Journal of Medicinal Chemistry

Article

Development of Mithramycin Analogues with Increased Selectivity towards ETS Transcription Factor Expressing Cancers

Prithiba Mitra, Joseph M. Eckenrode, Abhisek Mandal, Amit K. Jha, Shaimaa M. Salem, Markos Leggas, and Jurgen Rohr

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.8b01107 • Publication Date (Web): 16 Aug 2018

Just Accepted

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Manuscript Title: Development of Mithramycin Analogues with Increased Selectivity towards ETS Transcription Factor Expressing Cancers

Prithiba Mitra,† Joseph M. Eckenrode,† Abhisek Mandal, Amit K. Jha, Shaimaa M. Salem, Markos Leggas* and Jürgen Rohr*

Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lee T. Todd, Jr. Building, 789 South Limestone Street, Lexington, KY 40536-0596, USA

ABSTRACT: Mithramycin A (**1**) was identified as the top potential inhibitor of the aberrant ETS transcription factor EWS-FLI1, which causes Ewing sarcoma. Unfortunately, **1** has a narrow therapeutic window compelling us to seek less toxic and more selective analogues. Here, we used MTMSA (**2**) to generate analogues via peptide coupling and fragment-based drug development strategies. Cytotoxicity assays in ETS and non-ETS dependent cell lines identified two dipeptide analogues, **60** and **61**, with 19.1- and 15.6-fold selectivity, respectively, compared to 1.5-fold for **1**. Importantly, the cytotoxicity of **60** and **61** is <100 nM in ETS cells. Molecular assays demonstrated the inhibitory capacity of these analogues against EWS-FLI1 mediated transcription in Ewing sarcoma. Structural analysis shows that positioning the tryptophan residue in a distal position improves selectivity, presumably via interaction with the ETS transcription factor. Thus, these analogues may present new ways to target transcription factors for clinical use.

INTRODUCTION

Mithramycin (MTM; **1**, Figure 1) is an aureolic acid-type polyketide drug produced by various soil bacteria of the genus *Streptomyces* and was found to possess activity against a wide variety of human cancers.[1-2] MTM (**1**) was clinically evaluated in the 1960s and 70s as an agent for the chemotherapy of various cancers. Despite some remarkable success using MTM (**1**) as a single agent, the results were mixed due to its narrow therapeutic index and considerable variation in patients' ability to tolerate the

drug.[3] Another concern was the lack of understanding of MTM's (**1**) mode-of-action. Taken together these limitations limited clinical use of MTM (**1**) as a chemotherapeutic agent and it has now been largely abandoned.[4] Interest in MTM (**1**) was renewed recently, after the drug was identified as the top inhibitor of the ETS transcription factor fusion, EWS-FLI1, in a screen of more than 50,000 natural products and synthetic compounds. FLI1 and ERG are ETS transcription factors that are expressed as fusions with EWS and are the primary cause of Ewing sarcoma.[5-6] Aside from Ewing sarcoma, aberrant ETS transcription factors contribute significantly to the malignancy of prostate cancer, leukemia and lymphoma. With respect to prostate cancer, approximately 50% of patients express a truncated form of ERG as a result of the TMPRSS2-ERG gene fusion.[7] Interestingly, the DNA binding domain of ERG and FLI1 is conserved and thus molecules that interfere with the activity of one should also inhibit the other. Given the importance of these aberrant transcription factors in driving malignancy, the clinical use of MTM (**1**) gave investigators hope for a "targeted" therapy. This was tested in a recent national cancer institute (NCI) conducted clinical study where Ewing sarcoma patients were enrolled to assess the utility of MTM (**1**) in a population of patients, all of whom express ETS fusions. Unfortunately, the results were inconclusive because the trial was terminated early, due to toxicities. As such, the development of less toxic and more selective analogues of MTM (**1**) is highly desirable.

