Putting MicroED to the test: an unabridged account of the evaluation of 30 diverse pharma compounds

Jessica E. Burch*, Austin G. Smith, Seb Caille, Shawn D. Walker, Ryan Wurz, Victor J. Cee, Jose Rodriguez, D. Gostovic, Kyle Quasdorf*, Hosea M. Nelson*

Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA 90095, USA.
Pivotal and Commercial Drug Substance Technologies, Process Development, Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320, USA.
MicroEDLab.com, 1623 Central Ave. Ste. 18, Cheyenne, Wyoming 82001, USA.

Supporting Information

Table of Contents

1. Materials and Methods

Samples prepared according to previously disclosed procedures outlined in **Supporting Table 1** below. Data was collected on a Thermo Fisher Talos F200C transmission electron microscope operating with an accelerating voltage of 200keV, corresponding to an electron wavelength of 0.0251 Å. Electron diffraction data was collected using a Thermo Fisher CetaD camera. Screening the TEM grid for microcrystals was performed at 2600x magnification in imaging mode. Crystals selected for data collection were isolated by a selected area aperture. Data was collected by taking images of the diffraction patterns generated by a continuously rotating crystal integrated continuously at a rate of 3 seconds per frame. This rotation was performed at a rate of 0.3° per second with a minimum and maximum tilt range of −65° to +65°. Crystals selected for data collection were isolated by a selected area aperture to reduce the background noise contributions and calibrated to eucentric height to stay in the aperture over the entire tilt range. Samples collected at cryogenic conditions were placed onto a Gatan 626 cryo holder. Slow cooling the sample includes inserting room temperature Gatan 626 cryo holder and cooling to cryogenic temperatures after insertion into the TEM. Plunge frozen samples were frozen in liquid nitrogen, placed onto a liquid nitrogen cooled Gatan 626 cryo holder, and inserted and maintained at cryogenic temperature for the duration of data collection on the electron microscope. All diffraction data was processed using the XDS suite of programs as controlled by a custom Python automation script.[1–3] Structure were solved *ab initio* by direct methods in SHELXT or SHELXD and refined with SHELXL using ShelXle.[4–7] Thermal parameters were refined anisotropically for all non-hydrogen atoms. Hydrogen atoms were assigned using the riding model.

Compound	Purification by Crystallization	Compound Source and References
1	Yes	Medicinal Chemistry[8]
2	No	Medicinal Chemistry[8,9]
3	Yes	Medicinal Chemistry[8]
4	No	Medicinal Chemistry[8,9]
5	Yes	Process Chemistry[10]
6	Yes	Medicinal Chemistry[8,9,11–15]
7	Yes	Process Chemistry[10]
8	Yes	Process Chemistry[10]
9	No	Process Chemistry[10]
10	Yes	Process Chemistry[10]
11	No*	Medicinal Chemistry[11,13]
12	No	Medicinal Chemistry[8,9,11–15]
13	No*	Medicinal Chemistry[11,14]
14	Yes*	Process Chemistry[3]
15	Yes*	Process Chemistry[16]
Figure 5a	No	Medicinal Chemistry[17–19]
Figure 5b	No	PROTACs[20]

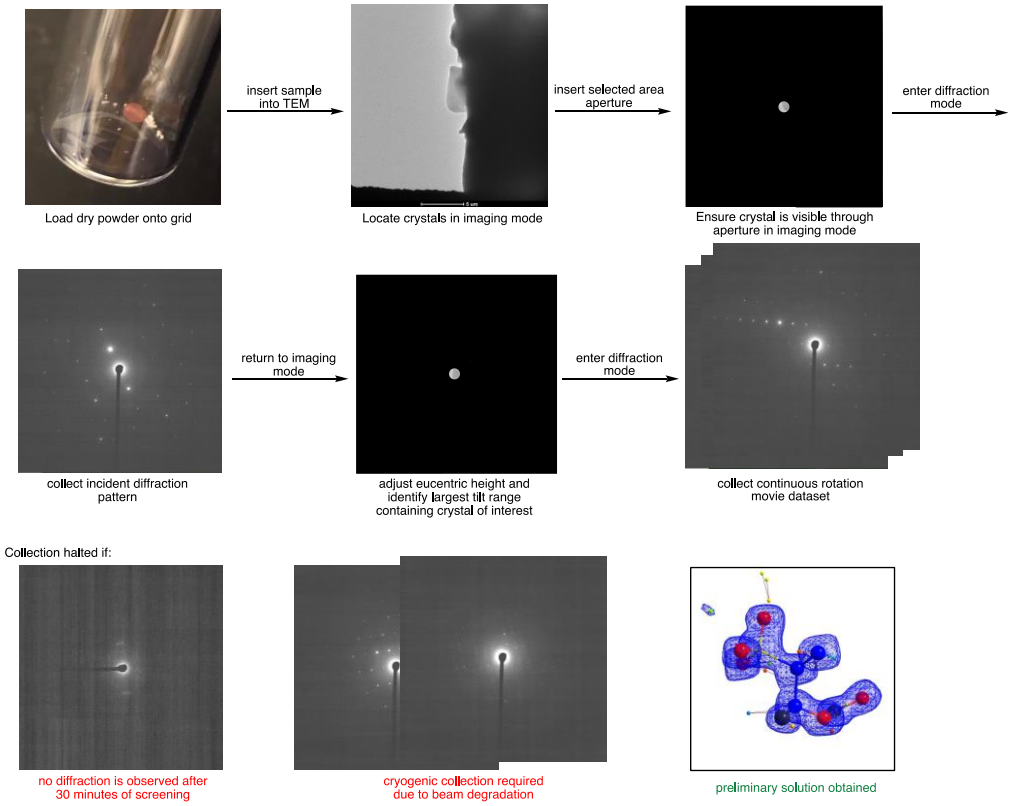
Supporting Table 1. Source of thirty pharmaceutical compounds analyzed in this study. * = samples were recrystallized for the purpose of obtaining a crystal structure.

2. Room Temperature TEM Screening Procedure

Milligram to sub-milligram quantities of dry powder were placed into a dram vial as received and manually ground with a glass pipette. A pure carbon 200 mesh Cu grid or lacey carbon Cu grid was placed inside of the vial and gently shaken together with the powder to "dry load" the grid (**1**, Supporting Figure 1). The grid was removed with Dumont straight self-closing tweezers and the tweezers were gently tapped against a lab bench while holding the grid to shake off excess powder. This sample was clipped into a single tilt holder and inserted into a well-aligned Thermo Fisher Scientific Talos F200C transmission electron microscopy operating at an accelerating voltage of 200keV.

After achieving suitable pressure, the column valves were opened and the grid was manually scanned at 2600x magnification in imaging mode (**2**). To screen for crystallinity, an incident diffraction pattern was recorded by isolating a region of the particle using a selected area aperture (**3**) and entering parallel-illuminated diffraction mode utilizing the low dose software on the Thermo Fisher microscope user interface.

A single image of the diffraction pattern was taken on a Thermo Fisher Scientific Ceta-D camera (**4**). If user inspection of the diffraction pattern suggested that the particle was monocrystalline and provided <1.2 Å resolution diffraction, the eucentric height of the sample was finely adjusted in imaging mode to ensure the crystal would remain within the selected area aperture throughout a tilt series with a maximum tilt range of ± 65° (**5**). Upon returning to diffraction mode at eucentric height, a continuously rotating electron diffraction movie was collected (**6**). The stage was rotated at a rate of 0.3° s^{-1} and a detector distance of 960mm. The Ceta-D CMOS 4k x 4k camera was operated using



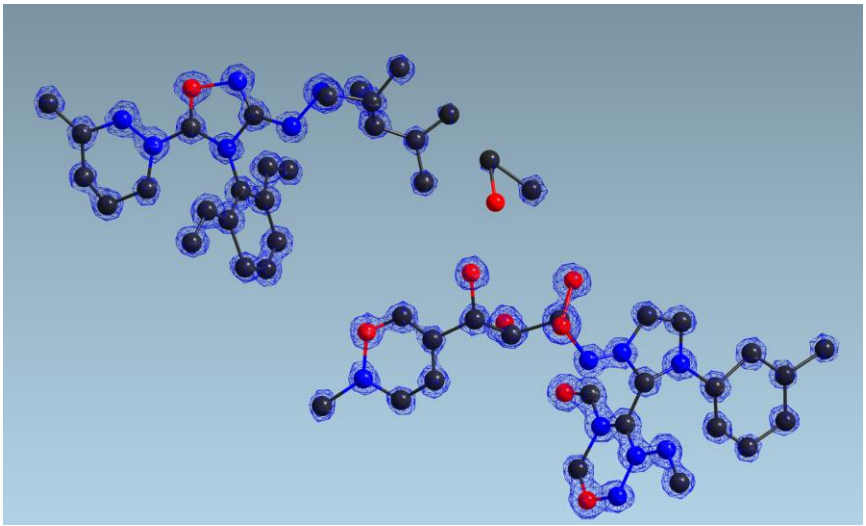
Supporting Figure 1. Representative data collection workflow.

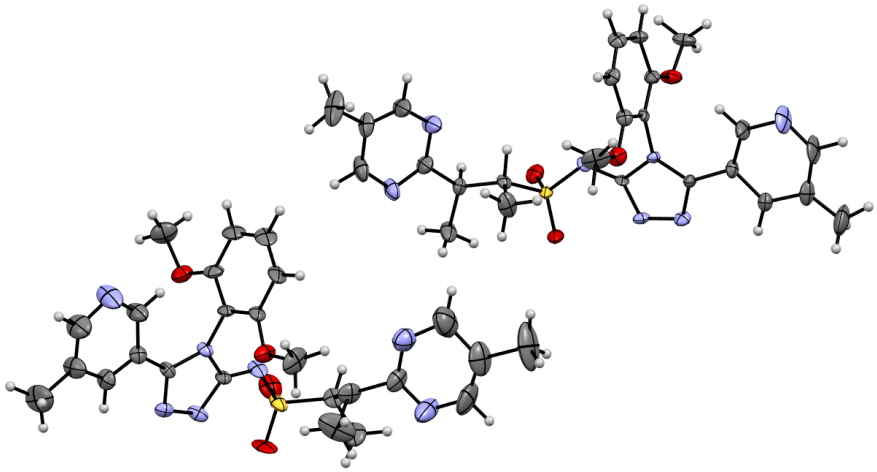
rolling shutter mode and continuously integrated at a rate of 3 seconds per frame with binning by 2 to produce 2k x 2k images. Diffraction movies were saved as SER files. Movies were saved with a standardized naming format and processed using the automated data workflow (See Supporting Information **9** and **10**) while additional movies were collected. These processed movies were manually re-indexed to different space groups and/or merged with other datasets as needed until preliminary solutions were obtained.

Screening was halted if no diffraction was observed after 30 minutes, the sample visibly lost resolution over the course of a single movie, or a preliminary solution with >90% of expected atoms was obtained.

3. Room Temperature Screening Crystal Structures

3.1 (2R,3S)-N-(4-(2,6-dimethoxyphenyl)-5-(5-methylpyridin-3-yl)-4H-1,2,4-triazol-3-yl)-3-(5-methylpyrimidin-2-yl)butane-2-sulfonamide (SI-1).





Initial direct methods solution of **SI-1** (top) with electron density map (F_{obs}) contoured at 1.41 e $Å^{-3}$ and ORTEP diagram of refined **SI-1** (bottom). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-1.

Empirical formula	$C_{25}H_{29}N_7O_4S$
Formula weight	523.61

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	294(2) K
Unit cell dimensions	a = 9.3100(10)
	b = 20.490(2)
	c = 12.650(4)
	β = 108.42
Volume	2289.5(8)
Z	2
Crystal system	Monoclinic
Space group	$P2_1$
Density (calculated)	1.515 Mg/m^3
F(000)	100
Measured reflections	6341
Reflections with I > 2σ(*I*)	3849
Resolution	0.90 Å
Completeness	81.4%
Index ranges	10 ≤ h ≤ -10, 23 ≤ k ≤
	-24, 13 ≤ l ≤ -13

Structure Solution and Refinement

Structure solution program	SHELXT (Uson & Sheldrick, 1999)

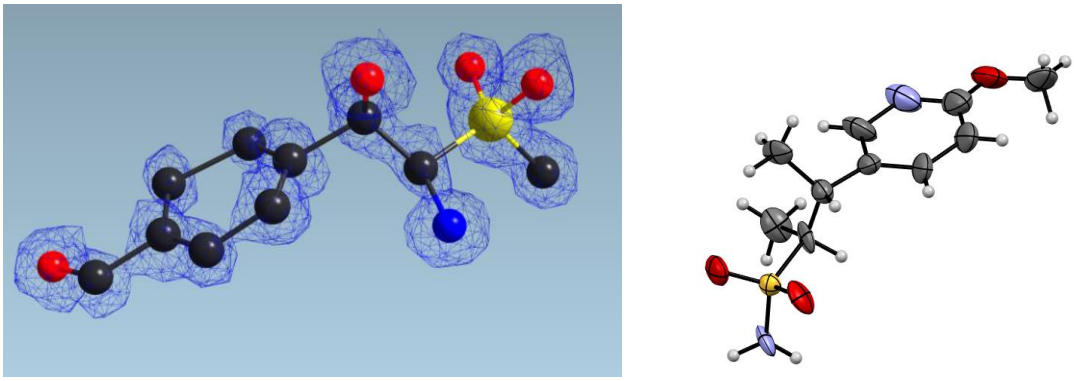
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	6341 / 1069 / 668
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.243
Final R indices [I>2s(I)]	R1 = 0.1228, *w*R2 = 0.2924
R indices (all data)	R1 = 0.1678, *w*R2 = 0.3187
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.044
Average shift/error	0.000
Largest diff. peak and hole	0.19 and -0.12 e.Å^{-3}

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

3.2 (2R,3S)-3-(5-methoxypyridin-2-yl)butane-2-sulfonamide (SI-2).



Initial direct methods solution of **SI-2** (left) with electron density map (F_{obs}) contoured at 1.03 e $Å^{-3}$ and ORTEP diagram of refined **SI-2** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-2.

