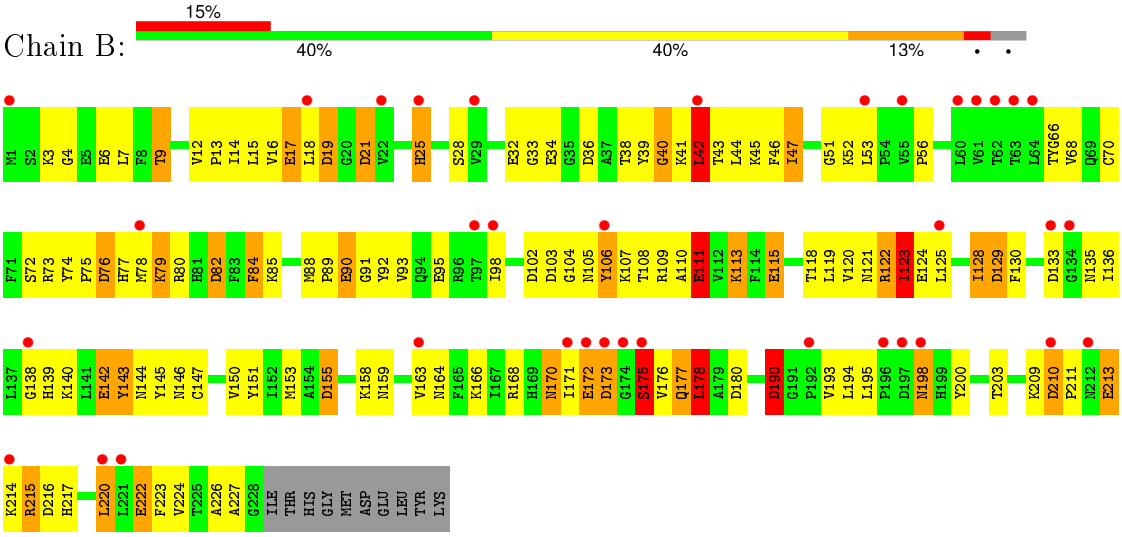
- Molecule 1: GREEN FLUORESCENT PROTEIN

Chain A:



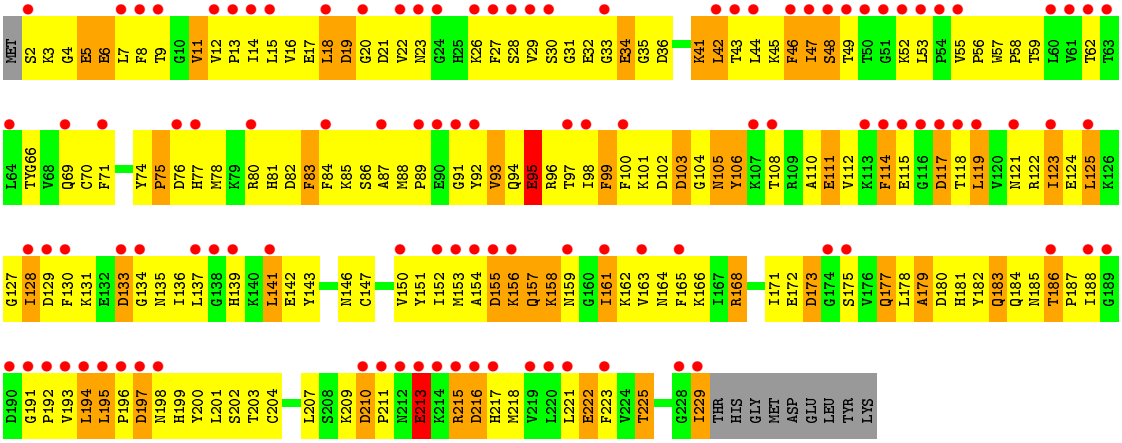
- Molecule 1: GREEN FLUORESCENT PROTEIN

Chain B:



- Molecule 1: GREEN FLUORESCENT PROTEIN

Chain C:


