wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 06:07 AM GMT

PDB ID	:	2VZ8
Title	:	CRYSTAL STRUCTURE OF MAMMALIAN FATTY ACID SYNTHASE
Authors	:	Maier, T.; Leibundgut, M.; Ban, N.
Deposited on	:	2008-07-31
Resolution	:	3.22 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:

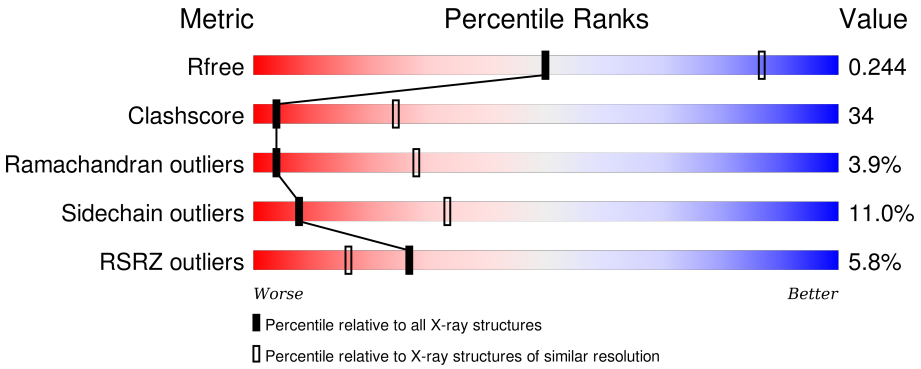
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.22 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	1095 (3.26-3.18)
Clashscore	102246	1046 (3.24-3.20)
Ramachandran outliers	100387	1026 (3.24-3.20)
Sidechain outliers	100360	1025 (3.24-3.20)
RSRZ outliers	91569	1100 (3.26-3.18)

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	2512	4% 35% 36% 6% 22%
1	B	2512	6% 36% 37% 6% 20%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 30281 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 FATTY ACID SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1962	Total	C	N	O	S	0	0	0
			14977	9466	2630	2803	78			
1	B	2004	Total	C	N	O	S	0	0	0
			15304	9671	2684	2869	80			

There are 2 discrepancies between the modelled and reference sequences:

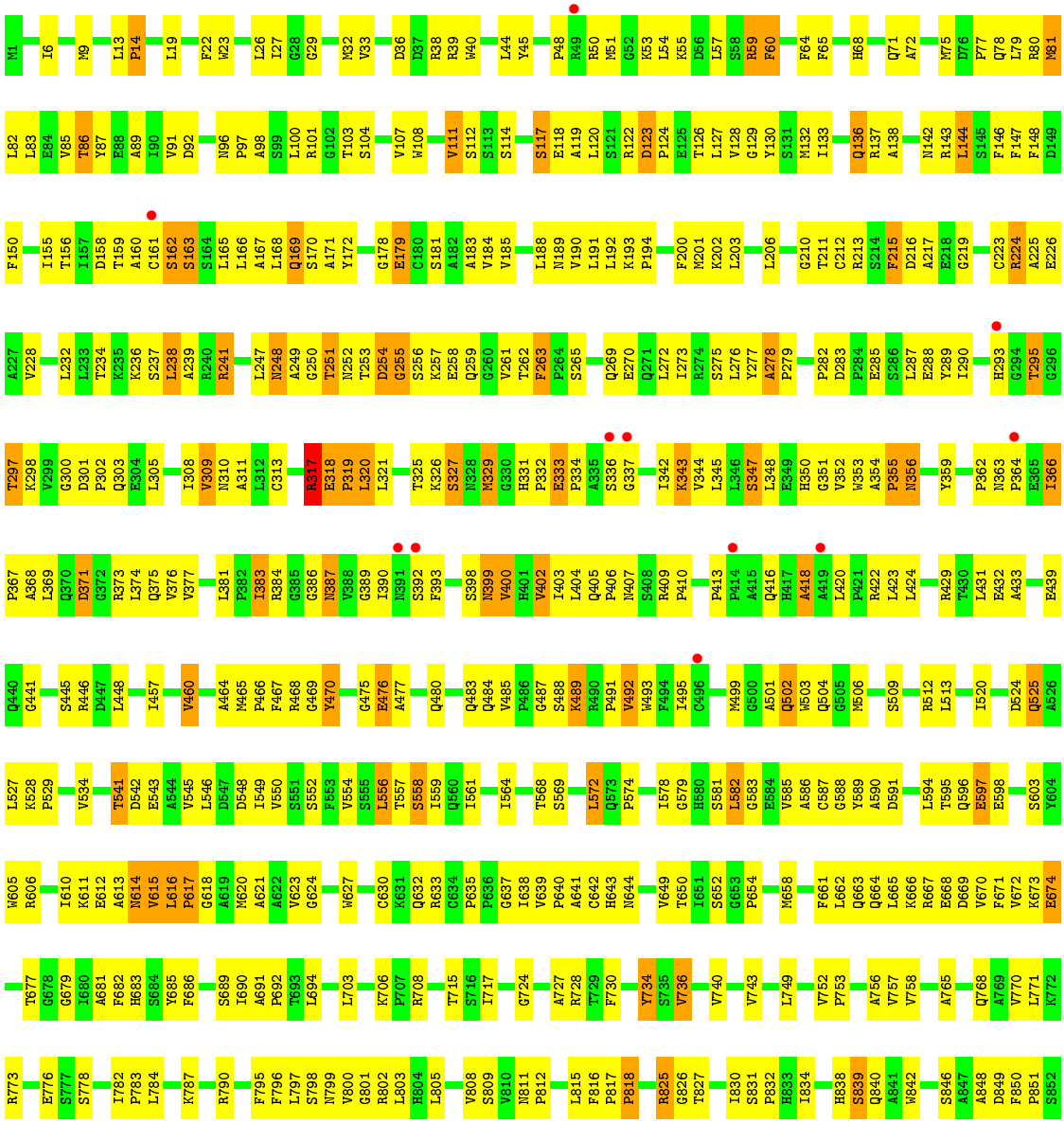
Chain	Residue	Modelled	Actual	Comment	Reference
A	834	ILE	UNK	CONFLICT SEE REMARK 9	UNP A5YV76
B	834	ILE	UNK	CONFLICT SEE REMARK 9	UNP A5YV76