Empirical formula	$C_{25}H_{29}N_7O_4S$
Formula weight	523.61

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	294(2) K
Unit cell dimensions	a = 22.830(4)
	b = 6.810(10)
	c = 6.980(2)
Volume	1085.2(4)
Z	1
Crystal system	Orthorhombic
Space group	$P2_12_12$
Density (calculated)	1.495 Mg/m^3
F(000)	103
Measured reflections	3816
Reflections with I > 2σ(*I*)	870

Resolution	0.95 Å
Completeness	82.9%
Index ranges	$7 \leq h \leq -7$, $25 \leq k \leq -25$, $7 \leq l \leq -7$

Structure Solution and Refinement

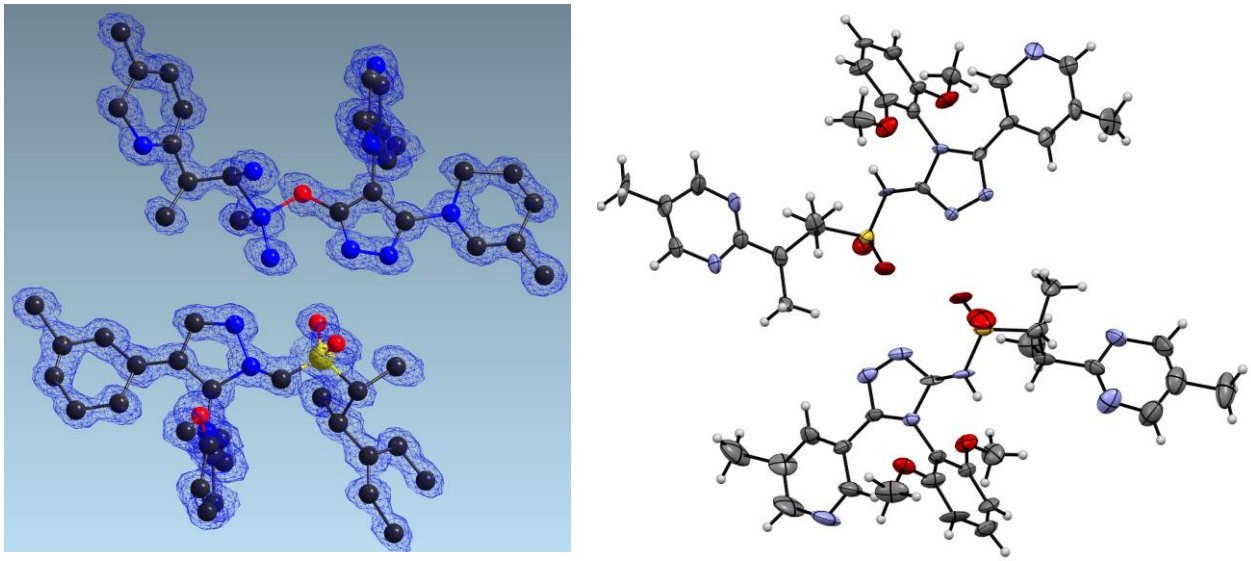
Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	1360 / 135 / 146
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.480
Final R indices [I>2s(I)]	R1 = 0.1347, *w*R2 = 0.3138
R indices (all data)	R1 = 0.1779, *w*R2 = 0.3349
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.051
Average shift/error	0.009
Largest diff. peak and hole	0.20 and -0.18 e.Å^{-3}

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

3.3 (2S,3R)-N-(4-(2,6-dimethoxyphenyl)-5-(5-methylpyridin-3-yl)-4H-1,2,4-triazol-3-yl)-3-(5-methylpyrimidin-2-yl)butane-2-sulfonamide (SI-3).



Initial direct methods solution of **SI-3** (left) with electron density map (F_{obs}) contoured at 1.03 e $Å^{-3}$ and ORTEP diagram of refined **SI-3** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-3.

Empirical formula	$C_{25}H_{29}N_7O_4S$
Formula weight	523.61

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	294(2) K
Unit cell dimensions	a = 9.3100(10)
	b = 20.450(2)
	c = 12.680(4)
	β = 108.40
Volume	2290.7(8)
Z	2
Crystal system	Monoclinic
Space group	$P2_1$

Density (calculated)	1.518 Mg/m^3
F(000)	100
Measured reflections	5615
Reflections with I > 2σ(*I*)	3218
Resolution	0.90 Å
Completeness	85.5%
Index ranges	10 ≤ h ≤ -10, 21 ≤ k ≤ -21, 14 ≤ l ≤ -14

Structure Solution and Refinement

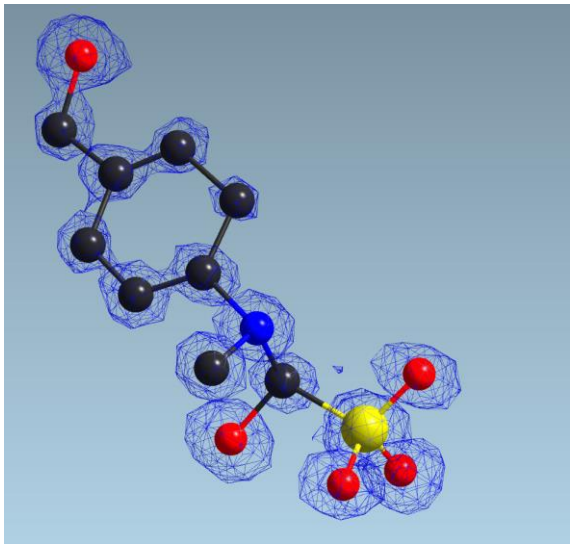
Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	5615 / 677 / 668
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.226
Final R indices [I>2s(I)]	R1 = 0.1294, *w*R2 = 0.3063
R indices (all data)	R1 = 0.1777, *w*R2 = 0.3408
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.016
Average shift/error	0.000
Largest diff. peak and hole	0.20 and -0.14 $e.Å^{-3}$

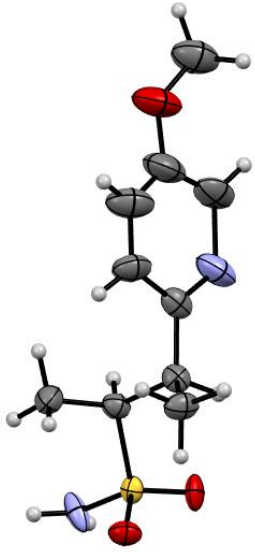
Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

3.4 (2S,3R)-3-(5-methoxypyridin-2-yl)butane-2-sulfonamide (SI-4)





Initial direct methods solution of **SI-4** (left) with electron density map (F_{obs}) contoured at 1.03 e $Å^{-3}$ and ORTEP diagram of refined **SI-4** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-4.

Empirical formula	$C_{10}H_{16}N_2O_3S$
Formula weight	244.31

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	294(2) K
Unit cell dimensions	a = 6.840(2)
	b = 22.820(4)
	c = 6.9800(10)
Volume	1089.5(4)
Z	4

Crystal system	Orthorhombic
Space group	$P2_12_12$
Density (calculated)	1.489 Mg/m^3
F(000)	103
Measured reflections	1212
Reflections with $I > 2\sigma(I)$	649
Resolution	0.95 Å
Completeness	89.4%
Index ranges	$24 \leq h \leq -24$, $7 \leq k \leq -7$, $7 \leq l \leq -7$

Structure Solution and Refinement

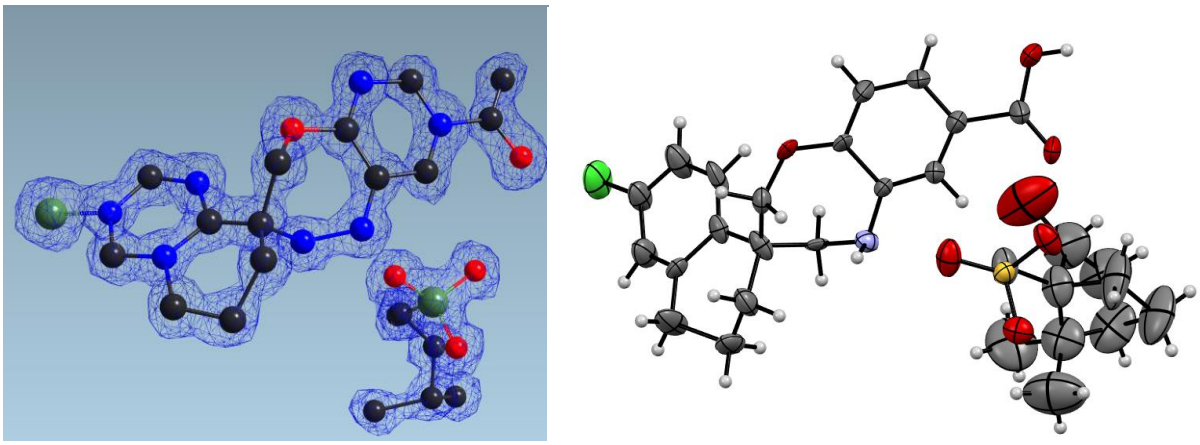
Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	1212 / 136 / 134
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.414
Final R indices [I>2s(I)]	R1 = 0.1706, *w*R2 = 0.3583
R indices (all data)	R1 = 0.2370, *w*R2 = 0.3910
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.030
Average shift/error	0.000
Largest diff. peak and hole	0.18 and -0.18 $e.Å^{-3}$

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

3.5 (R)-6'-chloro-3',4,4',5-tetrahydro-2H,2'H-spiro[benzo[b][1,4]oxazepine-3,1'-naphthalene]-7-carboxylic acid ((1R,4S)-7,7-dimethyl-2-oxobicyclo[2.2.1]heptan-1-yl)methansulfonate (SI-5)



Initial direct methods solution of **SI-5** (left) with electron density map (F_{obs}) contoured at 1.03 e $Å^{-3}$ and ORTEP diagram of refined **SI-5** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-5.

Empirical formula	$C_{29}H_{34}NO_7SCl$
Formula weight	574.07

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	294(2) K
Unit cell dimensions	a = 10.5200(10)
	b = 10.220(2)

	c = 12.660(4)
	β = 110.33
Volume	1276.4(5)
Z	2
Crystal system	Monoclinic
Space group	$P2_1$
Density (calculated)	1.494 Mg/m^3
F(000)	99
Measured reflections	2438
Reflections with I > 2σ(*I*)	1379
Resolution	1.0 Å
Completeness	95.7%
Index ranges	9 ≤ h ≤ -9, 10 ≤ k ≤ -10, 12 ≤ l ≤ -12

Structure Solution and Refinement

Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	2438 / 612 / 357
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.239
Final R indices [I>2s(I)]	R1 = 0.1198, *w*R2 = 0.2721
R indices (all data)	R1 = 0.1823, *w*R2 = 0.3103
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.061
Average shift/error	0.001
Largest diff. peak and hole	0.13 and -0.13 $e.Å^{-3}$

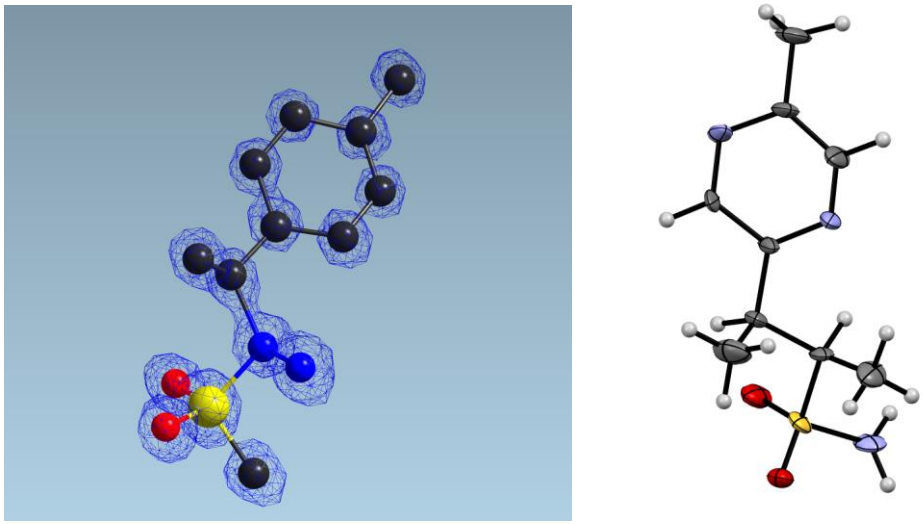
Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

Significant disorder of the CSA moiety prevents assignment of absolute stereochemistry.

3.6 (2S,3R)-3-(5-methylpyrazin-2-yl)butane-2-sulfonamide (SI-6)



Initial direct methods solution of **SI-6** (left) with electron density map (F_{obs}) contoured at 1.41 e $Å^{-3}$ and ORTEP diagram of refined **SI-6** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-6.

Empirical formula	$C_9H_{15}N_3O_2S$
Formula weight	229.30

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	294(2) K
Unit cell dimensions	a = 7.4500(10)
	b = 8.130(2)
	c = 16.240(4)
Volume	983.6(4)
Z	4
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Density (calculated)	1.548 Mg/m^3
F(000)	46
Measured reflections	1652
Reflections with I > 2σ(*I*)	975
Resolution	0.80 Å
Completeness	82.6%
Index ranges	8 ≤ h ≤ -8, 9 ≤ k ≤
	-9, 19 ≤ l ≤ -19

Structure Solution and Refinement

Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	1652 / 127 / 137
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.267
Final R indices [I>2s(I)]	R1 = 0.1808, *w*R2 = 0.4048
R indices (all data)	R1 = 0.2167, *w*R2 = 0.4398
Type of weighting scheme used	Sigma

Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.079
Average shift/error	0.001
Largest diff. peak and hole	0.21 and -0.38 e.Å^{-3}

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

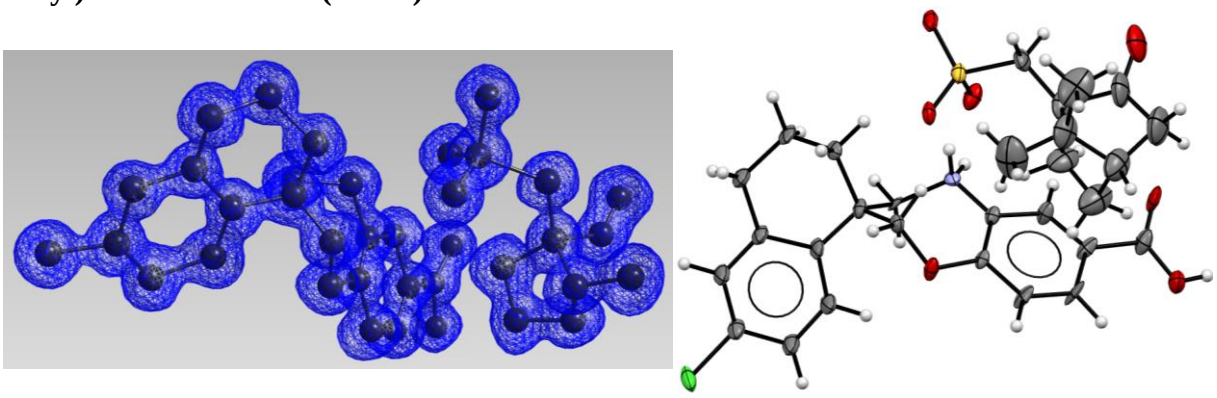
All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

4. Cryogenic TEM Screening

Milligram to sub-milligram quantities of dry powder were placed into a dram vial as received and manually ground with a glass pipette. A pure carbon 200 mesh Cu grid or lacey carbon Cu grid was placed inside of the vial and gently shaken together with the powder to “dry load” the grid. The grid was removed with Dumont straight self-closing tweezers and the tweezers were gently tapped against a lab bench while holding the grid to shake off excess powder. This sample was clipped into a Gatan 626 cryo holder at room temperature and inserted into a well-aligned Thermo Fisher Scientific Talos F200C transmission electron microscopy operating at an accelerating voltage of 200keV. After successful insertion, the cryo holder was cooled with liquid nitrogen until reaching a stable temperature of ~–177 °C. After achieving stable temperature and low vacuum pressure, incident diffraction screening and movie collection were performed as described in **Supporting Information Section 2**. Screening was halted after 3 hours, or if a preliminary solution with >90% of expected atoms was obtained.