MET S2 K3 G4 E5 E6 L7 F8 T9 G10 V11 V12 P13 I14 L15 V16 E17 L18 D19 G20 D21 V22 N23 G24 H25 K26 F27 S28 V29 S30 G31 E32 G33 E34 G35 D36 K41 L42 T43 L44 K45 F46 I47 S48 T49 T50 G51 K52 L53 P54 V55 P56 W57 P58 T59 L60 V61 T62 T63
L64 TYG66 V68 Q69 C70 F71 Y74 P75 D76 H77 M78 K79 R80 H81 D82 F83 F84 K85 S86 A87 M88 P89 E90 G91 Y92 V93 Q94 E96 R96 T97 I98 F99 F100 K101 D102 D103 G104 N105 Y106 K107 T108 R109 A110 E111 V112 K113 F114 E115 G116 D117 T118 L119 V120 N121 R122 I123 E124 L125 K126
G127 I128 D129 F130 K131 E132 D133 G134 N135 I136 L137 G138 H139 K140 L141 E142 Y143 N146 C147 V150 Y151 I152 M153 A154 D155 K156 Q157 K158 N159 G160 I161 K162 V163 N164 F165 K166 I167 R168 I171 E172 D173 G174 S175 V176 Q177 L178 A179 D180 H181 Y182 Q183 Q184 N185 T186 P187 I188 G189
D190 G191 P192 V193 L194 L195 P196 D197 N198 H199 Y200 L201 S202 T203 C204 L207 S208 K209 D210 P211 N212 E213 K214 R215 D216 H217 M218 V219 L220 L221 E222 F223 V224 T225 G228 I229 THR HIS GLY MET ASP GLU LEU TYR LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	186.84Å 67.61Å 56.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.90 29.71 – 1.90	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-1.90) 99.7 (29.71-1.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.11 (at 1.91Å)	Xtriage
Refinement program	TNT	Depositor
R, R_{free}	0.229 , (Not available) 0.241 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	28.3	Xtriage
Anisotropy	0.378	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 125.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 56783 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5394	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.35% 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, 0.375 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: CRO

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	1.38	14/1800 (0.8%)	1.78	39/2437 (1.6%)
1	B	1.28	13/1778 (0.7%)	1.82	42/2411 (1.7%)
1	C	1.26	11/1695 (0.6%)	1.81	39/2303 (1.7%)
All	All	1.31	38/5273 (0.7%)	1.80	120/7151 (1.7%)

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	3	0
1	B	4	2
1	C	8	0
All	All	15	2

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	32	GLU	CD-OE2	11.45	1.38	1.25
1	B	222	GLU	CD-OE2	8.48	1.34	1.25
1	A	132	GLU	CD-OE2	8.05	1.34	1.25
1	C	213	GLU	CD-OE2	8.02	1.34	1.25
1	A	115	GLU	CD-OE2	7.89	1.34	1.25

The worst 5 of 120 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	21	ASP	CB-CG-OD2	-11.19	108.23	118.30
1	B	180	ASP	CB-CG-OD2	-10.87	108.52	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	36	ASP	CB-CG-OD1	10.24	127.52	118.30
1	C	180	ASP	CB-CG-OD2	-10.22	109.10	118.30
1	A	180	ASP	CB-CG-OD2	-9.60	109.66	118.30

5 of 15 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	9	THR	CB
1	A	15	LEU	CA
1	A	229	ILE	CB
1	B	9	THR	CB
1	B	128	ILE	CB

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	190	ASP	Sidechain
1	B	198	ASN	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	1782	0	1685	112	0
1	B	1760	0	1647	127	0
1	C	1678	0	1488	223	0
2	A	93	0	0	4	1
2	B	58	0	0	8	0
2	C	23	0	0	1	0
All	All	5394	0	4820	460	1

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 46.