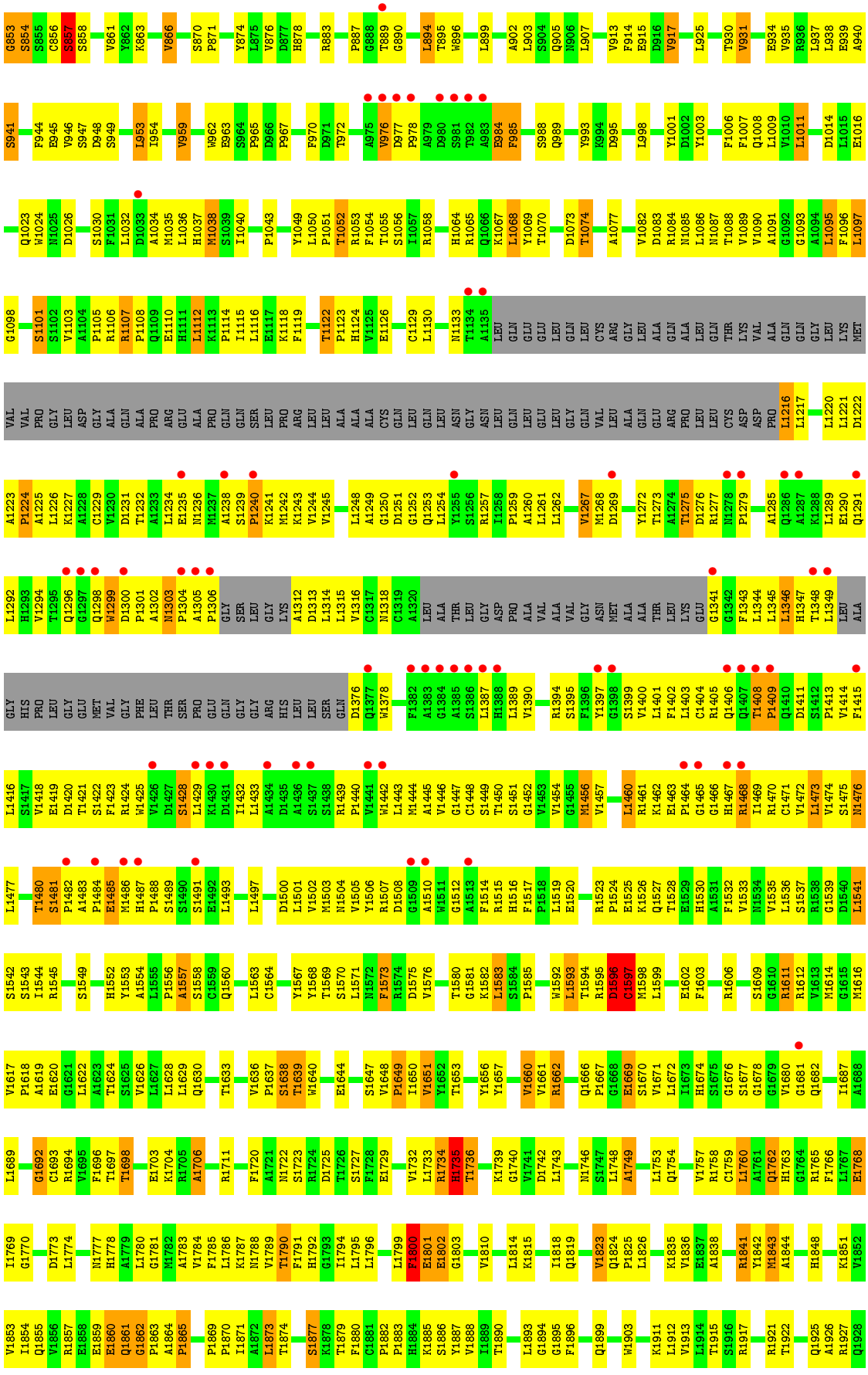
3 Residue-property plots (i)