5. Crystal Structure and Refinement Information of Cryogenically Cooled Samples

5.1 (R)-6'-chloro-3',4,4',5-tetrahydro-2H,2'H-spiro[benzo[b][1,4]oxazepine-3,1'-naphthalene]-7-carboxylic acid ((1R,4S)-7,7-dimethyl-2-oxobicyclo[2.2.1]heptan-1-yl)methansulfonate (SI-5.1)



Initial direct methods solution of **SI-5** (left) with electron density map (F_{obs}) contoured at 1.03 e $Å^{-3}$ and ORTEP diagram of refined **SI-5** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-5.

Empirical formula	$C_{29}H_{34}NO_7SCl$
Formula weight	574.07

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	294(2) K
Unit cell dimensions	a = 10.4700(10)
	b = 10.260(2)
	c = 12.440(4)
	β = 109.92
Volume	1256.4(5)
Z	2
Crystal system	Monoclinic
Space group	$P2_1$
Density (calculated)	1.523 Mg/m^3

F(000)	99
Measured reflections	2474
Reflections with I > 2σ(*I*)	1787
Resolution	1.0 Å
Completeness	92.6%
Index ranges	10 ≤ h ≤ -10, 10 ≤ k ≤ -10, 12 ≤ l ≤ -12

Structure Solution and Refinement

Structure solution program	SHELXD (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	2474 / 703 / 354
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.143
Final R indices [I>2s(I)]	R1 = 0.1127, *w*R2 = 0.2643
R indices (all data)	R1 = 0.1456, *w*R2 = 0.2863
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.000
Average shift/error	0.000
Largest diff. peak and hole	0.17 and -0.16 e.Å^{-3}

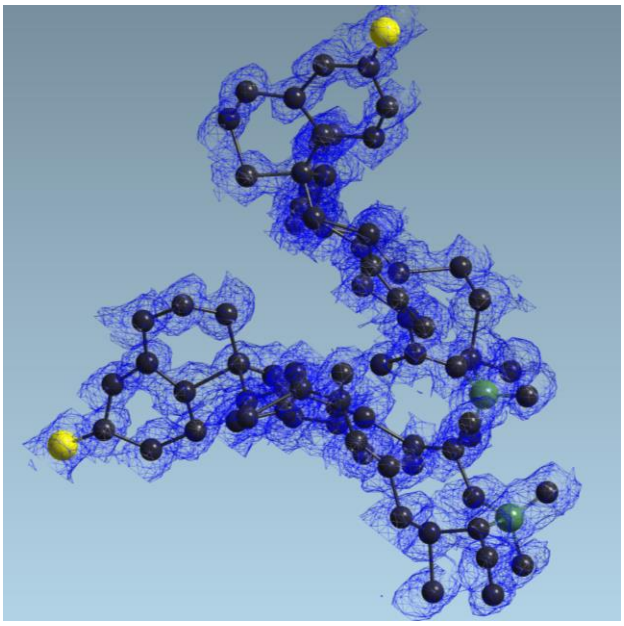
Special Refinement Details

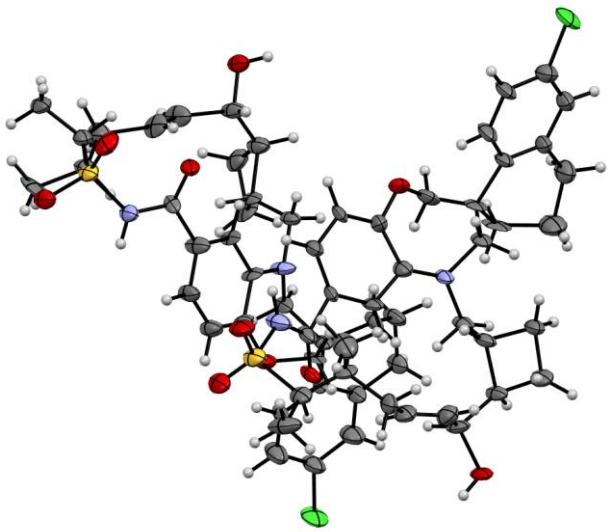
Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account

individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

5.2 (1S,3'R,6'R,7'S,8'E,11'S,12'R)-6-chloro-7'-hydroxy-11',12'-dimethyl-3,4-dihydro-2H,15'H-spiro[naphthalene-1,22'-[20]oxa[13]thia[1,14]diazatetracyclo [14.7.2.0^{3,6}.0^{19,24}]pentacosa[8,16,18,24]tetraen]-15'-one 13',13'-dioxide (SI-7)





Initial direct methods solution of **SI-7** (top) with electron density map (F_{obs}) contoured at 1.01 e $Å^{-3}$ and ORTEP diagram of refined **SI-7** (bottom). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-7.

Empirical formula	$C_{32}H_{39}N_2O_5SCl$
Formula weight	599.16

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	96(4) K
Unit cell dimensions	a = 11.3400(10)
	b = 11.340(2)
	c = 12.500(4)
	α = 73.74
	β = 69.36
	γ = 71.13
Volume	1398.2(5)
Z	2
Crystal system	Monoclinic
Space group	P_1
Density (calculated)	1.423 Mg/m^3
F(000)	243
Measured reflections	6522
Reflections with I > 2σ(*I*)	3782
Resolution	0.90 Å
Completeness	82.5%
Index ranges	$12 \leq h \leq -12$, $12 \leq k \leq -12$, $13 \leq l \leq -13$

Structure Solution and Refinement

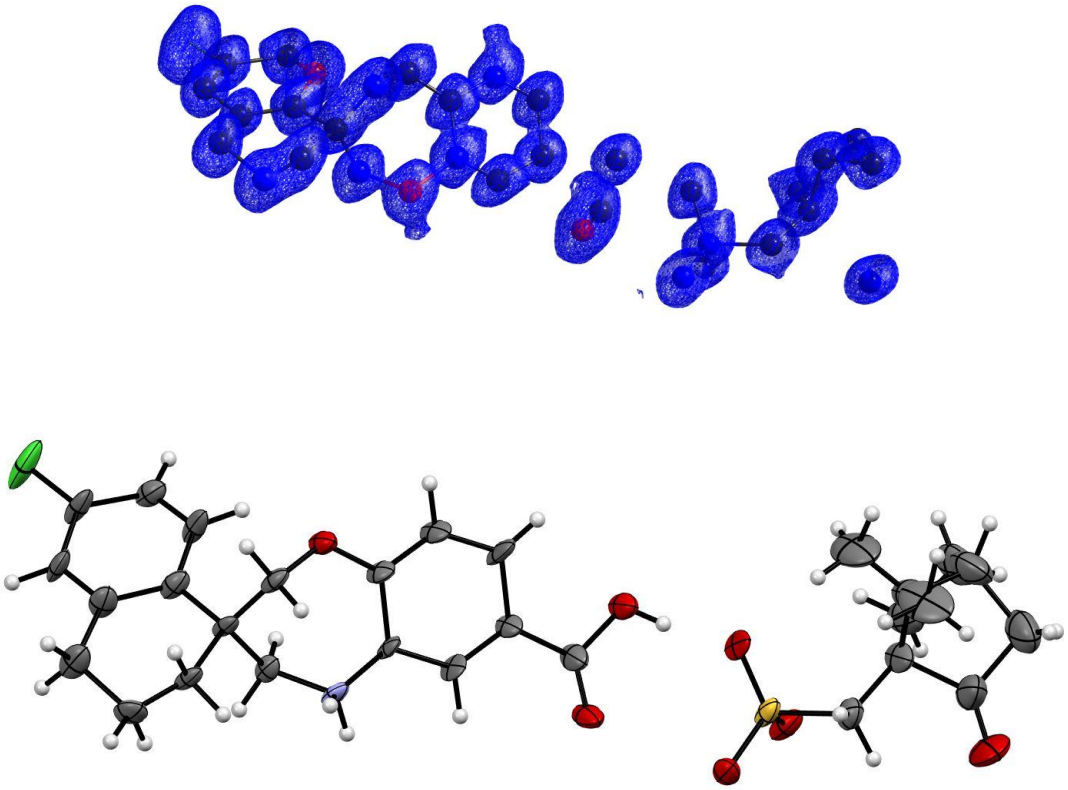
Structure solution program	SHELXD (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	6522 / 1253 / 740
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.242
Final R indices [I>2s(I)]	R1 = 0.1465, *w*R2 = 0.3107
R indices (all data)	R1 = 0.1954, *w*R2 = 0.3439
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.083
Average shift/error	0.000
Largest diff. peak and hole	0.25 and -0.26 e.Å^{-3}

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

5.3 (S)-6'-chloro-3',4,4',5-tetrahydro-2H,2'H-spiro[benzo[b][1,4]oxazepine-3,1'-naphthalene]-7-carboxylic acid ((1S,4R)-7,7-dimethyl-2-oxobicyclo[2.2.1]heptan-1-yl)methansulfonate (SI-8)



Initial direct methods solution of **SI-8** (top) with electron density map (F_{obs}) contoured at 1.03 e $Å^{-3}$ and ORTEP diagram of refined **SI-8** (bottom). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-8.

Empirical formula	$C_{29}H_{34}NO_7SCl$
Formula weight	576.08

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	96(4) K
Unit cell dimensions	a = 10.6900(10)
	b = 10.220(2)
	c = 12.680(4)

	β = 111.22
Volume	1291.4(5)
Z	2
Crystal system	Monoclinic
Space group	$P2_1$
Density (calculated)	1.482 Mg/m^3
F(000)	99
Measured reflections	3824
Reflections with I > 2σ(*I*)	2409
Resolution	0.85 Å
Completeness	85.2%
Index ranges	12 ≤ h ≤ -12, 11 ≤ k ≤ -11, 14 ≤ l ≤ -14

Structure Solution and Refinement

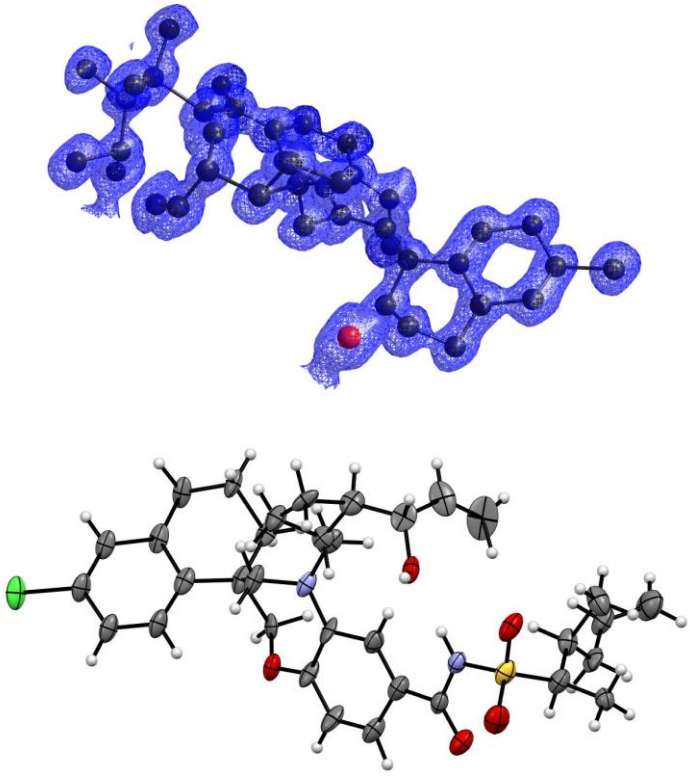
Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	3824 / 377 / 354
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.244
Final R indices [I>2s(I)]	R1 = 0.1415, *w*R2 = 0.3278
R indices (all data)	R1 = 0.1794, *w*R2 = 0.3523
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.044
Average shift/error	0.002
Largest diff. peak and hole	0.15 and -0.16 $e.Å^{-3}$

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

5.4 (S)-6'-chloro-5-(((1R,2R)-2-((S)-1-hydroxyallyl)cyclobutyl)methyl)-N-(((2R,3S)-3-methylhex-5-en-2-yl)sulfonyl)-3',4,4',5-tetrahydro-2H,2'H-spiro[benzo[B][1,4]oxazepine-3,1'-naphthalene]-7-carboxamide (SI-9)



Initial direct methods solution of **SI-9** (top) with electron density map (F_{obs}) contoured at 1.03 e $Å^{-3}$ and ORTEP diagram of refined **SI-9** (bottom). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-9.

Empirical formula	$C_{34}H_{43}N_2O_5SCl$
Formula weight	626.26

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	96(4) K
Unit cell dimensions	a = 10.5200(10)
	b = 15.050(2)
	c = 17.020(4)
Volume	2694.7(8)
Z	4
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Density (calculated)	1.546 Mg/m^3
F(000)	105
Measured reflections	2393
Reflections with I > 2σ(*I*)	1774
Resolution	1.0 Å
Completeness	83.1%
Index ranges	10 ≤ h ≤ -10, 15 ≤ k ≤
	-15, 15 ≤ l ≤ -14

Structure Solution and Refinement

Structure solution program	SHELXD (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	2393 / 748 / 390

Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.399
Final R indices [I>2s(I)]	R1 = 0.1111, *w*R2 = 0.2457
R indices (all data)	R1 = 0.1460, *w*R2 = 0.2596
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.045
Average shift/error	0.000
Largest diff. peak and hole	0.15 and -0.13 e.Å^{-3}