The worst 5 of 460 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:C:96:ARG:HG3	1:C:183:GLN:HB2	1.31	1.10
1:B:42:LEU:HD12	1:B:222:GLU:HB3	1.33	1.10
1:B:171:ILE:HD11	1:B:177:GLN:HB2	1.30	1.09
1:C:18:LEU:HD12	1:C:123:ILE:HB	1.33	1.08
1:C:13:PRO:HG2	1:C:118:THR:HA	1.39	1.04

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:312:HOH:O	2:A:312:HOH:O[2_755]	1.63	0.57

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	222/236 (94%)	214 (96%)	7 (3%)	1 (0%)	34 21
1	B	221/236 (94%)	205 (93%)	15 (7%)	1 (0%)	34 21
1	C	221/236 (94%)	189 (86%)	24 (11%)	8 (4%)	4 0
All	All	664/708 (94%)	608 (92%)	46 (7%)	10 (2%)	13 3

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	135	ASN
1	C	192	PRO
1	B	79	LYS
1	C	75	PRO
1	A	132	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	188/206 (91%)	169 (90%)	19 (10%)	9 3
1	B	184/206 (89%)	164 (89%)	20 (11%)	8 2
1	C	164/206 (80%)	140 (85%)	24 (15%)	4 1
All	All	536/618 (87%)	473 (88%)	63 (12%)	6 2

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	128	ILE
1	B	178	LEU
1	C	195	LEU
1	B	155	ASP
1	B	170	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	144	ASN
1	B	146	ASN
1	C	69	GLN
1	A	184	GLN
1	B	170	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

3 non-standard protein/DNA/RNA residues 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	$\#\|Z\|>2$
1	CRO	A	66	1	23,23,24	2.63	10 (43%)	29,32,34	1.84	7 (24%)
1	CRO	B	66	1	23,23,24	2.39	9 (39%)	29,32,34	1.86	6 (20%)
1	CRO	C	66	1	23,23,24	2.43	8 (34%)	29,32,34	1.77	8 (27%)

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
1	CRO	A	66	1	-	0/12/31/32	0/2/2/2
1	CRO	B	66	1	-	0/12/31/32	0/2/2/2
1	CRO	C	66	1	-	0/12/31/32	0/2/2/2

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	CRO	CA3-N3	-5.06	1.38	1.47
1	B	66	CRO	CG2-CB2	-4.02	1.38	1.46
1	C	66	CRO	CG2-CB2	-3.48	1.39	1.46
1	B	66	CRO	OH-CZ	-3.14	1.29	1.37
1	C	66	CRO	OH-CZ	-2.57	1.30	1.37

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	66	CRO	CE2-CZ-CE1	-4.45	113.44	119.79
1	C	66	CRO	CE2-CZ-CE1	-4.08	113.98	119.79
1	B	66	CRO	CE2-CZ-CE1	-3.79	114.39	119.79
1	C	66	CRO	C1-CA1-N1	-2.15	104.45	108.91
1	C	66	CRO	C3-CA3-N3	2.09	117.59	113.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	66	CRO	3	0
1	B	66	CRO	2	0
1	C	66	CRO	2	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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	226/236 (95%)	1.12	31 (13%)	4	4	22, 37, 60, 84	0
1	B	225/236 (95%)	1.22	36 (16%)	3	3	24, 41, 62, 81	0
1	C	225/236 (95%)	2.23	110 (48%)	0	0	31, 57, 88, 98	0
All	All	676/708 (95%)	1.52	177 (26%)	1	1	22, 44, 77, 98	0

The worst 5 of 177 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	229	ILE	6.8
1	C	55	VAL	6.7
1	C	91	GLY	6.4
1	C	48	SER	6.0
1	B	173	ASP	5.9

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

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^{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	LLDF	B-factors(Å^2)	Q<0.9
1	CRO	A	66	22/23	0.92	0.21	-	14,28,38,70	0
1	CRO	B	66	22/23	0.93	0.20	-	19,25,33,52	0
1	CRO	C	66	22/23	0.87	0.22	-	22,41,56,60	0

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.