

Supporting Information

Redesign of Rifamycin Antibiotics to Overcome ADP-Ribosylation-Mediated Resistance

T. Lan, U. S. Ganapathy, S. Sharma, Y.-M. Ahn, M. Zimmerman, V. Molodtsov, P. Hegde, M. Gengenbacher, R. H. Ebright, V. Dartois, J. S. Freundlich, T. Dick, C. C. Aldrich**

Table of Contents

Part I: Supplementary Tables and Figures

Table S1: Trials of C-25 esterification methods. Benzoic acid was used as the model acid in all the entries except Entry 5, where benzoyl fluoride was used

Table S2: Activity of the selected rifabutin analogs against *M. abscessus* clinical isolates. Tested synthetic rifabutin analogs showed equal potency against the clinical isolates^[a]

compound	Mab 19977 MIC	Mab Bamboo MIC	Mab K21 MIC
rifabutin	1150	1100	2300
5a	53	65	98
5b	43	54	66
5j	55	68	102
5m	25	23	33

[a] All MIC values were determined as the concentrations that result in 90% inhibition of bacterial growth. All MIC values are given in nM.

compound	RFB-R1 M. abscessus MIC
clarithromycin	2.1
rifampicin	>100
rifabutin	>100
5a	>100
5b	>100
5j	>100
5m	>100

Table S3: Activity of the synthetic rifabutin analogs against RFB-R1 *M. abscessus*^[a]

[a] All MIC values were determined as the concentrations that result in 90% inhibition of bacterial growth. All MIC values are given in µM.

Table S4: Data-collection and refinement statistics for crystal structure of RNAP*Mtb*(*M.* tuberculosis σ^A RPo)-5a complex. ^a Numbers in parentheses refer to highest-resolution shell.

Table S5: MS validation of the peaks observed in LCMS-based *in vitro* validation of ADPribosylation. For compounds converted to ADP-ribosyl adducts upon incubation, the mass-tocharge ratio (m/z) of the expected ADP-ribosyl adducts ([M+2H]⁺⁺) and the observed m/z of the post-incubation LC signals are labeled in red.

compound	MW (g/mol)	expected ADP- ribosyl adduct m/z		observed m/z (0 min) observed m/z (40 min)
rifampicin	822.94	682.74	822.50	682.80
rifabutin	847.02	694.76	847.50	694.90
5a	909.09	725.76	909.50	909.50
5b	923.12	732.77	923.50	923.50
5k	889.10	715.78	889.50	715.90
51	903.13	722.79	903.50	722.90
5 _m	910.08	726.29	910.50	910.50

Table S6. PK parameters of rifabutin and selected analogs with standard deviation.

[a] SD% = standard deviation/mean×100%.

[b] *f*AUC/MIC was measured based on p.o. doses of 10 mg/kg.

[c] For both administrative routes, deacylation % = AUC**6**/AUCcompound×100%, n.d.: undetectable.

Figure S1: Binding mode of rifampicin (green) in RNAP*Mtb* (PDB: 5UHB). The binding pocket is shown in pale cyan. The space accommodating C25-OAc is contoured in orange dashed lines.

Figure S2: Proposed binding modes of rifabutin (magenta) and **5a** (cyan) in the RNAP*Mab* homology model (light green). High overlap was observed for the two molecules.

Figure S3: Comparison of co-crystal structures of **5a**-RNAP*Mtb* (red/pink) and rifampicin-RNAP*Mtb* (black/gray, PDB: 5UHB). Residues with direct interactions with the ligands are labeled.

	Substitution	Binding Affinity (kcal/mol)	Color
5a	C_6H_5 -	-8.5	Cyan
5b	$O-MeC_6H_4$ -	-8.1	Blue
5e	O -FC $_6$ H ₄ -	-8.2	Magenta
5f	$O-CIC6H4$ -	-8.1	Orange
5g	o -OMe C_6H_4 -	-8	Yellow

Figure S4: Proposed binding modes and binding affinities of analogs **5b**-**5g** in the RNAP*Mab* homology model (gray).

Figure S5: *In vitro* characterization of ADP-ribosylation of rifamycins and synthetic compounds using overexpressed Arr*Mab*. All the peaks were further confirmed by MS (Table S4). The extension time difference of the peaks is labelled. Controls rifampicin and rifabutin were fully ADPribosylated upon 40-minute incubations. No ADP-ribosylated adducts were observed for the synthetic compounds **5a**, **5b** and **5m** bearing bulky C-25 substituents. Synthetic compounds **5k** and **5l** with small alkyl substituents on C-25 failed to block ADP-ribosylation and transformed into the respective ADP-ribosylated adducts upon incubation in a 40-minute time course.

Part II: Synthetic Chemistry and analog characterizations General methods for chemical synthesis

Reagents and solvents were purchased from commercial sources (Fisher Scientific, MilliporeSigma, A2B Chem, Oakwood Chemical) and used as received unless otherwise noted. Rifabutin was purchased from WuXi AppTec (Tianjin). Reactions were monitored using Macherey-Nagel® ALUGRAM® SIL G/UV254 aluminum TLC plates. LC were visualized under visible light or UV fluorescence (254 nm). Flash chromatography was performed using Sorbtech[®] Silica Gel [porosity: 60Å; particle size: 40-63 µm (230×400 mesh)]. Preparative TLC were performed on Silicyle® glass-backed TLC plates (thickness: 1000 µm; indicator: F-254). NMR Spectra were recorded on a Bruker 600-MHz Avance NEO. ¹H frequency is at 601 MHz. ¹³C frequency is at 151 MHz. Chemical shifts (δ) were reported in parts per million (ppm) relative to residual solvent peaks $[CDCl_3$ (¹H: 7.26, ¹³C: 77.2) or CD_2Cl_2 (¹H: 5.32, ¹³C: 53.8)]. Peak multiplicity was indicated as follows: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet) and br (broad). High resolution mass spectra (HRMS) were recorded on Bruker BioTOF II ESI/TOF-MS. Analytical HPLC analysis was performed on an Agilent® 1260 Infinity Quaternary LC system with a reversed-phase C18 column (Gemini-Nx 5 micron, 150 \times 4.60 mm, Phenomenex).

