Conformational-Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies

Joshua S. Alford, John W. Lampe, Dorothy Brach, Richard Chesworth, Kat Cosmopoulos, Kenneth W. Duncan, Sean T. Eckley, Jeffrey L. Kutok, Alejandra Raimondi, Thomas V. Riera, Brian Shook, Cuyue Tang, Jennifer Totman, and Neil A. Farrow*

*Contact email: nfarrow@epizyme.com

Table of Contents

Synthetic Procedures........................................................................................................................................... 2
Assay Protocols .................................................................................................................................................... 38
  SETD2 (1434-1711) Assay ............................................................................................................................ 38
  In-cell western (ICW) assay of SETD2 Inhibitors (A549 assay) ................................................................. 38
  14 Day Long Term Proliferation (LTP) Assay .............................................................................................. 41
In Vitro ADME and Pharmacokinetic Assays .................................................................................................... 43
  Stability in Hepatocytes .................................................................................................................................. 43
  CYP Inhibition in Human Liver Microsomes ................................................................................................. 45
  Caco-2 Permeability ....................................................................................................................................... 46
  Pharmacokinetics ........................................................................................................................................... 47
In vivo efficacy analysis in mice .......................................................................................................................... 48
Molecular Modeling Methods ............................................................................................................................. 49
Crystallography Protocols ................................................................................................................................. 50
  Protein Production ......................................................................................................................................... 51
  X-ray crystallography .................................................................................................................................... 51
References: ......................................................................................................................................................... 53
Synthetic Procedures

Proton NMR spectra were recorded on Agilent 400 or 300 MHz spectrometers. Proton chemical shifts are reported in ppm (δ) relative to tetramethylsilane and calibrated using the residual solvent resonance (chloroform-d, δ 7.26 ppm; methanol-d4, δ 3.31 ppm). Data are reported as follows: chemical shift (multiplicity [singlet (s), doublet (d), doublet of doublets (dd), doublet of doublets of doublets (ddd), doublet of doublets of triplets (ddt), doublet of triplets (dt), doublet of triplets of doublets of doublets (dddt), triplet of doublets of triplets (tt), triplet of quartets (tq), quartet (q), quartet of doublets (qd), quartet of triplets (qt), pentet (p), pentet of doublets (pd), septet (sept), multiplet (m)], coupling constants [Hz], integration, specific proton assignment).

Unless otherwise noted, all NMR spectra were acquired at ambient temperature. Analytical thin-layer chromatography (TLC) was performed using EMD Millipore silica gel 60 F254 precoated plates (0.25 mm thickness) and were visualized by irradiation with UV light (254 nm) and staining with KMnO4. TLC Rf values are reported. Normal phase flash chromatography was performed using Silicycle silica gel (particle size 32-63 μm). Reversed phase chromatography used C-18 silica and was performed on a Biotage Isolera One purification system. All chemicals were purchased commercially and used as received unless otherwise noted. Building blocks that are commercially available from local vendors (in China) were purchased and used as received. All derivatives tested for biological activity showed >95% purity in LC-MS. No unexpected or unusually high safety hazards were encountered doing this work.

**LC-MS analytical methods:** The instrument used for LC-MS analysis was a Shimadzu LCMS-2000 equipped with 2 LC-30AD pumps for a binary gradient, degasser, autosampler, PDA model SPD-M20A, D2 lamp, 190-400 nm, MS parameter: pos/neg 90-900, nebulizing gas flow: 1.5 mL/min, gas temp: 250 °C.

**Method 1:** Kinetex EVO C18, 50x2.1 mm, 2.6 μm, Temp 35 °C, flow 1.0 mL/min, gradient: t0 = 10% B, t0.0min = 95% B, t2.0min = 95% B, t2.70min = 95% B, t2.80min = 10% B. Mobile Phase A: Water +6.5 mM NH4HCO3 (pH = 10)/Mobile Phase B: MeCN.

**Method 2:** Shim-pack XR-ODS, 50x3.0 mm, 2.2 μm, Temp 35 °C, flow 1.2 mL/min, gradient: t0 = 5% B, t0.0min = 95% B, t2.70min = 95% B, t2.80min = 10% B. Mobile Phase A: Water + 0.05% TFA/Mobile Phase B: MeCN + 0.05% TFA.

**Method 3:** Halo C18, 50x3.0 mm, 2.7 μm, Temp 45 °C, flow 1.5 mL/min, gradient: t0 = 5% B, t1.25min = 45% B, t1.00min = 100% B, t1.60min = 100% B, t4.70min = 5% B Mobile Phase A: Water + 0.05% TFA/Mobile Phase B: MeCN + 0.05% TFA.

**Method 4:** Cortecs C18, 50x2.1 mm, 2.7 μm, Temp 40 °C, flow 1.0 mL/min, gradient: t0 = 10% B, t1.00min = 100% B, t1.50min = 100% B, t1.60min = 10% B, Mobile Phase A: Water + 0.05% FA/Mobile Phase B: MeCN + 0.05% FA.
Overview for Figure 1 and Table 1 Compounds: The preparation of compounds EPZ-719 (1) and 2 have been previously described. Compound 3 was prepared by coupling of 4-fluoro-7-methyl-1H-indole-2-carboxylic acid 8 with commercially available 1-(4-(3-aminophenyl)piperazin-1-yl)ethan-1-one. The preparation of 4-fluoro-7-methyl-1H-indole-2-carboxylic acid 8 is as follows:

**Supplementary Scheme S1:** Preparation of 4-fluoro-7-methyl-1H-indole-2-carboxylic acid 13.

To a reaction vessel containing a solution of (5-fluoro-2-methylphenyl)hydrazine hydrochloride S1 (100 g, 572.7 mmol, 1.0 equiv) in ethanol (400 mL) was added sulfuric acid (10 mL) followed by ethyl 2-oxopropanoate (66 g, 1.2 equiv) under a nitrogen atmosphere at 25 °C. The resulting mixture was aged with stirring for 2 hours, at which time the reaction was shown to be complete by LC-MS. The product mixture was then concentrated under reduced pressure (12 mbar), and the solids were collected by filtration furnishng ethyl (2E)-2-[5-fluoro-2-methylphenyl)hydrazin-1-ylidene]propanoate S2 (120 g, 88% yield) as a yellow solid. **LCMS:** Method 2 (ESI, m/z): RT = 1.40 min, Calc'd for C_{13}H_{15}FN_{2}O_{2} 238.1, found m/z = 239.0 [M+H]^+; **^1H NMR** (400 MHz, DMSO-d6) δ 11.96 (d, J = 2.0 Hz, 1H), 7.17 (d, J = 1.0 Hz, 3H), 6.98 (dd, J = 4.8, 7.6 Hz, 1H), 4.25 (q, J = 7.1 Hz, 2H), 2.12 (d, J = 9.3 Hz, 6H), 1.29 (t, J = 7.1 Hz, 3H).

To a reaction vessel containing a solution of ethyl (2E)-2-[5-fluoro-2-methylphenyl)hydrazin-1-ylidene]propanoate S2 (40 g, 167.9 mmol, 1.0 equiv) in toluene (400 mL) was added 4-methylbenzene-1-sulfonic acid (50 g, 290.4 mmol, 1.70 equiv) under a nitrogen atmosphere. The resulting mixture was aged for 18 hours at 100 °C with stirring, at which time the reaction was shown to be complete by LC-MS. The product mixture was then concentrated under reduced pressure (12 mbar), diluted with ethyl acetate (100 mL) followed by saturated aqueous ammonium chloride (300 mL) and transferred to a separatory funnel. The layers were separated, and the aqueous layer was re-extracted with additional ethyl acetate (2 x 150 mL). The organic layers were combined and dried over magnesium sulfate. The dried solution was filtered through a plug of celite and the filtrate was concentrated under reduced pressure (12 mbar). The residue obtained was purified by flash-column chromatography (eluting with 20% ethyl acetate/petroleum ether), fractions collected and concentrated, and the product residue was recrystallized from boiling ethanol to afford ethyl 4-fluoro-7-methyl-1H-indole-2-carboxylate S3 (9.0 g, 24% yield) as a yellow solid. **LCMS:** Method 2 (ESI, m/z): RT = 1.35 min. Anal. Calc’d for C_{13}H_{15}FN_{2}O_{2} 221.1, found m/z = 222.0 [M+H]^+; **^1H NMR** (400 MHz, DMSO-d6) δ 12.07 (s, 1H), 7.17 (d, J = 2.0 Hz, 1H), 7.00 (m, 1H), 6.77 (m, 1H), 4.36 (q, J = 7.1 Hz, 2H), 2.49 (d, J = 1.0 Hz, 3H), 1.35 (t, J = 7.1 Hz, 3H).

Sodium hydroxide (8 g, 200.0 mmol, 5.0 equiv) dissolved in water (50 mL) was added to a solution of ethyl 4-fluoro-7-methyl-1H-indole-2-carboxylate S3 (9.1 g, 41.1 mmol, 1.0 equiv) in tetrahydrofuran (150 mL) and methanol (2 mL). The reaction mixture was stirred at 21 °C for 6 hours. The product mixture was then concentrated under reduced pressure (12 mbar). The residue solution obtained was adjusted to pH = 4 with hydrochloric acid (3 M). The product was collected by filtration to afford 4-fluoro-7-methyl-1H-indole-2-carboxylic acid 13 (8.0 g, 81%) as a brown solid. **LCMS:** Method 4 (ESI) Anal. Calc’d for C_{13}H_{15}FN_{2}O_{2} 193.1, found m/z [M-H]^+: 192.1, RT = 0.65 min; **^1H NMR** (400 MHz, DMSO-d6) δ 13.12 (br s, 1H), 11.96 (s, 1H), 7.11 (s, 1H), 6.98 (dd, J = 4.8, 7.6 Hz, 1H), 6.75 (dd, J = 7.8, 10.6 Hz, 1H), 2.45 (s, 3H).
**General reaction procedure (GP1):** To a reaction vessel containing the appropriate carboxylic acid (1.0 equiv), N-ethyl-N-isopropylpropan-2-amine (5.0 equiv), and the amine partner (1.5 equiv) in DMF (5 mL) was added HATU (0.46 mmol, 1.3 equiv) at 21 °C. The reaction mixture was aged for 18 h with stirring, at which time the reaction was shown to be complete by TLC. The product mixture was then diluted with deionized water (10 mL) and transferred to a separatory funnel containing ethyl acetate (25 mL). The layers were separated, and the aqueous layer was re-extracted with additional ethyl acetate (2 x 25 mL). The organic layers were combined and dried over magnesium sulfate. The dried solution was filtered through a plug of celite and the filtrate was concentrated under reduced pressure (12 mbar). The residue was purified by Prep-HPLC with the following conditions: Column: XBridge Prep OBD C18 Column, 30×150 mm 5 um; Mobile Phase A: Water (10 mmol/L NH₄HCO₃), Mobile Phase B: MeCN; Flow rate: 10 mL/min; Gradient: 35 B to 65 B in 7 min; 254 nm.

![Supplementary Scheme S2](image)

**Supplementary Scheme S2:** Preparation of compound 3.

Prepared by using GP1 with 4-fluoro-7-methyl-1H-indole-2-carboxylic acid 8 (550 mg, 2.85 mmol, 1.1 equiv) and 1-(4-(3-aminophenyl)piperazin-1-yl)ethan-1-one (548 mg, 2.52 mmol, 1.0 equiv) resulting in 4-fluoro-7-methyl-N-(3-(4-methylpiperazin-1-yl)phenyl)-1H-indole-2-carboxamide 3 (171 mg, 17% yield) as a white solid. Spectral data: LCMS: Method 2 (ESI, m/z): RT=1.95 min, Anal. Calc’d for C₂₂H₂₃FN₄O₃, found m/z =395.1 [M +H]⁺; ¹H NMR (300 MHz, methanol-đ4) δ 7.46 (d, J = 1.5 Hz, 1H), 7.38 (s, 1H), 7.32 – 7.21 (m, 2H), 6.98 (dd, J = 7.9, 5.0 Hz, 1H), 6.80 (dt, J = 6.1, 2.7 Hz, 1H), 6.67 (dd, J = 10.4, 7.8 Hz, 1H), 3.73 (dt, J = 12.7, 5.2 Hz, 4H), 3.22 (dt, J = 17.2, 5.3 Hz, 4H), 2.52 (t, J = 1.0 Hz, 3H), 2.15 (s, 3H); ¹⁹F NMR (376 MHz, methanol-đ4) δ -127.9.
1H NMR spectrum of 3 (d4-methanol):

The aryl central core for 3-F was prepared via the route shown in Supplementary Scheme S3. 3,5-Difluoronitrobenzene S5 was reacted with the 1-(piperazin-1-yl)ethan-1-one to provide the displacement product S6. Reduction of the nitro group was accomplished by either Pd/C in the presence of hydrogen gas.

**Supplementary Scheme S3:** Preparation of intermediates for compound 3-F.

*Preparation of S7:* 3,5-Difluoronitrobenzene S5 (500 mg, 3.14 mmol, 1.0 equiv), 1-(piperazin-1-yl)ethan-1-one (402 mg, 3.14 mmol, 1.0 equiv), and potassium carbonate (868 mg, 6.28 mmol, 2.0 equiv) were added to dry DMSO (3 mL) under an atmosphere of nitrogen. The reaction mixture was aged for 12 hours with stirring at 100 °C, at which time the reaction was shown to be complete by TLC. The product mixture was then diluted with aqueous sodium chloride solution (50 mL) and transferred to a separatory funnel containing ethyl acetate (50 mL). The layers were separated, and the aqueous layer was re-extracted with additional ethyl acetate (3 x 50 mL). The organic layers were combined and dried over sodium sulfate. The dried solution was filtered through a plug of celite and the filtrate was concentrated under reduced pressure (12 mbar) furnishing 1-(4-(3-fluoro-5-nitrophenyl)piperazin-1-yl)ethan-1-one S6 (550 mg, 65% yield) as a yellow oil. **LCMS:** Method 3 (ESI, m/z): RT=1.16 min, Anal. Calc’d for C_{12}H_{14}FN_{3}O_{3} 267.1, found m/z =268.0 [M +H]+.
A 2-neck reaction vessel was charged with 1-(4-(3-fluoro-5-nitrophenyl)piperazin-1-yl)ethan-1-one S6 (150 mg, 0.56 mmol, 1 equiv), Pd/C (70 mg), and ethyl acetate (3 mL). The solids were stirred for 1 minute and the vessel was evacuated and backfilled with nitrogen (x3). Then a balloon of hydrogen gas was attached to the reaction vessel with an adapter that allowed for the balloon to be closed off to the reaction flask. With the hydrogen balloon closed off, the reaction vessel was evacuated and then the balloon was opened to the reaction vessel (x2). After 1 hour, the flask was evacuated and filled with nitrogen, and the solids were filtered through a plug of celite and the filtrate was concentrated under reduced pressure (12 mbar) furnishing 1-(4-(3-amino-5-fluorophenyl)piperazin-1-yl)ethan-1-one S7 (120 mg, 90%) as a colorless oil. **LCMS:** Method 3 (ESI, m/z): RT=0.82 min, Anal. Calc’d for C12H16F3N3O 237.1, found m/z =238.0 [M +H]+.