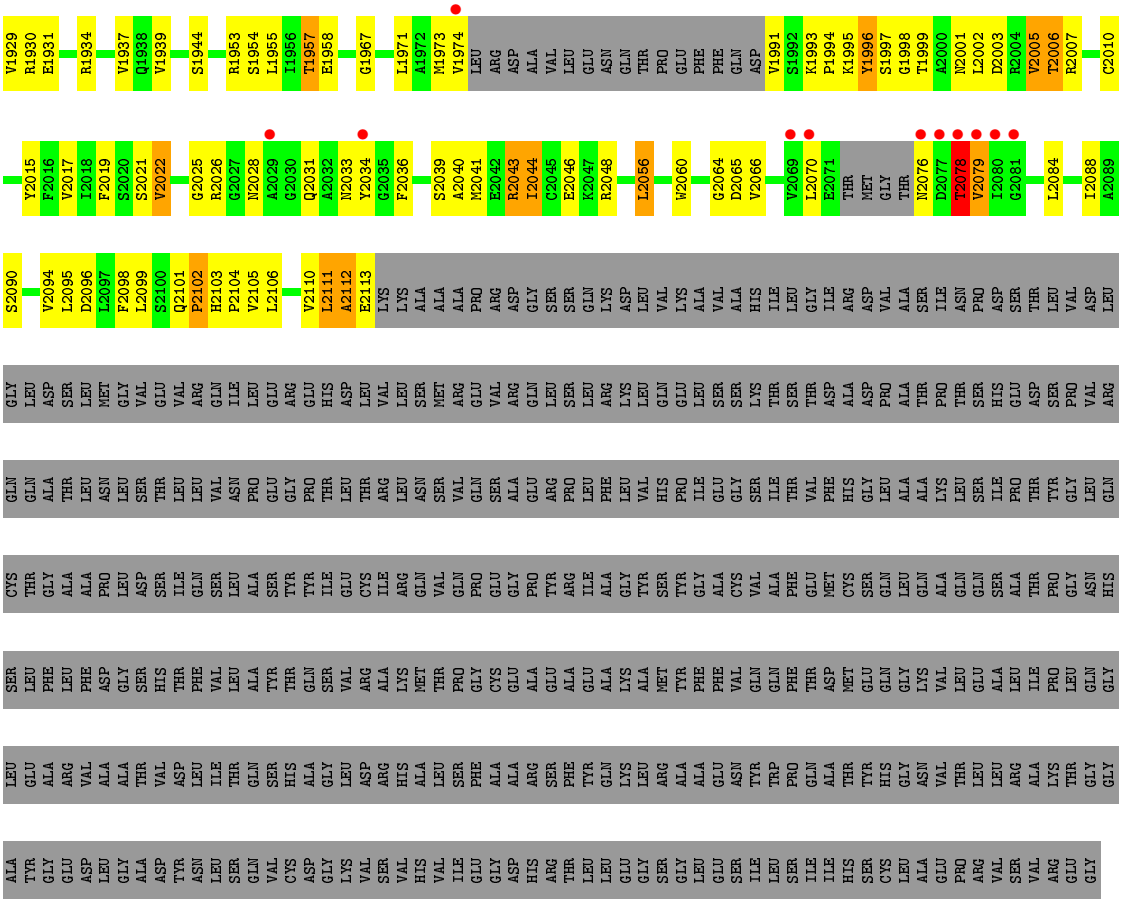
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: FATTY ACID SYNTHASE