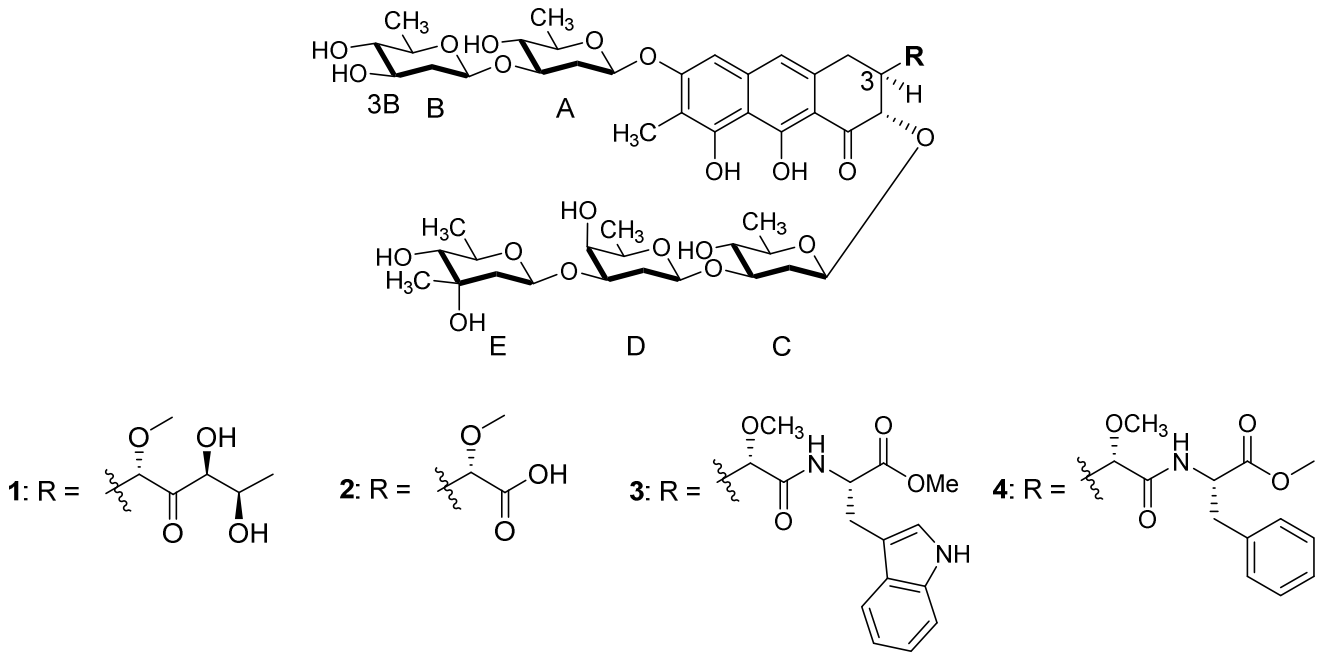


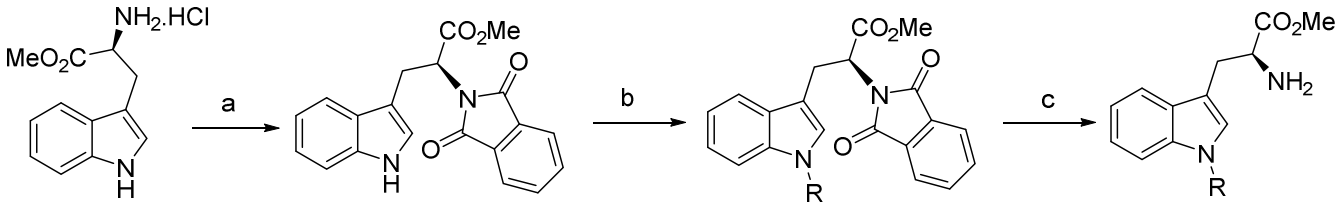
Figure 1. Structure of Mithramycin (MTM; **1**), Mithramycin SA (MTMSA; **2**), Mithramycin SA-Tryptophan (MTMSA-Trp; **3**) and Mithramycin SA Phenylalanine (MTMSA-Phe; **4**)