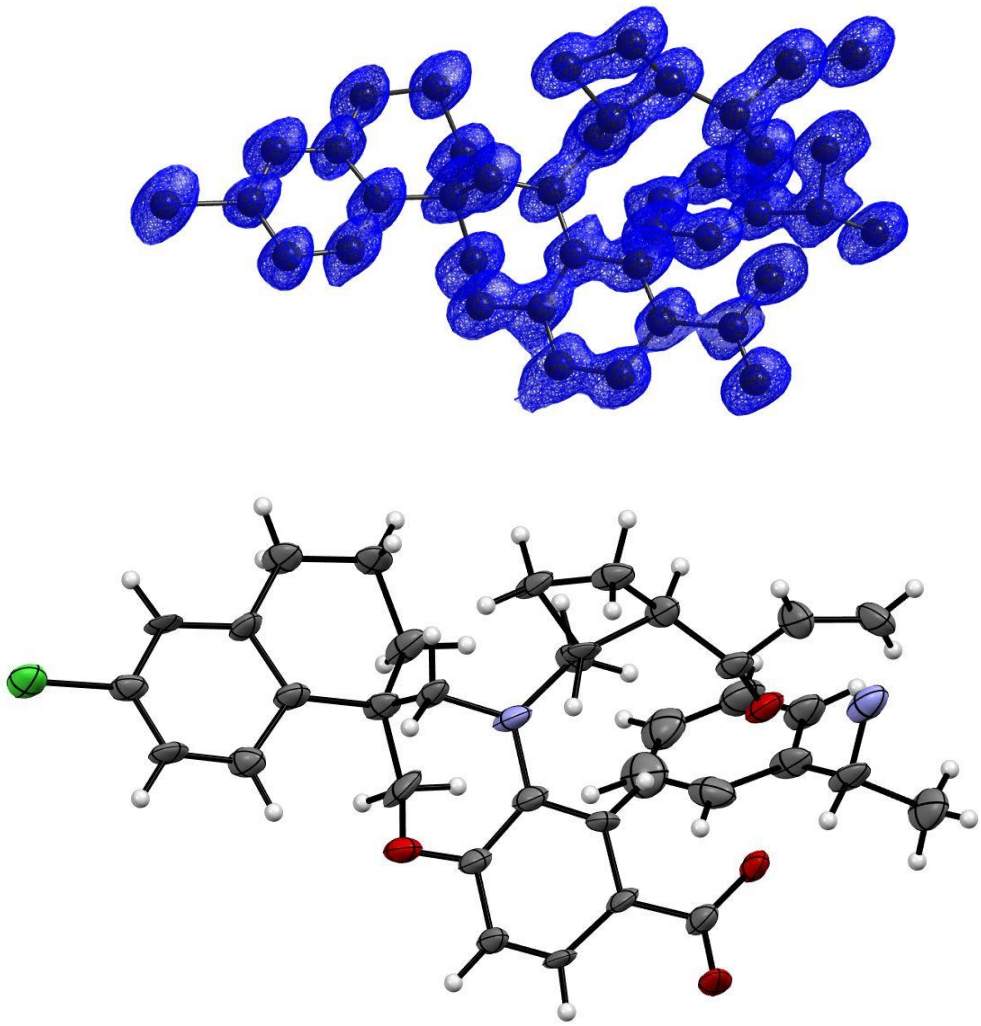
Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

5.5 (R)-1-phenylethan-1-aminium (S)-6'-chloro-5-(((1R,2R)-2-((S)-1-hydroxyallyl)

cyclobutyl)methyl)-3',4,4',5-tetrahydro-2H,2'H-spiro[benzo[b][1,4]oxazepine-3,1'-naphthalene]-7-carboxylate (SI-10)



Initial direct methods solution of **SI-10** (top) with electron density map (F_{obs}) contoured at 1.41 e $Å^{-3}$ and ORTEP diagram of refined **SI-10** (bottom). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-10.

Empirical formula	$C_{35}H_{36}N_2O_4Cl$
Formula weight	584.11

Data Collection

Type of instrument	Talos F200C

Wavelength	0.0215 Å
Data collection temperature	96(4) K
Unit cell dimensions	a = 7.9800(10)
	b = 11.730(2)
	c = 28.850(4)
Volume	2700.5(7)
Z	4
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Density (calculated)	1.437 Mg/m^3
F(000)	103
Measured reflections	3263
Reflections with I > 2σ(*I*)	2351
Resolution	0.85 Å
Completeness	68.6%
Index ranges	$9 \leq h \leq -9$, $13 \leq k \leq -13$, $25 \leq l \leq -25$

Structure Solution and Refinement

Structure solution program	SHELXD (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	3263 / 390 / 380
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.469
Final R indices [I>2s(I)]	R1 = 0.1206, *w*R2 = 0.2740
R indices (all data)	R1 = 0.1536, *w*R2 = 0.2834
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$

Max shift/error	0.026
Average shift/error	0.000
Largest diff. peak and hole	0.17 and -0.13 e.Å^{-3}

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

6. Additional Screening and Recrystallization of Samples

Sample **SI-12** was screened in the same manner as described in **Supporting Information Section 2** for ~4 additional hours at room temperature to locate monocrystalline domains in a largely polycrystalline sample.

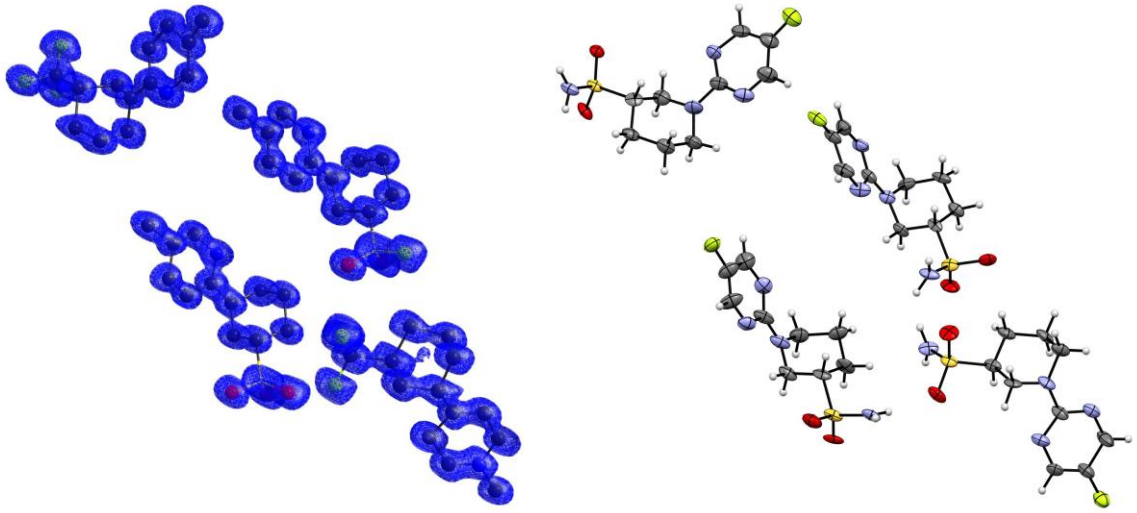
Crystallization of **SI-11**, **SI-13**, **SI-14**, and **SI-15** was performed by placing ~1 mg of powder as received into 6 x 50 mm borosilicate culture tubes purchased from VWR. Samples were dissolved in approximately 500 uL of solvent and allowed to slowly evaporate at room temperature. Higher boiling solvents were evaporated from open containers, while low boiling solvents required placing the culture tube inside an empty dram vial with a slightly loosened cap. If the initial solvent failed to produce a solid after fully evaporating based on visual inspection, the amorphous samples were re-dissolved in the same culture tube with a new solvent mixture. Evaporation occurred until precipitation was observed. Sample crystallization time spanned from overnight to 3 days. **SI-11** and **SI-13** were obtained from slow evaporation from a 50/50 mixture of MeCN and H_2O. The crystals were dried under reduced pressure and screened at cryogenic temperatures as outlined in **Supporting Information Section 4**.

SI-15 was generated from slow evaporation from H_2O with a small amount of DMSO. The crystals were blotted with a kimwipe and dried under reduced pressure to remove excess solvent before being brought into the TEM as described in **Supporting Information Section 4**.

SI-14 was crystallized from slow evaporation of diethyl ether. The crystals were placed onto a grid as a dry powder, and screened by a modified procedure of **Supporting Information Section 4**. Before typical screening, the prepared grid was plunge frozen in liquid nitrogen and transferred into the TEM while the holder was maintained at cryogenic temperatures.

7. Crystal Structure and Refinement Information of Additional Samples

7.1 (S)-1-(5-fluoropyrimidin-2-yl)piperidine-3-sulfonamide (SI-11).



Initial direct methods solution of **SI-11** (left) with electron density map (F_{obs}) contoured at 1.41 e $Å^{-3}$ and ORTEP diagram of refined **SI-11** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-11.

Empirical formula	$C_9H_{13}N_4O_2SF$
Formula weight	260.29

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	96(4) K
Unit cell dimensions	a = 22.990(2)
	b = 37.240(4)
	c = 4.6400(10)
Volume	3972.5(10)
Z	16
Crystal system	Orthorhombic

Space group	$P2_12_12$
Density (calculated)	1.741 Mg/m^3
F(000)	43
Measured reflections	4964
Reflections with I > 2σ(*I*)	3031
Resolution	0.90 Å
Completeness	83.1%
Index ranges	25 ≤ h ≤ -25, 37 ≤ k ≤ -36, 5 ≤ l ≤ -5

Structure Solution and Refinement

Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	4964 / 608 / 614
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.373
Final R indices [I>2s(I)]	R1 = 0.1254, *w*R2 = 0.2743
R indices (all data)	R1 = 0.1848, *w*R2 = 0.2955
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.000
Average shift/error	0.000
Largest diff. peak and hole	0.20 and -0.16 $e.Å^{-3}$

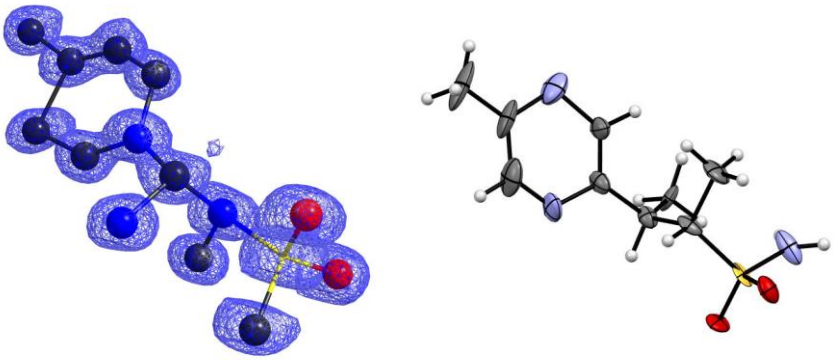
Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for

refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

7.2 (2R,3S)-3-(5-methylpyrazin-2-yl)butane-2-sulfonamide (SI-12).



Initial direct methods solution of **SI-12** (left) with electron density map (F_{obs}) contoured at 1.41 e $Å^{-3}$ and ORTEP diagram of refined **SI-12** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-12.

Empirical formula	$C_9H_{15}N_3O_2S$
Formula weight	229.30

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	294(4) K
Unit cell dimensions	a = 22.0000(10)
	b = 6.410(2)
	c = 7.060(4)
	β = 91.18

Volume	995.4(4)
Z	4
Crystal system	Monoclinic
Space group	C2
Density (calculated)	1.530 Mg/m^3
F(000)	49
Measured reflections	1265
Reflections with $I > 2\sigma(I)$	789
Resolution	0.90 Å
Completeness	86.5%
Index ranges	$24 \leq h \leq -24$, $7 \leq k \leq -7$, $7 \leq l \leq -7$

Structure Solution and Refinement

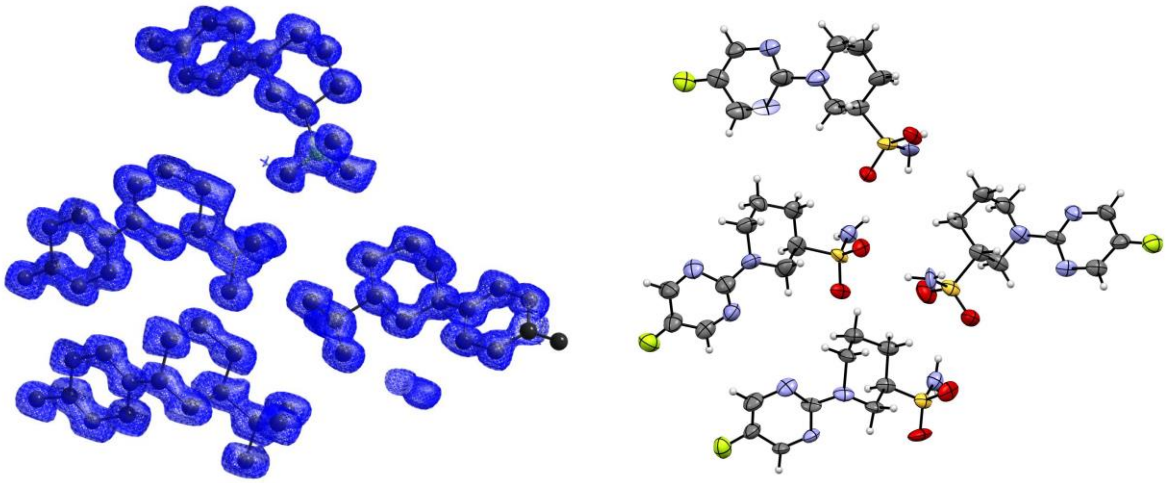
Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	1265 / 239 / 137
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.372
Final R indices [I>2s(I)]	R1 = 0.1348, *w*R2 = 0.3273
R indices (all data)	R1 = 0.1786, *w*R2 = 0.3500
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.034
Average shift/error	0.000
Largest diff. peak and hole	0.14 and -0.18 $e.Å^{-3}$

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

7.3 (R)-1-(5-fluoropyrimidin-2-yl)piperidine-3-sulfonamide (SI-13).



Initial direct methods solution of **SI-13** (left) with electron density map (F_{obs}) contoured at 1.41 e $Å^{-3}$ and ORTEP diagram of refined **SI-13** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-13.

Empirical formula	$C_9H_{13}N_4O_2SF$
Formula weight	260.29

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	96(4) K
Unit cell dimensions	a = 23.000(2) b = 38.090(4) c = 4.6000(10)
Volume	4029.9(10)
Z	16
Crystal system	Orthorhombic
Space group	$P2_12_12$
Density (calculated)	1.716 Mg/m^3
F(000)	170
Measured reflections	4939
Reflections with I > 2σ(*I*)	3181
Resolution	0.90 Å
Completeness	82.6%
Index ranges	25 ≤ h ≤ -25, 38 ≤ k ≤ -38, 5 ≤ l ≤ -5

Structure Solution and Refinement

Structure solution program	SHELXD (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	4939 / 608 / 614
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.323
Final R indices [I>2s(I)]	R1 = 0.1373, *w*R2 = 0.3069
R indices (all data)	R1 = 0.1752, *w*R2 = 0.3300
Type of weighting scheme used	Sigma

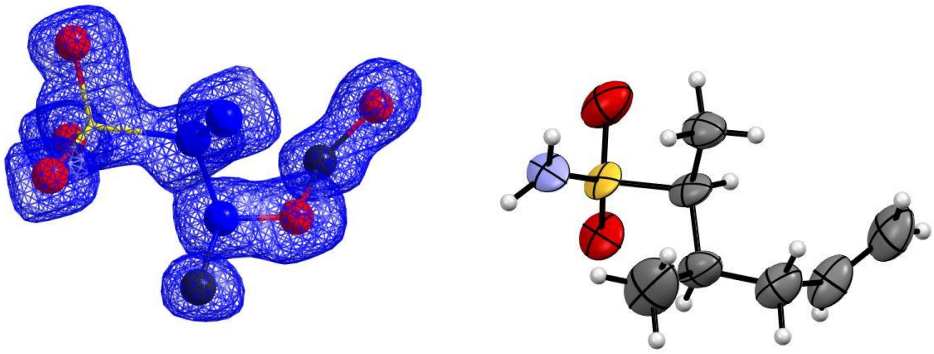
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.095
Average shift/error	0.000
Largest diff. peak and hole	0.18 and -0.13 e.Å^{-3}

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

7.4 (2R,3S)-3-methylhex-5-ene-2-sulfonamide (SI-14).