Chain A:

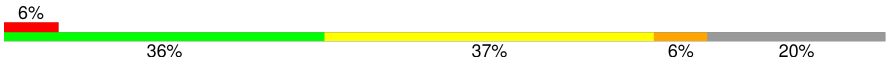


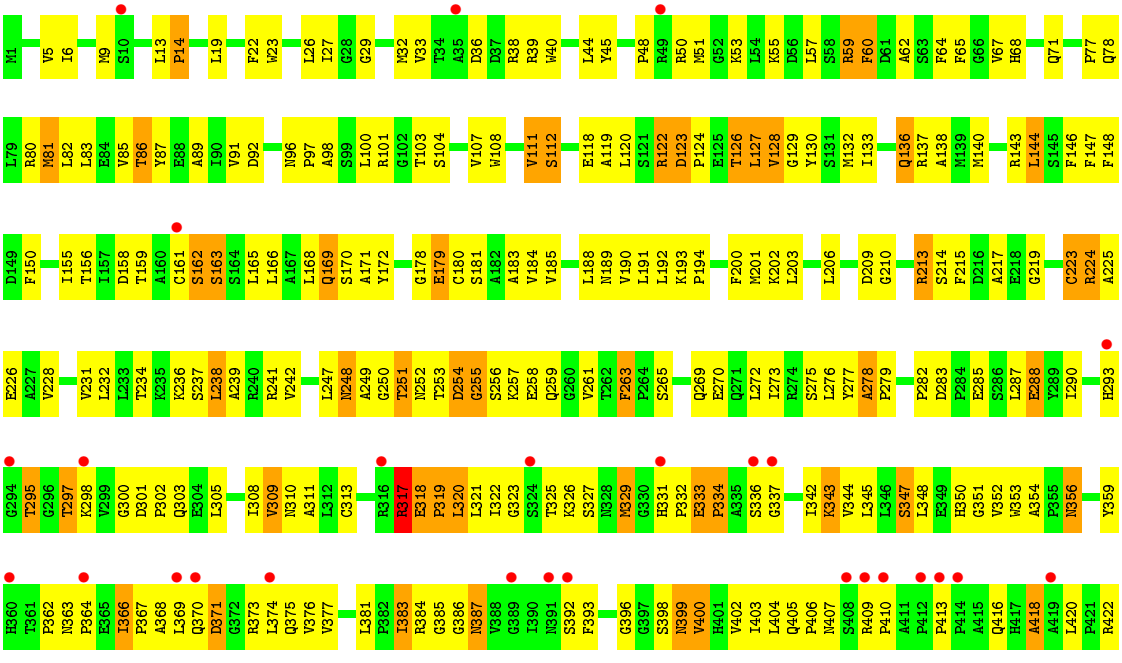