**Supplementary Scheme S4:** Preparation of compound 3-F.

Prepared by using GP1 with 4-flouro-7-methyl-1H-indole-2-carboxylic acid 8 (105 mg, 0.54 mmol, 1.0 equiv) and 1-[4-(3-amino-5-fluorophenyl)piperazin-1-yl]ethan-1-one S7 (130 mg, 0.55 mmol, 1.0 equiv) resulting in N-(3-(4-acetyl)piperazin-1-yl)-5-fluorophenyl)-4-fluoro-7-methyl-1H-indole-2-carboxamide 3-F (80.1 mg, 35% yield) as a white solid. Spectral data: **LCMS:** Method 2 (ESI, m/z): RT=1.85 min, Anal. Calc’d for C22H22F2N4O4 412.2, found m/z =413.1 [M +H]+; **1H NMR** (300 MHz, DMSO-d6) δ 11.93 (s, 1H), 10.23 (s, 1H), 7.48 (d, J = 1.9 Hz, 1H), 7.27 (dd, J = 10.9, 1.9 Hz, 1H), 7.16 (d, J = 2.1 Hz, 1H), 7.02 – 6.91 (m, 1H), 6.75 (dd, J = 10.4, 7.8 Hz, 1H), 6.57 (d, J = 12.5 Hz, 1H), 3.57 (t, J = 5.1 Hz, 4H), 3.19 (dd, J = 20.4, 5.5 Hz, 4H), 2.47 (s, 3H, overlapped in d6-dmso solvent peak) 2.04 (s, 3H); **19F NMR** (282 MHz, DMSO-d6) δ -125.6, -111.6.

**1H NMR spectrum of 3-F (d6-dmso):**
Experimental preparation of stereoisomers 4-7:

**Supplementary Scheme S5: Preparation of compounds 4-6, EZM0414.**

**Procedure for 4-6:** To a reaction vessel containing 4-fluoro-7-methyl-1H-indole-2-carboxylic acid 8 (4.5 g, 23.3 mmol, 1.0 equiv) was added HATU (10.6 g, 28 mmol, 1.2 equiv), 1-(4-(3-aminocyclohexyl)piperazin-1-yl)ethan-1-one (5.25 g, 23.3 mmol, 1.0 equiv) in DMF (50 mL) followed by N,N'-diisopropylethylamine (9.0 g, 69.9 mmol, 3.0 equiv) at 21 °C. After stirring for 2 hours, the reaction mixture was diluted with deionized water (50 mL) and transferred to a separatory funnel containing ethyl acetate (100 mL). The layers were separated, and the aqueous layer was re-extracted with additional ethyl acetate (2 x 100 mL). The organic layers were combined and dried over sodium sulfate, filtered, and concentrated under reduced pressure (12 mbar). Compounds 4-6 were isolated by chiral prep-HPLC and characterized by 1D and 2D NMR (¹H, ¹³C, gCOSY, HSQC, HMBC, DEPT, NOESY) techniques, HRMS, FT-IR, and optical rotation. Chiral Prep-HPLC method: Agilent 1260; Column: CHIRALPAK IF-3, 3.0*100 mm, 3.0 μm; Mobile Phase A: 0.05% DEA in (n-Hex:EtOH 95:5)/Mobile Phase B: 0.05% DEA in (n-Hex:EtOH 92:8); Temp 40 °C, flow 1.5 mL/min, gradient: t₀ = 0% B, t₅₀.0min = 0% B, t₅₀.1min = 100% B, t₉₀.0min = 100% B; Wavelength: 290 nm.
Chiral HPLC chromatogram of mixture:

Retention time: 36.1 min
Absolute configuration (1S,3R)

Retention time: 19.4 min
Absolute configuration (1S,3S)

Retention time: 65.5 min
Absolute configuration (1R,3S)

Retention time: 21.8 min
Absolute configuration (1R,3R)
**Characterization for 4:** Off-white solid; XRPD spectrum of compound 4 should it was crystalline. The HRMS spectrum showed the molecular formula is C\textsubscript{22}H\textsubscript{29}FN\textsubscript{4}O\textsubscript{2}. The IR spectrum displayed the absorption bands for C=O, an aromatic ring, and C-H bonds; UV spectrum gave the absorption at \(\lambda_{\text{max}}\) 288, 234 and 205 nm are \(\pi-\pi^*\) transition absorption of substituent benzene ring and the \(n-\pi^*\) transition absorption of conjugated system.

The \(^1\text{H}\)-NMR spectrum showed 29 proton signals, including 6 methyl prosots, 16 methylene prosots, 5 methine prosots and 2 active prosots. The \(^{13}\text{C}\)-NMR spectrum together with the DEPT135 spectrum showed 22 carbon signals, including 2 methyl carbons, 8 methylene carbons, 5 methine carbons and 7 quaternary carbons. In combination with the comprehensive analysis of 2D-NMR (HSQC, HMBC, COSY and NOESY) spectra indicated two fragments in the structure: Fragment A and Fragment B. In the COSY spectrum, the correlation from H-14 to NH-13, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

The 2D NMR data confirmed that the relative stereochemistry of 4 was \(-\text{cis}\). Based on the absolute configuration assignment of EZM0414, the absolute configuration of C-14 and C-16 is inferred as S and R configuration. The specific optical rotation was +15.35°(MeOH). In summary, the results of IR, UV, NMR and HRMS analysis were all consistent with the proposed structure of 4.

**FT-IR data:** Bands at 2936, 2855 and 2814 cm\(^{-1}\) derived from C-H stretching vibration absorption, 1247 cm\(^{-1}\) derived from C-H bending vibration absorption, in accordance with CH\(_3\), CH\(_2\) and CH groups in the structure; 1629 cm\(^{-1}\) derived from C=O stretching vibration absorption, in accordance with C=O group in the structure; 1546, 1527, 1504 and 1442 cm\(^{-1}\) derived from benzene ring skeletal vibration, 800, 776 and 746 cm\(^{-1}\) derived from =C-H out-of-plane bending vibration absorption, in accordance with aromatic ring group in the structure.

**HRMS data:** Anal. Calc’d for C\textsubscript{22}H\textsubscript{29}FN\textsubscript{4}O\textsubscript{2} 401.2353, found \(m/z\) =401.2358 [M +H]\(^+\), deviation of 1.2 ppm.

NMR data were collected using a Bruker AVANCE III @ 400MHz, the sample was dissolved in d-chloroform. The following is a summary of the results:

![Chemical Structure](image)

**\(^1\text{H}\) NMR data for 4 (d-chloroform):**

| Chemical Shift (ppm) | Splitting      | Integration | Assignment |
|---------------------|----------------|-------------|------------|
| 9.93                | s              | 1H          | NH-8       |
| 7.05                | overlap        | 1H          | NH-13      |
| 7.01                | d (J = 2.0 Hz) | 1H          | H-10       |
| 6.93                | ddd (J = 8.0, 5.2, 0.8 Hz) | 1H | H-2       |
| 6.69                | dd (J = 10.4, 8.0 Hz) | 1H | H-1       |
| 4.16-4.08           | m              | 1H          | H-14       |
### Chemical Shift (ppm) | Splitting | Integration | Assignment
---|---|---|---
3.64-3.60 | m | 2H | H-22/24
3.45 | t (J = 4.8 Hz) | 2H | H-24/22
2.55 | overlap | 2H | H-21/25
2.52 | overlap | 2H | H-25/21
2.48 | overlap | 1H | H-16
2.45 | s | 3H | H-7
2.15 | d (J = 12.0 Hz) | 1H | H-15a
2.09 | s | 3H | H-28
2.01, 1.30 | overlap | 2H | H-19
1.82, 1.33 | overlap | 2H | H-17
1.88, 1.40 | overlap | 2H | H-18
1.45-1.25 | m | 1H | H-15b

**1H NMR spectrum of 4 (d-chloroform):**

**1H NMR analysis:** The 1H-NMR spectrum showed 29 proton signals, including 6 methyl protons, 16 methylene protons, 5 methine protons and 2 active protons. The signal at 9.93 (s, 1H) ppm showed no correlation in the HSQC spectrum, was assigned to NH-8 according to its chemical shift and HMBC correlations with C-4, C-5, C-9 and C-10; 7.05 (overlap, 1H) ppm showed no correlation in the HSQC spectrum, was assigned to NH-13 according to its chemical shift; the signal at 7.01 (d, J = 2.0 Hz, 1H) ppm showed HMBC correlations with C-4, C-5, C-6, C-9 and C-11, was assigned to H-10; 6.93 (dd, J = 8.0, 5.2, 0.8 Hz, 1H) ppm showed COSY correlations (Table 6) with 6.69 (dd, J = 10.4, 8.0 Hz, 1H), in combination with HMBC correlations, were assigned to H-2 and H-1, respectively; 4.16-4.08 (m, 1H) ppm showed COSY correlations with 2.15 (d, J = 12.0 Hz, 1H) ppm, 1.45-1.25 (m, 1H) ppm and 2.01 ppm, 2.30 (overlap, 2H), in combination with HMBC correlations, were assigned to H-14, H-15a, 15b and H-
19, respectively; 3.64-3.60 (m, 2H) ppm, 3.45 (t, J = 4.8 Hz, 2H) ppm showed COSY correlations with 2.55 (overlap, 2H) ppm, 2.52 (overlap, 2H) ppm, in combination with HMBC correlations, were assigned to H-22/24, H-24/22, H-21/25 and H-25/21, respectively; 2.45 (s, 3H) ppm showed HMBC correlations with C-1, C-2, C-3, C-4, C-5 and C-6, was assigned to H-7; 2.09 (s, 3H) ppm showed HMBC correlations with C-22/24 and C-26, was assigned to H-28; 1.82, 1.33 (overlap, 2H) ppm showed COSY correlations with 2.48 (overlap, 1H) ppm and 1.88, 1.40 (overlap, 2H) ppm, in combination with HMBC correlations, were assigned to H-17, H-16 and H-18, respectively.

$^{13}$C NMR data for 4 (d-chloroform):

| Chemical Shift (ppm) | Hybridization | Assignment |
|----------------------|---------------|------------|
| 168.94               | C=O           | C-26       |
| 160.60               | C=O           | C-11       |
| 156.60, 154.15 (d, J = 245 Hz) | C   | C-6       |
| 137.99, 137.89 (d, J = 10 Hz) | C   | C-4       |
| 131.11               | C             | C-9       |
| 124.61, 124.53 (d, J = 8 Hz) | CH  | C-2       |
| 117.33, 117.29 (d, J = 4 Hz) | C   | C-3       |
| 116.78, 116.56 (d, J = 22 Hz) | C   | C-5       |
| 104.84, 104.66 (d, J = 18 Hz) | CH  | C-1       |
| 98.85                | CH            | C-10       |
| 61.72                | CH            | C-16       |
| 49.81, 48.54         | CH2           | C-21, 25   |
| 47.84                | CH            | C-14       |
| 46.70, 41.85         | CH2           | C-22, 24   |
| 34.53                | CH2           | C-15       |
| 32.73                | CH2           | C-19       |
| 28.00                | CH2           | C-17       |
| 22.34                | CH2           | C-18       |
| 21.29                | CH3           | C-28       |
| 16.33                | CH3           | C-7        |
$^{13}$C NMR spectrum of 4 (d-chloroform):

$^{13}$C NMR analysis: The $^{13}$C-NMR spectrum together with the DEPT135 spectrum showed 22 carbon signals, including 2 methyl carbons, 8 methylene carbons, 5 methine carbons and 7 quaternary carbons. Based on the HSQC spectrum data analysis, all the CH$_3$, CH$_2$ and CH carbon signals were assigned to related proton resonances. The rest of the carbons not connected to hydrogen were assigned by HMBC correlation, chemical shift and coupling constant. The carbonyl signal at 168.94 ppm showed HMBC correlations with H-22/24, H-24/22 and H-28, was assigned to C-26; 160.60 ppm showed HMBC correlations with H-10 and NH-13, was assigned to C-11; 156.60, 154.15 (d, J = 245 Hz) ppm showed HMBC correlations with H-1, H-2, H-7 and H-10, in combination with its coupling constant and chemical shift, was assigned to C-6; 137.99, 137.89 (d, J = 10 Hz) ppm showed HMBC correlation with H-1, H-2, H-7, NH-8 and H-10, in combination with its coupling constant and chemical shift, was assigned to C-4; 131.11 ppm showed HMBC correlations with NH-8 and H-10, was assigned to C-9; 117.33, 117.29 (d, J = 4 Hz) ppm showed HMBC correlations with H-1 and H-7, in combination with its coupling constant, was assigned to C-3; 116.78, 116.56 (d, J = 22 Hz) ppm showed HMBC correlations with H-2, NH-8 and H-10, in combination with its coupling constant, was assigned to C-5.

HSQC and HMBC data for 4:

| $^1$H NMR (ppm) | Assignment | HSQC | HMBC |
|-----------------|------------|------|------|
| 9.93            | NH-8       | -    | C-4, C-5, C-9, C-10 |
| 7.05            | NH-13      | -    | C-11 |
| 7.01            | H-10       | 98.85| C-4, C-5, C-6, C-9, C-11 |
| 6.93            | H-2        | 124.61, 124.53| C-1, C-4, C-5, C-6, C-7 |
| 6.69            | H-1        | 104.84, 104.66| C-2, C-3, C-4, C-6, C-10 |
| 4.16-4.08       | H-14       | 47.84| -    |
HSQC and HMBC analysis: The carbons connect to hydrogen in structure have been assigned accurately based on the HSQC spectrum. In the HMBC spectrum, from NH-8 to C-4, C-5, C-9 and C-10, from H-10 to C-4, C-5, C-6, C-9 and C-11, from NH-13 to C-11, from H-2 to C-1, C-4, C-5, C-6 and C-7, from H-1 to C-2, C-3, C-4, C-6 and C-10, from H-7 to C-1, C-2, C-3, C-4 and C-6, which were consistent with the Fragment A of the structure; from H-22/24 and H-24/22 to C-21, C-25 and C-26, from H-21/25 and H-25/21 to C-16, C-22 and C-24, from H-16 to C-21 and C-25, from H-28 to C-22/24 and C-26, from H-18a to C-14, C-16, C-17 and C-19, which were consistent with the Fragment B of the structure.

gCOSY and NOESY data for 4:

| $^1H$ NMR (ppm) | Assignment | gCOSY          | NOESY          |
|------------------|------------|----------------|----------------|
| 9.93             | NH-8       | -              | H-7            |
| 6.93             | H-2        | H-1            | H-7            |
| 4.16-4.08        | H-14       | NH-13, H-15a, 15b, H-19 | H-15a, 15b, H-17, H-18, H-19 |
| 3.64-3.60        | H-22/24    | H-21/25        | H-21/25        |
| 3.45             | H-24/22    | H-25/21        | H-21/25, H-28  |
| 2.48             | H-16       | H-15a, 15b, H-17 | H-15a, 15b, H-17 |
| 1.88, 1.40       | H-18       | H-17, H-19     | H-16           |

**gCOSY and NOESY analysis:** In the COSY spectrum, the correlations from H-2 to H-1, further confirmed Fragment A exist in the structure, the correlations from H-14 to H-15a, 15b and H-19, from H-22/24 to H-21/25, from H-24/22 to H-25/21, from H-16 to H-15 and H-17, from H-18 to H-17 and H-19, further confirmed Fragment B existed in the structure; the correlations from H-14 to NH-13, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

In the NOESY spectrum, the correlations from H-7 to H-2 and NH-8, further confirmed Fragment A exist in the structure; the correlations from H-14 to H-15a, 15b, H-17, H-18 and H-19, from H-22/24 to H-21/25, from H-24/22
to H-25/21 and H-28, from H-16 to H-15 a, 15b, and H-17, from H-18 to H-16, further confirmed Fragment B existed in the structure.

