Supplementary Materials for

The CryoEM method MicroED as a powerful tool for small molecule structure determination

Christopher G. Jones[1†], Michael W. Martynowycz[2†], Johan Hattne[2], Tyler J. Fulton[3], Brian M. Stoltz[3$], Jose A. Rodriguez[1$], Hosea M. Nelson[1$], Tamir Gonen[2$]

[1]Department of Chemistry and Biochemistry, University of California, Los Angeles, 90095, USA

[2]Howard Hughes Medical Institute, Departments of Biological Chemistry and Physiology, University of California, Los Angeles, 90095, USA

[3]The Warren and Katharine Schlinger Laboratory of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, 91125, USA

[$]Correspondence to: stoltz@caltech.edu, jrodriguez@chem.ucla.edu, hosea@chem.ucla.edu, tgonen@ucla.edu

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S11
Captions for Movies S1
References

Other Supporting Online Material for this manuscript includes the following:

Movies S1

Table of Contents

1. Materials and methods

All commercial samples were used as received with no additional crystallization or chemical modification. Ethisterone, cinchonine, carbamazepine, and biotin were purchased from Sigma-Aldrich. Brucine was purchased from the The Matheson Company, Inc. Progesterone was purchased from Preparations Laboratories Inc. Thiostrepton was purchased from EMD Millipore. CVS®-brand acetaminophen and Kroger® brand ibuprofen were used as over-the-counter medications. (+)-Limaspermadine and HKL-I-029 were synthesized according to previously reported literature procedures (*1*, *2*).

1.1 Sample Preparation

To prepare commercial compounds for MicroED, approximately 1 mg of product as received was placed between two microscope slides and ground to a fine powder. The ground powder was placed into an Eppendorf tube along with a pre-clipped Quantifoil R2/2 Cu300 or Quantifoil R1/4 Cu300 mesh grid. The TEM grid was then removed from the Eppendorf tube and gently tapped against a filter paper to remove excess powder. Non-commercial samples of HKL-I-029 and (+)-limaspermadine were concentrated under vacuum to yield a dry film and solid powder respectively. Sample grids of HKL-I-029 were prepared by adding a TEM grid directly to a 20 mL scintillation vial with gentle shaking. (+)-Limaspermadine grids were prepared by scraping the residue off the side of a 20 mL scintillation vial over a TEM grid. Once sample grids were prepared, they were subsequently plunged into liquid nitrogen, placed into the sample cartridge, and loaded into the microscope for analysis. Heterogenous sample mixtures were prepared by adding ~1 mg of biotin, carbamazepine, cinchonine, and brucine to a glass cover slide and grinding to a fine powder. The heterogenous powder was then added to an Eppendorf tube and the grid was prepared in the same manner as the homogeneous samples.

1.2 Instrument Parameters

All data were collected on a Thermo-Fischer Talos Artica electron cryomicroscope operating at an acceleration voltage of 200keV, corresponding to a wavelength of ~0.0251 Å. Screening of

the TEM grids for micro crystals was done by operating the microscope in over focused diffraction mode to minimize diffraction and hysteresis between screening and diffraction operational modes.

1.3 Data Collection Procedure

MicroED data collection was collected in rolling shutter using a Thermo-Fischer CetaD CMOS 4k x 4k camera. Images were collected as a movie as the crystal was continuously rotated in the electron beam (*3*). Typical data collection was performed using a constant tilt rate of ~0.6 ° s^{-1} d over an angular wedge of ~60° between the minimum and maximum tilt ranges of -72° to +72° degrees, respectively. During continuous rotation the camera integrated frames continuously at a rate of 1-3s per frame. The dose rate was calibrated to <0.03 e^{-} $Å^{-2}$ s^{-1}. 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.