Our effort to develop such analogues is largely based on mechanistic studies that focused on understanding MTM's (**1**) mechanism of action. At the molecular level, it is known that MTM (**1**) binds to GC-rich DNA as a Mg^{2+}coordinated dimer and modulates the activity of the transcription factor Sp1 and presumably others.[8] We previously found that mithramycin SA (MTMSA; **2,** Figure 1), which is a combinatorial biosynthetic analogue of MTM (**1**) produced by *S. argillaceus*, upon inactivation of the *mtmW* gene,[9] has no cytotoxicity (Table 1, Entry 19). Using the free carboxylic acid group in the 3-side chain of MTMSA, we coupled natural amino acids and small molecules to generate a series of analogues,[10] out of which MTMSA-Trp (**3**, Figure 1) and MTMSA-Phe (**4**, Figure 1) were found to have cytotoxicity, comparable to MTM (**1**).[11] Most importantly, we demonstrated with crystallography studies that the 3-side chain of the MTMSA analogues can interact with FLI1.[12] In such complexes, aromatic 3-side chain MTMSA derivatives have sufficient length to directly interact with the FLI1 DNA binding domain of EWS-FLI1, reflecting the in vitro potency of MTMSA-Trp (**3**) and MTMSA-Phe (**4**) against Ewing sar-

coma. Moreover, these studies pose a new mode-of-action hypothesis, which requires a ternary MTM (**1**)-DNA-FLI1 (or MTM (**1**)-DNA-ERG) complex. In this article, we describe our efforts to find more selective MTM (**1**) analogues for the treatment of cancers expressing aberrant ETS fusions or ETS factors. Our approach combined fragment-based drug development (FBDD) with structure-activity relationship (SAR) studies starting from of MTMSA-Trp (**3**).

RESULTS AND DISCUSSION

Initial pharmacological studies showed that both, MTMSA-Trp (**3**) and MTMSA-Phe (**4**), appeared to have the potential to overcome the limitation of MTM (**1**), with MTMSA-Trp (**3**) being the more cytotoxic of the two derivatives (Table 3, Entry 10).[11] However, their selectivity towards Ewing sarcoma cell lines were only slightly improved compared to MTM (**1**) (Table 3). We chose MTMSA-Trp (**3**) as a starting point to better understand and improve its potential binding properties to EWS-FLI1, by varying its electronic, steric and hydrogen bonding properties of tryptophan residue, with the major objective to increase the selectivity of it towards Ewing sarcoma cell lines while maintaining a cytotoxic activity comparable to MTM (**1**).



5 **6**

7: R = Me, 60%
8: R = Bn, 70%
9: R = Allyl, 85%
10:R = Prenyl, 65%

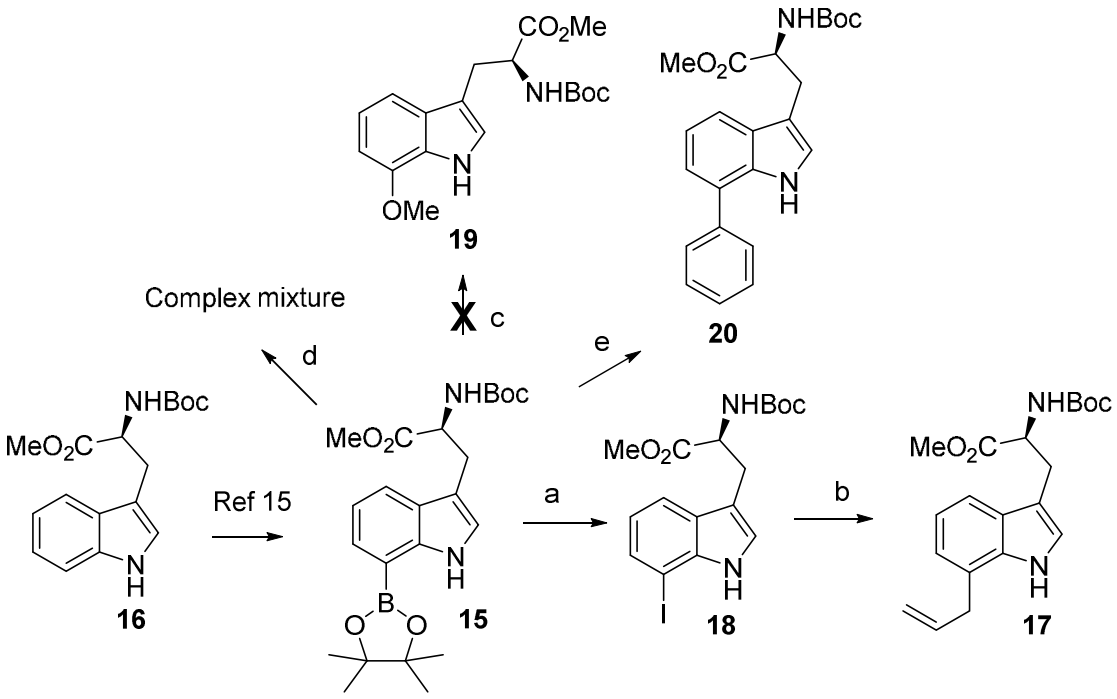
11: R = Me,65%
12: R = Bn, 51%
13: R = Allyl, 58%
14: R = Prenyl, 50%