The key HMBC, gCOSY and NOESY correlations are shown in Supplementary Figure S1.

**Supplementary Figure S1:** The key HMBC, gCOSY and NOESY correlations of compound 4.

**Characterization for 5:** Off-white solid; XRPD showed compound 5 was amorphous. The HRMS spectrum showed the molecular formula of compound 5 is C_{22}H_{30}FN_{4}O_{2}. The IR spectrum displayed the absorption bands for C=O, aromatic ring, and C-H bonds; UV spectrum gave the absorption at \( \lambda_{\text{max}} \) 288, 234 and 205 nm are \( \pi-\pi^* \) transition absorption of substituent benzene ring and the \( n-\pi^* \) transition absorption of conjugated system.

The \(^1\)H-NMR spectrum showed 29 proton signals, including 6 methyl protons, 16 methylene protons, 5 methane protons and 2 active protons. The \(^{13}\)C-NMR spectrum together with the DEPT135 spectrum showed 22 carbon signals, including 2 methyl carbons, 8 methylene carbons, 5 methine carbons and 7 quaternary carbons. In combination with the comprehensive analysis of 2D-NMR (HSQC, HMBC, gCOSY and NOESY) spectra indicated two fragments in the structure: Fragment A and Fragment B. In the HMBC spectrum, the correlations from NH-13 to C-15 and C-19, from H-14 to C-11, indicated that N-13 of Fragment A was linked to C-14 of Fragment B. In the NOESY spectrum, the correlations from H-14 to NH-13, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

The 2D NMR data confirmed that the relative stereochemistry of 5 was -trans. Based on the absolute configuration assignment of EZM0414 and co-crystal assignment of 6, the absolute configuration of C-14 and C-16 were determined as S and S configuration. The specific optical rotation was +43.51°(MeOH). In summary, the results of IR, UV, NMR and HRMS analysis were all consistent with the proposed structure of 5.

**FT-IR data:** Bands at 2931 and 2859 cm\(^{-1}\) derived from C-H stretching vibration absorption, 1247 cm\(^{-1}\) derived from C-H bending vibration absorption, in accordance with CH\(_3\), CH\(_2\) and CH groups in the structure; 1622 cm\(^{-1}\) derived from C=O stretching vibration absorption, in accordance with C=O group in the structure; 1527, 1503 and 1440 cm\(^{-1}\) derived from benzene ring skeletal vibration, 800, 771 and 742 cm\(^{-1}\) derived from \( =\text{C}-\text{H} \) out-of-plane bending vibration absorption, in accordance with aromatic ring group in the structure.

**HRMS data:** Anal. Calc'd for C_{22}H_{30}FN_{4}O_{2} 401.2353, found m/z =401.2356 [M +H]\(^+\), deviation of 0.7 ppm.

NMR data were collected using a Bruker AVANCE III @ 400MHz, the sample was dissolved in \( d \)-chloroform. The following is a summary of the results:
**$^1$H NMR data for 5 (d-chloroform):**

| Chemical Shift (ppm) | Splitting           | Integration | Assignment |
|----------------------|---------------------|-------------|------------|
| 10.09                | s                   | 1H          | NH-8       |
| 7.01                 | d (J = 2.0 Hz)      | 1H          | H-10       |
| 6.92                 | ddd (J = 7.6, 4.8, 0.4 Hz) | 1H          | H-2        |
| 6.69                 | overlap             | 1H          | H-1        |
| 6.68                 | overlap             | 1H          | NH-13      |
| 4.48-4.46            | m                   | 1H          | H-14       |
| 3.69-3.67, 3.52-3.51 | m                   | 2H          | H-22/24    |
| 3.42-3.40            | m                   | 2H          | H-24/22    |
| 2.62, 2.41           | overlap             | 2H          | H-21/25    |
| 2.53                 | overlap             | 2H          | H-25/21    |
| 2.49                 | overlap             | 1H          | H-16       |
| 2.43                 | overlap             | 3H          | H-7        |
| 2.06                 | s                   | 3H          | H-28       |
| 1.99-1.95            | m                   | 1H          | H-15a      |
| 1.82, 1.61           | overlap             | 2H          | H-19       |
| 1.81, 1.55           | overlap             | 2H          | H-18       |
| 1.73                 | overlap             | 1H          | H-15b      |
| 1.60                 | overlap             | 2H          | H-17       |
**1H NMR spectrum of 5 (d-chloroform):**

![1H NMR spectrum of 5](image)

**1H NMR analysis:** The 1H NMR spectrum showed 29 proton signals, including 6 methyl protons, 16 methylene protons, 5 methine protons and 2 active proton. The signal at δH 10.09 (s, 1H) showed no correlation in the HSQC spectrum, was assigned to NH-8 according to its chemical shift and HMBC correlations with C-4, C-5, C-9 and C-10; the signal at δH 7.01 (d, J = 2.0 Hz, 1H) showed HMBC correlations with C-4, C-5, C-6, C-9 and C-11, was assigned to H-10; δH 6.92 (ddd, J = 7.6, 4.8, 0.4 Hz, 1H) showed COSY correlations (Table 6) with 6.69 (overlap, 1H), in combination with HMBC correlations, were assigned to H-2 and H-1, respectively; δH 6.68 (overlap, 1H) showed HMBC correlations with C-11, C-15 and C-19, was assigned to NH-13; δH 4.48-4.46 (m, 1H) showed COSY correlations with δH 1.99-1.95 (m, 1H), δH 1.73 (overlap, 1H), δH 1.82, 1.61 (overlap, 2H), in combination with HMBC correlations, were assigned to H-14, H-15a, 15b and H-19, respectively; δH 1.60 (overlap, 2H) showed COSY correlations with δH 2.49 (overlap, 1H), δH 1.81, 1.55 (overlap, 2H), in combination with HMBC correlations, were assigned to H-17, H-16 and H-18, respectively; δH 3.69-3.67, 3.52-3.51 (m, 2H) and δH 3.42-3.40 (m, 2H) showed COSY correlations with δH 2.62, 2.41 (overlap, 2H), δH 2.53 (overlap, 2H), in combination with HMBC correlations, were assigned to H-22/24, H-24/22, H-21/25 and H-25/21, respectively; δH 2.43 (overlap, 3H) showed HMBC correlations with C-1, C-2, C-3, C-4, C-5 and C-6, was assigned to H-7; δH 2.06 (s, 3H) showed HMBC correlations with C-22/24 and C-26, was assigned to H-28.

**13C NMR data for 5 (d-chloroform):**

| Chemical Shift (ppm) | Hybridization | Assignment |
|----------------------|---------------|------------|
| 168.95               | C=O           | C-26       |
| 160.96               | C=O           | C-11       |
| 156.59, 154.14 (d, J = 245 Hz) | C | C-6 |
| 138.07, 137.97 (d, J = 10 Hz) | C | C-4 |
| 131.07               | C             | C-9        |
| 124.60, 124.52 (d, J = 8 Hz) | CH | C-2 |
**Chemical Shift (ppm)** | **Hybridization** | **Assignment**
--- | --- | ---
117.38, 117.34 (d, J = 4 Hz) | C | C-3
116.74, 116.52 (d, J = 22 Hz) | C | C-5
104.82, 104.63 (d, J = 19 Hz) | CH | C-1
99.04 | CH | C-10
58.57 | CH | C-16
49.63, 49.12 | CH2 | C-21, 25
46.63, 41.81 | CH2 | C-22, 24
45.05 | CH | C-14
34.25 | CH2 | C-15
31.45 | CH2 | C-19
27.80 | CH2 | C-17
21.27 | CH3 | C-28
20.11 | CH2 | C-18
16.31 | CH3 | C-7

**13C NMR spectrum of 5 (d-chloroform):**

**13C NMR analysis:** The 13C NMR spectrum together with the DEPT 135 spectrum showed 22 carbon signals, including 2 methyl carbons, 8 methylene carbons, 5 methine carbons and 7 quaternary carbons. Based on the HSQC spectrum data analysis, all the CH3, CH2 and CH carbon signals were assigned to related proton resonances. The rest
of the carbons not connected to hydrogen were assigned by HMBC correlation, chemical shift and coupling constant. The carbonyl signal at δC 168.95 showed HMBC correlations with H-22/24, H-24/22 and H-28, was assigned to C-6; δC 160.96 showed HMBC correlations with H-10, NH-13 and H-14, was assigned to C-11; δC 156.59, 154.14 (d, J = 245 Hz) showed HMBC correlations with H-1, H-2, H-7 and H-10, in combination with its coupling constant and chemical shift, was assigned to C-6; δC 138.07, 137.97 (d, J = 10 Hz) showed HMBC correlation with H-1, H-2, H-7, NH-8 and H-10, in combination with its coupling constant and chemical shift, was assigned to C-4; δC 131.07 showed HMBC correlations with NH-8 and H-10, was assigned to C-9; δC 117.38, 117.34 (d, J = 4 Hz) showed HMBC correlations with H-1, H-2 and H-7, in combination with its coupling constant, was assigned to C-3; δC 116.74, 116.52 (d, J = 22 Hz) showed HMBC correlations with H-1, H-7, NH-8 and H-10, in combination with its coupling constant, was assigned to C-5.

**HSQC and HMBC data for 5:**

| $^1$H NMR (ppm) | Assignment | HSQC | HMBC |
|-----------------|------------|------|------|
| 10.09           | NH-8       | -    | C-4, C-5, C-9, C-10 |
| 7.01            | H-10       | 99.04| C-4, C-5, C-6, C-9, C-11 |
| 6.92            | H-2        | 124.60, 124.52 | C-1, C-3, C-4, C-6, C-7 |
| 6.69            | H-1        | 104.82, 104.63 | C-2, C-3, C-4, C-5, C-6, C-10 |
| 6.68            | NH-13      | -    | C-11, C-15, C-19 |
| 4.48-4.46       | H-14       | 45.05| C-11 |
| 3.69-3.67, 3.52-3.51| H-22/24 | 46.63, 41.81 | C-21, C-25, C-26 |
| 3.42-3.40       | H-24/22    |      |      |
| 2.62, 2.41      | H-21/25    | 49.63, 49.12 | C-16, C-22, C-24 |
| 2.53            | H-25/21    |      |      |
| 2.49            | H-16       | 58.57| C-21, C-25 |
| 2.43            | H-7        | 16.31| C-1, C-2, C-3, C-4, C-5, C-6 |
| 2.06            | H-28       | 21.27| C-22/24, C-26 |
| 1.99-1.95       | H-15a      | 34.25| C-14, C-16, C-17, C-19 |
| 1.82, 1.61      | H-19       | 31.45| C-14, C-16 |
| 1.81, 1.55      | H-18       | 20.11| - |
| 1.73            | H-15b      | 34.25| C-14, C-16, C-17, C-19 |
| 1.60            | H-17       | 27.80| C-15, C-18, C-19 |

**HSQC and HMBC analysis:** The carbons connect to hydrogen in structure have been assigned accurately based on the HSQC spectrum. In the HMBC spectrum, the correlations from NH-8 to C-4, C-5, C-9 and C-10, from H-10 to C-4, C-5, C-6, C-9 and C-11, from H-2 to C-1, C-3, C-4, C-6 and C-7, from H-1 to C-2, C-3, C-4, C-5, C-6 and C-10, from NH-13 to C-11, from H-7 to C-1, C-2, C-3, C-4, C-5 and C-6, which were consistent with the Fragment A of the structure; from H-22/24 and H-24/22 to C-21, C-25 and C-26, from H-21/25 and H-25/21 to C-16, C-22 and C-24, from H-16 to C-21 and C-25, from H-28 to C-22/24 and C-26, from H-15a, 15b to C-14, C-16, C-17 and C-19, from H-19 to C-14 and C-16, from H-17 to C-15, C-18 and C-19, which were consistent with the Fragment B of the structure; the correlations from NH-13 to C-15 and C-19, from H-14 to C-11, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.
**gCOSY and NOESY data for 5:**

| $^1$H NMR (ppm) | Assignment | gCOSY | NOESY |
|-----------------|------------|-------|-------|
| 10.09           | NH-8       | -     | H-7   |
| 6.92            | H-2        | H-1   | H-7   |
| 4.48–4.46       | H-14       | NH-13, H-15a, 15b, H-19 | NH-13, H-15a, 15b, H-16, H-19 |
| 3.69–3.67, 3.52–3.51 | H-22/24 | H-21/25 | H-21/25 |
| 3.42–3.40       | H-24/22    | H-25/21| H-21/25, H-28 |
| 2.49            | H-16       | H-15a, 15b, H-17 | H-15a, 15b, H-17, H-18 |
| 1.81, 1.55      | H-18       | H-17, H-19 | H-16 |

**gCOSY and NOESY analysis:** In the gCOSY spectrum, the correlations from H-2 to H-1, further confirmed Fragment A existed in the structure, the correlations from H-14 to NH-13, H-15a, 15b and H-19, from H-22/24 to H-21/25, from H-24/22 to H-25/21, from H-16 to H-15a, 15b and H-17, from H-18 to H-17 and H-19, further confirmed Fragment B existed in the structure.

In the NOESY spectrum, the correlations from NH-8 to H-7, from H-2 to H-7, further confirmed Fragment A existed in the structure; the correlations from H-14 to H-15a, 15b, H-16 and H-19, from H-22/24 to H-21/25, from H-24/22 to H-25/21 and H-28, from H-16 to H-15a, 15b and H-17, from H-18 to H-16, further confirmed Fragment B existed in the structure; the correlations from H-14 to NH-13, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

The key HMBC, gCOSY and NOESY correlations are shown in **Supplementary Figure S2.**

**Supplementary Figure S2:** The key HMBC, gCOSY and NOESY correlations of compound 5.

**Characterization for EZM0414:** Off-white solid; XRPD spectrum showed compound EZM0414 was crystalline. The HRMS spectrum showed the molecular formula is C$_{22}$H$_{29}$FN$_{4}$O$_{2}$. The IR spectrum displayed the absorption bands for C=O, an aromatic ring, and C-H bonds; UV spectrum gave the absorption at $\lambda_{\text{max}}$ 288, 234 and 205 nm are $\pi$-$\pi^*$ transition absorption of substituent benzene ring and the $n$-$\pi^*$ transition absorption of conjugated system.