Diffraction movies saved as SER files were converted to SMV format using in-house software developed for the CetaD and made freely available online (https://cryoem.ucla.edu/pages/MicroED). Frames were indexed and integrated in XDS, and multiple datasets were scaled and merged using XSCALE (*4*, *5*). The intensities were converted to SHELX format using XDSCONV(*5*). All structures except thiostrepton (see below) were solved by *ab initio* direct methods in SHELXT, and refined in SHELXL as previously described (*6*, *7*).

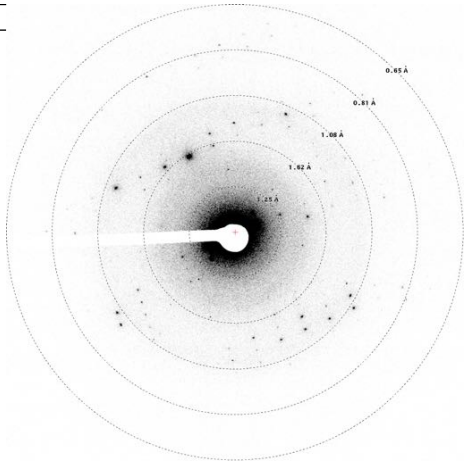
Four datasets from thiostrepton were indexed and integrated in MOSFLM through its graphical user interface, iMosflm (*8*, *9*). Data were merged in AIMLESS, and phased by molecular replacement in MOLREP using 1E9W as a search model (*10*, *11*). The solution was refined using REFMAC5 with electron scattering factors to a resolution of 1.9Å with the free *R* set copied from the initial search model (*12*).

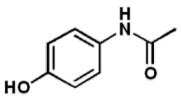
2. Compound Data and Statistics

Individual integration and refinement statistics can be found for each compound in SI Figures 1-11 along with corresponding densities.

3. **Movie S1.** Continuous rotation MicroED data from a carbamazepine nanocrystal with corresponding resolution rings.

acetaminophen	
Stoichiometric formula	$C_8 N_1 O_2$
Temperature (K)	100
Space group	$P 2_1/n$
Unit cell lengths a, b, c (Å)	6.630(2), 8.620(2), 10.790(2)
angles α, β, γ (°)	90.00(3), 97.56(3), 90.00(3)
Reflections (#)	2300 (380)
Unique reflections (#)	874 (141)
R_{obs}	18.3 (34.7)
R_{meas}	22.8 (43.2)
$CC_{1/2}$	95.2 (83.6)
Resolution (Å)	0.8
Completeness (%)	69.9 (70.1)
Total exposure ($e^- Å^{-2}$)	~3
R	0.22
wR2	0.4462
GooF	2.003





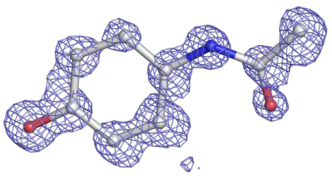
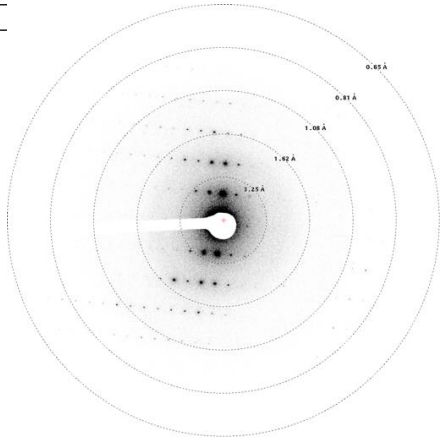
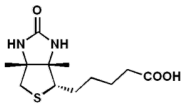


Figure S1. Data processing statistics and final structure of acetaminophen.