Initial direct methods solution of **SI-14** (left) with electron density map (F_{obs}) contoured at 1.41 e Å^{-3} and ORTEP diagram of refined **SI-14** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-14.

Empirical formula	$C_7H_{15}NO_2S$
Formula weight	177.26

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	96(4) K
Unit cell dimensions	a = 7.4100(10)
	b = 9.270(2)
	c = 12.490(4)
Volume	857.9(4)
Z	4
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Density (calculated)	1.372 Mg/m^3
F(000)	29
Measured reflections	1002
Reflections with I > 2σ(*I*)	461
Resolution	0.90 Å
Completeness	80.3%
Index ranges	$8 \leq h \leq -8$, $9 \leq k \leq$
	-9, $12 \leq l \leq -12$

Structure Solution and Refinement

Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	1002 / 84 / 101
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.326
Final R indices [I>2s(I)]	R1 = 0.1396, *w*R2 = 0.3178
R indices (all data)	R1 = 0.2144, *w*R2 = 0.3575
Type of weighting scheme used	Sigma

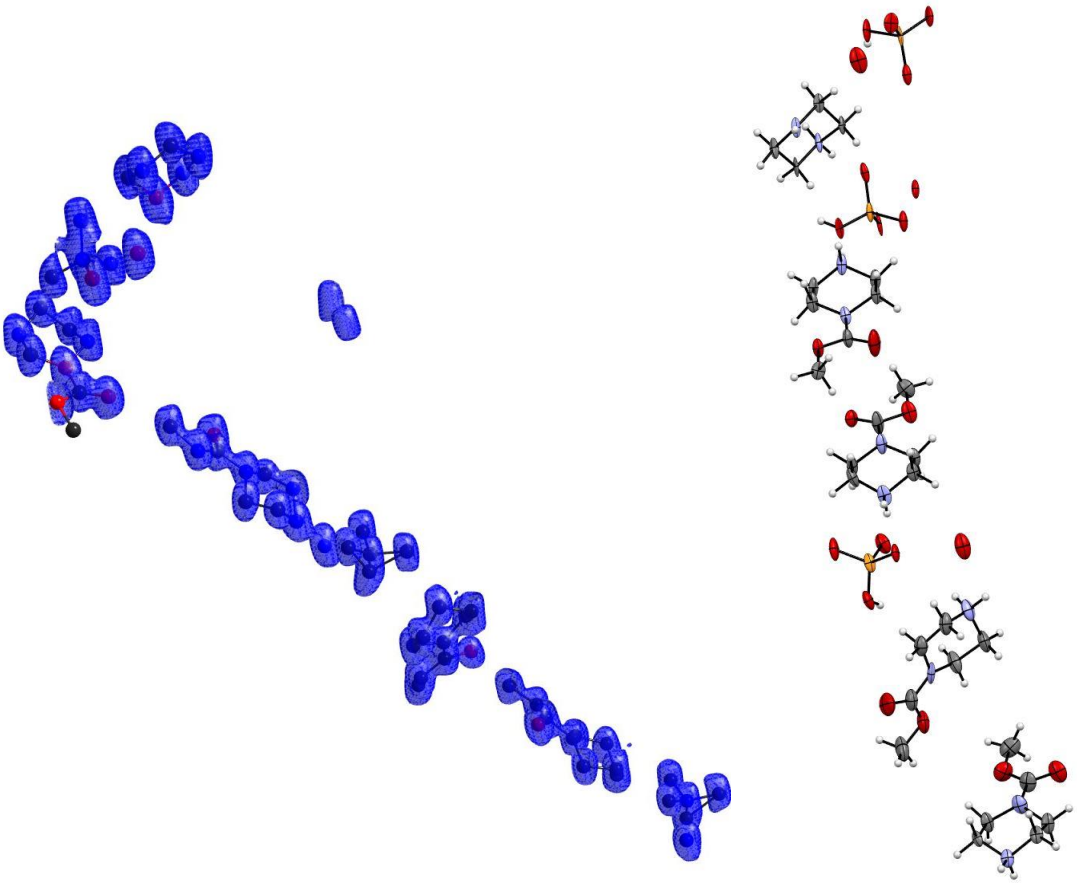
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.099
Average shift/error	0.002
Largest diff. peak and hole	0.12 and -0.12 e.Å^{-3}

Special Refinement Details

Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

7.5 methyl piperazine-1-carboxylate phosphate hydrate (SI-15)



Initial direct methods solution of **SI-15** (left) with electron density map (F_{obs}) contoured at 1.41 e $Å^{-3}$ and ORTEP diagram of refined **SI-15** (right). Thermal ellipsoids shown as shaded octants at 30% probability.

Crystal data and structure refinement for SI-15.

Empirical formula	$C_{28}H_{66}N_{10}O_{23}P_3$
Formula weight	1003.82

Data Collection

Type of instrument	Talos F200C
Wavelength	0.0215 Å
Data collection temperature	96(4) K
Unit cell dimensions	a = 66.2000(10)

	b = 6.220(2)
	c = 9.940(4)
	β = 92.14
Volume	4090(2)
Z	4
Crystal system	Monoclinic
Space group	Cc
Density (calculated)	1.630 Mg/m^3
F(000)	169
Measured reflections	5976
Reflections with I > 2σ(*I*)	3714
Resolution	0.85 Å
Completeness	84.6%
Index ranges	70 ≤ h ≤ -71, 7 ≤ k ≤ -7, 11 ≤ l ≤ -11

Structure Solution and Refinement

Structure solution program	SHELXT (Uson & Sheldrick, 1999)
Primary solution method	Direct methods
Secondary solution method	Difference Fourier map
Hydrogen placement	Geometric positions
Structure refinement program	SHELXL-2018/3 (Sheldrick, 2018)
Refinement method	Full matrix least-squares on F^2
Data / restraints / parameters	5976 / 873 / 581
Treatment of hydrogen atoms	Riding
Goodness-of-fit on F^2	1.345
Final R indices [I>2s(I)]	R1 = 0.1383, *w*R2 = 0.2795
R indices (all data)	R1 = 0.1901, *w*R2 = 0.2995
Type of weighting scheme used	Sigma
Weighting scheme used	$w=1/s^2(Fo^2)$
Max shift/error	0.001
Average shift/error	0.000

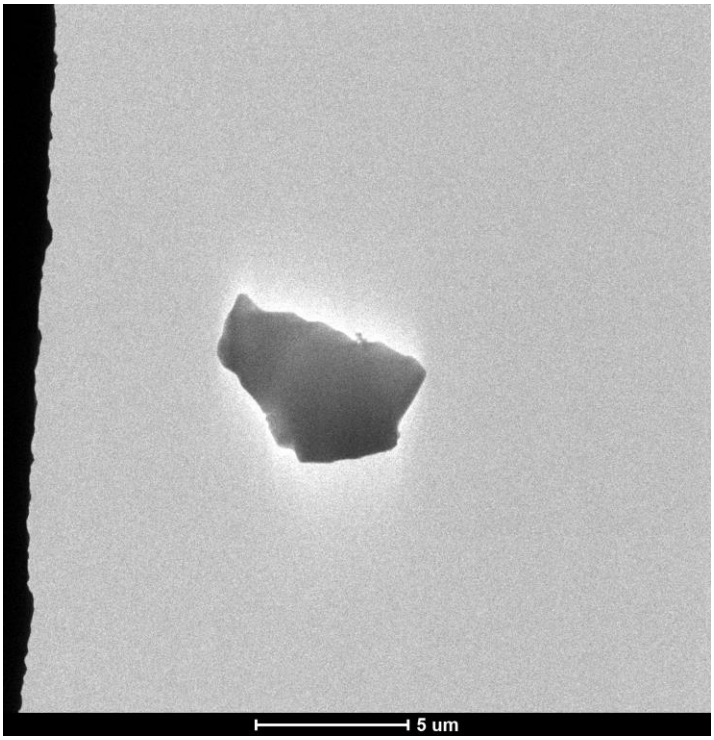
Largest diff. peak and hole 0.29 and -0.28 e.Å^{-3}

Special Refinement Details

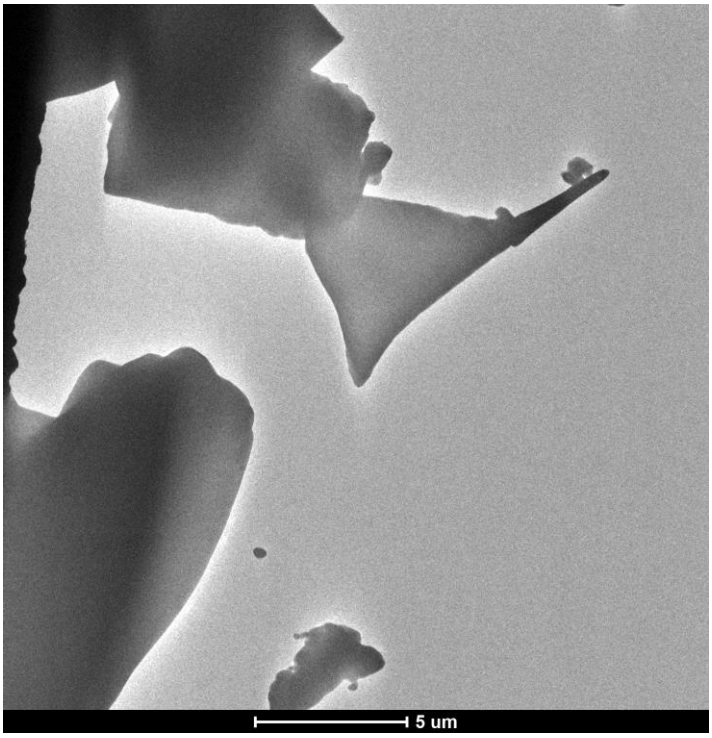
Refinement of F^2 against ALL reflections. The weighted R-factor (wR) and goodness of fit (S) are based on F^2, conventional R-factors (R) are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > 2s(F^2)$ is used only for calculating R-factors(gt) etc. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R-factors based on ALL data will be even larger.

All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

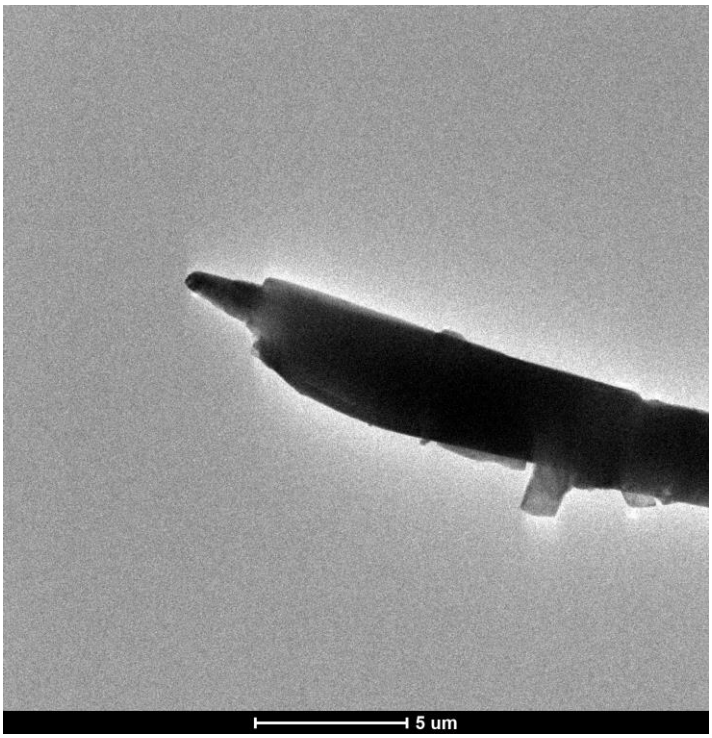
8. Transmission Electron Microscope Images of Crystals



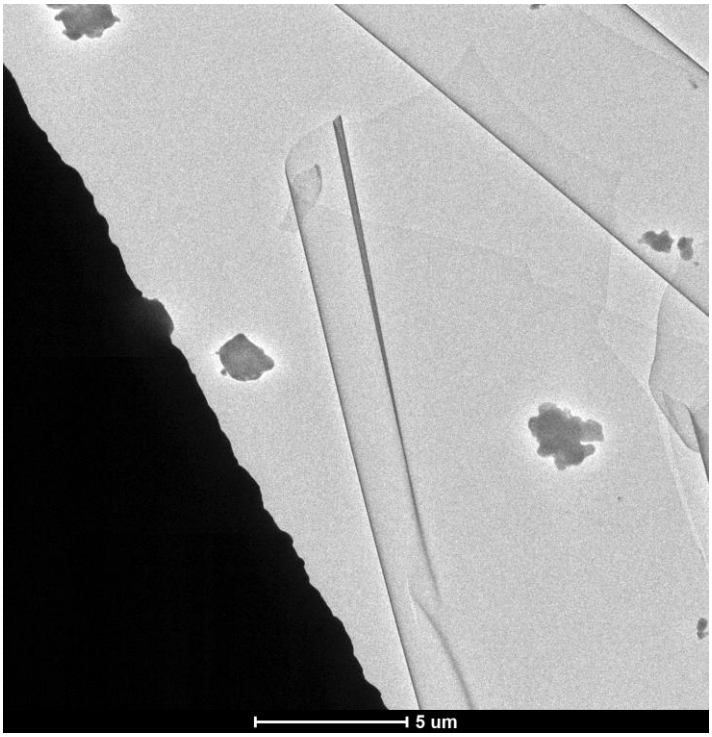
TEM image of **SI-1** crystal at 2600x magnification



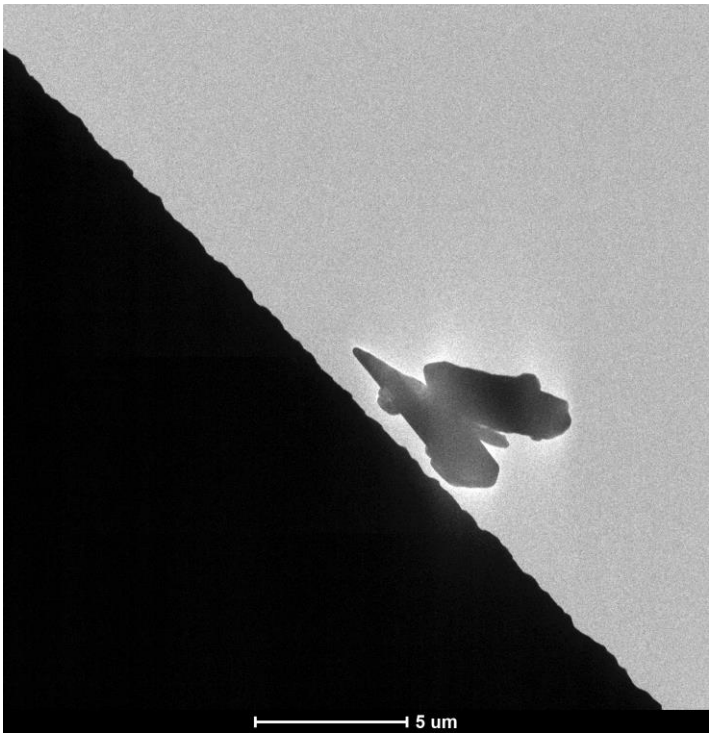
TEM image of **SI-2** crystal at 2600x magnification