Reagents and conditions: a) Phthalic anhydride, Et_3N, toluene, reflux, 15 h, 80%; b) NaH, Alkyl bromide, DMF, 0 °C to rt, 10 h; c) $NH_2NH_2{\cdot}H_2O$, MeOH-DCM, rt.

Scheme 1. Selective indole *N*-alkylation of tryptophan

To evaluate a potential hydrogen bonding of the indole-NH, it was protected by *N*-alkylation. After initial protection of the primary amine group of L-tryptophan hydrochloride (**5**) with phthalimide using phthalic anhydride in the presence of excess triethylamine in refluxing toluene, which gave the protected tryptophan **6**[13] in 80% yield, the indole *N*-methylation of **6** with MeI-NaH in DMF provided **7**[19] in 60% yield, and benzylation with benzyl bromide under similar conditions furnished *N*-benzylated tryptophan **8** in 70% yield (Scheme 1). Likewise, treatment with allyl bromide and prenyl bromide gave *N*-allyl tryptophan **9** and *N*-prenyl tryptophan **10**,[14] respectively. Treatment with hydrazine hydrate in MeOH-DCM at room temperature ensured the removal of the phthalimide protection group in all cases (**7** - **10**), to provide the corresponding amines (**11** - **14**) in 50 - 65% yield (Scheme 1).