The $^1$H-NMR spectrum showed 29 proton signals, including 6 methyl protons, 16 methylene protons, 5 methine protons and 2 active protons. The $^{13}$C-NMR spectrum together with the DEPT135 spectrum showed 22 carbon signals,
including 2 methyl carbons, 8 methylene carbons, 5 methine carbons and 7 quaternary carbons. In combination with the comprehensive analysis of 2D-NMR (HSQC, HMBC, COSY and NOESY) spectra indicated two fragments in the structure: In the HMBC spectrum, the correlation from H-14 to C-11, indicated that N-13 of Fragment A was linked to C-14 of Fragment B. In the NOESY spectrum, the correlation from H-14 to NH-13, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

The absolute configuration of C-14 and C-16 were determined as R and S configuration by X-ray single crystal diffraction and confirmed upon resynthesis from compound 15. The specific optical rotation was -11.74°(MeOH). In summary, the results of IR, UV, NMR and HRMS analysis were all consistent with the proposed structure.

**FT-IR data:** Bands at 3292 cm\(^{-1}\) was the stretching vibration of N-H, 1628 cm\(^{-1}\) was the stretching vibration of C=O, in accordance with the amide group in the structure; The absorption peaks at 2934 and 2860 cm\(^{-1}\) were the stretching vibration of C-H, 1248 cm\(^{-1}\) derived from C-H bending vibration absorption, illustrated the CH\(_3\), CH\(_2\) and CH groups in the structure; 1528, 1508 and 1449 cm\(^{-1}\) derived from benzene ring skeletal vibration, 791, 777 and 745 cm\(^{-1}\) derived from =C-H out-of-plane bending vibration absorption, in accordance with aromatic ring group in the structure.

**HRMS data:** Anal. Calc’d for C\(_{22}\)H\(_{30}\)FN\(_{4}\)O\(_{2}\) 401.2353, found m/z = 401.2357 [M +H]\(^+\), deviation of 1.0 ppm.

NMR data were collected using a Bruker AVANCE III @ 400MHz, the sample was dissolved in d-chloroform. The following is a summary of the results:

\(^1\)H NMR data (d-chloroform peak is 7.19 ppm) for EZM0414:

| Chemical Shift (ppm) | Splitting | Integration | Assignment |
|----------------------|-----------|-------------|------------|
| 9.71                 | s         | 1H          | NH-8       |
| 6.91                 | d (J = 2.4 Hz) | 1H          | H-10       |
| 6.86                 | ddd (J = 8.0, 5.2, 1.2 Hz) | 1H          | H-2        |
| 6.85                 | overlap   | 1H          | NH-13      |
| 6.62                 | dd (J = 10.0, 7.6Hz) | 1H          | H-1        |
| 4.06-4.04            | m         | 1H          | H-14       |
| 3.57-3.53            | m         | 2H          | H-22/24    |
| 3.38                 | t (J = 4.8 Hz) | 2H          | H-24/22    |
| 2.47                 | overlap   | 2H          | H-21/25    |
| 2.44                 | overlap   | 2H          | H-25/21    |
| 2.40                 | overlap   | 1H          | H-16       |
| 2.38                 | overlap   | 3H          | H-7        |
| 2.09-2.06            | m         | 1H          | H-15a      |
### Chemical Shift (ppm) | Splitting | Integration | Assignment
--- | --- | --- | ---
2.02 | overlap | 3H | H-28
1.92, 1.21 | overlap | 2H | H-19
1.73, 1.25 | overlap | 2H | H-17
1.80, 1.31 | overlap | 2H | H-18
1.28 | overlap | 1H | H-15b

**1H NMR spectrum of EZM0414 (d-chloroform):**

**1H NMR analysis (d-chloroform):** The 1H NMR spectrum showed 29 proton signals, including 6 methyl protons, 16 methylene protons, 5 methine protons and 2 active proton. The signal at δH 9.71 (s, 1H) showed no correlation in HSQC spectrum, was assigned to NH-8 according to its chemical shift and HMBC correlations with C-4, C-5, C-9 and C-10; the signal at δH 6.91 (d, J = 2.4 Hz, 1H) showed HMBC correlations (Table 5) with C-4, C-5, C-6, C-9 and C-11, was assigned to H-10; δH 6.86 (ddd, J = 8.0, 5.2, 1.2 Hz, 1H) showed COSY correlation (Table 6) with 6.62 (dd, J = 10.0, 7.6 Hz, 1H), in combination with HMBC correlations, were assigned to H-2 and H-1, respectively; δH 6.85 (overlap, 1H) according to its chemical shift was assigned to NH-13; δH 4.06-4.04 (m, 1H) showed COSY correlations with δH 2.09-2.06, (m, 1H), δH 1.28 (overlap, 1H) and δH 1.92, 1.21 (overlap, 2H), in combination with HMBC correlations, were assigned to H-14, H-15a, 15b and H-19, respectively; δH 1.73, 1.25 (overlap, 2H) showed COSY correlations with δH 2.40 (overlap, 1H) and δH 1.80, 1.31 (overlap, 2H), in combination with HMBC correlations, were assigned to H-17, H-16 and H-18, respectively; δH 3.57-3.53 (m, 2H) and δH 3.38 (t, J = 4.8 Hz, 2H) showed COSY correlations with δH 2.47 (overlap, 2H) and δH 2.44 (overlap, 2H), in combination with HMBC correlations, were assigned to H-22, 24, H-21, H-25, respectively; δH 2.38 (overlap, 3H) showed HMBC correlations with C-1, C-2, C-3, C-4, C-5 and C-6, was assigned to H-7; δH 2.02 (s, 3H) showed HMBC correlations with C-22, 24 and C-26, was assigned to H-28.
### ¹³C NMR data for EZM0414 (d-chloroform):

| Chemical Shift (ppm) | Hybridization | Assignment |
|----------------------|---------------|------------|
| 168.92               | C=O           | C-26       |
| 160.53               | C=O           | C-11       |
| 156.61, 154.16 (d, J = 245 Hz) | C           | C-6        |
| 137.93, 137.83 (d, J = 10 Hz) | C           | C-4        |
| 131.06               | C             | C-9        |
| 124.64, 124.57 (d, J = 7 Hz) | CH          | C-2        |
| 117.29, 117.25 (d, J = 4 Hz) | C           | C-3        |
| 116.79, 116.56 (d, J = 23 Hz) | C           | C-5        |
| 104.88, 104.70 (d, J = 18 Hz) | CH          | C-1        |
| 98.71                | CH            | C-10       |
| 61.70                | CH            | C-16       |
| 49.82, 48.61         | CH₂           | C-21, 25   |
| 47.82                | CH            | C-14       |
| 46.70, 41.84         | CH₂           | C-22, C-24 |
| 34.55                | CH₂           | C-15       |
| 32.73                | CH₂           | C-19       |
| 28.01                | CH₂           | C-17       |
| 22.29                | CH₂           | C-18       |
| 21.29                | CH₃           | C-28       |
| 16.31                | CH₃           | C-7        |
$^{13}$C NMR spectrum of EZM0414 (d-chloroform):

$^{13}$C NMR analysis (d-chloroform): The $^{13}$C-NMR spectrum together with DEPT 135 spectrum showed 22 carbon signals, including 2 methyl carbons, 8 methylene carbons, 5 methine carbons and 7 quaternary carbons. Based on the HSQC spectrum data analysis, all the CH$_3$, CH$_2$ and CH carbon signals were assigned to related proton resonances. The rest of the carbons are not connect to hydrogen were assigned by the HMBC correlation, chemical shift and coupling constant. The carbonyl signal at $\delta$C 168.92 was assigned to C-26 based on its HMBC correlations to H-22/24, H-24/22 and H-28 and chemical shift; $\delta$C 160.53 showed HMBC correlations with H-10, in combination with its chemical shift, was assigned to C-11; $\delta$C 156.61, 154.16 (d, $J = 245$ Hz) showed HMBC correlations with H-1, H-2, H-7 and H-10, in combination with its coupling constant and its chemical shift, was assigned to C-6; $\delta$C 137.93, 137.83 (d, $J = 10$ Hz) showed HMBC correlations with H-1, H-2, H-7, NH-8 and H-10, in combination with its coupling constant and its chemical shift, was assigned to C-5. $\delta$C 117.29, 117.25 (d, $J = 4$ Hz) showed HMBC correlations with H-1, H-2 and H-7, in combination with its coupling constant and its chemical shift, was assigned to C-3; $\delta$C 116.79, 116.56 (d, $J = 23$ Hz) showed HMBC correlations with H-1, H-2, H-7, NH-8 and H-10, in combination with its coupling constant and its chemical shift, was assigned to C-5.

HSQC and HMBC data (d-chloroform set at 7.19 ppm) for EZM0414:

| $^1$H NMR (ppm) | Assignment | HSQC | HMBC |
|-----------------|------------|------|------|
| 9.71            | NH-8       |      | C-4, C-5, C-9, C-10 |
| 6.91            | H-10       | 98.71| C-4, C-5, C-6, C-9, C-11 |
| 6.86            | H-2        | 124.64, 124.57 | C-1, C-3, C-4, C-5, C-6, C-7 |
**HSQC and HMBC analysis:** The carbons connect to hydrogen in structure have been assigned accurately based on the HSQC spectrum. In the HMBC spectrum, the correlations from H-1 to C-2, C-3, C-4, C-5, C-6 and C-10, from H-2 to C-1, C-3, C-4, C-5, C-6 and C-7, H-7 to C-1, C-2, C-3, C-4, C-5 and C-6, from H-10 to C-4, C-5, C-6, C-9 and C-11, from NH-8 to C-4, C-5, C-9 and C-10, which were consistent with the Fragment A of the structure (Figure 1); the correlation from H-14 to C-15, from H-18 to C-14, C-16, C-17 and C-19, from H-28 to C-22/24 and C-26 which were consistent with the Fragment B of the structure; the correlation from H-14 to C-11, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

**gCOSY and NOESY data for EZM0414:**

| 1H NMR (ppm) | Assignment | gCOSY | NOESY |
|--------------|------------|-------|-------|
| 9.71         | NH-8       | -     | H-7   |
| 6.86         | H-2        | H-1   |       |
| 4.06-4.04    | H-14       | H-15a, 15b, H-19 | NH-13, H-15a, 15b, H-16, H-19 |
| 3.57-3.53    | H-22/24    | H-21/25 | H-21/25 |
| 3.38         | H-24/22    | H-25/21 | H-21/25, H-28 |
| 2.40         | H-16       | H-15a,15b, H-17 | H-15, H-17 |
| 1.80, 1.31   | H-18       | H-17, H-19 | -     |

**gCOSY and NOESY analysis:** In the gCOSY spectrum, the cross peaks from H-1 to H-2, further confirmed Fragment A existed in the structure; the correlations from H-14 to NH-13, H-15a, 15b and H-19, from H-22/24 to H-21/25, from H-24/22 to H-25/21, from H-16 to H-15 and H-17, from H-18 to H-17 and H-19, further confirmed Fragment B existed in the structure.
In the NOESY spectrum, the cross peaks from H-7 to H-2 and NH-8 further confirmed Fragment A existed in the structure; the correlations from H-14 to H-15a, 15b, H-16 and H-19, from H-22/24 to H-21/25, from H-24/22 to H-25/21 and H-28, from H-16 to H-15 and H-17, further confirmed Fragment B existed in the structure. The correlation from H-14 to NH-13, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

The key HMBC, gCOSY and NOESY correlations are shown in Supplementary Figure S3.

**Supplementary Figure S3:** The key HMBC, gCOSY and NOESY correlations of compound EZM0414.

**Characterization for 6:** Off-white solid, XRPD spectrum showed compound 7 was amorphous. The HRMS spectrum showed the formula is C_{22}H_{29}FN_{4}O_{2}. The IR spectrum displayed the absorption peaks of C=O, aromatic ring, and C-H bonds; UV spectrum gave the absorption at \( \lambda_{\text{max}} \) 288, 234, 205 and 194 nm were the \( \pi-\pi^* \) and \( n-\pi^* \) transition absorption of substituent benzene ring.

The \('H NMR spectrum showed 29 proton signals, including 6 methyl protons, 16 methylene protons, 5 methine protons and 2 active protons. The \('C NMR spectrum together with the DEPT135 spectrum showed 22 carbon signals, including 2 methyl carbons, 8 methylene carbons, 5 methine carbons and 7 quaternary carbons. In combination with the comprehensive analysis of 2D NMR (HSQC, HMBC, COSY and NOESY) spectra indicated two fragments in the structure: Fragment A and Fragment B. In the HMBC spectrum, the correlations from H-14 to C-11, indicated that N-13 of Fragment A was linked to C-14 of Fragment B. In the NOESY spectrum, the correlations from H-14 to NH-13, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

The 2D NMR data confirmed the relative configuration to be \(-\text{trans}\). The absolute configuration of C-14 and C-16 were determined as \(R\) and \(R\) configuration by X-ray co-crystallization with SETD2 (vide infra) and upon resynthesis using compound 15 as a starting material. The specific optical rotation is \(-43.14^\circ\) (MeOH). In summary, the results of IR, UV, NMR and HRMS analysis were all consistent with the proposed structure.

**FT-IR data:** Bands at 2931 and 2859 cm\(^{-1}\) derived from C-H stretching vibration absorption, 1246 cm\(^{-1}\) derived from C-H bending vibration absorption, in accordance with CH\(_3\), CH\(_2\) and CH groups in the structure; 1621 cm\(^{-1}\) derived from C=O stretching vibration absorption, in accordance with C=O group in the structure; 1526, 1503 and 1441 cm\(^{-1}\) derived from benzene ring skeletal vibration, 800, 770 and 741 cm\(^{-1}\) derived from \(=\text{C-H out-of-plane} \) bending vibration absorption, in accordance with aromatic ring group in the structure.

**HRMS data:** Anal. Calc’d for C\(_{22}\)H\(_{30}\)FN\(_{4}\)O\(_{2}\) 401.2353, found \(m/z = 401.2357\) [M +H]\(^+\), deviation of 1.0 ppm.