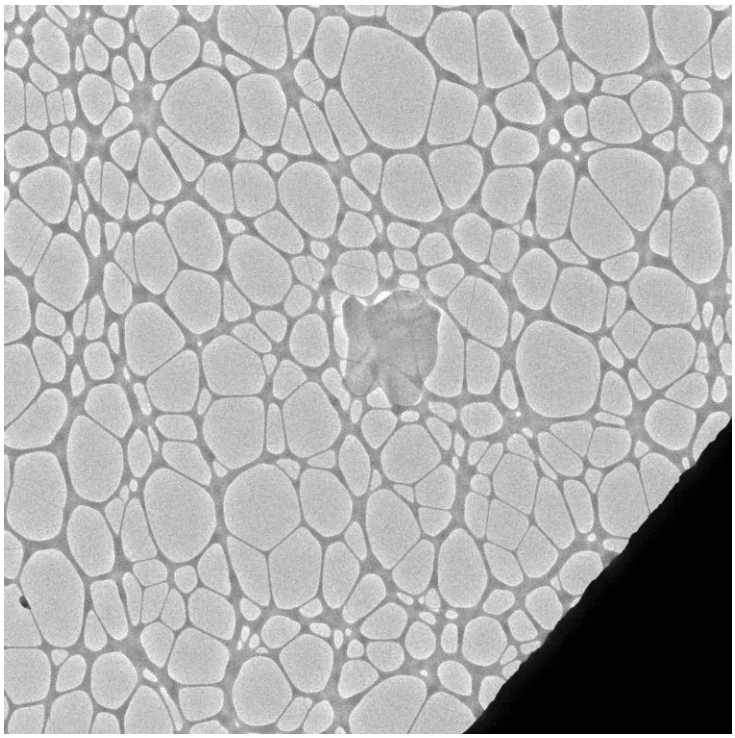
TEM image of **SI-3** crystal at 2600x magnification



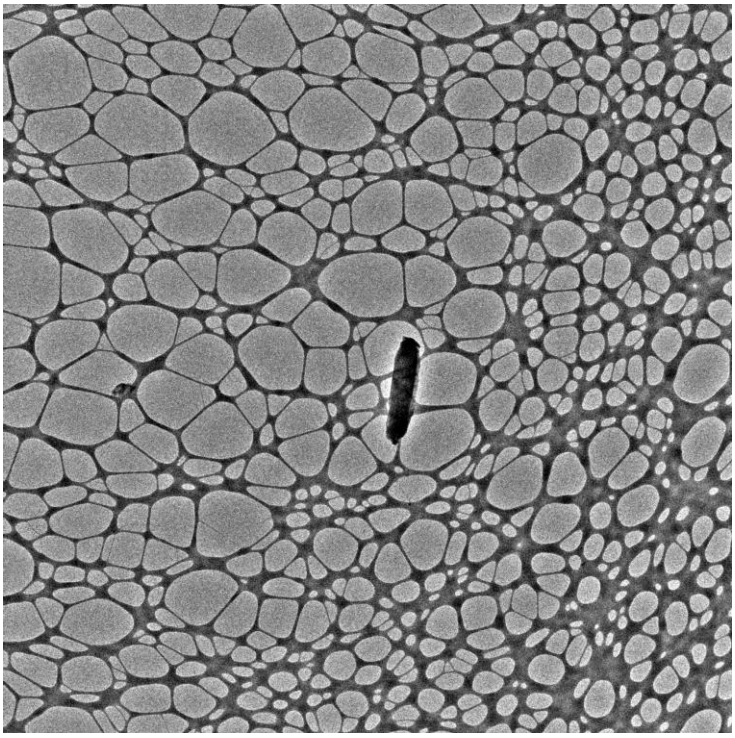
TEM image of **SI-4** crystal at 2600x magnification



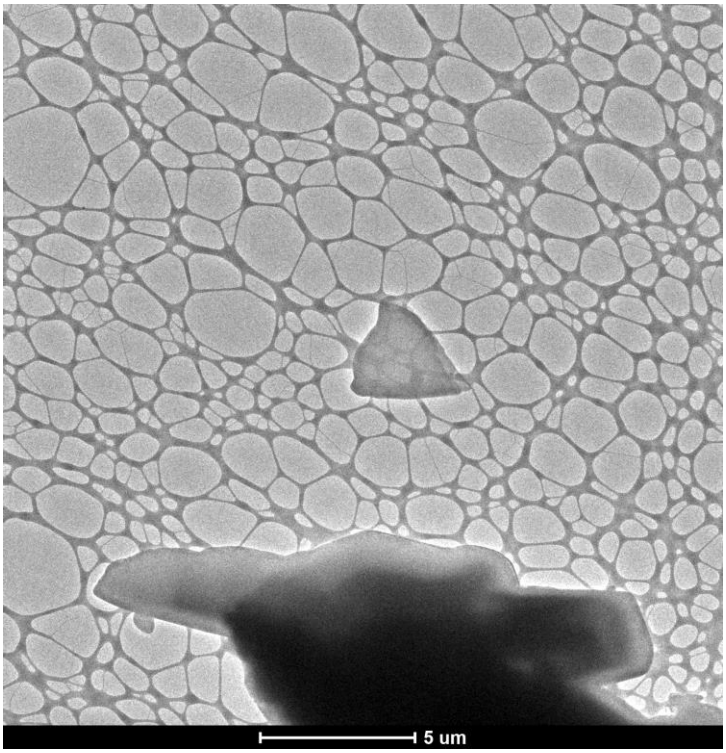
TEM image of **SI-5** crystal at 2600x magnification



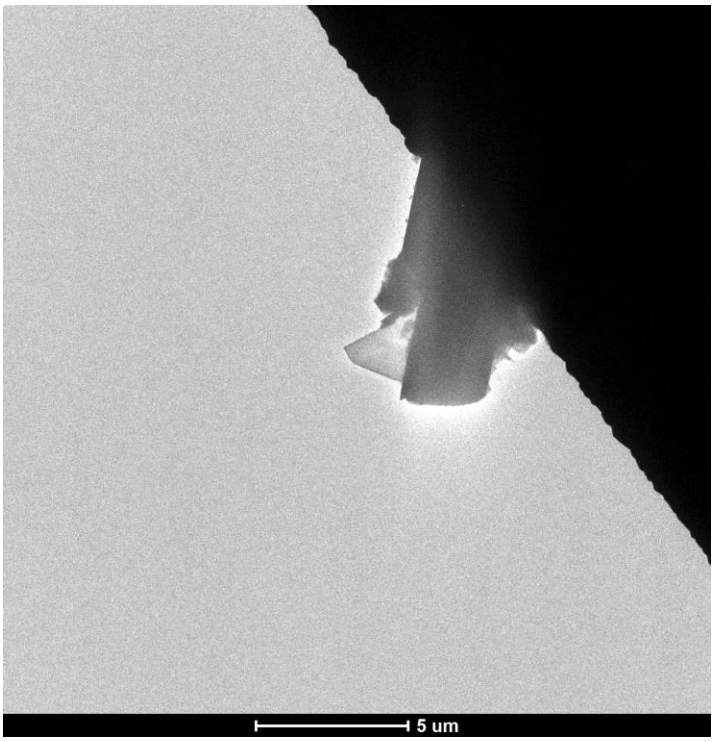
TEM image of **SI-6** crystal at 2600x magnification.



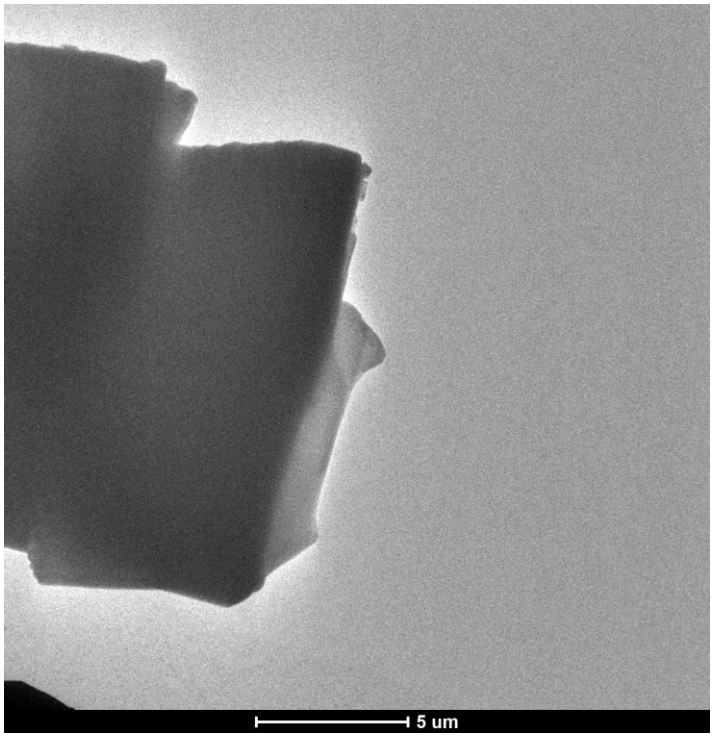
TEM image of **SI-7** crystal at 2600x magnification.



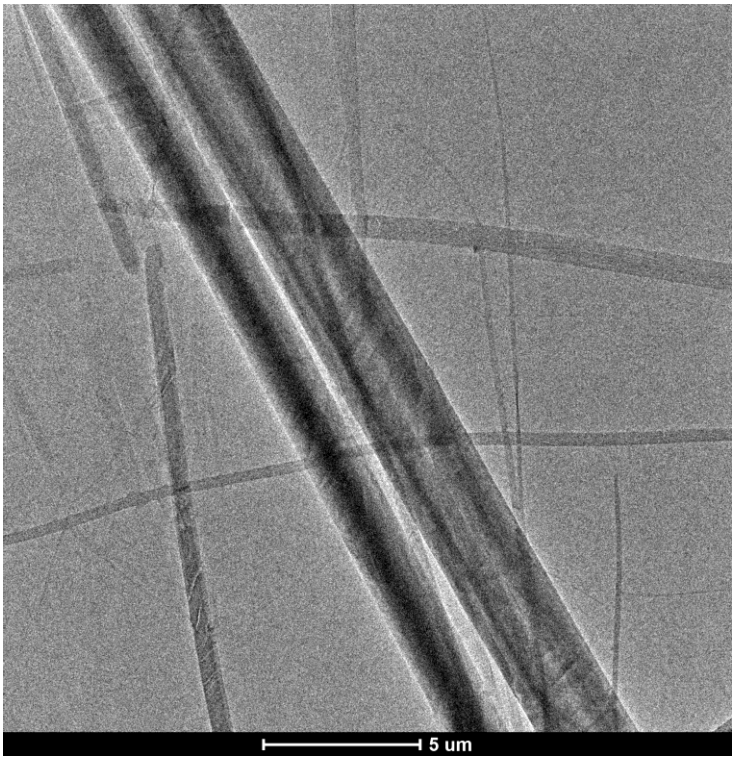
TEM image of **SI-8** crystal at 2600x magnification.



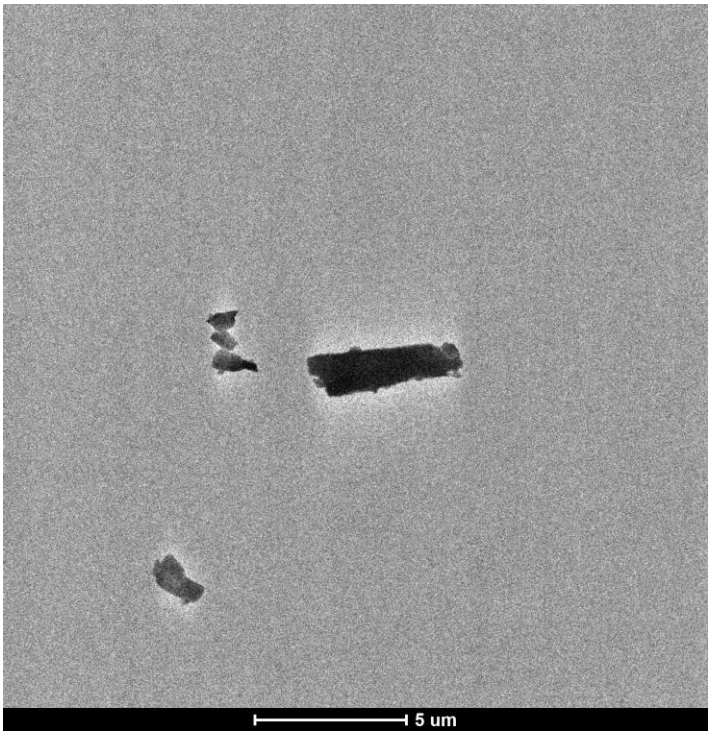
TEM image of **SI-9** crystal at 2600x magnification.



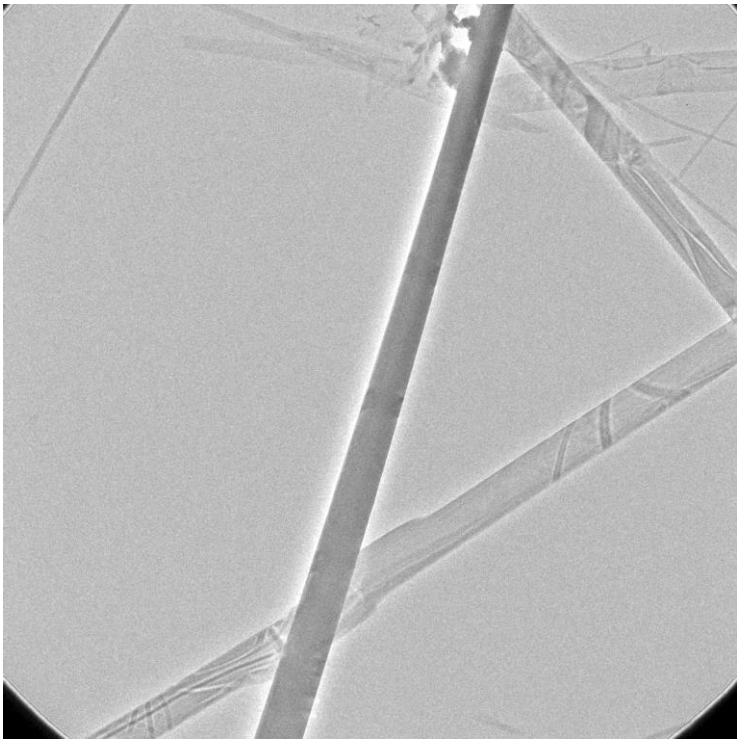
TEM image of **SI-10** crystal at 2600x magnification.