biotin	
Stoichiometric formula	$C_{10} N_2 O_3 S$
Temperature (K)	100
Space group	$P 2_1 2_1 2_1$
Unit cell lengths a, b, c (Å)	5.200(2), 10.310(2), 20.910(4)
angles α, β, γ (°)	90.00(3), 90.00(3), 90.00(3)
Reflections (#)	5498 (1081)
Unique reflections (#)	1323 (246)
R_{obs}	20.3 (37.1)
R_{meas}	23.3 (42.1)
$CC_{1/2}$	95.5 (78.4)
Resolution (Å)	0.9
Completeness (%)	82.6 (84.8)
Total exposure ($e^- Å^{-2}$)	~3
R	0.186
wR2	0.3458
GooF	1.818





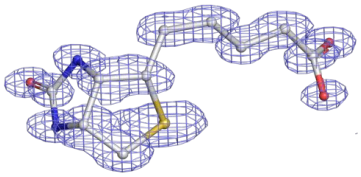
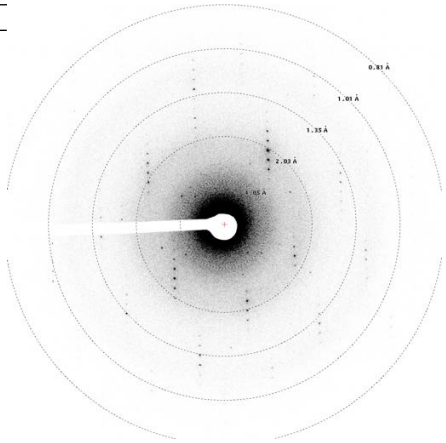
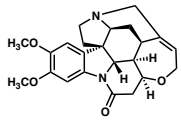


Figure S2. Data processing statistics and final structure of biotin.

brucine	
Stoichiometric formula	$C_{23} N_2 O_4$
Temperature (K)	100
Space group	$P 2_1$
Unit cell lengths a, b, c (Å)	15.340(3), 7.540(2), 20.010(4)
angles α, β, γ (°)	90.00(3), 112.49(3), 90.00(3)
Reflections (#)	12427 (814)
Unique reflections (#)	5858 (416)
R_{obs}	18.2 (56.1)
R_{meas}	24.2 (74.9)
$CC_{1/2}$	95.1 (25.9)
Resolution (Å)	0.9
Completeness (%)	95.3 (96.1)
Total exposure ($e^- Å^{-2}$)	~3
R	0.2244
wR2	0.4468
GooF	1.711





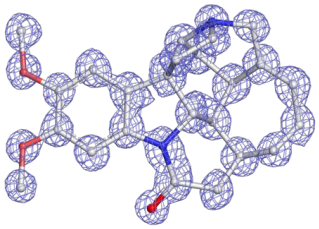
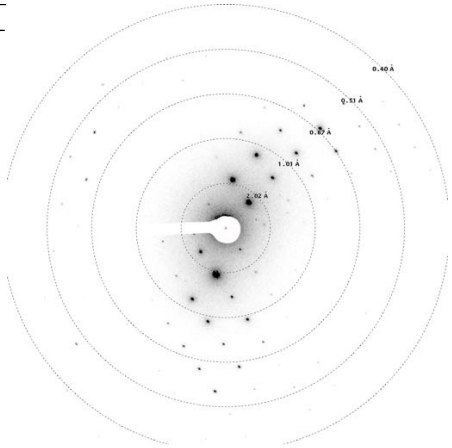
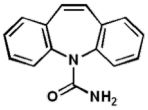


Figure S3. Data processing statistics and final structure of brucine.

carbamazepine	
Stoichiometric formula	$C_{15}\,N_2\,O$
Temperature (K)	100
Space group	$P\,2_1/n$
Unit cell lengths a, b, c (Å)	7.460(2), 11.040(2), 13.760(3)
angles α, β, γ (°)	90.00(3), 92.61(3), 90.00(3)
Reflections (#)	4682 (678)
Unique reflections (#)	1044 (146)
R_{obs}	17.3 (22.1)
R_{meas}	19.5 (24.7)
$CC_{1/2}$	97.3 (93.8)
Resolution (Å)	1.0
Completeness (%)	88.3 (84.9)
Total exposure (e^- Å^{-2})	~3
R	0.1931
wR2	0.3902
GooF	2.398