**Conformational-Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies** SI-25
NMR data were collected using a Bruker AVANCE III @ 400MHz, the sample was dissolved in $d$-chloroform. The following is a summary of the results:

$^1$H NMR data for 6 ($d$-chloroform):

| Chemical Shift (ppm) | Splitting          | Integration | Assignment |
|----------------------|--------------------|-------------|------------|
| 9.76                 | s                  | 1H          | NH-8       |
| 6.97                 | d (J = 2.0)        | 1H          | H-10       |
| 6.94                 | ddd (J = 7.6, 4.8, 0.8 Hz) | 1H | H-2    |
| 6.70                 | dd (J = 10.4, 8.0 Hz)  | 1H          | H-1        |
| 6.47                 | d (J = 7.2 Hz)      | 1H          | NH-13      |
| 4.48-4.47            | m                  | 1H          | H-14       |
| 3.74-3.70, 3.55-3.48 | m                  | 2H          | H-22/24    |
| 3.48-3.40            | m                  | 2H          | H-24/22    |
| 2.66, 2.46           | overlap            | 2H          | H-21/25    |
| 2.56                 | overlap            | 2H          | H-25/21    |
| 2.53                 | overlap            | 1H          | H-16       |
| 2.44                 | overlap            | 3H          | H-7        |
| 2.07                 | s                  | 3H          | H-28       |
| 2.00-1.94            | m                  | 1H          | H-15a      |
| 1.78                 | overlap            | 1H          | H-15b      |
| 1.85, 1.56           | overlap            | 2H          | H-18       |
| 1.84, 1.64           | overlap            | 2H          | H-19       |
| 1.64                 | overlap            | 2H          | H-17       |
**1H NMR spectrum of 6 (d-chloroform):**

\[ \text{H NMR spectrum of 6 (d-chloroform):} \]

**1H NMR analysis (d-chloroform):** The \(^1\)H NMR spectrum showed 29 proton signals, including 6 methyl protons, 16 methylene protons, 5 methine protons and 2 active protons. The signal at \(\delta H 9.76\) (s, 1H) showed no correlation in the HSQC spectrum, was assigned to NH-8 according to HMBC correlations with C-4, C-5, C-9 and C-10; the signal at \(\delta H 6.97\) (d, J = 2.0 Hz, 1H) showed HMBC correlations with C-4, C-5, C-6, C-9 and C-11, was assigned to H-10; \(\delta H 6.94\) (ddd, J = 7.6, 4.8, 0.4 Hz, 1H) showed COSY correlations with \(\delta H 6.70\) (overlap, 1H), in combination with HMBC correlations, were assigned to H-2 and H-1, respectively; \(\delta H 6.47\) (overlap, 1H) showed HMBC correlation with C-11, was assigned to NH-13; \(\delta H 4.48-4.47\) (m, 1H) showed COSY correlations with \(\delta H 2.00-1.94\) (m, 1H), 1.78 (overlap, 1H) and \(\delta H 1.84, 1.64\) (overlap, 2H), in combination with HMBC correlations, were assigned to H-14, H-15a, 15b and H-19, respectively; \(\delta H 1.64\) (overlap, 2H) showed COSY correlations with \(\delta H 2.53\) (overlap, 1H), \(\delta H 1.85, 1.56\) (overlap, 2H), in combination with HMBC correlations, were assigned to H-17, H-16 and H-18, respectively; \(\delta H 3.74-3.70, 3.55-3.48\) (m, 2H) and \(\delta H 3.48-3.40\) (m, 2H) showed COSY correlations with \(\delta H 2.66, 2.46\) (overlap, 2H), \(\delta H 2.56\) (overlap, 2H), in combination with HMBC correlations, were assigned to H-22/24, H-24/22, H-21/25 and H-25/21, respectively; \(\delta H 2.44\) (overlap, 3H) showed HMBC correlations with C-1, C-2, C-3, C-4, C-5 and C-6, was assigned to H-7; \(\delta H 2.07\) (s, 3H) showed HMBC correlations with C-22/24 and C-26, was assigned to H-28.

**13C NMR data for 6 (d-chloroform):**

| Chemical Shift (ppm) | Hybridization | Assignment |
|---------------------|--------------|------------|
| 168.89              | C=O          | C-26       |
| 160.79              | C=O          | C-11       |
| 156.60, 154.15 (d, J = 245 Hz) | C           | C-6        |
| 137.93, 137.83 (d, J = 10 Hz) | C           | C-4        |
| 130.92              | C            | C-9        |
| 124.70, 124.63 (d, J = 7 Hz) | CH          | C-2        |
### Chemical Shift (ppm) | Hybridization | Assignment
--- | --- | ---
117.29, 117.26 (d, J = 3 Hz) | C | C-3
116.77, 116.54 (d, J = 23 Hz) | C | C-5
104.93, 104.75 (d, J = 18 Hz) | CH | C-1
98.75 | CH | C-10
58.65 | CH | C-16
49.67, 49.08 | CH2 | C-21, 25
46.59, 41.73 | CH2 | C-22, 24
45.06 | CH | C-14
34.23 | CH2 | C-15
31.42 | CH2 | C-19
27.78 | CH2 | C-17
21.29 | CH3 | C-28
20.14 | CH2 | C-18
16.28 | CH3 | C-7

**13C NMR spectrum of 6 (d-chloroform):**

**13C NMR analysis:** The 13C NMR spectrum together with the DEPT135 spectrum showed 22 carbon signals, including 2 methyl carbons, 8 methylene carbons, 5 methine carbons and 7 quaternary carbons. Based on HSQC spectrum data analysis, all the CH3, CH2 and CH carbon signals were assigned to related proton resonances. The rest of the carbons are not connected to hydrogen were assigned by HMBC correlations, chemical shift and coupling constant. The

*Conformational-Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies*  SI-28
carbonyl signal at δC 168.89 showed HMBC correlations with H-22/24, H-24/22 and H-28, was assigned to C-26; δC 160.79 showed HMBC correlations with H-10, NH-13 and H-14, was assigned to C-11; δC 156.60, 154.15 (d, J = 245 Hz) showed HMBC correlations with H-1, H-2, H-7 and H-10, in combination with its coupling constant and chemical shift, was assigned to C-6; δC 137.93, 137.83 (d, J = 10 Hz) showed HMBC correlations with H-1, H-2, H-7, NH-8 and H-10, in combination with its coupling constant and chemical shift, was assigned to C-4; δC 130.92 showed HMBC correlations with NH-8 and H-10, was assigned to C-9; δC 117.29, 117.26 (d, J = 3 Hz) showed HMBC correlations with H-1, H-2 and H-7, in combination with its coupling constant, was assigned to C-3; δC 116.77, 116.54 (d, J = 23 Hz) showed HMBC correlations with H-1, H-2, H-7, NH-8 and H-10, in combination with its coupling constant, was assigned to C-5.

HSQC and HMBC data for 6:

| Assignment | C-4, C-5, C-9, C-10 | C-4, C-5, C-6, C-9, C-11 | C-1, C-3, C-4, C-5, C-6, C-7 | C-2, C-3, C-4, C-5, C-6, C-10 |
|------------|------------------|---------------------|------------------|---------------------|
| NH-8       | 9.76             | 98.71               | 124.70, 124.63   | 104.93, 104.75      |
| H-10       | 6.97             | 6.94                | 6.70             | 6.47                |
| H-2        | 4.48-4.47        | 3.74-3.70, 3.55-3.48| 2.66, 2.46       | 2.56                |
| H-14       | 2.53             | 2.44                | 2.07             | 2.00-1.94           |
| H-22/24    | 2.56             | 2.53                | 2.07             | 2.00-1.94           |
| H-21/25    | 2.00-1.94        | 1.78                | 1.84, 1.64       | 1.64                |

**HSQC and HMBC analysis:** The carbons connect to hydrogen in structure have been assigned accurately based on the HSQC spectrum. In the HMBC spectrum, the correlations from NH-8 to C-4, C-5, C-9 and C-10, from H-10 to C-4, C-5, C-6, C-9 and C-11, from H-2 to C-1, C-3, C-4, C-5, C-6 and C-7, from H-1 to C-2, C-3, C-4, C-5, C-6 and C-10, from NH-13 to C-11, from H-7 to C-1, C-2, C-3, C-4, C-5 and C-6, which were consistent with the Fragment A of the structure; the correlation from H-14 to C-15, from H-16 to C-21 and C-25, from H-18 to C-14 and C-16, from H-19 to C-15, C-16, C-17, and C-18, from H-22 and H-24 to C-21, C-25 and C-26, from H-21 and H-25 to C-16, C-22, C-24, from H-28 to C-22, 24 and C-26 which were consistent with the Fragment B of the structure; the correlation from H-14 to C-11, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.
**Supporting Information**

**gCOSY and NOESY data for 6:**

| \(^1H\) NMR (ppm) | Assignment | gCOSY | NOESY          |
|---------------------|------------|-------|----------------|
| 9.76                | NH-8       | -     | H-7            |
| 6.94                | H-2        | H-1   | H-7            |
| 4.48-4.47           | H-14       | NH-13, H-15, H-19 | NH-13, H-15a, 15b, H-19 |
| 3.74-3.70, 3.55-3.48| H-22/24    | H-21/25 | H-21/25, H-28 |
| 3.48-3.40           | H-24/22    | H-25/21 | H-21/25, H-28 |
| 2.53                | H-16       | H-15, H-17 | H-15a, 15b, H-17 |
| 1.85, 1.56          | H-18       | H-17, H-19 | -              |

**gCOSY and NOESY analysis:** In the gCOSY spectrum, the cross peaks from H-2 to H-1, further confirmed Fragment A existed in the structure; the cross peaks from H-14 to NH-13, H-15a, 15b and H-19, from H-22/24 to H-21/25, from H-24/22 to H-25/21, from H-16 to H-15a, 15b and H-17, from H-18 to H-17 and H-19, further confirmed Fragment B existed in the structure.

In the NOESY spectrum, the cross peaks from NH-8 to H-7, from H-2 to H-7, further confirmed Fragment A existed in the structure; the cross peaks from H-14 to H-15a, 15b and H-19, from H-22/24 to H-21/25 and H-28, from H-24/22 to H-25/21 and H-28, from H-16 to H-15a, 15b and H-17 further confirmed Fragment B existed in the structure. The correlation from H-14 to NH-13, indicated that N-13 of Fragment A was linked to C-14 of Fragment B.

The key HMBC, gCOSY and NOESY correlations were shown in **Supplementary Figure 4**.

**Supplementary Figure S4:** The key HMBC, gCOSY and NOESY correlations of compound 6.

Overview for Table 2 analogs: **Supplementary Scheme S6** shows the method used for the preparation of compounds with a saturated central core. We determined early in the project that the only active isomer in this series of compounds was the *cis* isomer with the absolute configuration as drawn in the Scheme 1. Since we knew the desired absolute configuration, we began our syntheses by coupling the necessary indole acid with (1R,3R)-3-aminocyclohexan-1-ol to provide the amide, which was oxidized to the corresponding ketone with PCC. Analogs in Table 4 are derived from this intermediate.
Supplementary Scheme S6: Preparation of compound 15.

**Preparation of 15:** To a reaction vessel containing 4-fluoro-7-methyl-1H-indole-2-carboxylic acid 13 (4.5 g, 23.3 mmol, 1.0 equiv) was added HATU (10.6 g, 28 mmol, 1.2 equiv), (1R,3R)-3-aminocyclohexan-1-ol (2.7 g, 23.3 mmol, 1.0 equiv) in DMF (50 mL) followed by N,N'-diisopropylethylamine (9.0 g, 69.9 mmol, 3.0 equiv) at 21 °C. After stirring for 2 hours, the reaction mixture was diluted with deionized water (50 mL) and transferred to a separatory funnel containing ethyl acetate (100 mL). The layers were separated, and the aqueous layer was re-extracted with additional ethyl acetate (2 x 100 mL). The organic layers were combined and dried over sodium sulfate, filtered, and concentrated under reduced pressure (12 mbar). The residue obtained was purified by flash-column chromatography (eluting with 50% ethyl acetate/petroleum ether to afford 4-fluoro-N-((1R,3R)-3-hydroxycyclohexyl)-7-methyl-1H-indole-2-carboxamide (6.5 g, 96%) as a light yellow solid. **LCMS:** (ESI): RT=0.651 min. Anal. Calc’d for C_{16}H_{15}FN_{2}O_{2} 291.3 [M+H]^+; ^1H NMR (400 MHz, methanol-d4) δ 7.21 (s, 1H), 6.99 – 6.93 (m, 1H), 6.66 (dd, J = 10.4, 7.8 Hz, 1H), 4.42 – 4.32 (m, 1H), 4.20 – 4.135 (m, 1H), 2.51 (s, 3H), 2.07 – 1.78 (m, 3H), 1.79 – 1.26 (m, 5H).

Pyridinium chlorochromate (PCC, 5.9 g, 27.6 mmol, 2.0 equiv) was added in a single portion under nitrogen to a round bottom flask equipped with a magnetic stir bar. Ethyl acetate (50 mL) was added to the vessel followed by 4-fluoro-N-((1R,3R)-3-hydroxycyclohexyl)-7-methyl-1H-indole-2-carboxamide (4.0 g, 13.8 mmol, 1.0 equiv). After stirring for 18 hours at 21 °C, the reaction mixture was diluted with aqueous sodium bicarbonate (50 mL) and transferred to a separatory funnel containing ethyl acetate (100 mL). The layers were separated, and the aqueous layer was re-extracted with additional ethyl acetate (2 x 100 mL). The organic layers were combined and dried over sodium sulfate, filtered, and concentrated under reduced pressure (12 mbar). The residue obtained was purified by flash-column chromatography (eluting with 50% ethyl acetate/petroleum ether to afford 4-fluoro-7-methyl-N-((1R)-3-oxocyclohexyl)-1H-indole-2-carboxamide 15 (2.9 g, 73% yield) as a yellow solid. **LCMS:** (ESI): RT=0.729 min. Anal. Calc’d for C_{16}H_{17}FN_{2}O_{3} 288.13, found m/z =289.13 [M+H]^+; ^1H NMR (400 MHz, DMSO-d6) δ 11.71 (s, 1H), 8.53 (d, J = 7.7 Hz, 1H), 7.25 (d, J = 2.1 Hz, 1H), 6.90 (dd, J = 7.9, 5.1 Hz, 1H), 6.70 (dd, J = 10.4, 7.8 Hz, 1H), 4.25 – 4.08 (m, 1H), 2.58 – 2.51 (m, 2H), 2.45 (s, 3H), 2.41 – 2.15 (m, 2H), 2.06 – 1.92 (m, 2H), 1.82 – 1.56 (m, 2H).

With intermediate 15 in hand, reductive amination with the requisite amine afforded the desired product as a mixture of cis and trans isomers. A cis/trans ratio of 2/1 to 1/1 was typically obtained. The reaction scheme is outlined in **Supplementary Scheme S7.**

**Supplementary Scheme S7:** Preparation of advanced analogs in Chart 1.

**General reaction procedure (GP2):** To a reaction vessel containing the amine hydrochloride (0.70 mmol, 2 equiv) suspended in 1,2-DCE (5.0 mL) was added N,N'-diisopropylethylamine (270 mg, 2.1 mmol, 6.0 equiv) at 21 °C.

**Conformational-Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies**
After the reaction had aged 2 hours, the mixture was concentrated under reduced pressure. To this residue, 4-fluoro-7-methyl-N-((1R)-3-oxocyclohexyl)-1H-indole-2-carboxamide (100 mg, 0.35 mmol, 1.0 equiv) was added to the flask followed by methanol (5.0 mL). Sodium cyanoborohydride (3 equiv) was added in one portion to the reaction mixture and stirred at 21 °C. After the reaction mixture had aged 18 hours, the product mixture was concentrated under reduced pressure (12 mbar). The residue was purified by Prep-HPLC (conditions listed).