Structures of all the evaluated rifabutin analogs

Experimental Procedures and Characterization Data

Synthesis of **5a**-**5h**, **5j**-**5o**

21,23-*O***-Isopropylidenerifabutin (2)**. To a solution of rifabutin **1** (5.01 g, 5.90 mmol, 1.00 equiv) and camphorsulfonic acid (1.66 g, 7.15 mmol, 1.20 equiv) in acetone (59.0 mL, dried with 4 Å molecular sieves) was added 2,2-dimethoxypropane (15.0 mL, 122 mmol, 20.7 equiv). The reaction mixture was stirred at 23 °C for 2 h; then NaHCO₃ (1.8 g, 21 mmol, 3.60 equiv) was added, and the mixture was further stirred for 30 min at room temperature. Then the reaction was partitioned between CH_2Cl_2 (80 mL) and H₂O (80 mL). The aqueous phase was back extracted with CH₂Cl₂ (3 \times 50 mL). The combined organic layers were dried (Na₂SO₄) and concentrated under reduced pressure. Purification by silica gel flash chromatography (CH_2Cl_2 :MeOH 50:1) on silica gel afforded the title compound $(3.94 \text{ g}, 75%)$ as a purple solid: $R_f = 0.13$ (50:1) CH₂Cl₂:MeOH); ¹H NMR (601 MHz, CD₂Cl₂) δ 14.88 (s, 1H), 8.80 (s, 1H), 7.77 (s, 1H), 6.29 (dd, *J* = 15.8, 10.8 Hz, 1H), 6.18–6.14 (m, 1H), 6.06 (dd, *J* = 15.8, 7.1 Hz, 1H), 5.91 (dd, *J* = 12.2, 1.1 Hz, 1H), 5.04 (dd, *J* = 12.2, 6.6 Hz, 1H), 4.89 (dd, *J* = 7.9, 1.5 Hz, 1H), 3.59 (dd, *J* = 10.7, 3.4 Hz, 1H), 3.38 (ddd, *J* = 6.6, 2.3, 1.1 Hz, 1H), 3.03 (dd, *J* = 10.3, 5.2 Hz, 1H), 3.01–2.87 (m, 2H), 2.82 (s, 3H), 2.68–2.53(m, 2H), 2.27 (s, 3H), 2.32–2.19 (m, 3H), 2.03 (s, 2H), 1.93 (s, 3H), 1.89–1.79 (m, 1H), 1.79–1.75 (m, 1H), 1.74 (s, 3H), 1.56–1.48 (m, 1H), 1.44 (pd, *J* = 7.1, 2.2 Hz, 1H), 1.20 (s, 1H), 1.17 (s, 3H), 0.94 (d, *J* = 6.6 Hz, 6H), 0.86 (s, 3H), 0.85 (d, *J* = 6.6 Hz, 2H), 0.82 (d, *J* = 6.8 Hz, 3H), 0.69 (d, J = 7.2 Hz, 3H), 0.36 (d, J = 7.1 Hz, 3H); ¹³C NMR (151 MHz, CD₂Cl₂) δ 192.6, 181.6, 172.4, 170.4, 169.0, 168.5, 155.4, 142.7, 141.2, 141.0, 132.5, 131.9, 125.8, 124.2, 115.5, 113.8, 112.1, 109.0, 106.3, 104.9, 100.2, 100.2, 95.4, 79.0, 77.3, 74.5, 71.3, 69.7, 66.8, 56.4, 51.9, 51.8, 41.5, 41.0, 36.6, 36.5, 35.9, 34.7, 32.0, 29.5, 26.2, 26.0, 24.0, 21.0, 21.0, 21.0, 20.4, 20.4, 17.9, 13.0, 9.9, 9.6, 7.8; MS (ESI): m/z [M+H]⁺ calcd for $C_{49}H_{66}N_4O_{11}$: 887.4801; found: 887.4786 (error 1.7 ppm).

25-Desacetyl-21,23-*O***-isopropylidenerifabutin (3)**. A mixture of **2** (3.56 g, 4.01 mmol, 1.00 equiv) and K_2CO_3 (3.33 g, 24.1 mmol, 6.00 equiv) in MeOH (40 mL) was heated under 50 °C for 48 h. The reaction was cooled to room temperature and partitioned between CH_2Cl_2 (50 mL) and brine (50 mL). The organic layer was separated and the aqueous phase was extracted with $CH₂Cl₂$ (3 x 40 mL). The combined organic extracts were dried (Na₂SO₄) and concentrated under reduced pressure. Purification by flash chromatography (CH₂Cl₂:MeOH 60:1) on silica gel afforded the title compound (2.17 g, 64%, containing trace amount of the starting material) as a purple solid: *R^f* = 0.19 (hexanes:EtOAc:MeOH:Et3N 8:1:1:0.1); ¹H NMR (600 MHz, CD2Cl2) δ 14.93 (s, 1H), 8.72 (s, 1H), 7.81 (s, 1H), 6.28 (dd, *J* = 15.8, 10.3 Hz, 1H), 6.16 (dd, *J* = 11.4, 1.6 Hz, 1H), 6.15 (d, *J* = 11.2 Hz, 1H), 5.96 (dd, *J* = 15.8, 6.6 Hz, 1H), 4.95 (dd, *J* = 12.2, 9.1 Hz, 1H), 3.60 (dd, *J* = 9.1, 3.4 Hz, 1H), 3.52 (dd, *J* = 10.4, 3.2 Hz, 1H), 3.44 (dt, *J* = 8.8, 3.0 Hz, 1H), 3.15 (d, *J* = 3.7 Hz, 1H), 3.14 (s, 3H), 3.11 (dd, *J* = 9.1, 5.0 Hz, 1H), 2.99–2.86 (m, 1H), 2.67–2.50 (m, 2H), 2.25 (dd, *J* = 7.4, 1.9 Hz, 2H), 2.23 (s, 3H), 2.03 (s, 3H), 2.09 – 1.87 (bs, 4H), 1.84 (dq, *J* = 13.6, 6.7 Hz, 1H), 1.75 (s, 3H), 1.66–1.58 (m, 1H), 1.59–1.52 (m, 1H), 1.35 (ddd, *J* = 9.5, 6.8, 2.7 Hz, 1H), 1.02 (s, 3H), 0.94 (d, *J* = 6.6 Hz, 6H), 0.85 (d, *J* = 7.0 Hz, 3H), 0.82 (d, *J* = 6.8 Hz, 3H), 0.80 (s, 3H), 0.72 (d, J = 6.8 Hz, 3H), 0.49 (d, J = 7.1 Hz, 3H); ¹³C NMR (151 MHz, CD₂Cl₂) δ 191.3, 181.9, 171.6, 169.2, 168.5, 155.5, 143.0, 142.9, 140.6, 132.5, 132.2, 125.9, 124.5, 114.4, 112.2, 112.1, 108.9, 105.9, 105.2, 99.7, 95.4, 83.2, 76.0, 71.6, 71.1, 66.8, 56.4, 52.1, 51.8, 41.2, 40.1, 36.7, 35.8, 35.2, 35.1, 26.2, 25.7, 24.5, 21.0, 21.0, 20.0, 20.0, 17.8, 13.0, 13.0, 8.5, 7.8; MS $(ESI): m/z [M+Na]^+$ calcd for $C_{49}H_{65}N_4O_{11}Na$: 867.4515; found: 867.4538 (error 2.6 ppm).