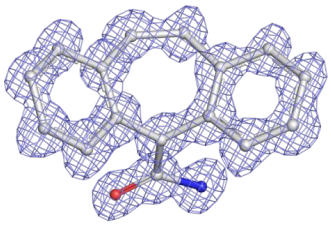
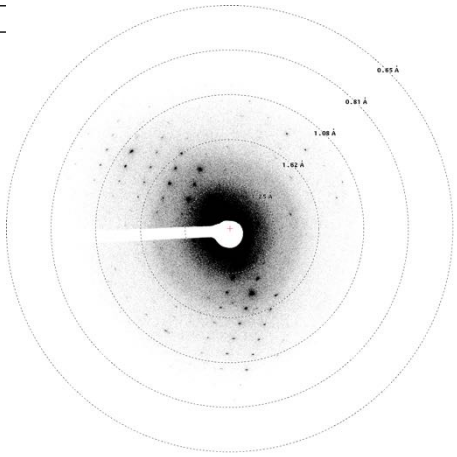
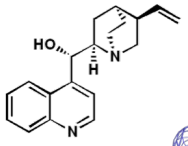


Figure S4. Data processing statistics and final structure of carbamazepine.

cinchonine	
Stoichiometric formula	$C_{19}\,N_2\,O$
Temperature (K)	100
Space group	$P\,2_1/n$
Unit cell lengths a, b, c (Å)	10.710(2), 7.060(2), 11.150(2)
angles α, β, γ (°)	90.00(3), 109.66(3), 90.00(3)
Reflections (#)	1933 (399)
Unique reflections (#)	1289 (262)
R_{obs}	11.0 (14.8)
R_{meas}	15.6 (21.0)
$CC_{1/2}$	95.0 (89.2)
Resolution (Å)	1.0
Completeness (%)	77.4 (78.9)
Total exposure (e^- Å^{-2})	~3
R	0.1793
wR2	0.3907
GooF	1.831





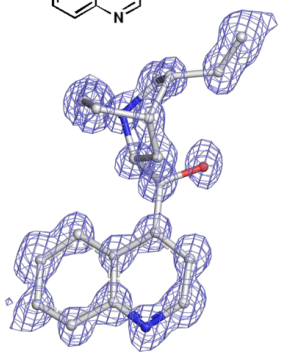
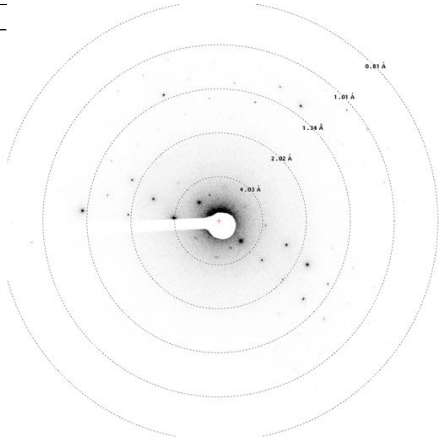
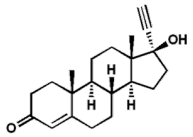


Figure S5. Data processing statistics and final structure of cinchonine.

ethisterone	
Stoichiometric formula	$C_{21}\,O_2$
Temperature (K)	100
Space group	$P\,2_1$
Unit cell lengths a, b, c (Å)	6.43(2), 21.17(4), 6.48(2)
angles α, β, γ (°)	90.00(3), 105.6(3), 90.00(3)
Reflections (#)	1811 (231)
Unique reflections (#)	1506 (197)
R_{obs}	10.0 (25.4)
R_{meas}	14.1 (35.9)
$CC_{1/2}$	97.3 (56.1)
Resolution (Å)	0.9
Completeness (%)	60.8 (54.6)
Total exposure (e^- Å^{-2})	~3
R	0.2481
wR2	0.5109
GooF	2.087