**Compound 7:**

Prepared by using GP2 with 4-fluoro-7-methyl-N-((1R)-3-oxocyclohexyl)-1H-indole-2-carboxamide 15 (150 mg, 0.52 mmol, 1.0 equiv) and (S)-hexahydropyrrolo[1,2-a]pyrazin-6(2H)-one hydrochloride (183 mg, 1.04 mmol, 2.0 equiv) resulting in 4-fluoro-7-methyl-N-((1R,3S)-3-((S)-6-oxohexahydropyrrolo[1,2-a]pyrazin-2(1H)-yl)cyclohexyl)-1H-indole-2-carboxamide 7 (15 mg, 7% yield) as an off-white solid. **Prep-HPLC conditions:** XBridge Shield RP18 OBD Column, 30*150 mm, 5μm; Mobile Phase A: Water (10 mmol/L NH₄HCO₃), Mobile Phase B: ACN; Flow rate: 60 mL/min; Gradient: 25% B to 42% B in 10 min; Wavelength: 254/220 nm; Spectral data: **LCMS:** Method 2 (ESI): RT=2.51 min, Anal. Calc’d for C₂₃H₂₉FN₄O₄ 412.23, found m/z =413.25 [M +H]+; **¹H NMR** (400 MHz, methanol-d4) δ 7.19 (s, 1H), 6.96 (dd, J = 7.9, 4.9 Hz, 1H), 6.66 (dd, J = 10.4, 7.9 Hz, 1H), 3.96 (td, J = 11.8, 5.8 Hz, 2H), 3.68 (ddt, J = 13.8, 10.3, 5.0 Hz, 1H), 3.11 (dd, J = 11.3, 3.6 Hz, 1H), 2.91 (q, J = 12.8 Hz, 2H), 2.70 – 2.56 (m, 1H), 2.50 (s, 3H), 2.47 – 2.15 (m, 5H), 2.14 – 1.85 (m, 4H), 1.74 – 1.57 (m, 1H), 1.54 – 1.19 (m, 4H); **¹⁹F NMR** (376 MHz, methanol-d4) δ -128.1.
$^1$H NMR spectrum of 7-\((1R,3S)\) \((d_4\text{-methanol})\):

\[
\begin{align*}
\text{H NMR spectrum of 7-\((1R,3R)\)-trans \((d_4\text{-methanol})\):}
\end{align*}
\]
Compound 8:

Prepared by using GP2 with 4-fluoro-7-methyl-N-((1R)-3-oxocyclohexyl)-1H-indole-2-carboxamide 15 (100 mg, 0.35 mmol, 1.0 equiv) and 1-((3aS,6aS)-hexahydropyrrolo[3,4-b]pyrrol-1(2H)-yl)ethan-1-one hydrochloride (99 mg, 0.52 mmol, 1.5 equiv) resulting in N-((1R,3S)-3-((3aS,6aS)-1-acetylhexahydropyrrolo[3,4-b]pyrrol-5(1H)-yl)cyclohexyl)-4-fluoro-7-methyl-1H-indole-2-carboxamide 8 (26 mg, 15% yield) as an off-white solid. **Prep-HPLC conditions:** XBridge Shield RP18 OBD Column, 30*150 mm, 5μm; Mobile Phase A: Water (10 mmol/L NH₄HCO₃), Mobile Phase B: ACN; Flow rate: 60 mL/min; Gradient: 34% B to 50% B in 8 min; Wavelength: 220 nm; Spectral data: **LCMS:** Method 2 (ESI): RT=1.32 min, Anal. Calc’d for C₂₄H₃₁FN₄O₂: 426.24, found m/z =427.30 [M +H]+; **¹H NMR** (400 MHz, methanol-d₄) δ 7.20 (d, J = 3.5 Hz, 1H), 6.96 (ddd, J = 7.8, 5.0, 1.1 Hz, 1H), 6.67 (dd, J = 10.4, 7.8 Hz, 1H), 4.37 (s, 1H), 3.98 (t, J = 11.1 Hz, 1H), 3.90 – 3.52 (m, 2H), 3.23 – 2.80 (m, 3H), 2.68 (dd, J = 10.6, 3.4 Hz, 1H), 2.58 (dd, J = 13.9, 4.0 Hz, 1H), 2.51 (t, J = 11.1 Hz, 3H), 2.26 (d, J = 10.5 Hz, 2H), 2.06 (m, 8H), 1.55 – 1.11 (m, 4H); **¹⁹F NMR** (376 MHz, methanol-d₄) δ -128.0.

**¹H NMR spectrum of 8-((1R,3S) (d₄-methanol):**
$^1$H NMR spectrum of 8-(1R,3R)-trans (d4-methanol):
Compound 9:

Prepared by using GP2 with 4-fluoro-7-methyl-N-((1R)-3-oxocyclohexyl)-1H-indole-2-carboxamide 15 (100 mg, 0.35 mmol, 1.0 equiv) and 1-((3aS,6aS)-hexahydropyrrol[3,2-b]pyrrol-1(2H)-yl)ethan-1-one hydrochloride (99 mg, 0.52 mmol, 1.5 equiv) resulting in N-((1R,3S)-3-((3aS,6aS)-4-acetylhexahydropyrrol[3,2-b]pyrrol-1(2H)-yl)cyclohexyl)-4-fluoro-7-methyl-1H-indole-2-carboxamide 9 (18 mg, 12% yield) as an off-white solid. **Prep-HPLC conditions**: XBridge Shield RP18 OBD Column, 30 x 150 mm, 5μm; Mobile Phase A: Water (10 mmol/L NH₄HCO₃), Mobile Phase B: ACN; Flow rate: 60 mL/min; Gradient: 34% B to 50% B in 8 min; Wavelength: 220 nm; Spectral data: **LCMS**: Method 2 (ESI): RT=0.99 min, Anal. Calc’d for C₂₄H₃₁FN₄O₂ 426.24, found m/z =427.20 [M +H]+; **¹H NMR** (300 MHz, methanol-d₄) δ 7.22 (s, 1H), 6.97 (ddd, J = 7.8, 5.0, 1.1 Hz, 1H), 6.67 (dd, J = 10.4, 7.8 Hz, 1H), 4.43 (dd, J = 8.2, 5.7 Hz, 1H), 4.00 (ddd, J = 11.6, 7.7, 3.8 Hz, 1H), 3.82 – 3.53 (m, 3H), 3.18 – 2.99 (m, 1H), 2.74 – 2.57 (m, 1H), 2.51 (d, J = 1.1 Hz, 3H), 2.45 – 2.18 (m, 2H), 2.13 – 1.82 (m, 8H), 1.65 (dt, J = 12.8, 5.9 Hz, 1H), 1.54 – 1.10 (m, 5H); **¹⁹F NMR** (282 MHz, methanol-d₄) δ -128.0.

**¹H NMR spectrum of 9-(1R,3S) (d₄-methanol):**
$^1$H NMR spectrum of 9-((1R,3R)-trans (d4-methanol):
SETD2 (1434-1711) Assay

The biochemical assay monitored the incorporation of the tritiated methyl group from S-adenosyl-methionine (SAM) into a biotinylated histone 3 peptide corresponding to residues 26-40. The sequence of the substrate peptide is biotin-Ahx-RKSAPATGKKVKPHR-NH₃ (Biopeptide Co., Inc., San Diego, CA) and ³H-SAM was purchased from American Radiolabeled Chemicals, Inc. (St. Louis, MO). For the assay, 40 µL of enzyme was incubated with 1 µL of compound or DMSO for 30 minutes before initiating the reaction with 10 µL of substrate solution in a 384-well assay plate. The assay was performed at room temperature in assay buffer composed of 25 mM bicine, pH 8.0, 7.5 mM β-mercaptoethanol, 0.002 % Tween-20, and 0.01 % bovine skin gelatin (BSG). The reaction was quenched during the linear portion of product formation with 10 µL of 1 mM S-adenosyl-homocysteine (SAH) and 1 mM SAM. From the quenched reaction, 50 µL was transferred to a streptavidin-coated Flashplate (Perkin Elmer) and incubated for at least 2h before washing once with 0.1 % Tween-20. Signal from the ³H-labeled peptide captured by the streptavidin-coated plates was counted by a Topcount plate reader (Perkin Elmer). Percent inhibition (%I) and IC₅₀ values were calculated using equations 1 and 2 respectively.

\[
\%I = \left(1 - \frac{s - \text{min}}{\text{max} - \text{min}}\right)
\]  

(eq 1)

\[
\%I = (100 - \text{bottom}) \frac{1}{1 + \left(\frac{\text{IC}₅₀}{I}\right)^n} + \text{bottom}
\]  

(eq 2)

Min is the signal from fully inhibited SETD2 from wells with a final concentration of 20 uM SAH and max is the signal from wells with DMSO instead of compound. For IC₅₀ calculation, bottom is the theoretical minimum %I, I is the concentration of inhibitor, and n is the Hill slope. Compound IC₅₀ determination was performed by testing 10 concentrations of compound diluted 3-fold in duplicate at final concentrations of 4 nM enzyme and substrate concentrations equal to their K₉ values of 0.7 µM peptide and 2 µM SAM.

In-cell western (ICW) assay of SETD2 Inhibitors (A549 assay)

Reagents and Consumables:

| Reagent Number | Reagent Name                                      | Reagent Source | Catalog Number |
|---------------|--------------------------------------------------|----------------|----------------|
| 1             | A549                                             | ATCC           | CCL-185        |
| 2             | F12/Glutamax Medium                              | ThermoFisher   | 31765035       |
| 3             | FBS                                              | ThermoFisher   | 10099141       |
| 4             | Penicillin-Streptomycin                          | ThermoFisher   | 15140122       |
| 5             | 0.25% Trypsin-EDTA                               | ThermoFisher   | 25200072       |
| 6             | Poly-D-Lysine Black/Clear Microtest (TM)         | BD Biosciences | 356663         |
| 7             | Methanol                                         | Concord        | R266           |

Conformational-Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies  SI-38
Conformational Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies

**Supplementary Information**

| Tool Number | Tool Name                     | Brand    | Model                        |
|-------------|-------------------------------|----------|------------------------------|
| 1           | Cell Counter                  | Invitrogen| Countess® Automated Cell Counter |
| 2           | CO₂ incubator                 | Thermo Scientific | 371                        |
| 3           | Biological Safety Cabinet (Class II) | Thermo Scientific | 1389                     |
| 4           | Centrifuge                    | Eppendorf| 5810R                        |
| 5           | ELx405 plate washer           | BioTek   | ELx405 Select CW             |
| 6           | Echo 550                      | Echo     | Echo 550                     |
| 7           | Plate Shaker                  | Eppendorf| MixMate                      |
| 8           | ODYSSEY CLx                   | LI-COR   | ODYSSEY CLx                  |
| 9           | Multiflo                      | Biotek   | MultiFlo FX                  |
| 10          | Liquid handle system          | Tecan    | Freedom EVO200               |

**Tools and Equipment:**

**Screening Method:** EZM0414 (10 nL) from 384-well source plates was added directly to Poly-D-Lysine (PDL) coated 384-well culture plates. A549 cells were seeded in assay medium at a concentration of 80,000 cells per mL and added to PDL coated plates with a volume of 50 µL per well. Plates were left on the bench top for ~20 minutes to allow cells to settle on the bottom of the well. Plates were incubated at 37°C, 5% CO₂ for 3 days. After three days of incubation, plates were removed from the incubator and allowed to come to room temperature. Media was blotted out of the plate and 50 µL per well of ice cold 100% methanol was added to each plate and incubated for 30 minutes. Methanol was removed by aspiration and plates were washed 3 times with 115 µL per well of wash buffer (1X PBS with 0.05% Tween-20 (v/v)). Next, 50 µL per well of Odyssey blocking buffer + 0.1% Tween-20 was
added to each plate and incubated for 1 hour at room temperature. Blocking buffer was removed, plates were washed 3 times, and 20 µL per well of primary antibody was added (anti-histone H3 tri-methyl K36 diluted 1:1000 in Odyssey buffer with 0.1% Tween-20 (v/v)) and plates were incubated overnight (16 hours) at 4°C. Plates were washed 5 times with 115 µL per well of wash buffer (1X PBS with 0.05% Tween-20 (v/v)). Next 20 µL per well of secondary antibody was added (1:500 800CW goat anti-rabbit IgG (H+L) antibody, 1:1000 DRAQ5 in Odyssey buffer with 0.1% Tween-20 (v/v)) and incubated for 1 hour at room temperature. The plates were washed 5 times with 115 µL per well wash buffer (1X PBS with 0.05% Tween-20 (v/v)) then 3 times with 115 µL per well of water. After rinsing, plates were centrifuged upside down on a thin bed of paper towels up to 1000 rpm for 1 minute to remove excess reagent then allowed to dry at room temperature, out of direct exposure to light. Plates were imaged on the Licor Odyssey system which measures integrated intensity at 700 nm and 800 nm wavelengths. Both 700 and 800 channels were scanned.

Data Analysis:

First, the ratio for each well was determined by:
\[
\frac{(\text{anti-histone H3 tri-methyl K36 800nm value})}{\text{DRAQ5 700nm value}}
\]

Each plate included twenty-eight control wells of DMSO only treatment (Minimum Inhibition) as well as twenty-eight control wells for maximum inhibition treated with 5 µM of EPZ-039698-2 (Background wells). EPZ-039698-2 is a tool compound which gives maximum inhibition of the anti-histone H3 tri-methyl K36 signal in the ICW assay without anti-proliferative effects.

The average of the ratio values for each control type was calculated and used to determine the percent inhibition for each test well in the plate. EPZ-039698-2 was serially diluted three-fold for a total of nine test concentrations from a final starting concentration of 2 µM. The final concentration of DMSO in the assay is 0.2% v/v. Percent inhibition was determined according to the following formula:

\[
\text{Percent Inhibition} = 100 - \left( \frac{\left( \frac{\text{Individual Test Sample Ratio}}{\text{Minimum Inhibition Ratio}} \right) - \left( \frac{\text{Background Avg Ratio}}{\text{Background Avg Ratio}} \right)}{100} \right)
\]

Half-maximal inhibitory concentration (IC₅₀) curves were generated using triplicate wells per each concentration of EZM0414 tested. The IC₅₀ is the concentration of compound at which measured methylation is inhibited by 50% and was calculated using Studies (Dotmatics) from dose response data using a non-linear regression (variable slope–four parameter fit model). The formula used in Studies (Dotmatics) to yield IC₅₀ values is:

\[
Y = \text{Bottom} + \frac{\text{Top-Bottom}}{(1+10^{\text{LogIC}_{50}-X})^{\text{HillSlope}}}
\]

Where

- **Top and Bottom** = plateaus of the Y axis
- **IC₅₀** = half-maximal inhibitory concentration
- **X** = logarithm of the concentration
- **Y** = response
14 Day Long Term Proliferation (LTP) Assay

**Reagents, Consumables, and Equipment:**

| Materials                  | Source     | Cat. No.          |
|----------------------------|------------|-------------------|
| KMS-34                    | JCRB       | JCRB1195          |
| KMS11                     | JCRB       | JCRB1179          |
| RPMI 1640                  | Invitrogen | 11875-119        |
| FBS                       | BI         | 04-002-1A         |
| 0.25% Trypsin-EDTA        | ThermoFisher| 25200-072        |
| 96-well plate, white-walled, tissue culture treated | Corning | 3903 |
| 96-well plate, black-walled, tissue culture treated | Corning | 3904 |
| 6-well, clear, tissue culture treated | Corning | 3506 |
| DMSO                      | Sigma      | 276855-1L         |
| Calcein-AM                | Invitrogen | C3099            |
| CellTiter-Glo             | Promega    | G7571             |

**Suspension:**

Cell lines in log phase growth were treated with EZM0414 in a concentration-dependent manner ranging from 0.5x10^-3 to 10 μmol/L, in a final DMSO concentration of 0.2% volume/volume (v/v). Cells were plated at the determined (1.25 x 10^5 cells/mL) plating densities, in triplicate, in 96-well plates. Assay plates were incubated in a humidified atmosphere of 5% CO₂ at 37°C. Every 3-4 days, viable cell numbers were determined using Calcein AM (Invitrogen, C3099) and analyzed on the Acumen high content imager. After cell counts, cells were split back to the original plating density, and growth media and compound replaced. The final split-adjusted number of viable cells/mL from day 14 were used to calculated percent inhibition. Averages of percent inhibition in technical triplicates were used to plot concentration response curves and calculate absolute half-maximal inhibitory concentration (IC₅₀) values at each time point.