Acylation procedure A: To a solution of **3** (1.00 equiv) in 1,2-dichloroethane (0.1 M) at room temperature or 50 \degree C, the acid anhydride (3.00–5.00 equiv) prepared from the respective acid^[1] and DMAP (0.50–0.80 equiv) were added every 12 h until the majority of **3** was converted. For most cases a total amount of 10.0–20.0 equivalents anhydride and 2.00–3.00 equivalents DMAP are required in 2–3 days. Then the reaction mixture was poured into CH_2Cl_2 (10 mL) and saturated aqueous NaHCO₃ (10 mL) and the biphasic mixture was stirred at 23 °C for 16 h to quench any remaining anhydride. The organic layers were separated and the aqueous phase was extracted with CH_2Cl_2 (2 x 10 mL). The combined organic layers were dried (Na₂SO₄) and concentrated under reduced pressure. Purification by preparative silica gel TLC afforded the 25-*O*-acyl-21,23 isopropylidenerifabutin analogs **4a-4h** and **4j**-**4l**, which were deprotected following the general procedure for deprotection.

Acylation procedure B. To a solution of the acid (4.00 equiv), triethylamine (5.00 equiv) and DMAP (2.00 equiv) in CH_2Cl_2 (0.4 M to the acid) at 0 °C was added pivaloyl chloride (4.00 equiv) and the reaction was allowed to warm to 23 °C over 1 h. Next, **3** (1.00 equiv) was added and the resulting solution was further stirred at 23 °C for 4 h. The reaction mixture was poured into CH_2Cl_2 (10 mL) and saturated NaHCO₃ aqueous solution (10 mL) and the biphasic mixture was stirred at 23 °C for 16 h to quench any remaining mixed anhydride. The organic layers were separated and the aqueous phase was extracted with CH_2Cl_2 (2 x 10 mL). The combined organic layers were dried (Na2SO4) and concentrated under reduced pressure. Purification by preparative silica gel TLC afforded the 25-*O*-acyl-21,23-isopropylidenerifabutin analogs **4m**-**4o**, which were deprotected following the general procedure for deprotection.

General deprotection procedure. A solution of a 25-*O*-acyl-21,23-isopropylidene rifabutin analog (1.00 equiv), prepared using acylation procedure A or B, and camphorsulfonic acid (2.00 equiv) in methanol (0.1 M) was stirred at room temperature for 30 min; then the reaction was partitioned between CH2Cl₂ (10 mL) and saturated NaHCO₃ aqueous solution (10 mL). The aqueous phase was then back extracted with CH_2Cl_2 (2 x 10 mL). The combined organic layers were dried (Na₂SO₄) and concentrated under reduced pressure. The resulting crude was purified by preparative TLC (hexane:ethyl acetate:methanol:triethylamine = 8:1:1:0.1 or hexane:ethyl acetate:methanol = 7:2:1) to obtain 25-*O*-acyl rifabutin analogs **5a-5h,** and **5j**-**5o** (see below for full characterizations).

Synthesis of **5i**

To a solution of NaOH (0.28 g, 6.00 equiv) and $ZnCl₂$ (0.16 g, 1.00 equiv) in MeOH (12 mL) was added **1** (1.00 g, 1.00 equiv), and the reaction mixture was stirred at room temperature overnight. Then the mixture was partitioned between CH_2Cl_2 (30 mL) and brine (30 mL). The aqueous phase was extracted with CH_2Cl_2 (3 x 30 mL). The combined organic phases were dried (Na₂SO₄) and concentrated under reduced pressure. Purification by flash column chromatography (CH₂Cl₂ : MeOH = 9:1) afforded the titled compound (0.68 g, 71%); ¹H NMR (601 MHz, CDCl₃) δ 14.57 (s, 1H), 9.68 (s, 1H), 8.26 (s, 1H), 6.34 (d, *J* = 12.8 Hz, 1H), 6.30–6.23 (m, 2H), 5.94–5.85 (m, 1H), 5.18 (dd, *J* = 12.8, 10.1 Hz, 1H), 4.15 (s, 1H), 3.70 (d, *J* = 9.8 Hz, 1H), 3.55 (ddd, *J* = 10.3, 7.9, 2.4 Hz, 1H), 3.39 (dd, *J* = 10.3, 4.2 Hz, 1H), 3.32 (dd, *J* = 10.0, 2.1 Hz, 1H), 3.16 (s, 3H), 3.03– 2.92 (m, 3H), 2.87 (dq, *J* = 8.1, 2.6 Hz, 1H), 2.68 (br, 2H), 2.42 (dt, *J* = 9.6, 6.6 Hz, 1H), 2.32 (s, 2H), 2.28 (s, 3H), 2.04 (s, 3H), 1.90–1.82 (m, 2H), 1.80–1.74 (m, 1H), 1.72 (s, 3H), 1.24 (s, 4H), 1.08 (d, *J* = 7.0 Hz, 3H), 0.95 (d, *J* = 6.5 Hz, 6H), 0.83 (d, *J* = 7.0 Hz, 3H), 0.55 (d, *J* = 6.8 Hz, 3H), -0.14 (d, *J* = 7.0 Hz, 3H). ¹³C NMR (151 MHz, CDCl3) δ 192.1, 180.5, 171.0, 168.3, 168.2, 155.0, 147.5, 141.6, 141.0, 132.9, 132.6, 124.9, 123.3, 114.7, 114.2, 111.5, 109.5, 108.1, 104.4, 94.8, 85.5, 76.8, 71.6, 70.8, 66.4, 56.1, 51.6, 51.6, 39.5, 38.9, 37.9, 36.1, 35.2, 32.8, 29.8, 25.9, 22.9, 21.0, 21.0, 20.2, 17.2, 12.2, 10.9, 8.4, 7.9; MS (ESI): m/z [M+Na]⁺ calcd for C₄₉H₆₅N₄O₁₁Na: 827.4202; found: 827.4198 (error 0.5 ppm).