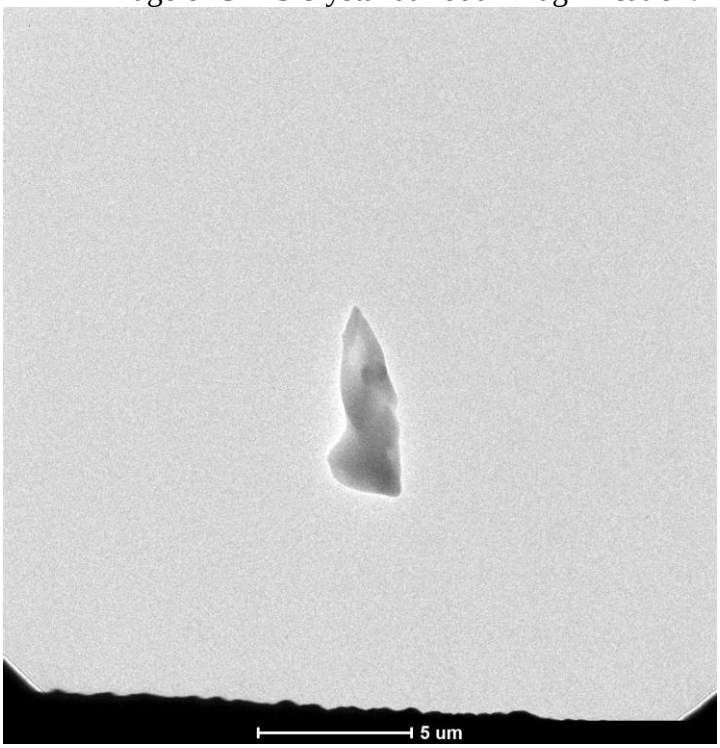
TEM image of **SI-11** crystal at 2600x magnification.



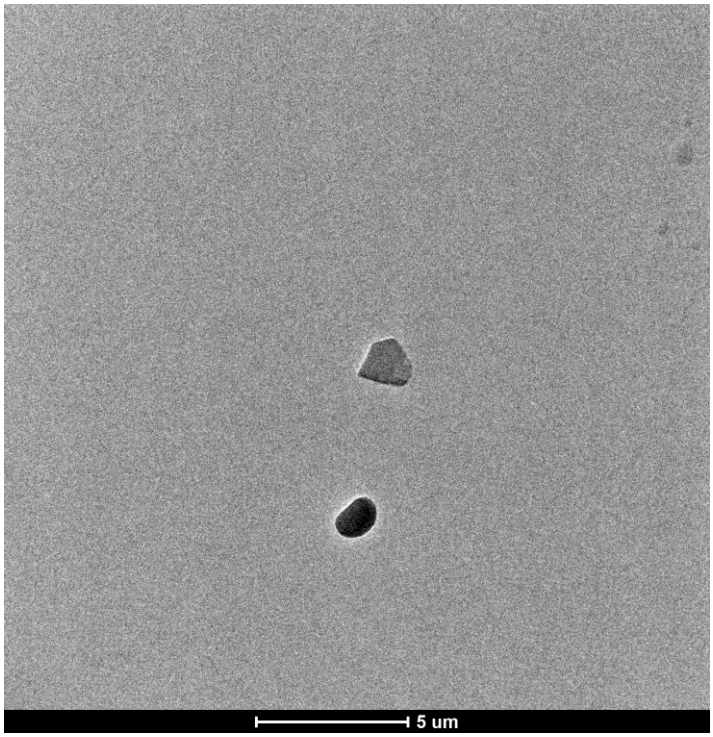
TEM image of **SI-12** crystal at 2600x magnification.



TEM image of **SI-13** crystal at 2600x magnification.

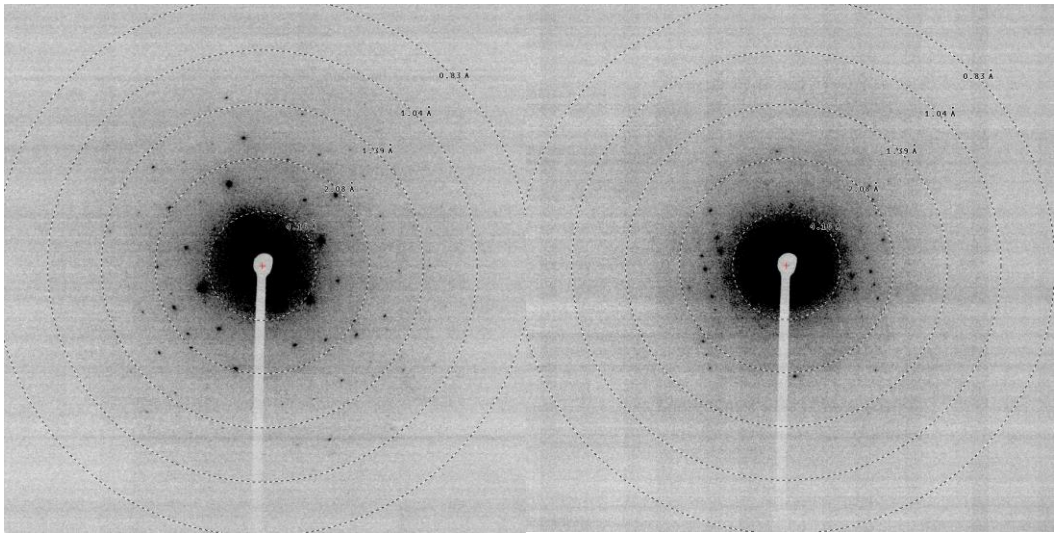


TEM image of **SI-14** crystal at 2600x magnification.

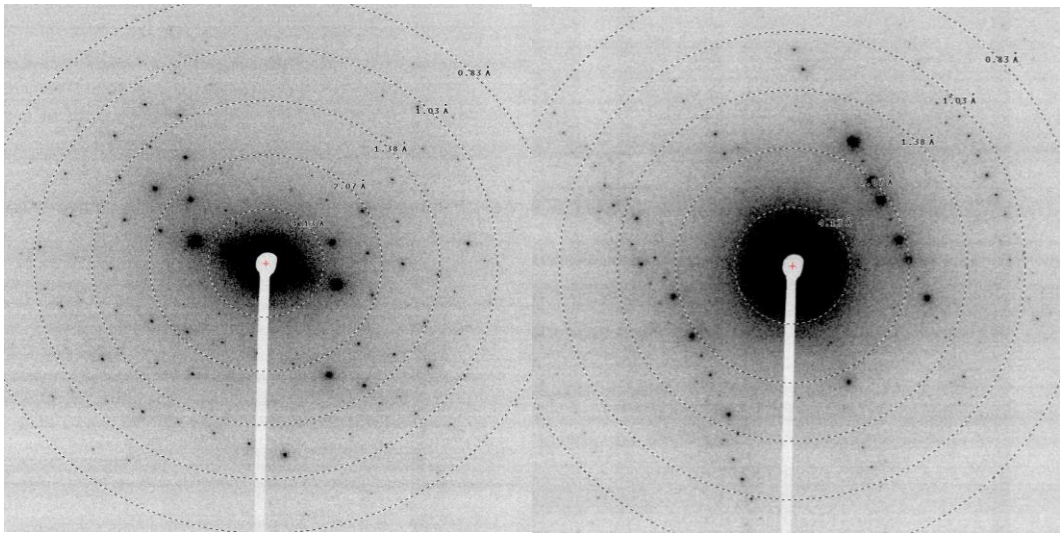


TEM image of **SI-15** crystal at 2600x magnificati

9. Images of Diffraction Resolution Loss



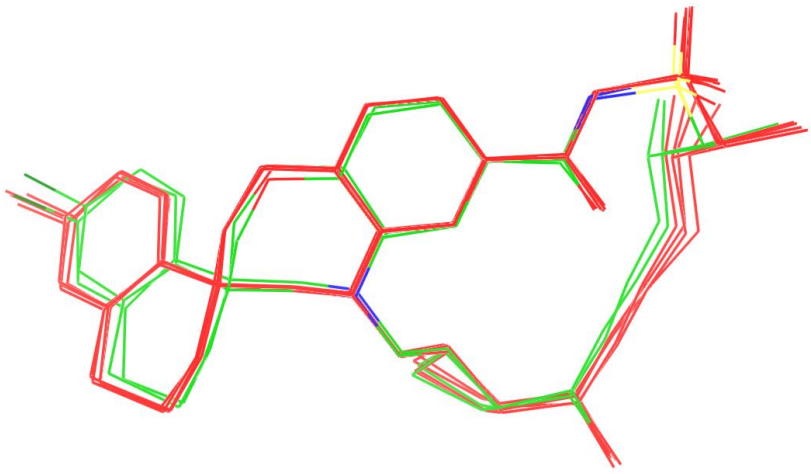
TEM diffraction movie frame 1 (left) and frame 50 (right) of **SI-9** crystal at 960 mm detector distance at room temperature.



TEM diffraction movie frame 1 (left) and frame 50 (right) of **SI-9** crystal at 960 mm detector distance at cryogenic temperature.

10. Comparison of MicroED Data to X-ray Structures

Each **SI-7** molecule from the asymmetric unit of two previously solved single crystal X-ray structures **AGX918A** and **AGX918B**, corresponding to solvated structures of **SI-7** containing two molecules in the asymmetric unit with differing solvation states, were overlayed with the microED structure of **SI-7**.



MicroED (green) and X-ray (red) crystallography data overlay of **SI-7**. Any solvent molecules observed in the crystal structure were removed for this analysis.

Structure Name	Source	RMS	Maximum Difference
AGX918A_1	X-ray	0.0	0.0
AGX918A_2	X-ray	0.5645	2.5284
AGX918C_1	X-ray	0.1869	0.5034
AGX918C_2	X-ray	0.2704	0.8914
SI-7_1	Electron	0.6067	2.6340
SI-7_2	Electron	0.2539	0.4747

RMS of structure overlay comparing one molecule of **AGX918A** to remaining five molecules in **AGX918A**, **AGX918B**, and **SI-7**.

11. Automated Data Processing Procedure

Movie files were saved in a standardized format separated by underscores to allow for automated data processing. An example format is provided below:

samplename-mov1_960_0.3_3_cryo.ser

samplename-mov1 can be any name not including an underscore or special character. This will become the name of the folder containing processed data.

960 is the detector distance used in mm. This can be set to any value.

0.3 is the rotation speed of the stage, in °/s.

3 is the image integration time.

cryo can be any additional notes about the sample and can include underscores.

On a computer running Ubuntu Windows Subsystem for Linux with properly installed XDS suite and free ser2smv[21] data conversion file, “python3 auto_indexing.py” is called to run Python3.8 in a folder containing an executable copy of ser2smv, the python scripts, and the .ser movie files to be processed. Merging and solutions obtained subsequent to autoprocessing were done by the user.

12. Automated Data Processing Python Code

auto_indexing.py

```
"""
Written by Jessica Burch, jessburch@g.ucla.edu
This is a script to batch process individual MicroED datasets using XDS.
version: 03/01/2021
"""
import os
import shutil

def main():
    stats = open("stats.LP","w")
    stats.write("Data summary: ")
    files = os.listdir(".")
    if os.path.isfile("ser2smv") == True:
        for name in files:
            if name.endswith(".ser"):
                newname = name.split("_")
                path = os.getcwd()
                os.mkdir(path + "/" + str(newname[0]))
                os.mkdir(path + "/" + str(newname[0]) + "/images")
                os.mkdir(path + "/" + str(newname[0]) + "/auto_process")
                shutil.move(name,str(path + "/" + newname[0] + "/" + name))
                shutil.copyfile('xds_for_me.py', str(path + "/" + newname[0]
                + "/xds_for_me.py"))
                print("Setting up files for " + newname[0] + ".")
                os.chdir(path + "/" + str(newname[0]))
                os.system("python3 xds_for_me.py")
                os.system("rm xds_for_me.py")
                with open('auto_process/XSCALE.LP', 'r') as f:
                    lines = f.readlines()
```

```
            for index, line in enumerate(lines):
               if " ========== STATISTICS OF INPUT DATA SET ==========" 
in line:
                  t = lines[index-3]
                  t1 = t.split()
                  completeness = t1[4]
                  Roverall = t1[5]

                  l = lines[index-13]
                  l1 = l.split()
                  t = l1[5]
                  t2 = t[:-1]
                  if float(t2) < 100 and float(t2) > 0:
                     with open('xscale_report.LP','w') as f1:
                        f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                              + " " + str(l1[5]))
                     l = lines[index-12]
                     l1 = l.split()
                     t = l1[5]
                     t2 = t[:-1]
                  if float(t2) < 100 and float(t2) > 0:
                     with open('xscale_report.LP','a') as f1:
                        f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                              + " " + str(l1[5]))
                     l = lines[index-11]
                     l1 = l.split()
                     t = l1[5]
                     t2 = t[:-1]
                  if 1 < float(t2) < 100 and float(t2) > 0:
                     with open('xscale_report.LP','a') as f1:
                        f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                              + " " + str(l1[5]))
                     l = lines[index-10]
                     l1 = l.split()
                     t = l1[5]
                     t2 = t[:-1]
                  if float(t2) < 100 and float(t2) > 0:
                     with open('xscale_report.LP','a') as f1:
                        f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                              + " " + str(l1[5]))
                     l = lines[index-9]
                     l1 = l.split()
                     t = l1[5]
                     t2 = t[:-1]
                  if float(t2) < 100 and float(t2) > 0:
                     with open('xscale_report.LP','a') as f1:
```

```
            f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                  + " " + str(l1[5]))
        l = lines[index-8]
        l1 = l.split()
        t = l1[5]
        t2 = t[:-1]
    if float(t2) < 100 and float(t2) > 0:
        with open('xscale_report.LP','a') as f1:
            f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                  + " " + str(l1[5]))
        l = lines[index-7]
        l1 = l.split()
        t = l1[5]
        t2 = t[:-1]
    if float(t2) < 100 and float(t2) > 0:
        with open('xscale_report.LP','a') as f1:
            f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                    + " " + str(l1[5]))
        l = lines[index-6]
        l1 = l.split()
        t = l1[5]
        t2 = t[:-1]
    if float(t2) < 100 and float(t2) > 0:
        with open('xscale_report.LP','a') as f1:
            f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                  + " " + str(l1[5]))
        l = lines[index-5]
        l1 = l.split()
        t = l1[5]
        t2 = t[:-1]
    if float(t2) < 100 and float(t2) > 0:
        with open('xscale_report.LP','a') as f1:
            f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                  + " " + str(l1[5]))
        l = lines[index-4]
        l1 = l.split()
        t = l1[5]
        t2 = t[:-1]
    if float(t2) < 100 and float(t2) > 0:
        with open('xscale_report.LP','a') as f1:
            f1.write("\n" + str(l1[0]) + " " + str(l1[4])
                  + " " + str(l1[5]))
    else:
        with open('xscale_report.LP','a') as f1:
            f1.write("\n ! low resolution data")
```

```
            with open('xscale_report.LP','r') as f2:
                xsc = f2.read()
            with open('auto_process/stats.LP','r') as f3:
                ind = f3.read()
            os.chdir("..")
            stats = open("stats.LP","a")
            stats.write("\n==========\n" + newname[0] + "\n" + ind + "\nXSCALE
stats")
            stats.write("\n" + str(t1[4]) + " " + str(t1[5]) + "\n" + xsc)

    if os.path.isfile("ser2smv") == False:
        print("Please add an executable copy of ser2smv to this folder!\n" +
              "Download from https://cryoem.ucla.edu/downloads/snapshots")

if __name__== "__main__":
    main()
```