To calculate growth for days 4,7,11, and 14:

1. Calculate the split factor for day 4 to 7, day 7 to 11, and day 11-14. The split factor is the viable cells/mL on Day X (either 4, 7, or 11) divided by the density the cells are being split back to 2. For growth of cells from day 4 to 7, multiply the day 7 viable cells/mL density by the split factor from day 4.

2. For growth of cells from day 7 to 11, multiply the day 11 viable cells/mL density by the days 4, and 7 split factors.

3. For growth of cells from Day 11 to 14, multiply the Day 14 viable cells/mL density by the days 4, 7, and 11 split factors. 5. Plot growth on semi-log chart (viable cells/mL on Y axis, in log, and days on X axis).
Adherent:
Cell lines in log phase growth were treated with EZM0414 in a concentration-dependent manner ranging from 0.5x10^3 to 10 μmol/L, in a final DMSO concentration of 0.2% v/v. Cells were plated at the determined plating densities in either 96-well plates (1.25 x 10^5 cells/mL) in triplicate (for the day 0–7-time course) or 6-well plates (3.75 x 10^5 cells/mL) (for re-plating on day 7 for the remainder of the time course). Assay plates were incubated in a humidified atmosphere of 5% CO_2 at 37°C. Plates were read on day 0, 4, and 7, with compound/media being replenished on day 4. On day 7, the 6-well plates were trypsinized, centrifuged, and re-suspended in fresh media for counting by Vi-CELL. Cells from each treatment were re-plated at the original density in 96-well plates in triplicate and re-treated with the same compound concentration. Plates were read on days 7 (another baseline reading for day 7 seeding), 11, and 14, with compound/media being replenished on day 11. Quantification of proliferation through measurement of cellular adenosine-5'-triphosphate (ATP) was performed via a luminescent cell viability assay using CellTiter-Glo® and luminescence was detected using an EnSpire multimode microplate reader. The final raw luminescence values from day 14 were used to calculate percent inhibition. Averages of percent inhibition in technical triplicates were used to plot concentration response curves and calculate absolute IC_{50} values at each time point.
In Vitro ADME and Pharmacokinetic Assays

Stability in Hepatocytes (Pharmaron, Beijing, China)

Materials and Reagents:

| Items                  | Supplier                     |
|------------------------|------------------------------|
| William’s E Medium     | Life Technologies            |
| Human recombinant insulin | Life Technologies          |
| GlutaMAX               | Life Technologies            |
| Isotonic Percoll       | General Electric             |
| Fetal bovine serum     | Corning                      |
| HEPES                  | Life Technologies            |
| Dexamethasone          | Local suppliers (Beijing, China) |
| Mouse Hepatocytes      | BioreclamationIVT            |
| Human Hepatocytes      | BioreclamationIVT            |

Study Design:

Preparation of Working Solutions: Prepare 10 mM stock solutions of test compound(s) and positive control in appropriate solvent (DMSO). Prepare working solutions of test compounds and reference compound in 50% acetonitrile/50% water at the concentration of 50 μM.

Preparation of Hepatocytes

1) Place incubation medium (William’s E Medium supplemented with GlutaMAX) and hepatocyte thawing medium in a 37°C water bath, and allow warming for at least 15 minutes prior to use. Remove a vial of cryopreserved hepatocytes from storage, ensuring that vials remain at cryogenic temperatures until thawing process ensues. Thaw the cells by placing the vial in a 37°C water bath and gently shaking the vials for 2 minutes. After thawing is completed, spray vial with 70% ethanol, transfer the vial to a biosafety cabinet.

2) Use wide-bore pipette tip to transfer hepatocytes into 50 mL conical tube containing thawing medium. Place the 50 mL conical tube into a centrifuge and spin at 100 × g for 10 minutes. Upon completion of spin, aspirate thawing medium and resuspend hepatocytes in enough incubation medium to yield ~1.5 × 10⁶ cells/mL.

3) Using AOPI (acridine orange / propidium iodide) staining solution, count cells and determine the viable cell density. Cells with poor viability (<75% viability) are not acceptable for use. Dilute cells with incubation medium to a working cell density of 0.5 × 10⁶ viable cells/mL.

Procedure for Stability Determination

1) Pipette 198 μL of hepatocytes into each wells of a 96-well non-coated plate. Place the plate in the incubator on an orbital shaker to allow the hepatocytes to warm for 10 minutes. Pipette 2 μL of working solution to mix with 198 μL of mouse or human hepatocytes to achieve final concentration of 0.5 μM. Return the plate to the incubator and place on an orbital shaker.
2) Remove well contents in 25 μL aliquots at time points of 0 and 120 minutes for human hepatocytes and 0 and 60 minutes for mouse hepatocytes. The aliquots are then mixed with 6 volumes (150 μL) of cold acetonitrile with IS (100 nM alprazolam, 200 nM labetalol, 200 nM caffeine and 2 μM ketoprofen) to terminate the reaction. Centrifuge for 30 minutes at 3,220 × g. Aliquots of 100 μL of the supernatants will be used for LC/MS/MS analysis. The supernatant may be diluted with ultrapure water according to the LC-MS/MS signal response and peak shape. All incubations will be performed in duplicate.

**Data analysis:** All calculations were carried out using Microsoft Excel. Peak areas were determined from extracted ion chromatograms. Determine the in vitro half-life \( t_{1/2} \) of parent compound by regression analysis of the percent parent disappearance vs. time curve.

a. The in vitro \( t_{1/2} \) was determined from the slope value: \( t_{1/2} = \frac{0.693}{k} \)

b. Conversion of the in vitro \( t_{1/2} \) (in min) into the scale-up intrinsic clearance (Scaled-up \( CL_{int} \), in mL/min/kg) was done using the following equation (mean of duplicate determinations):

\[
\text{Scaled-up } CL_{int} = \frac{kV}{N} \times \text{scaling factor}, \text{ where } V = \text{incubation volume (0.2 mL)}; N = \text{number of hepatocytes per well (0.1} \times 10^6 \text{ cells}).
\]

Scaling factors for in vivo intrinsic clearance prediction using different species of hepatocytes are listed below:

| Species | Liver Weight (g liver/kg body weight) | Hepatocyte Concentration (10⁶ cells/g liver) | Scaling Factor |
|---------|--------------------------------------|---------------------------------------------|----------------|
| Human   | 25.7                                 | 99                                          | 2544           |
| Mouse   | 87.5                                 | 135                                         | 11812          |

**LC-MS/MS analysis:** All samples were analyzed using similar methodology, however different instruments were used over course of time to gather all data. The table below lists the instruments that were used to analyze the samples:

| Instrument Type | Instrument (Manufacturer) | Related Details |
|-----------------|---------------------------|------------------|
| HPLC            | Shimadzu (Shimadzu Corporation; Kyoto, Japan) | Mobile Phase: 0.1% formic acid in Water (A) or Acetonitrile (B), elution by gradient Column(s): Waters XSelect HSS T3 C18, 2.5μm (2.1 x 30mm) - or - Phenomenex Kinetex 2.6 μm C18 100R (2.1×50 mm) Column Temperature: 40°C |
| MS/MS           | API4000|TripleQuad 5500|TripleQuad6500+ (Sciex; Framingham, MA, USA) | Ionization source: Electron Spray Ionization (ESI) Ionization mode: positive (+) |
CYP Inhibition in Human Liver Microsomes (Pharmaron, Beijing, China)

**Study Procedure:** The master solution was prepared according to table below, and then 1 μL of compound solution or 1 μL of DMSO was added to the above master solution. The final concentration of control compounds was 10 μM. The final concentration of test compounds was 50 μM.

| Reagent                        | Stock Concentration | Volume | Final Concentration |
|--------------------------------|---------------------|--------|---------------------|
| MgCl₂ solution                 | 50 mM               | 20 μL  | 5 mM                |
| Phosphate buffer               | 200 mM              | 100 μL | 100 mM              |
| Ultra-pure H₂O                 | -                   | 56 μL  | -                   |
| Human liver microsomes         | 20 mg/mL            | 2 μL   | 0.2 mg/mL           |

For CYP1A2 inhibition, 1 μL of specific drug substrate (Phenacetin: 8 mM) was added at the final concentration of 40 μM to the above solution. For CYP2B6 inhibition, 1 μL of specific drug substrate (Bupropion: 10 mM) was added at the final concentration of 50 μM to the above solution. For CYP2C8 inhibition, 1 μL of specific drug substrate (Paclitaxel: 1 mM) was added at the final concentration of 5 μM to the above solution. For CYP2C9 inhibition, 1 μL of specific drug substrate (Tolbutamide: 40 mM) was added at the final concentration of 200 μM to the above solution. For CYP2C19 inhibition, 1 μL of specific drug substrate ((s)-Mephenytoin: 10 mM) was added at the final concentration of 50 μM to the above solution. For CYP2D6 inhibition, 1 μL of specific drug substrate (Dextromethorphan: 2 mM) was added at the final concentration of 10 μM to the above solution. For CYP3A4 inhibition, 1 μL of specific drug substrate (Midazolam: 1 mM) was added at the final concentration of 5 μM to the above solution.

The mixture was pre-warmed at 37°C for 5 min. The reaction was started by the addition of 20 μL of 10 mM NADPH solution at the final concentration of 1 mM and carried out at 37°C. The reaction was stopped by addition of 400 μL of cold quench solution (methanol containing internal standards [IS: 100 nM alprazolam, 500 nM labetalol and 2 μM ketoprofen]) at the designated time points (Phenacetin: 20 min; Bupropion: 20 min; Paclitaxel: 10 min; Tolbutamide: 20 min; (s)-Mephenytoin: 20 min; Dextromethorphan: 20 min; Midazolam: 5 min). Samples were vortexed for 5 minutes and centrifuged at 3220 g for 40 minutes at 4°C. And then 100 μL of the supernatant was transferred to a new 96-well plate with 100 μL water (depends on the LC-MS signal response and peak shape) for LC-MS/MS analysis. All experiments were performed in duplicate.

**LC-MS/MS analysis:** All samples were analyzed using traditional methodology. For data interpretation, select metabolites of probe substrates noted above were monitored. Selected LC-MS/MS information is noted below:

| Instrument Type | Instrument (Manufacturer) |
|-----------------|---------------------------|
| HPLC            | Shimadzu (Shimadzu Corporation; Kyoto, Japan) |

**Related Details**

*Mobile Phase: 0.1% formic acid in Water (A) or Acetonitrile (B), elution by gradient (0.6mL/min)*

*Column: XSelect Hss T3 2.5μ (2.1×50 mm) \(250 μm\) Column XPColumn*

*Temperature: 40°C*

*1 or 2 μL injection volume*
Related Details

Ionization source: Electron Spray Ionization (ESI)
Ionization mode: positive (+)

| Marker Metabolite   | Q1 (m/z) | Q3 (m/z) | DP (v) | EP (v) | CE (v) | CXP (v) |
|---------------------|----------|----------|--------|--------|--------|---------|
| Acetaminophen       | 152.3    | 110.2    | 80     | 8      | 25     | 10      |
| OH Bupropion        | 256.1    | 238.2    | 51     | 10     | 10     | 12      |
| OH Paclitaxel       | 870.3    | 286.2    | 70     | 10     | 20     | 15      |
| OH Tobutamide       | 287.2    | 188.1    | 51     | 10     | 17     | 12      |
| OH-mephenytoin      | 235.1    | 150.1    | 50     | 10     | 26     | 12      |
| Dextrophan          | 258.1    | 157.1    | 60     | 10     | 35     | 15      |
| OH-midazolam        | 342.0    | 203.0    | 95     | 10     | 37     | 15      |

Data Analysis: Percent inhibition was calculated by comparing peak area of marker metabolite in DMSO control samples to that of test samples. Percent inhibition was then used to estimate an IC$_{50}$ using the following equation:

$$\text{Estimated IC}_{50} = \frac{[\text{Drug}] \times (100\% - \text{Obs'd Inh} \%)}{\text{Obs'd Inh} \%}$$

Caco-2 Permeability (Pharmaron, Beijing, China)

Material and Reagents:

| Items                           | Supplier   |
|--------------------------------|------------|
| DMEM                           | Corning    |
| FBS                            | Corning    |
| Penicillin, Streptomycin mixture | Solarbio   |
| 96-well HTS transwell plate     | Corning    |

Preparation of Caco-2 Cells: 50 μL and 25 mL of cell culture medium were added to each well of the Transwell insert and reservoir, respectively. And then the HTS transwell plates were incubated at 37 °C, 5% CO$_2$ for 1 hour before cell seeding. Caco-2 cells were diluted to 6.86x10$^5$ cells/mL with culture medium and 50 μL of cell suspension were dispensed into the filter well of the 96-well HTS Transwell plate. Cells were cultivated for 14-18 days in a cell culture incubator at 37 °C, 5% CO$_2$, 95% relative humidity. Cell culture medium was replaced every other day, beginning no later than 24 hours after initial plating.

Preparation of Stock Solutions: 10 mM stock solutions of test compounds were prepared in DMSO. The stock solutions of positive controls were prepared in DMSO at the concentration of 10 mM. Digoxin and propranolol were used as control compounds in this assay.