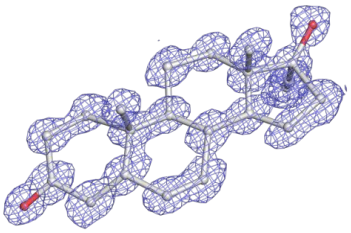
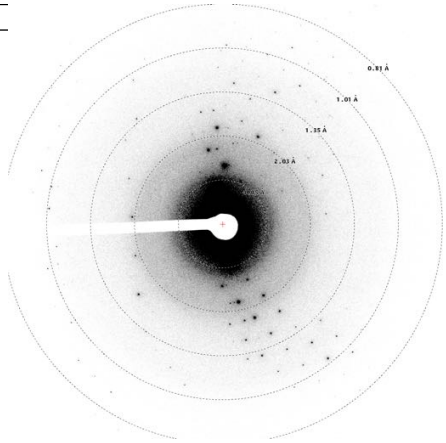
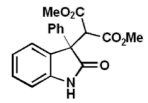


Figure S6. Data processing statistics and final structure of ethisterone.

HKL-I-029	
Stoichiometric formula	C_{21} N O_3
Temperature (K)	100
Space group	*P* 2_1/n
Unit cell lengths a, b, c (Å)	8.280(2), 24.370(5), 8.810(2)
angles α, β, γ (°)	90.00(3), 108.80(3), 90.00(3)
Reflections (#)	3369 (446)
Unique reflections (#)	1970 (262)
R_{obs}	14.1 (22.8)
R_{meas}	18.4 (29.5)
$CC_{1/2}$	94.5 (84.8)
Resolution (Å)	1.0
Completeness (%)	55.3 (55.7)*
Total exposure (e$^-$ Å^{-2})	~3
R	0.2366
wR2	0.4762
GooF	2.656





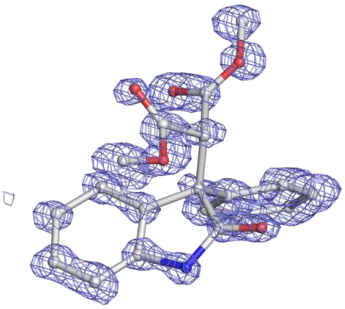
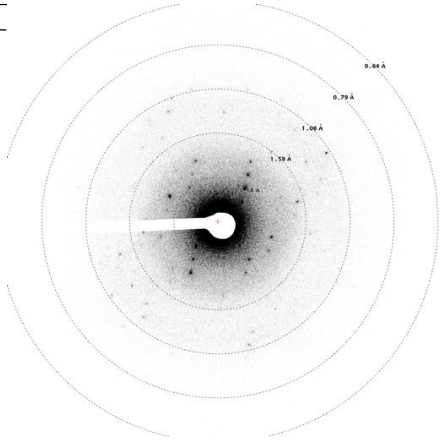
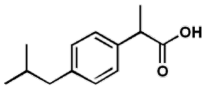


Figure S7. Data processing statistics and final structure of HKL-I-029.
*The completeness of this compound was limited due to preferred orientation.

ibuprofen	
Stoichiometric formula	C_{13} O_2
Temperature (K)	100
Space group	*P* 2_1/c
Unit cell lengths a, b, c (Å)	14.65(3), 7.88(2), 10.73(2)
angles α, β, γ (°)	90.00(3), 99.7(3), 90.00(3)
Reflections (#)	1452 (402)
Unique reflections (#)	506 (138)
R_{obs}	14.7 (20.8)
R_{meas}	17.8 (25.2)
$CC_{1/2}$	97.8 (89.9)
Resolution (Å)	1.1
Completeness (%)	54.3 (53.1)*
Total exposure (e$^-$ Å^{-2})	~3
R	0.2559
wR2	0.5282
GooF	2.686