xds_for_me.py

```
"""
Written by Jessica Burch, jessburch@g.ucla.edu
This is a script to automate indexing of MicroED data using XDS.
version: 03/01/2021
"""
import os
from subprocess import run
import random
"""
This portion reads your file name and converts .ser files to images.
Must have ser2smv program in the folder containing these scripts
This can be downloaded from https://cryoem.ucla.edu/downloads/snapshots
"""
def main():
    files = os.listdir(".")
    for name in files:
        if name.endswith(".ser"):
            newname = name.split("_")
            path = os.getcwd()
            newpath = str(path + "/images")
            os.chdir(newpath)
            #This is where data collection information such as
            conversion = str(path + "/../ser2smv -P 0.014 -B 2 -r " + newname[2]
                        + " -w 0.0251 -d " + newname[1] + " -E " + newname[3] +
                        " -M 200 -v -o " + newname[0] + "_###.img " + path
                        + "/" + name)
            print("Converting your .ser file to .img frames.")
```

```
            os.system(conversion + ' > summary.LP')
            global movname
            movname = newname[0]
            with open('summary.LP') as f1:
               lines = f1.readlines()
            with open('summary.LP', 'w') as f2:
               f2.writelines(lines[-15:])
            with open('summary.LP', 'r') as f:
               line = f.readline()
               for line in f:
                  if "+++" in line:
                     element = str.split(line)
                     frame = str(element[2])
                     print("You have " + frame + " images.")

      path = os.getcwd()
      newpath = str(path + "/../auto_process")
      os.chdir(newpath)
      f = open("XDS.INP","w+")
      if os.path.isfile("*.LP") == True:
         os.remove("*.LP")
      if os.path.isfile("*.XDS") == True:
         os.remove("*.XDS")
      if os.path.isfile("*.HKL") == True:
         os.remove("*.HKL")
      f.write("JOB= XYCORR INIT COLSPOT IDXREF DEFPIX INTEGRATE
CORRECT" +
            "\n!JOB=DEFPIX INTEGRATE CORRECT")
      #These are estimates for our beam center. The beam may be slightly off in
      #actuality, but XDS does a good job of refining the beam center if the
      #values are close.
      x = str("1018")
      y = str("1000")
      osc = str(float(newname[3]) * float(newname[2]))
      #This corrected distance value arises from indexing diffraction data of
      #standard samples on our TEM and adjusting the detector distance value
      #until these standards agree with the X-ray unit cell.
      corrected_distance = float(newname[1]) * 0.943
      data_path = str(path + "/" + newname[0])
      f.write("\nORGX= " + x + " ORGY= " + y + " ! check these using adxv" +
            "\nDETECTOR_DISTANCE= " + str(corrected_distance) +
            "\nOSCILLATION_RANGE= " + osc + "\nX-RAY_WAVELENGTH=
0.0251000002")
      f.write("\nNAME_TEMPLATE_OF_DATA_FRAMES=" + data_path + "_???.img" +
           "\nBACKGROUND_RANGE=1
10\n!DELPHI=15\n!SPACE_GROUP_NUMBER=0"
```

```
        + "\n!UNIT_CELL_CONSTANTS= 1 1 1 90 90 90")
   #"res" is the high resolution cutoff based on our detector distance.
   if newname[1] == str("1050"):
      res = str("0.8")
   elif newname[1] == str("1100"):
      res = str("0.9")
   elif newname[1] == str("960"):
      res = str("0.8")
   elif newname[1] == str("850"):
      res = str("0.65")
   elif newname[1] == str("1350"):
      res = str("1.1")
   elif newname[1] == str("670"):
      res = str("0.45")
   elif newname[1] == str("420"):
      res = str("0.25")
   elif newname[1] == str("2200"):
      res = str("1.7")
   elif newname[1] == str("330"):
      res = str("0.15")

   #An important value to change based on your microscope is "ROTATION_AXIS"
   f.write("\nINCLUDE_RESOLUTION_RANGE= 40 " + res +
         "\nTEST_RESOLUTION_RANGE= 40 " + res + "\nTRUSTED_REGION=0.0
1.2"+
         "\nVALUE_RANGE_FOR_TRUSTED_DETECTOR_PIXELS=6000. 30000." +
         "! parameters for detector and beamline:" +
         "\nDETECTOR= ADSC MINIMUM_VALID_PIXEL_VALUE= 1
OVERLOAD= 65000" +
         "\nSENSOR_THICKNESS= 0.01" + "\nNX= 2048 NY= 2048 QX=
0.0280000009"
         + " QY= 0.0280000009" + "\nROTATION_AXIS=0 -1 0" +
         "\nDIRECTION_OF_DETECTOR_X-AXIS=1 0 0" +
         "\nDIRECTION_OF_DETECTOR_Y-AXIS=0 1 0" +
         "\nINCIDENT_BEAM_DIRECTION=0 0
1\nFRACTION_OF_POLARIZATION=0.98"
         + "\nPOLARIZATION_PLANE_NORMAL=0 1 0" +
         "\nREFINE(IDXREF)=CELL BEAM ORIENTATION AXIS ! DISTANCE" +
         "\nREFINE(INTEGRATE)= DISTANCE BEAM ORIENTATION ! AXIS
CELL" +
         "\nREFINE(CORRECT)=CELL BEAM ORIENTATION AXIS ! DISTANCE !"
+
         "\n\nDATA_RANGE= 1 " + str(element[2]) + "\nSPOT_RANGE= 1 "
         + str(element[2]))
   sp = "4"
   minpix = "7"
```

```
    f.write("\nSTRONG_PIXEL= " + sp +
"\nMINIMUM_NUMBER_OF_PIXELS_IN_A_SPOT= "
          + minpix + "\n!\n!")
    f.close()
    xds_out = open("XDS.LP", "w+")
    print("XDS is running...")
    run("xds", stdout= xds_out)

def autoprocessing():
    if os.path.isfile('X-CORRECTIONS.cbf') == False:
        xds_out = open("XDS.LP", "w+")
        print("XDS is running...")
        run("xds", stdout= xds_out)

    if os.path.isfile('XPARM.XDS') == False:
        for i in range(10):
            with open('XDS.INP') as f1:
                lines = f1.readlines()
            with open('XDS.INP', 'w') as f2:
                strong = random.randrange(3,9,1)
                mpix = random.randrange(4,9,1)
                f2.writelines(lines[:-4])
                f2.write("STRONG_PIXEL= " + str(strong) +
                        "\nMINIMUM_NUMBER_OF_PIXELS_IN_A_SPOT= " + str(mpix) +
                        "\n!\n!")
                f2.close()
                print("Screening new indexing values.")
                xds_out = open("XDS.LP", "w+")
                run("xds",stdout= xds_out)
            if os.path.isfile('XPARM.XDS') == True:
                if os.path.isfile('DEFPIX.LP') == False:
                    with open('XDS.INP') as f1:
                        lines = f1.readlines()
                    with open('XDS.INP', 'w') as f2:
                        f2.write("!JOB=XYCORR INIT COLSPOT IDXREF DEFPIX
INTEGRATE CORRECT"
                                + "\nJOB=DEFPIX INTEGRATE CORRECT\n")
                        f2.writelines(lines[2:])
                        f2.close()
                        print("Less than 70% of spots went through. Running with JOB= DEFPIX
"
                             + "INTEGRATE CORRECT.")
                        xds_out = open("XDS.LP", "w+")
                        run("xds",stdout= xds_out)
                        if os.path.isfile('XPARM.XDS') == False:
                            print("Unable to autoprocess " + movname + "!")
```

```
                    exit()
                else:
                    return autoprocessing()
            else:
                return autoprocessing()
        else:
            print("Unable to autoprocess " + movname + "!")
            f2.close()
            exit()

    if os.path.isfile('DEFPIX.LP') == False:
        with open('XDS.INP') as f1:
            lines = f1.readlines()
        with open('XDS.INP', 'w') as f2:
            f2.write("!JOB=XYCORR INIT COLSPOT IDXREF DEFPIX INTEGRATE
CORRECT"
                + "\nJOB=DEFPIX INTEGRATE CORRECT\n")
            f2.writelines(lines[2:])
            f2.close()
        print("Less than 70% of spots went through. Running with JOB= DEFPIX "
            + "INTEGRATE CORRECT.")
        xds_out = open("XDS.LP", "w+")
        run("xds",stdout= xds_out)
        return autoprocessing()

    if os.path.isfile("INTEGRATE.HKL") == False:
        with open('XDS.INP') as f1:
            lines = f1.readlines()
        with open('XDS.INP', 'w') as f2:
            f2.writelines(lines)
            f2.write("\nBEAM_DIVERGENCE= 0.03 BEAM_DIVERGENCE_E.S.D.=
0.003" +
            "\nREFLECTING_RANGE=1.0 REFLECTING_RANGE_E.S.D.= 0.2")
            f2.close()
            print("Adding beam divergence values to correct a common error.")
            xds_out = open("XDS.LP", "w+")
            run("xds",stdout= xds_out)
            return autoprocessing()
    if os.path.isfile("CORRECT.LP") == True:
        print ("Successful indexing!")
        return mosaicity()
def mosaicity():
    with open('XDS.INP') as f1:
        lines = f1.readlines()
    with open('XDS.INP', 'w') as f2:
```

```
        f2.write("!JOB=XYCORR INIT COLSPOT IDXREF DEFPIX INTEGRATE
CORRECT"
                 + "\nJOB=DEFPIX INTEGRATE CORRECT\n")
        f2.writelines(lines[2:-2])
    with open('INTEGRATE.LP', 'r') as l1:
        f2 = open('XDS.INP', 'a')
        line = l1.readline()
        for line in l1:
            if "BEAM_DIVERGENCE= " in line:
                f2.write(line)
            if "REFLECTING_RANGE=" in line:
                f2.write(line)
        f2.close()
        return iterate_opt()

def iterate_opt():
    with open('XDS.LP') as f1:
        lines = f1.readlines()
    with open('XDS.LP', 'w') as f2:
        f2.writelines(lines[-26:])
    with open('XDS.LP', 'r') as f:
        line = f.readline()
        for line in f:
            if "     a        b          ISa" in line:
                next_line = f.readline()
                stats = str.split(next_line)
                Isa1 = float(stats[2])
                print("Isa: " + str(Isa1) + ". Testing new values now.")
    xds_out = open("XDS.LP", "w+")
    run("xds",stdout= xds_out)
    with open('XDS.LP') as f1:
        lines = f1.readlines()
    with open('XDS.LP', 'w') as f2:
        f2.writelines(lines[-26:])
    with open('XDS.LP', 'r') as f:
        line = f.readline()
        for line in f:
            if "     a        b          ISa" in line:
                new_next_line = f.readline()
                new_stats = str.split(new_next_line)
                Isa2 = float(new_stats[2])
                print("Isa: " + str(Isa2))
            if "SPACE_GROUP_NUMBER=" in line:
                number = str.split(line)
                space_group = number[1]
            if "UNIT_CELL_CONSTANTS=" in line:
```

```
            cell = str.split(line)
            temp = cell[-6:]
            temp_str = str(temp).strip("]['")
            temp_str2 = temp_str.replace(",","")
            unit_cell = temp_str2.replace("'","")
    Isa_change = abs(Isa2 - Isa1)
    if Isa_change > 0.5:
        print("Optimizing beam divergence values.")
        return iterate_opt()
    else:
        print("Optimized beam divergence values.")
        f = open('stats.LP','w')
        f.write(str(space_group) + "\n" + unit_cell)

        f.close
        print("Autoprocessing found space group " + str(space_group) + " and a unit cell of
"
            + "\n" + unit_cell)


def scale_conv():
    xscale = open('XSCALE.INP','w')
    xscale_out = open("xscale.LP","w+")
    m = movname
    xscale.write("OUTPUT_FILE= " + m +".ahkl"+"\nINPUT_FILE= XDS_ASCII.HKL"
              + "\nRESOLUTION_SHELLS= 10 8 5 3 2.3 2.0 1.7 1.5 1.3 " +
              "1.2 1.1 1.0 0.90 0.80")
    xscale.close()
    run("xscale", stdout= xscale_out)
    print("Data scaled with XSCALE.")
    xdsconv_out = open("xdsconv.LP", "w+")
    xdsconv = open('XDSCONV.INP','w')
    xdsconv.write("INPUT_FILE= " + m + ".ahkl" + "\nOUTPUT_FILE= " +
              m + ".hkl" + " SHELX" +
              "\nGENERATE_FRACTION_OF_TEST_REFLECTIONS=0.10"
              + "\nFRIEDEL'S_LAW=FALSE")
    xdsconv.close()
    run("xdsconv",stdout= xdsconv_out)
    print("Data converted for use in shelx!")
```

```
if __name__== "__main__":
    main()
    autoprocessing()
    scale_conv()
"""

"""
```

13. References

(1) Kabsch, W. *Acta Cryst.* **2010**, *D66*, 125–132.

(2) Kabsch, W. *Acta Cryst.* **2010,** *D66*, 133–144.

(3) Hattne, J., *et al. Acta Cryst.* ***2015***, *71*, 353–360.

(4) Sheldrick, G. M. A short history of SHELX. *Acta Cryst.* **2008**, *A64*, 112–122.

(5) Sheldrick, G. M. *Acta Cryst.* **2015** *A71*, 3–8.

(6) Sheldrick, G. M. *Acta Cryst.* **2015**, *C71*, 3–8.

(7) Hübschle, C. B., Sheldrick, G. M. & Dittrich, B. *J. Appl. Cryst.* **2011,** *44*, 1281–1284.

(8) *PCT Int. Appl.* **2016**, WO 2016187308 A1 20161124.

(9) *PCT Int. Appl.* **2018**, WO 2018097945 A1 20180531.

(10) PCT/US2015/047472, WO 2016033486 A1.

(11) *PCT Int. Appl.* **2018**, WO 2018093576 A1 20180524.

(12) *PCT Int. Appl.* **2018**, WO 2018093577 A1 20180524.

(13) *PCT Int. Appl.* **2018**, WO 2018093579 A1 20180524.

(14) *PCT Int. Appl.* **2018**, WO 2018097944 A1 20180531.

(15) *PCT Int. Appl.* **2018**, WO 2018097945 A1 20180531.

(16) *PCT Int. Appl.* **2019**, WO 2019006231 A1 20190103.

(17) *PCT Int. Appl.* **2016**, WO 2016141035 A1.

(18) *J. Org. Chem.* **2014**, *79*, 3684–3687.

(19) *J. Med. Chem.* **2014**, *57*, 9796–9810.

(20) *J. Med. Chem.* **2018**, *61*, 453–461.

(21) Program ser2smv obtained from https://cryoem.ucla.edu/downloads/snapshots.