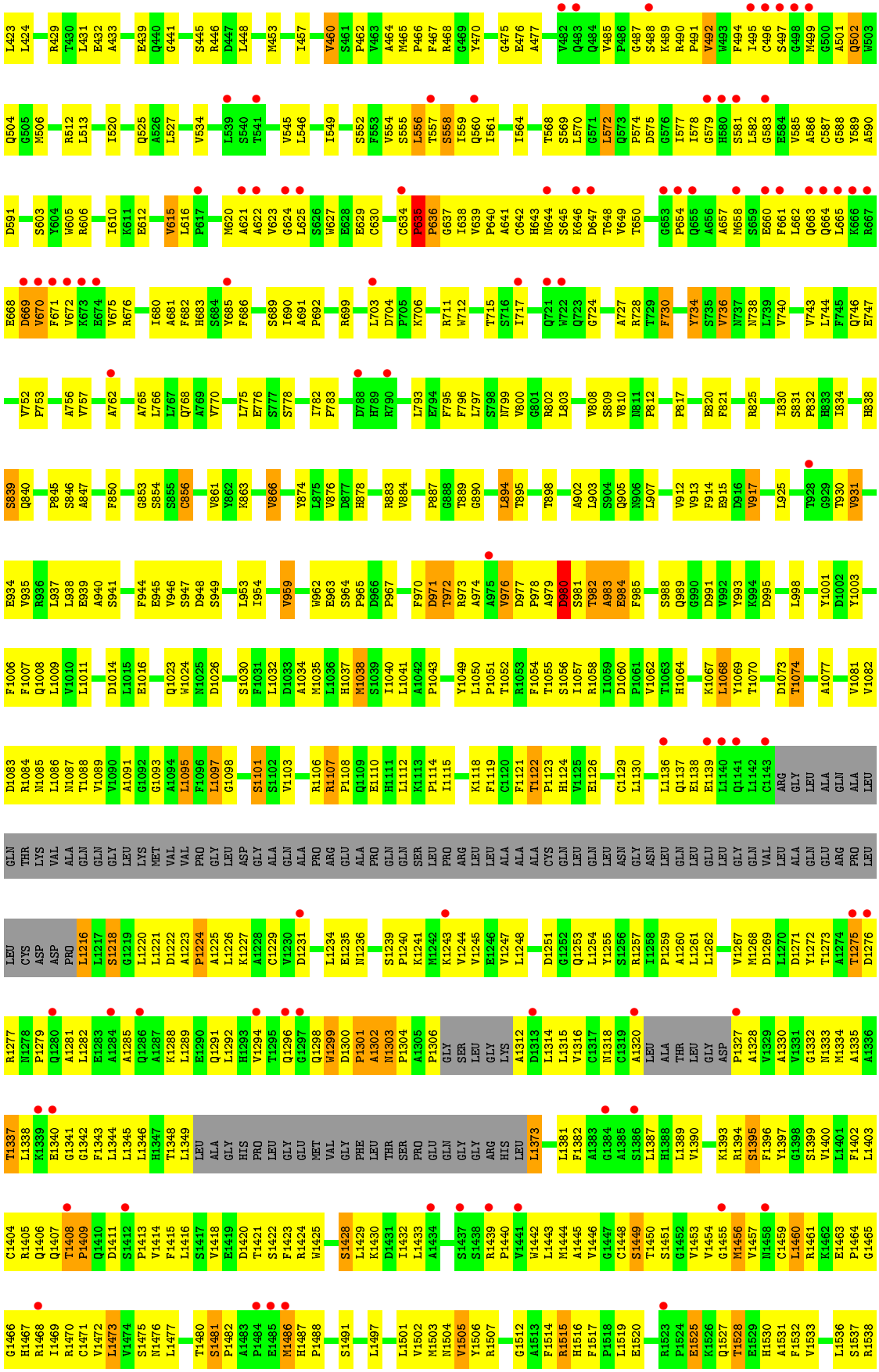


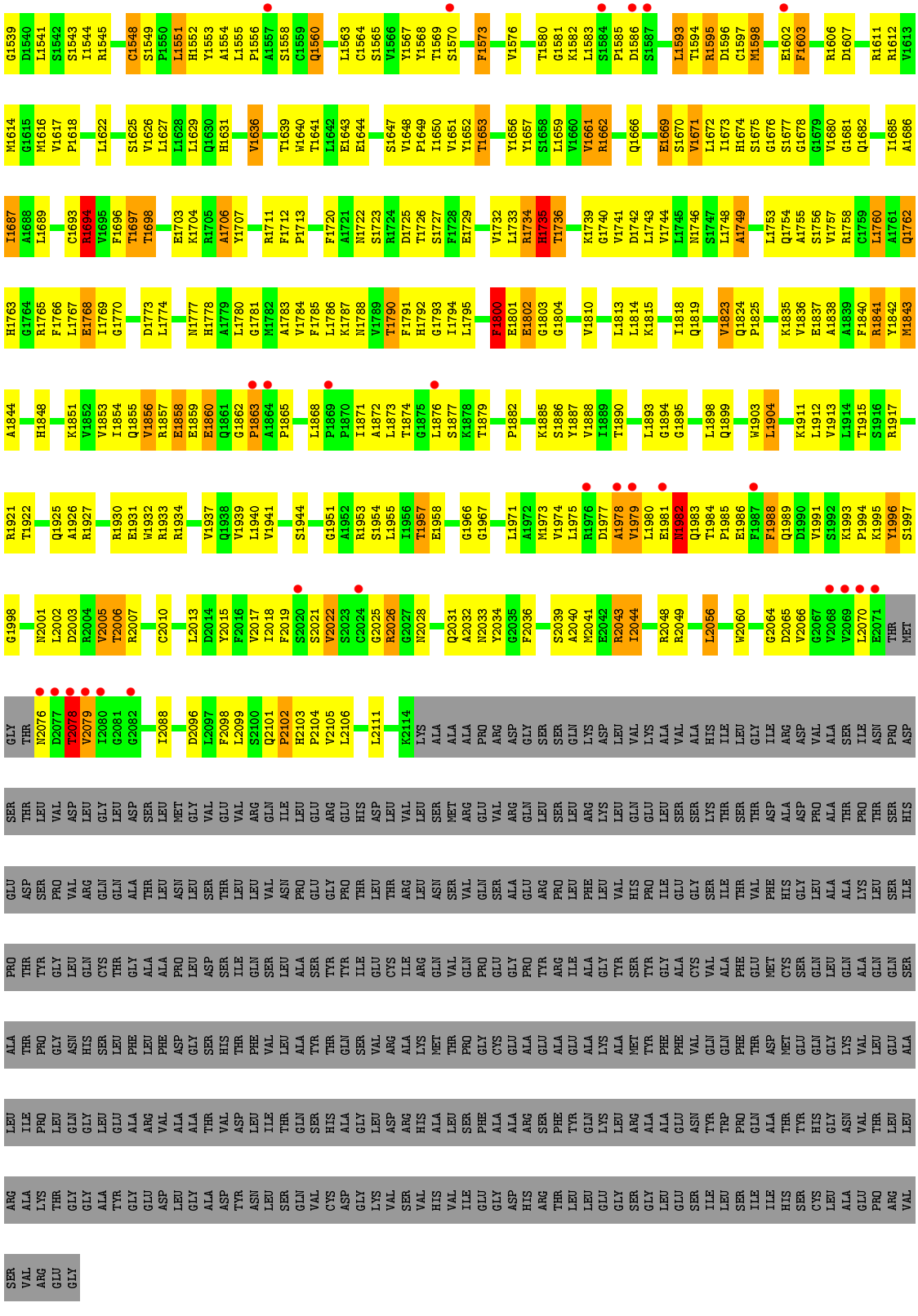
• Molecule 1: FATTY ACID SYNTHASE

Chain B:








G1539 D1540 L1541 S1542 S1543 I1544 R1545 C1548 S1549 P1550 L1551 H1552 Y1553 A1554 L1555 P1556 A1557 S1558 C1559 Q1560 L1563 C1564 S1565 V1566 Y1567 Y1568 T1569 S1570 F1573 V1576 T1580 G1581 K1582 L1583 S1584 P1585 D1586 S1587 L1593 T1594 R1595 D1596 C1597 M1598 E1602 F1603 R1606 D1607 R1611 R1612 V1613
M1614 G1615 M1616 V1617 P1618 L1622 S1625 V1626 L1627 L1628 L1629 Q1630 H1631 V1636 T1639 W1640 T1641 L1642 E1643 E1644 S1647 V1648 P1649 I1650 V1651 Y1652 T1653 Y1656 Y1657 S1658 L1659 V1660 V1661 R1662 Q1666 E1669 S1670 V1671 L1672 I1673 H1674 S1675 G1676 S1677 G1678 G1679 V1680 G1681 Q1682 I1685 A1686
I1687 A1688 L1689 C1693 R1694 V1695 F1696 T1697 T1698 E1703 K1704 R1705 A1706 Y1707 R1711 F1712 P1713 F1720 A1721 N1722 S1723 R1724 D1725 T1726 S1727 F1728 E1729 V1732 L1733 R1734 H1735 T1736 K1739 G1740 V1741 D1742 L1743 V1744 L1745 N1746 S1747 L1748 A1749 L1753 Q1754 A1755 S1756 V1757 R1758 C1759 L1760 A1761 Q1762
H1763 G1764 R1765 F1766 L1767 E1768 I1769 G1770 D1773 L1774 N1777 H1778 A1779 L1780 G1781 M1782 A1783 V1784 F1785 L1786 K1787 N1788 V1789 T1790 F1791 H1792 G1793 I1794 L1795 P1800 E1801 E1802 G1803 G1804 V1810 L1813 L1814 K1815 I1818 Q1819 V1823 Q1824 P1825 K1835 V1836 E1837 A1838 A1839 F1840 R1841 Y1842 M1843
A1844 H1848 K1851 V1852 V1853 I1854 Q1855 V1856 R1857 E1858 E1859 E1860 Q1861 G1862 P1863 A1864 P1865 L1868 P1869 P1870 I1871 A1872 L1873 T1874 G1875 L1876 S1877 K1878 T1879 P1882 K1885 S1886 Y1887 V1888 I1889 T1890 L1893 G1894 G1895 L1898 Q1899 W1903 L1904 K1911 L1912 V1913 L1914 T1915 S1916 R1917
R1921 T1922 Q1925 A1926 R1927 R1930 E1931 W1932 R1933 R1934 V1937 Q1938 V1939 L1940 V1941 S1944 G1951 A1952 R1953 S1954 L1955 I1956 T1957 E1958 G1966 G1967 L1971 A1972 M1973 V1974 L1975 R1976 D1977 A1978 V1979 L1980 E1981 N1982 Q1983 T1984 P1985 E1986 F1987 F1988 Q1989 D1990 V1991 S1992 K1993 P1994 K1995 Y1996 S1997
G1998 N2001 L2002 D2003 R2004 V2005 T2006 R2007 C2010 L2013 D2014 Y2015 F2016 V2017 I2018 F2019 S2020 S2021 V2022 S2023 C2024 G2025 R2026 G2027 N2028 Q2031 A2032 N2033 Y2034 G2035 F2036 S2039 A2040 M2041 E2042 R2043 I2044 R2048 R2049 L2056 W2060 G2064 D2065 V2066 G2067 V2068 V2069 L2070 E2071 THR MET
GLY THR N2076 D2077 T2078 V2079 I2080 G2081 G2082 I2088 D2096 L2097 F2098 L2099 S2100 Q2101 P2102 H2103 P2104 V2105 L2106 L2111 K2114 LYS ALA ALA ALA PRO ARG ASP GLY SER SER GLN LYS ASP LEU VAL LYS ALA VAL ALA HIS ILE LEU GLY ILE ARG ASP VAL ALA SER ILE ASN PRO ASP