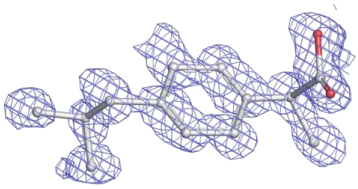
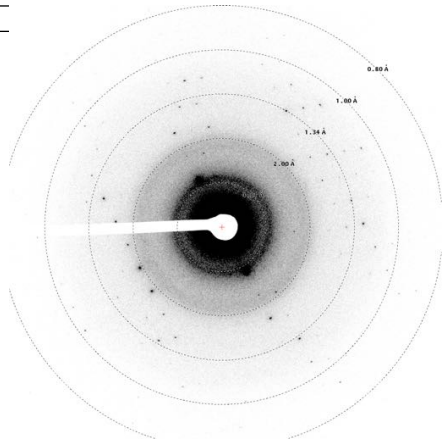
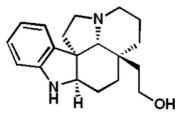


Figure S8. Data processing statistics and final structure of ibuprofen.
*The completeness of this compound was limited due to preferred orientation.

(+)-limaspermidine	
Stoichiometric formula	C_{19} N_2 O
Temperature (K)	100
Space group	*P* 2_1 2_1 2_1
Unit cell lengths a, b, c (Å)	7.620(2), 13.880(3), 15.200(3)
angles α, β, γ (°)	90.00(3), 90.00(3), 90.00(3)
Reflections (#)	8252 (387)
Unique reflections (#)	3430 (185)
R_{obs}	16.7 (68.2)
R_{meas}	21.6 (88.6)
$CC_{1/2}$	97.0 (34.2)
Resolution (Å)	0.77
Completeness (%)	93.0 (69.3)
Total exposure (e$^-$ Å^{-2})	~3
R	0.2422
wR2	0.4309
GooF	1.541





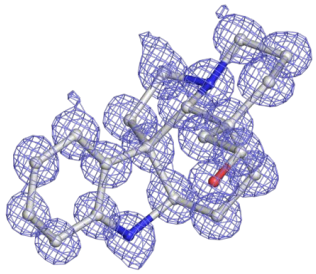
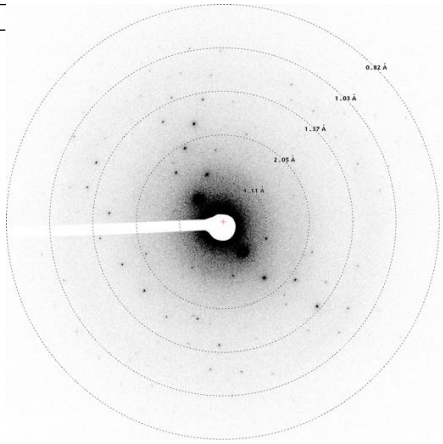
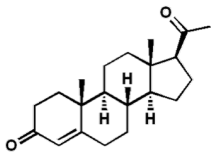


Figure S9. Data processing statistics and final structure of (+)-limaspermidine.

progesterone	
Stoichiometric formula	$C_{21}O_2$
Temperature (K)	100
Space group	$P\,2_1\,2_1\,2_1$
Unit cell lengths a, b, c (Å)	10.380(2), 12.810(3), 13.890(3)
angles α, β, γ (°)	90.00(3), 90.00(3), 90.00(3)
Reflections (#)	4487(577)
Unique reflections (#)	1871 (238)
R_{obs}	14.7 (44.6)
R_{meas}	18.0 (53.9)
$CC_{1/2}$	98.0 (66.1)
Resolution (Å)	0.9
Completeness (%)	72.1 (68.6)
Total exposure ($e^- Å^{-2}$)	~3
R	0.2045
wR2	0.4155
GooF	1.888