Acylation procedure C. To a solution of **6** (1.00 equiv) in 1,2-dichloroethane (0.1 M) under room temperature, the acid anhydride (2.00 equiv) and DMAP (0.50 equiv) were added every 12 hours until completion of conversion. The reaction mixture was then concentrated to give a residue. The residue was purified by flash column chromatography (hexane:ethyl acetate:methanol = $7:2:1$) and then by Prep-TLC (hexane:ethyl acetate:methanol = 7:2:1) to afford 25-*O*-acyl rifabutin analog **5i** (see below for the full characterization).

25-*O***-benzoyl-25-***O***-desacetyrifabutin (5a).** Prepared from **3** (50.0 mg) using acylation procedure A and the general deprotection procedure to afford the title compound (30.6 mg, 57% over two steps) as a purple solid. HRMS $(ESI-TOF)$ m/z $[M+H]^+$ calcd for $C_{51}H_{65}N_4O_{11}$ 909.4644, found 909.4665 (error 2.3 ppm).

25-*O***-(2-methylbenzoyl)-25-***O***-desacetyrifabutin (5b).** Prepared from **3** (30.0 mg) using acylation procedure A and the general deprotection procedure to afford **5b** (3.9 mg, 12% over two steps) as a purple solid. HRMS (ESI-TOF) *m/z* [M+H]⁺ calcd for C52H67N4O11: 923.4801, found: 923.4829 (error 3.0 ppm).

[a] In the NMR spectra of **5b** and **5h**, we observed disappearance or serious broadening of ¹H and ¹³C signals of the piperidine unit, which could be attributed to conformational mobility, the coupling with geminal and vicinal proton couplings, or quadrupolar coupling with the piperidine nitrogen. This phenomenon has been observed in natural products and intensively investigated.[2-5]

25-*O***-(3-methylbenzoyl)-25-***O***-desacetyrifabutin (5c).**

Prepared from **3** (25.0 mg) using acylation procedure A and the general deprotection procedure to afford **5c** (2.5 mg, 9% over two steps) as a purple solid. HRMS (ESI-TOF) m/z [M+H]⁺ calcd for C₅₂H₆₇N₄O₁₁: 923.4801, found: 923.4815 (error 1.5 ppm).

25-*O***-(4-methylbenzoyl)-25-***O***-desacetyrifabutin (5d).**

Prepared from **3** (18.9 mg) using acylation procedure A and the general deprotection procedure to afford **5d** (8.1 me are general dependence precedure to anonounce (solid mg, 39% over two steps) as a purple solid. HRMS (ESI-TOF) *m/z* [M+H]⁺ calcd for C52H67N4O11: 923.4801, found: 923.4821 (error 2.2 ppm).

25-*O***-(2-fluorobenzoyl)-25-***O***-desacetyrifabutin (5e).**

Prepared from **3** (30.0 mg) using acylation procedure A and the general deprotection procedure to afford **5e** (17.5 mg, 53% over two steps) as a purple solid. HRMS $(ESI-TOF)$ m/z $[M+Na]$ ⁺ calcd for $C_{51}H_{63}FN_4O_{11}Na$: 949.4370, found: 949.4341 (error 3.0 ppm).

25-*O***-(2-chlorobenzoyl)-25-***O***-desacetyrifabutin (5f).**

Prepared from **3** (25.0 mg) using acylation procedure A and the general deprotection procedure to afford **5f** (19.0 mg, 68% over two steps) as a purple solid. HRMS $(ESI-TOF)$ m/z $[M+H]^+$ calcd for $C_{51}H_{64}CIN_4O_{11}$: 943.4255, found: 943.4245 (error 1.0 ppm).

25-*O***-(2-methoxybenzoyl)-25-***O***-desacetyrifabutin (5g).**

Prepared from **3** (600 mg) using acylation procedure A and the general deprotection procedure to afford **5g** (116.2 mg, 26% over two steps) as a purple solid. HRMS (ESI-TOF) *m*/z [M+H]⁺ calcd for C₅₂H₆₇N₄O₁₂: 939.4750, found: 939.4777 (error 2.9 ppm).

25-*O***-(2-trifluoromethylbenzoyl)-25-***O***-**

desacetyrifabutin (5h). Prepared from **3** (30.0 mg) using acylation procedure A and the general deprotection procedure to afford **5h** (10.4 mg, 30% over two steps) as a purple solid. HRMS (ESI-TOF) *m/z* [M+H]⁺ calcd for C52H64F3N4O11: 977.4518, found: 977.4519 (error 0.1 ppm).

[a] See the footnote of the tabulated NMR data of **5b**.

25-*O***-(2-phenylbenzoyl)-25-***O***-desacetyrifabutin (5i).** Prepared from **6** (415.5 mg) using acylation procedure C to afford **5i** (29.8 mg, 6%) as a purple solid. HRMS (ESI-TOF) m/z [M+H]⁺ calcd for C₅₇H₆₉N₄O₁₁: 985.4957, found: 985.4952 (error 0.5 ppm).

25-*O***-(3-fluorobenzoyl)-25-***O***-desacetyrifabutin**

(5j). Prepared from **3** (20.0 mg) using acylation procedure A and the general deprotection procedure to afford **5j** (12.0 mg, 55% over two steps) as a purple solid. HRMS (ESI-TOF) *m/z* [M+H]⁺ calcd for $C_{49}H_{69}N_4O_{11}$: 927.4550, found: 927.4556 (error 0.6 ppm).

25-*O***-(2-methylbutyl)-25-***O***-desacetyrifabutin**

(5k). Prepared from **3** (32.8 mg) using acylation procedure A and the general deprotection procedure to afford **5k** (22.2 mg, 64% over two steps) as a purple solid. HRMS (ESI-TOF) *m/z* [M+H]⁺ calcd for $C_{49}H_{69}N_4O_{11}$: 889.4957, found: 889.4995 (error 4.3 ppm).