SER THR LEU VAL ASP LEU GLY LEU ASP SER LEU MET GLY VAL GLU VAL ARG GLU HIS ASP LEU VAL LEU SER MET ARG GLU VAL ARG GLN LEU SER LEU ARG LYS LEU VAL GLN LEU GLU LEU SER SER LYS THR SER THR VAL ASP ALA ASP PRO ALA THR PRO THR SER HIS
GLU ASP SER PRO VAL ARG GLN ALA THR LEU ASN LEU SER THR LEU LEU VAL ASN PRO GLU GLY PRO THR LEU THR ARG LEU ASN SER VAL GLN SER ALA GLU ARG PRO LEU PHE LEU VAL HIS PRO ILE GLU GLY SER ILE THR VAL PHE HIS GLY LEU ALA ALA LYS LEU SER ILE
PRO THR TYR GLY LEU GLN CYS THR GLY ALA ALA PRO LEU ASP SER ILE GLN LEU ALA SER TYR TYR ILE GLU CYS ILE ARG GLN VAL GLN PRO GLU GLY PRO TYR ARG ILE ALA GLY TYR SER TYR GLY ALA CYS VAL ALA PHE GLU MET CYS SER GLN LEU GLN ALA GLN GLN SER
ALA THR PRO GLY ASN HIS SER LEU PHE LEU PHE ASP GLY SER HIS THR PHE VAL LEU ALA TYR THR GLN SER VAL ARG ALA LYS MET THR PRO GLY CYS GLU ALA ARG GLU ALA ALA LYS ALA MET TYR PHE PHE VAL GLN GLN PHE THR ASP MET GLU GLN GLY LYS VAL LEU GLU GLU
LEU ILE PRO LEU GLN GLY LEU GLU ALA ARG VAL ALA ALA THR VAL ASP LEU ILE THR GLN SER HIS ALA GLY LEU ASP ARG HIS ALA LEU SER PHE ALA ALA ARG SER PHE TYR GLN LYS LEU ARG ALA ALA GLU GLU ASN TYR TRP PRO GLN PHE THR ASP ILE ILE HIS LEU LEU
ARG ALA LYS THR GLY GLY GLU ASP LEU ASN LEU SER GLN VAL CYS ASP GLY LYS VAL SER VAL HIS VAL ILE GLU GLY ASP HIS ARG THR LEU LEU GLU GLY SER GLY LEU GLU SER ILE LEU SER ILE ILE HIS SER CYS LEU ALA GLU PRO ARG VAL
SER VAL ARG GLU GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	96.32Å 244.70Å 135.25Å 90.00° 101.65° 90.00°	Depositor
Resolution (Å)	29.50 – 3.22 29.50 – 3.22	Depositor EDS
% Data completeness (in resolution range)	94.8 (29.50-3.22) 97.6 (29.50-3.22)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 3.24Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.217 , 0.259 0.202 , 0.244	Depositor DCC
R_{free} test set	4841 reflections (5.02%)	DCC
Wilson B-factor (Å^2)	95.2	Xtriage
Anisotropy	0.218	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 73.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Outliers	0 of 96482 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	30281	wwPDB-VP
Average B, all atoms (Å^2)	143.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.45	1/15302 (0.0%)	0.63	0/20792
1	B	0.41	0/15634	0.60	1/21243 (0.0%)
All	All	0.43	1/30936 (0.0%)	0.61	1/42035 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1759	CYS	CB-SG	-5.79	1.72	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1694	ARG	NE-CZ-NH1	6.90	123.75	120.30