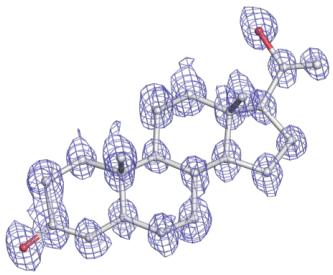
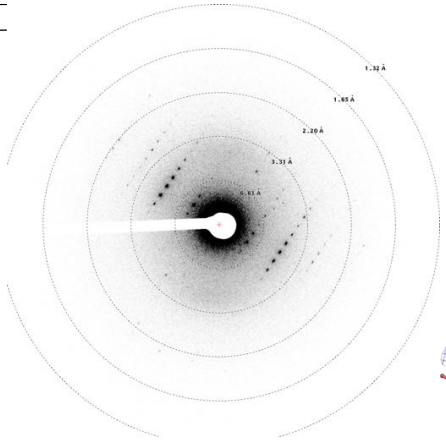
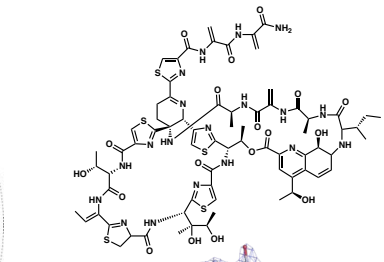


Figure S10. Data processing statistics and final structure of progesterone.

thiostrepton	
Resolution range (Å)	18.99–1.91 (2.13–1.91)
Space group	$P\,4_3\,2_1\,2$
Unit cell lengths a, b, c (Å)	26.219, 26.219, 27.534
angles α, β, γ (°)	90, 90, 90
Total reflections	5578 (458)
Unique reflections	686 (93)
Multiplicity	8.1 (4.9)
Completeness (%)	78.6 (40.3)
Mean $I/\sigma(I)$	5.1 (3.4)
Wilson *B*-factor	2.6
R_{merge}	0.236 (0.320)
R_{meas}	0.251 (0.353)
$CC_{1/2}$	0.985 (0.813)
Reflections used in refinement	620 (92)
R_{work}	0.1818 (0.2191)
R_{free}	0.2396 (0.1766)





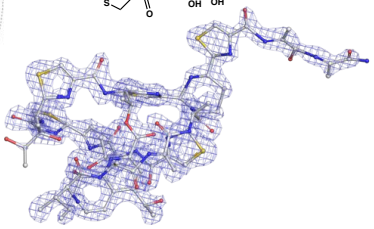


Figure S11. Data processing statistics and final structure of thiostrepton.

References

1. Y. Liu, S.-J. Han, W.-B. Liu, B. M. Stoltz, Catalytic Enantioselective Construction of Quaternary Stereocenters: Assembly of Key Building Blocks for the Synthesis of Biologically Active Molecules. *Acc. Chem. Res.* **48**, 740–751 (2015).

2. B. P. Pritchett, E. J. Donckele, B. M. Stoltz, Enantioselective Catalysis Coupled with Stereodivergent Cyclization Strategies Enables Rapid Syntheses of (+)-Limaspermidine and (+)-Kopsihainanine A. *Angew. Chemie Int. Ed.* **56**, 12624–12627 (2017).

3. B. L. Nannenga, D. Shi, A. G. W. Leslie, T. Gonen, High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods*. **11**, 927–930 (2014).

4. W. Kabsch, Xds. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).

5. W. Kabsch, Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 133–144 (2010).

6. G. M. Sheldrick, SHELXT - Integrated space-group and crystal-structure determination.

Acta Crystallogr. Sect. A Found. Crystallogr. **71**, 3–8 (2015).

7. G. M. Sheldrick, Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **71**, 3–8 (2015).

8. A. G. W. Leslie, H. R. Powell, (2007; http://link.springer.com/10.1007/978-1-4020-6316-9_4), pp. 41–51.

9. T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. W. Leslie, iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 271–281 (2011).

10. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).

11. A. Vagin, A. Teplyakov, MOLREP : an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).

12. G. N. Murshudov *et al.*, REFMAC 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).