[a] A diastereomeric mixture of the acid anhydride was used in this synthesis.

25-*O***-(2-ethylbutyl)-25-***O***-desacetyrifabutin (5l).** Prepared from **3** (20.9 mg) using acylation procedure A and the general deprotection procedure to deliver **5l** (16.0 mg, 72% over two steps) as a purple solid. HRMS (ESI-TOF) *m/z* $[M+H]^+$ calcd for $C_{50}H_{71}N_4O_{11}$: 903.5114, found: 903.5105 (error 1.0 ppm).

25-*O***-(3-pyridinecarbonyl)-25-***O***-desacetyrifabutin**

(5m). Prepared from **3** (59.7 mg) using acylation procedure A and the general deprotection procedure to afford **5m** (29.3 mg, 46% over two steps) as a purple solid. HRMS (ESI-TOF) m/z [M+H]⁺ calcd for $C_{50}H_{64}N_5O_{11}$: 910.4597, found: 910.4606 (error 1.0 ppm).

25-*O***-(5-pyrimidinecarbonyl)-25-***O***-desacetyrifabutin (5n).** Prepared from **3** (40.0 mg) using acylation procedure B and the general deprotection procedure to afford **5n** (29.3 mg, 56% over two steps) as a purple solid. HRMS (ESI-TOF) m/z [M+Na]⁺ calcd for $C_{49}H_{62}N_6O_{11}Na$: 933.4369, found: 933.4395 (error 2.8 ppm).

25-*O***-(2-thiazolecarbonyl)-25-***O***-desacetyrifabutin**

(5o). Prepared from **3** (40.0 mg) using acylation procedure B and the general deprotection procedure to afford **5o** (29.3 mg, 51% over two steps) as a purple solid. HRMS (ESI-TOF) *m/z* [M+H]⁺ calcd for $C_{48}H_{62}N_5O_{11}S$: 916.4161, found: 916.4204 (error 4.7 ppm).

Notes on the acylation process of the compound synthesis

The acylation procedure reported above using excessive amounts of anhydrides and DMAP was the only method we found out that could afford the desired C25-OH acylation products. Under the conditions of using anhydride along or with substoichiometric amount of DMAP, the starting material had almost no conversion (see below, Case 1 and Case 2). This method was proved effective for most of our synthesis with reasonable to high yield. In certain cases, the relatively low yield was due to extremely slow conversion and, once heated up, the massive side reactions (see below, Case 3 and Case 4).

[a] TLC plates were developed at least twice using a mixture of solvents hexanes:EtOAc:MeOH:Et₃N = 8:1:1:0.1; S: starting material; R: reaction mixture (after a miniextraction between H₂O and EtOAc); SR: co-spots of the starting material and the reaction mixture.

Copies of the ¹H, ¹³C, 1H-13C HMBC, and representative 1H-¹H NOESY NMR Spectra

(In the ¹H-¹³C HMBC NMR spectra, correlations of CH-25 and C-35 were labeled to clarify the acylation site.)

25-*O*-benzoyl-25-*O*-desacetyrifabutin (**5a**)

¹H NMR spectrum

25-*O*-(2-methylbenzoyl)-25-*O*-desacetyrifabutin (**5b**)

25-*O*-(3-methylbenzoyl)-25-*O*-desacetyrifabutin (**5c**)

25-*O*-(4-methylbenzoyl)-25-*O*-desacetyrifabutin (**5d**)

C NMR spectrum

25-*O*-(2-fluorobenzoyl)-25-*O*-desacetyrifabutin (**5e**)

¹H-¹H NOESY NMR spectrum

25-*O*-(2-chlorobenzoyl)-25-*O*-desacetyrifabutin (**5f**)

C NMR spectrum

25-*O*-(2-methoxybenzoyl)-25-*O*-desacetyrifabutin (**5g**)

¹H-¹H NOESY NMR spectrum

25-*O*-(2-trifluoromethylbenzoyl)-25-*O*-desacetyrifabutin (**5h**)

25-*O*-(2-phenylbenzoyl)-25-*O*-desacetyrifabutin (**5i**)

C NMR spectrum

¹H-¹H NOESY NMR spectrum

25-*O*-(3-fluorobenzoyl)-25-*O*-desacetyrifabutin (**5j**).

25-*O*-(2-methylbutyl)-25-*O*-desacetyrifabutin (**5k**)

C NMR spectrum

25-*O*-(2-ethylbutyl)-25-*O*-desacetyrifabutin (**5l**)

C NMR spectrum

25-*O*-(3-pyridinecarbonyl)-25-*O*-desacetyrifabutin (**5m**)

¹H NMR spectrum

¹³C NMR spectrum

25-*O*-(5-pyrimidinecarbonyl)-25-*O*-desacetyrifabutin (**5n**)

¹H NMR spectrum

C NMR spectrum

25-*O*-(2-thiazolecarbonyl)-25-*O*-desacetyrifabutin (**5o**)

¹H NMR spectrum

C NMR spectrum

Part III: Biological procedures

Ethical approval statements for the animal experiments

All experiments involving live mice were approved by the Institutional Animal Care and Use Committee of the Center for Discovery and Innovation, Hackensack Meridian Health. Accreditation number: A4278-01. IACUC number: 269.00 (*in vivo* pharmacokinetic studies); 287.00 (*in vivo* animal efficacy studies).

Bacterial strains, culture media and compounds.

M. abscessus Bamboo was isolated from the sputum of a patient with amyotrophic lateral sclerosis and bronchiectasis and was provided by Wei Chang Huang, Taichung Veterans General Hospital, Taichung, Taiwan. *M. abscessus* Bamboo whole-genome sequencing showed that the strain belongs to *M. abscessus* subsp. *abscessus* and harbors an inactive clarithromycinsensitive *erm*(41) C28 sequevar [6-7] . *Mycobacterium abscessus* subsp. *abscessus* ATCC 19977, harboring the inducible clarithromycin resistance-conferring *erm*(41) T28 sequevar^[8], was purchased from the American Type Culture Collection (ATCC). *M. abscessus* subsp. *abscessus* K21 was isolated from a patient and provided by Sung Jae Shin (Department of Microbiology, Yonsei University College of Medicine, Seoul, South Korea) and Won-Jung Koh (Division of Pulmonary and Critical Care Medicine, Samsung Medical Center, Seoul, South Korea). This strain harbors the inactive, clarithromycin-sensitive erm(41) C28 sequevar as determined previously^[9].