Assay Procedures: The Caco-2 plate was removed from the incubator and washed twice with pre-warmed HBSS (10 mM HEPES, pH 7.4), and then incubated at 37 °C for 30 minutes. The stock solutions of control compounds were diluted in DMSO to get 1 mM solutions and then diluted with HBSS (10 mM HEPES, pH 7.4) to get 5 μM working solutions. The stock solutions of the test compounds were diluted in DMSO to get 1 mM solutions, then diluted with
HBSS (10 mM HEPES, pH 7.4) to get 5 μM working solutions. The final concentration of DMSO in the incubation system was 0.5%. To determine the rate of drug transport in the apical to basolateral direction. 75 μL of working solution of test compound and control compounds was added to the Transwell insert (apical compartment) and the wells in the receiver plate (basolateral compartment) were filled with 235 μL of HBSS (10 mM HEPES, pH 7.4). To determine the rate of drug transport in the basolateral to apical direction, 235 μL of working solution of test compound and control compounds was to the receiver plate wells (basolateral compartment) and then the Transwell inserts (apical compartment) were filled with 75 μL of HBSS (10 mM HEPES, pH 7.4). Time 0 samples were prepared by transferring 10 μL of working solution to 40 μL HBSS (10 mM HEPES, pH 7.4) in a new 96-well plate, followed by the addition of 200 μL cold acetonitrile or methanol containing appropriate internal standards (IS). The plates were incubated at 37 °C for 2 hours. At the end of the incubation, 10 μL samples from donor sides (apical compartment for A→B flux, and basolateral compartment for B→A) to 40 μL HBSS (10 mM HEPES, pH 7.4) and 50 μL samples from receiver sides (basolateral compartment for A→B flux, and apical compartment for B→A) were transferred to wells of a new 96-well plate, followed by the addition of 4 volume of cold acetonitrile or methanol containing appropriate internal standards (IS). Samples were vortexed for 5 minutes and then centrifuged at 3,220 g for 40 minutes. An aliquot of 100 μL of the supernatant was mixed with an appropriate volume of ultra-pure water before LC-MS/MS analysis.

Data Analysis: The apparent permeability coefficient (Papp), in units of centimeter per second, can be calculated for Caco-2 drug transport assays using the following equation:

\[
Papp = \frac{(VA \times [\text{drug}]_{\text{acceptor}})}{(\text{Area} \times \text{Time} \times [\text{drug}]_{\text{initial, donor}})}
\]

Where VA is the volume (in mL) in the acceptor well, Area is the surface area of the membrane (0.143 cm² for Transwell-96 Well Permeable Supports), and time is the total transport time in seconds.

The efflux ratio will be determined using the following equation:

\[
\text{Efflux Ratio} = \frac{Papp(B \rightarrow A)}{Papp(A \rightarrow B)}
\]

Where Papp (B→A) indicates the apparent permeability coefficient in basolateral to apical direction, and Papp (A→B) indicates the apparent permeability coefficient in apical to basolateral direction.

Pharmacokinetics

Study Design: A single intravenous (IV) dose was administered to the animal. Fasted animals were also dosed orally (PO). At the designated time points, blood was collected via the dorsal metatarsal vein in mice. Blood was transferred into collection tubes containing K₂-EDTA. For plasma analysis, blood was immediately processed for plasma by centrifugation and stored in a freezer set to be maintained at approximately -80°C until analysis.

Sample Preparation: The desired serial concentrations of working solutions were achieved by diluting stock solution (1 mg/mL in DMSO) of analyte with 50% acetonitrile in water. Ten microliters of working solutions were added to 10 μL of the blank animal plasma to achieve calibration standards of 0.5-1000 ng/mL in a total volume of 20 μL. The resulting 20 μL standard samples were added to 200 μL of acetonitrile for protein precipitation. All samples were then vortexed for 30 seconds. After centrifugation at 4°C and 4000 rpm (ca. 3740 x g) for 15 minutes, the supernatant was diluted with water.

Analytical Method: Concentrations in extracted samples were determined by liquid chromatography–tandem mass spectrometry (LC-MS/MS) using reversed-phase liquid chromatography. Analytes were monitored using Electron Spray Ionization (ESI) with multiple reaction monitoring in positive ion mode. Peak areas were integrated by Analyst® (Sciex) where concentrations were determined by a weighted (1/x²) linear or quadratic regression of peak area ratios (peak area of analyte/peak area of IS) versus the theoretical concentrations of the plasma calibration standards.

Conformational-Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies  SI-47
**Conformational-Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies**

**SUPPORTING INFORMATION**

**Instruments and Analytical Details**

| Method  | Details |
|---------|---------|
| HPLC    | Instrument: Shimadzu (Degasser DGU-20A5; LC-30AD Pumps, CBM-20A Communication Bus; CTO-30A Column oven; Rack Changer II) |
| MS      | API 4000 or API5500 |
| Column  | Phenomenex Kinetex 5µ C18 100A (50×2.1 mm) |
| Polarity| Positive ion |

**Mobile Phase**

| Flow rate (0.6 mL/min) | Mobile Phase A: Water (0.1% formic acid, 5% Acetonitrile) |
|------------------------|------------------------------------------------------------|
|                       | Mobile Phase B: Acetonitrile (0.1% formic acid, 5 % Water) |

**Pharmacokinetic Analysis**: Individual plasma concentration-time data of mice were analyzed by non-compartmental methods using the Linear/Log trapezoidal method (IV) or the Linear-up/Log-down trapezoidal method (PO) (Phoeniix WinNonlin 6.1, Certara, Princeton, NJ). After IV dosing, clearance (CL), steady-state volume of distribution (Vss), terminal elimination half-life (t1/2), area under the curve from time zero to infinity (AUCINF), mean resonance time from time zero to infinity (MRTINF), and terminal phase volume of distribution (Vz) were calculated. After PO dosing, maximum observed concentration (Cmax), time of Cmax (tmax), terminal elimination half-life (t1/2), area under the curve from time zero to infinity (AUCINF), mean resonance time from time zero to infinity (MRTINF), apparent total clearance (CL/F), bioavailability (%F), and estimated fraction absorbed (Fa) were calculated.

**In vivo efficacy analysis in mice**

Animals were quarantined for 7 days before the study. The treatments were started for the efficacy study when the mean tumor volume reached about 119 mm³. Based on the tumor volume, mice were randomly assigned to respective groups such that the mean starting tumor volume was the same between groups. EZM0414 dissolved in methyl cellulose/Tween-80 in water (0.5% CMC (Sigma-Aldrich)/0.1% Tween-80 (Sigma-Aldrich)), or vehicle alone was administered p.o. to NOD SCID mice (n = 10 per dose group).

| Preparation | Concentration (mg/ml) | Dosing volume (ml/kg) | Storage |
|-------------|-----------------------|-----------------------|---------|
| 0.5% CMC + 0.1% Tween 80 in water | -- | 10 | 4°C |
| Dispersed 9.19 mg of EPZ-040414-6 in 6.0 mL of vehicle with vortex and sonication to obtain a uniform suspension. | 1.5 | 10 | 3.0 mL used immediately, 3.0 mL stored at RT for BID dosing |
| Dispersed 18.39 mg of EPZ-040414-6 in 6.0 mL of vehicle with vortex and sonication to obtain a uniform suspension. | 3 | 10 | 3.0 mL used immediately, 3.0 mL stored at RT for BID dosing |

All study animals were monitored not only for tumor growth but also behavior such as mobility, food and water consumption, body weight, eye/hair matting, and any other abnormal effects. The measurement of tumor volume was conducted twice a week throughout the efficacy study with a caliper and the tumor volume (mm³) was estimated using the formula, Tumor Volume = a x b³/2, where “a” and “b” are the long and short diameters of a tumor. The tumor volumes were used for calculating the tumor growth inhibition (TGI, an indicator of antitumor effectiveness) value using the formula: TGI = (1-T/C) × 100%, where “T” and “C” is the mean relative volumes (%)...
tumor growth) of the tumors in the treated (T) and the control (C) groups, respectively, on a given day after tumor inoculation.

Molecular Modeling Methods

**Protein Preparation:** Among our in-house co-crystal structures for SETD2, the co-crystal structure for 2 was selected for docking based on structure similarity and high X-ray resolution (R = 1.82 Å). Crystallization solvent and ions were removed, whereas water molecules and 2 were retained for the hydrogen atoms assignment step and then removed. Ionization states and hydrogen positions were assigned with the “QuickPrep” tool, as implemented in the Molecular Operating Environment (MOE, version 2019.01, default settings). Then, compound 2 was subjected to energy minimization with Amber10 force field. Finally, ligand and water molecules were removed, and protein atoms partial charges computed with the Amber10 force field.

**Molecular Docking:** In this study, the saturated isomers of compound 2 were drawn in 2-dimensional format using ChemDraw software, which later was saved in .sdf format. The saturated isomers were further optimized in MOE using 3-dimensional (3D) structure-energy minimization and stored in .mdb database format. The process of molecular docking simulation between the prepped protein structure and stereoisomers was then performed by removing the ligand 2 from the binding site. The Induced Fit Docking tool (MOE, version 2019.01) was employed, using triangle matcher (scoring function: London dG; 30 poses) and induced fit refinement (Cutoff of 6 residues; tethered side chains: 10; GB/VI energy field) with GBVI/WSA dG scoring function. Maximum of 5 poses were generated, and the top scoring pose was selected for each isomer.

Supplementary Figure S5 is a sample of the simulations with a cis- and trans- isomer described in Figure 2, with higher docking scores associated with the -cis isomer. The structure in green is compound 2 from co-crystal structure in Figure 1. A clear preference for the cis isomer being able to mimic a similar binding pose to 2 can be seen when comparing the highest-ranking poses for both isomers.

**Supplementary figure S5:** Docking pose comparison for a -cis (left) and -trans (right) isomer (grey), compound 2 has been superimposed (green).
Conformer Energetics: Lowest energy conformer was calculated using a conformer generation plugin in Chemaxon (MD simulation). In the top 100 conformers for the cis orientation, none of the (a,a)-cis isomer were observed. For the (e,a)-trans and (a,e)-trans, preference for the classical chair conformation was still observed, at least in the MD simulation. A-values would suggest the dialkylamine (~2.1) would adopt the equatorial position; however, an estimation or ratio of which (e,a)-trans/(e,a)-trans would be preferred in the context of the cellular binding site is too complex in that setting.

Crystallography Protocols

Crystallographic data collection and refinement statistics

| Compound | 6          | EZM0414    |
|----------|------------|------------|
| PDB code | 7TY3       | 7TY2       |
| Cofactor | SAM        | SAM        |
| Space group | P 2 1 2 1 21 | P 2 1 2 1 21 |
| a, b, c (Å) | 52.09,76.70,75.46 | 48.69,75.74,75.92 |
| α, β, γ (°) | 90.0,90.0,90.0 | 90.0,90.0,90.0 |
| Resolution range (Å) | 43.13-2.30 | 41.02-2.44 |
| (Highest resolution shell) | 2.36-2.30 | 2.50-2.44 |
| R<sub>merge</sub>, overall | 0.085 | 0.122 |
| Completeness, overall (%) | 95.1 | 99.2 |
| Reflections, unique | 12603 | 10292 |
| Multiplicity | 4.1 | 6.5 |
| I/σ | 5.8 | 7.3 |
| R<sub>value</sub> work (%) | 21.4 | 21.2 |
| R<sub>value</sub> free (%) | 26.9 | 27.0 |
| R.M.S. deviations from ideal |
| Bond lengths (Å) | 0.005 | 0.006 |
| Bond angles (°) | 1.40 | 1.55 |
| Φ,Ψ angle distributions for residues | 95 | 94 |
| In preferred regions (%) | 12 | 12 |
Conformational-Design-Driven Discovery of EZM0414: A Selective, Potent SETD2 Inhibitor for Clinical Studies

SI-51

Outliers (%) | 0 | 2

1\(R_{\text{merge}} = \sum_{hkl} \left[ \sum |I_i - <I>| / \sum I_i \right]

2\(R_{\text{value}} = \sum_{hkl} \frac{||F_{\text{obs}}|| - |F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}

3\text{Ramachandran statistics as defined Engh \\& Huber (2001)}

**Supplementary Table S1:** Crystallographic collection parameters.

Protein Production: SETD2 (1434-1711) was purified from an insect cell expression system. The protein sequence was incorporated in a pFastBacHTb-lic with N-His-TEV and expressed in HF cells for 72 hours. Post-harvest the cells were resuspended in 25mM HEPES, 200mM NaCl, 5mM β-ME, pH 7.5 (buffer A) at a concentration of 1g cell per 8ml buffer. Roche protease inhibitor cocktail and 0.5% CHAPS was added. The mixture was incubated on ice for 30 min. The suspension was sonicated at 200W with a 3sec on, 5sec off cycle seventy times. The sample was centrifuged at 13,000rpm for 30 min at 4°C, the supernatant was collected.

A 20ml CV XK Qiagen Ni affinity column was used. Resin was preincubated with buffer A, then incubated with the sonication supernatant for 2 hours at 4C. The resin was washed with 25mM HEPES, 200mM NaCl, 5mM β-ME, 20mM imidazole, pH 7.5 (buffer B) until no UV signal was observed. The protein was eluted with 10CV of 25mM HEPES, 200mM NaCl, 5mM β-ME, 250mM imidazole, pH 7.5 (buffer C). A 12% SDS-PAGE gel was used to identify the fractions containing the protein.

TEV protease was added at a ratio of 1:20 relative to SETD2 and the sample was incubated overnight.

The cleaved sample was applied to a 10ml CV XK Qiagen Ni affinity column. The resin was pre-equilibrated with buffer A, and the sample was loaded onto the column. The column was washed with 25mM HEPES, 200mM NaCl, 5mM β-ME, 10mM imidazole, pH 7.5 (buffer D) for 6CV and buffer C for 5CV.

The flowthrough from the affinity column was concentrated and applied to a Superdex 200 column preincubated with 25mM HEPES, 200mM NaCl, 5mM β-ME, 10mM imidazole, pH 7.5 (buffer E). Fractions containing the protein were collected, concentrated to 36mg/ml and flash frozen on liquid nitrogen.

X-ray crystallography: SETD2 (1434-1711) protein was concentrated to 10 mg/ml in a protein buffer containing 25mM HEPES, 200mM NaCl, 5mM β-ME, pH 7.5, 20mg/ml SETD2, 3mM SAM and 2mM compound. A crystallization reservoir solution of 0.1 M Potassium thiocyanate, 0.1M Tris pH7.8, 25% w/v PEG2000MME was used. Using a hanging drop method with seeding 1 µL of protein solution was combined with 1 µL of reservoir solution at 18°C. Crystals were cryo-protected with the addition of 10% glycerol. Diffraction data was collected on BL19U1 at the SSRF or MX2 at AS. Data was reduced with either HKL3000 or XDS and refined with Refmac5. Data collection details and structure details are in the accompanying table.
Supplementary figure S6: Superimposition of X-ray co-crystal structures of EZM0414 (pale green, PDD:7TY2) and 6 (wheat, PDB 7TY3).
References:

1. Lampe, J. W.; Alford, J. S.; Boriak-Sjodin, P. A.; Brach, D.; Cosmopoulos, K.; Duncan, K. W.; Eckley, S. T.; Foley, M. A.; Harvey, D. M.; Motwani, V.; Munchhof, M. J.; Raimondi, A.; Riera, T. V.; Tang, C.; Thomenius, M. J.; Totman, J.; Farrow, N. A., Discovery of a First-in-Class Inhibitor of the Histone Methyltransferase SETD2 Suitable for Preclinical Studies. *ACS Med Chem Lett* **2021**, *12* (10), 1539-1545.