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	14977	0	14938	1049	0
1	B	15304	0	15266	1083	0
All	All	30281	0	30204	2085	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 34.

The worst 5 of 2085 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:1694:ARG:HH11	1:B:1694:ARG:HG3	1.17	1.10
1:B:1303:ASN:H	1:B:1304:PRO:HD3	1.25	1.02
1:A:1473:LEU:HD21	1:A:1503:MET:HG2	1.36	1.02
1:A:165:LEU:HD23	1:A:400:VAL:HG22	1.37	1.02
1:B:1456:MET:HG2	1:B:2036:PHE:HB2	1.41	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	1948/2512 (78%)	1586 (81%)	282 (14%)	80 (4%)	3 27
1	B	1992/2512 (79%)	1622 (81%)	296 (15%)	74 (4%)	4 29
All	All	3940/5024 (78%)	3208 (81%)	578 (15%)	154 (4%)	4 28

5 of 154 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	179	GLU
1	A	255	GLY
1	A	278	ALA
1	A	333	GLU
1	A	413	PRO

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	1624/2072 (78%)	1450 (89%)	174 (11%)	8	34
1	B	1660/2072 (80%)	1473 (89%)	187 (11%)	7	32
All	All	3284/4144 (79%)	2923 (89%)	361 (11%)	8	33

5 of 361 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1996	TYR
1	B	318	GLU
1	B	1823	VAL
1	A	2028	ASN
1	B	127	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 70 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2086	GLN
1	B	350	HIS
1	B	1777	ASN
1	A	2103	HIS
1	B	136	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	1962/2512 (78%)	-0.10	91 (4%)	36	24	63, 118, 226, 276	0
1	B	2004/2512 (79%)	0.12	141 (7%)	19	12	54, 158, 230, 276	0
All	All	3966/5024 (78%)	0.01	232 (5%)	26	16	54, 136, 229, 276	0

The worst 5 of 232 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	581	SER	8.1
1	B	496	CYS	7.6
1	B	579	GLY	7.6
1	A	1297	GLY	7.5
1	B	498	GLY	7.3

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.