Generation of Δ*arrMab* in the *M. abscessus* ATCC 19977 genetic background by recombineering was described previously.^[10]

The selection and characterization of the rifamycin-resistant *M. abscessus* mutant RFB-R1 was described previously.[10] RFB-R1 carries a *rpoB* (RNAP) c1339t nucleotide mutation, which corresponds to a H447Y missense mutation that was previously reported in M. tuberculosis.[11-12]

For general bacteria culturing and MIC experiments, Middlebrook 7H9 broth (BD Difco) supplemented with 0.5% albumin, 0.2% glucose, 0.085% sodium chloride, 0.0003% catalase, 0.2% glycerol, and 0.05% Tween 80.

Clarithromycin was purchased from Sigma-Aldrich. Rifampicin was purchased from GoldBio. Rifabutin was purchased from Acros Organics. All drugs were prepared as 10 mM stocks in 100% DMSO.

MIC Assay in 96-well Plate Format.

MIC determination was carried out in 96-well plate format as previously described.^[13-14] 96-well plates were initially set up with 100 µL of 7H9 per well. For each compound, a ten-point two-fold dilution series starting at twice the desired highest concentration was dispensed onto the 96-well plates using a Tecan D300e Digital Dispenser, with the DMSO concentration normalized to 2%. *M. abscessus* culture grown to mid-log phase ($OD_{600} = 0.4-0.6$) was diluted to $OD_{600} = 0.1$ (1 \times 10⁷ CFU/mL). 100 μ L of the resulting bacteria suspension was dispensed onto the 96-well plates containing compounds to give a final volume of 200 µL per well with an initial OD₆₀₀ = 0.05 (5 \times $10⁶ CFU/mL)$ and final DMSO concentration of 1%. Final compound concentration ranges were typically 50–0.098 µM, 6.25–0.012 µM, 0.006–3.13 µM, or 0.003–1.56 µM. Untreated control wells are included on each plate that contain bacteria suspension and 1% DMSO. Plates were sealed with parafilm, stored in boxes with wet paper towels and incubated at 37 °C with shaking (110 RPM). Plates were incubated for 3 days. To determine growth, OD_{600} was measured using a Tecan Infinite M200 plate reader on Day 0 and Day 3. Two biological replicates were performed. Clarithromycin was included in each experiment as a positive control.

For each well on the 96-well plate, bacterial growth was calculated by subtracting the Day 0 $OD₆₀₀$ value from the Day 3 $OD₆₀₀$ value. For each compound series, the bacterial growth values for the untreated control wells were averaged to give the average drug-free bacterial growth. For compound-containing wells, percentage growth was calculated by dividing their growth values by the average drug-free bacterial growth for the compound series and multiplying by 100. For each compound series, we plotted percentage growth *versus* compound concentration. By visual inspection of the dose-response curve, we determined the MIC of a compound as the compound concentrations that would result in 90% growth inhibition.

Pharmacokinetics studies.

CD-1 female mice (22-25 g) were used in oral pharmacokinetic studies. Rifabutin, **5b**, **5j**, and **5m** were administered as a single intravenous (IV) or oral (PO) dose gavage at 10 mg/kg in a solution formulation composed of 5% DMSO:95% (4% Cremophor EL). Aliquots of 50 μL of blood were taken by puncture of the lateral tail vein from each mouse (n = 3 per route and dose) at 30 minutes, 1, 3, 5, 7, and 24 hours post-dose for oral dosing and at 1 minute, 15 minutes, 1, 3, 7, and 24

hours for IV dosing. Blood was captured in CB300 blood collection tubes containing K₂EDTA and stored on ice. Plasma was recovered after centrifugation and stored at -80°C until analyzed by high pressure liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). Pharmacokinetic parameters were calculated using non-compartmental pharmacokinetic analysis.

LC-MS/MS analytical methods for the pharmacokinetic studies

Neat 1 mg/mL DMSO stocks of rifabutin, C25-desacetyl rifabutin, **5b**, **5j**, and **5m** were serial diluted in 50/50 Acetonitrile (ACN)/Milli-Q water to create neat standard solutions. Plasma standards were created by adding 10 µL of spiking solutions to 90 µL of drug free plasma (CD-1 K2EDTA Mouse, Bioreclamation IVT). 5 µL of control, standard, or study sample were added to 100 µL of ACN protein precipitation solvent containing 10 ng/mL of the internal standards Verapamil (Sigma Aldrich) and rifabutin-d7 (Toronto Research Chemical). Extracts were vortexed for 5 minutes and centrifuged at 4000 RPM for 5 minutes. 75 µL of supernatant was transferred for LC-MS/MS analysis and diluted with 75 µL of Milli-Q deionized water. Rifabutin was purchased from Carbosynth. C25-desacetyl rifabutin and rifabutin-d7 were purchased from Toronto Research Chemical. Verapamil was purchased from Sigma-Aldrich.

LC-MS/MS analysis was performed on a Sciex Applied Biosystems Qtrap 6500+ triple-quadrupole mass spectrometer coupled to a Shimadzu Nexera X2 UHPLC system to quantify each drug in plasma. Chromatography was performed on an Agilent SB-C8 (2.1x30 mm; particle size, 3.5 µm) using a reverse phase gradient. Milli-Q deionized water with 0.1% formic acid was used for the aqueous mobile phase and 0.1% formic acid in ACN for the organic mobile phase. Multiplereaction monitoring of parent/daughter transitions in electrospray positive-ionization mode was used to quantify all analytes. The following MRM transitions were used for rifabutin (847.60/755.60), rifabutin-d7 (854.60/762.60), C25-desacetyl rifabutin (805.48/773.50), **5b** (923.48/891.40), **5j** (927.39/895.30), **5m** (910.45/878.40) and Verapamil (455.40/165.00). Sample analysis was accepted if the concentrations of the quality control samples were within 20% of the nominal concentration. Data processing was performed using Analyst software (version 1.6.2; Applied Biosystems Sciex).

Plasma protein binding assays

84

DMSO stocks were spiked into plasma to a concentration of 10,000 ng/mL. 200 µL of spiked plasma was pipetted into the sample chamber of the rapid equilibrium dialysis (RED) cartridge. 350 µL of PBS was added to the adjacent cartridge. The plate containing the RED was sealed and incubated at 37°C on the thermomixer at 300 RPM for 4 h. After incubation, 50 µL aliquots of plasma were removed and added to 50 µL of blank plasma in a deep well plate (1:1). Similarly, 50 µL aliquots of PBS were removed and added to 50 µL of blank plasma. This created an identical matrix between buffer and non-buffer samples. Samples were processed and quantified as specified in the LC-MS/MS analytical method.

Crystal structure determination

Crystals of *M. tuberculosis* RNAP (RNAP*Mtb*) *M. tuberculosis* σ ^A RPo were prepared as described.^[15] Crystals were soaked overnight at 23 °C in cryoprotection solution (20 mM Tris-HCl, pH 8.2, 200 mM potassium chloride, 20 mM magnesium chloride, 7% (m/v) PEG-3350, 20% (v/v) (2R,3R)-(−)-2,3-butanediol, 1 mM CHAPSO) supplemented with 0.5 mM **5a** and were flash-frozen in liquid nitrogen.

X-ray diffraction data were collected at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12-2 and processed using HKL3000.^[16] The structure was solved by molecular replacement using the structure of *M. tuberculosis* σ^A RPo (PDB: 5UHA)^[15] as the search model. Iterative cycles of model building and refinement were performed using $Coot^[17]$ and Phenix Refine^[18]. The final model was obtained by a refinement with secondary-structure restraints and individual and group B-factors. The atomic model and structure factors were deposited in the Protein Data Bank (PDB) with accession number 7U22.

Arr*Mab* **expression and purification**

Arr*Mab* was codon optimized for *E. Coli* and cloned into pET-28b (+) expression vector. Clone was transformed in *E. coli* BL21 (DE3) cells. Single colony was picked and allowed to grow in 50 mL LB having desired antibiotic overnight at 37 °C shaking at 250 RPM. Overnight grown Primary cultures were transferred to 1 L LB broth with kanamycin 50 µg/mL and grown at 37 °C shaking at 250 RPM till OD (600 nm) reaches 0.6. Protein expressions were induced by adding 0.5 mM IPTG and reducing the temperature to 20 °C for 16 hours. Cells were harvested by centrifugation at 8000*g* for 10 minutes and resuspended in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, and

1.0 M sorbitol) containing complete protease inhibitor tablet, PMSF (5 mM) and hen egg white lysozyme (0.5 mg/mL). Cells were lysed by Avastin C3 ultra high-pressure liquidizer at 15000- 20000 psi for 15 minutes. Supernatant was obtained after high-speed centrifugation and incubated with equilibrated Ni-NTA beads for 90 minutes at 4 °C on rotator. Both proteins were eluted with gradient of 50-200 mM imidazole.

HPLC analysis for in vitro validation of ADP-ribosylation

Endpoint reactions were set up for rifampicin, rifabutin and rifabutin analogs in the presence of 5 µM enzyme in 50 mM HEPES buffer. Reactions were quenched using methanol and subjected to analysis on HPLC. Reverse-phase LC was performed on a Kinetex C8 column (100 mm \times 2.1 mm, 2.6 µm; Phenomenex, Torrance, CA) using LC 1200 Infinity Series, Agilent Technologies instrument with wavelength monitored at 260 nm (for rifampicin) and 277 nm (for rifabutin and rifabutin analogs). The elution gradient was carried out with binary solvent system consisting of 0.1% formic acid in H_2O (solvent A) and 0.1% formic acid in MeCN (solvent B). A linear gradient profile with the following proportions (v/v) of solvent B was applied (t (min), %B): (0, 5), (0.5, 5), (9, 100), (10.5, 100), (12, 10) with 5 min for re-equilibration to provide a total run time of 17 min. The flow rate was 0.3 mL/min and the column oven was maintained at 28 °C. The injection volume was 10 µL.

Efficacy evaluation in *M. abscessus* **Mouse Infection Model**

Eight-week-old female NOD.CB17-*Prkdcscid*/NCrCrl (NOD SCID) mice (Charles River Laboratories) were infected intranasally with ∼10⁶ CFU of *M. abscessus* subsp. *abscessus* K21 as described previously.^[9] Acute infection was achieved within one day. Drugs or the vehicle control was administered once daily for 10 consecutive days by oral gavage, starting 1 day postinfection. Clarithromycin (250 mg/kg, Sandoz clinical tablets), rifabutin (10 mg/kg, Carbosynth) and **5j** (10 mg/kg) were formulated in 0.5% carboxy-methyl-cellulose/ 0.5% tween 80 at 8 mL/kg dosing volume. All mice were euthanized 24 h after the last dose, and lungs and spleen were aseptically removed prior to homogenization. The bacterial load in these organs was determined by plating serial dilutions of the organ homogenates onto Middlebrook 7H11 agar (BD Difco) supplemented with 0.2% (v/v) glycerol and 10% (v/v) OADC. The agar plates were incubated for 5 days at 37°C prior to counting of colonies.

Part IV: Computational procedures

Sequence retrieval and Homology modeling of RNAP*Mab* **and Arr***Mab*

The sequences of RNAP*Mab* and Arr*Mab* were retrieved from the UniProt database (entry B1MH62 & B1MH05). A protein-protein BLAST (Blastp) search was performed to find a suitable homologous sequence (template) with known 3D structure to this amino acid sequence.^[19] The sequences of the best template, RNAP*Mtb* (PDB: 5UHB) and Arr*Msm* (PDB: 2HW2), and the studied targets were then aligned using Clustal Omega.^[20] The initial homology models of RNAP_{Mab} and Arr*Mab* based on the selected templates were generated with the automated homology modeling software, MODELLER (Version 9), in which the program is based on comparative structure modeling.^[21] The catalytic domain of Arr_{Mab} was further aligned to the similar poly ADP-ribose polymerase (PARP) domains of *Pseudomonas aeruginosa* exotoxin (PDB: 1AER) and *Gallus gallus* poly ADP-ribose polymerase (PDB: 1A26) to predict the NAD⁺ binding site of Arr*Mab*. The ADP-ribosyl-oxocarbenium intermediate was docked at the predicted binding site of Arr*Mab* and the best binding mode was selected to build an Arr_{Mab} homology model with the ADP-ribosyloxocarbenium intermediate bound. Then, both the RNAP*Mab* and Arr*Mab* models were further energy minimized with Discovery Studio 3.5 using the CHARMM force field. The quality and stability of the homology models were validated by checking the stereochemical parameters using PROCHECK, VERIFY3D, and ERRAT at SAVES server (http://nihserver.mbi.ucla.edu/SAVES).

Rational for ADP-ribosyl-oxocarbenium intermediate in the Arr*Mab* **model**

The Arr*Mab* homology model with the ADP-ribosyl-oxocarbenium intermediate was used to dock **5a** and rifabutin after confirming that both compounds showed better binding affinity values when nicotinamide moiety of NAD⁺ was not present in the model. In this model, it was predicted that the hydroxyl groups at C23 position of rifabutin and **5a** were more closely located to C1′ of the ribose ring, compared to NAD⁺ bound Arr*Mab* homology model. This prediction also supports the proposed mechanism that the oxocarbenium transition state enables the hydroxyl group at position 23 of the antibiotic to attack C1' of the ribose.^[22] Thus, the Arr_{Mab} homology model with the ADP-ribosyl-oxocarbenium intermediate was a better model to predict the interaction between rifabutin analogs and Arr_{Mab}.

Ligand generation

The 2D structures of rifabutin and synthetic rifabutin analogs were drawn in Chemdraw and their SMILES notation was obtained. The 3D structures were obtained and converted into SDF files after energy minimization with Discovery Studio 3.5.

In silico **molecular docking analysis**

The molecular docking was carried out by Pyrx with Autodock Vina engine. The AutoDockTools package was employed to generate the docking input files in pdbqt. Both protein targets and ligands were opened in PyRx virtual screening tool as a starting protein structure in pdbqt format. Docking was carried out taking the rifampicin binding site residues for RNAP_{Mab} and for Arr_{Mab} inside a grid box with co-ordinates along X, Y, and Z-axis and dimensions conformed to 162.8 Å \times 163.38 Å \times 20.22 Å and 20 Å \times 20 Å \times 20 Å, and 66.89 Å \times 66.62 Å \times 10.54 Å and 20 Å \times 20 Å × 20 Å, respectively. Lamarckian Genetic Algorithm (LGA) was used as ligand conformation search process, and the other parameters were default settings. The quality of docking was validated by re-docking ligand to observe the precision of the docking condition. All the docking simulation using various ligands was performed with the exhaustiveness of 24. The best binding mode and affinity values were obtained in PyRx virtual screening GUI and log files.

REFERENCES

- [1] V. Novohradsky, L. Zerzankova, J. Stepankova, O. Vrana, R. Raveendran, D. Gibson, J. Kasparkova, V. Brabec, *J. Inorg. Biochem.* **2014**, *140*, 72-79.
- [2] M. C. McCoy, D. J. Faulkner, *J. Nat. Prod.* **2001**, *64*, 1087-1089.
- [3] H. Booth, *Chem. Commun. (London)* **1968**, 802b-803.
- [4] F. A. Anet, I. Yavari, I. J. Ferguson, A. R. Katritzky, M. Moreno-Mañas, M. J. Robinson, *J. Chem. Soc. Chem. Commun.* **1976**, 399-400.
- [5] E. L. Eliel, D. Kandasamy, C. Yen, K. D. Hargrave, *J. Am. Chem. Soc.* **1980**, *102*, 3698-3707.
- [6] S. Bastian, N. Veziris, A. L. Roux, F. Brossier, J. L. Gaillard, V. Jarlier, E. Cambau, *Antimicrob. Agents Chemother.* **2011**, *55*, 775-781.
- [7] M. Yee, D. Klinzing, J. R. Wei, M. Gengenbacher, E. J. Rubin, T. Dick, *Genome Announc* **2017**, *5*, e00388-17.
- [8] F. Ripoll, S. Pasek, C. Schenowitz, C. Dossat, V. Barbe, M. Rottman, E. Macheras, B. Heym, J. L. Herrmann, M. Daffe, R. Brosch, J. L. Risler, J. L. Gaillard, *PLoS One* **2009**, *4*, e5660.
- [9] T. Dick, S. J. Shin, W. J. Koh, V. Dartois, M. Gengenbacher, *Antimicrob. Agents Chemother.* **2020**, *64*, e01943-19.
- [10] U. S. Ganapathy, T. Lan, P. Krastel, M. Lindman, M. D. Zimmerman, H. Ho, J. P. Sarathy, J. C. Evans, V. Dartois, C. C. Aldrich, T. Dick, *Antimicrob. Agents Chemother.* **2021**, *65*, e00978-21.
- [11] A. Rominski, A. Roditscheff, P. Selchow, E. C. Böttger, P. Sander, *J. Antimicrob. Chemother.* **2017**, *72*, 376-384.
- [12] S. Ramaswamy, J. M. Musser, *Tubercle and lung disease* **1998**, *79*, 3-29.
- [13] D. B. Aziz, J. L. Low, M. L. Wu, M. Gengenbacher, J. W. P. Teo, V. Dartois, T. Dick, *Antimicrob. Agents Chemother.* **2017**, *61*, e00155-17.
- [14] J. L. Low, M. L. Wu, D. B. Aziz, B. Laleu, T. Dick, *Front. Microbiol.* **2017**, *8*, 1539.
- [15] W. Lin, S. Mandal, D. Degen, Y. Liu, Y. W. Ebright, S. Li, Y. Feng, Y. Zhang, S. Mandal, Y. Jiang, S. Liu, M. Gigliotti, M. Talaue, N. Connell, K. Das, E. Arnold, R. H. Ebright, *Mol. Cell* **2017**, *66*, 169- 179.e168.
- [16] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307-326.
- [17] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126-2132.
- [18] P. V. Afonine, M. Mustyakimov, R. W. Grosse-Kunstleve, N. W. Moriarty, P. Langan, P. D. Adams, *Acta Crystallogr. Sect. D* **2010**, *66*, 1153-1163.
- [19] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403-410.
- [20] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins, *Mol. Syst. Biol.* **2011**, *7*, 539.
- [21] A. Fiser, R. K. Do, A. Sali, *Protein Sci.* **2000**, *9*, 1753-1773.
- [22] J. Baysarowich, K. Koteva, D. W. Hughes, L. Ejim, E. Griffiths, K. Zhang, M. Junop, G. D. Wright, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 4886-4891.