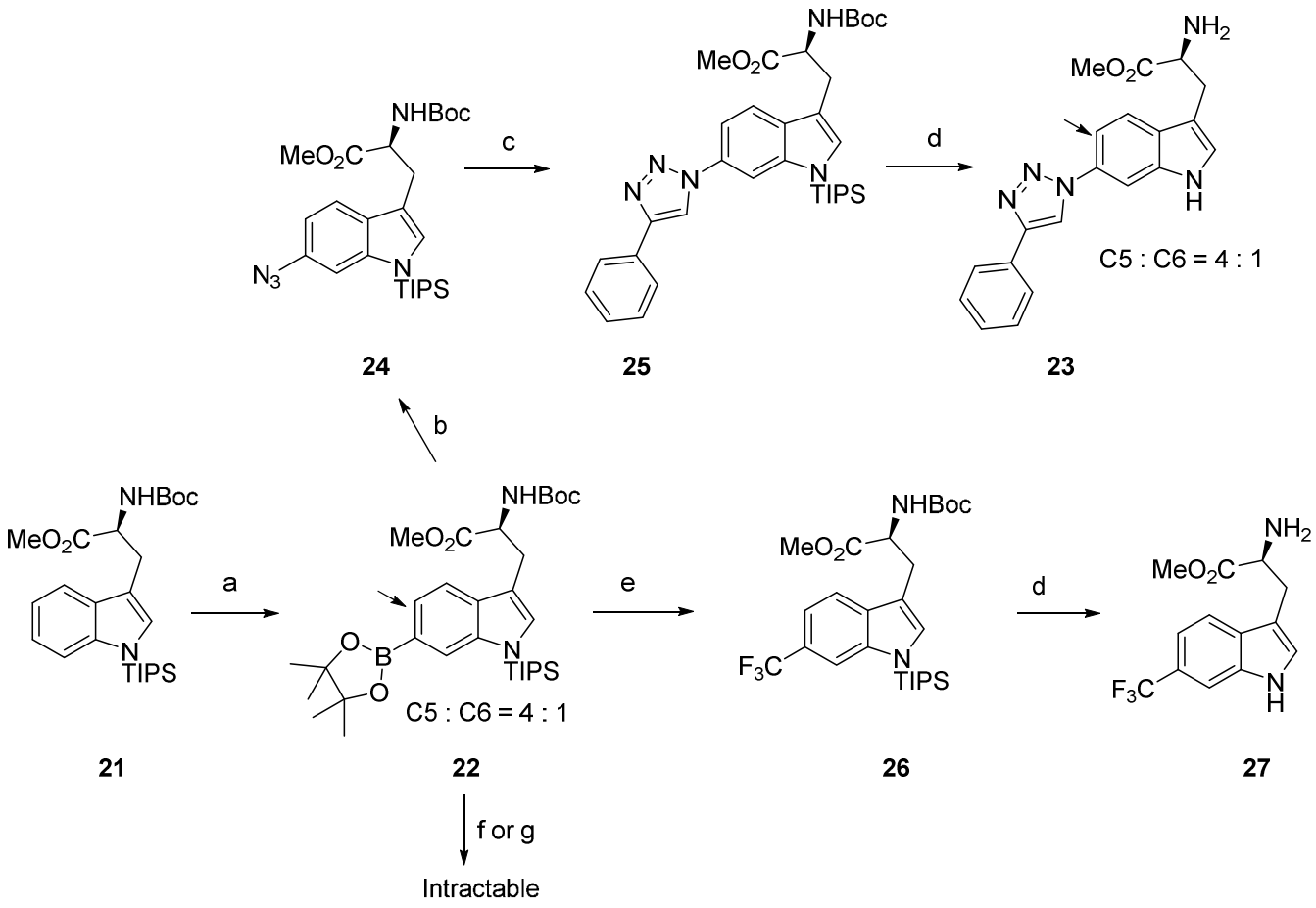
To diversify the tryptophan residue, iridium-catalyzed borylation and palladium catalyzed cross-coupling reactions were applied. Borylation allowed the introduction of various functionalities into the indole core of tryptophan, since the carbon-boron bond can be easily modified. For the C7 diversification of tryptophan, the required borylated tryptophan **15**[15] was prepared following the method developed by Movagasshi and co-workers starting from protected tryptophan **16** (Scheme 2). To incorporate an allyl residue into tryptophan, which could serve as a handle for further modifications through Grubbs chemistry, 7-allyltryptophan **17** was chosen.



Reagents and conditions: a) CuI, 1,10-Phen, KI, MeOH-H_2O, 40%; b) Allyl(n-Bu)$_3$Sn, Pd(PPh_3)$_4$, PhMe, 120 °C, 60%; c) MeOH, $Cu(OAc)_2 \cdot H_2O$, Et_3N, O_2, rt 12 h; d) CuTC, Togni reagent, 1,10-Phen, DCM, LiOH, H_2O, rt; e) $Pd_2(dba)_3$, SPhos, K_3PO_4, PhI, PhMe, 80 °C, 70%.

Scheme 2. C7 Functionalization of tryptophan using iridium catalyzed borylation chemistry

Borylated tryptophan **15** was treated with potassium iodide in the presence of CuI as the catalyst and 1,10-phenanthroline as the ligand[16] to avail 7-iodotryptophan **18**, which upon Stille coupling reaction with allyltributylstannane furnished the allyltryptophan derivative **17** in 60% yield (Scheme 2). Chen-Lam coupling of **15** failed to provide an electron enriched methoxytryptophan **19**. Trifluoromethylation of **15** using the Togni reagent produced a complex mixture, possibly because of the interference of the free indole-NH. However, Suzuki coupling reaction of **15** with iodobenzene in the presence of $Pd_2(dba)_3$ as the catalyst resulted in 7-phenyltryptophan **20** in 70% yield (Scheme 2).

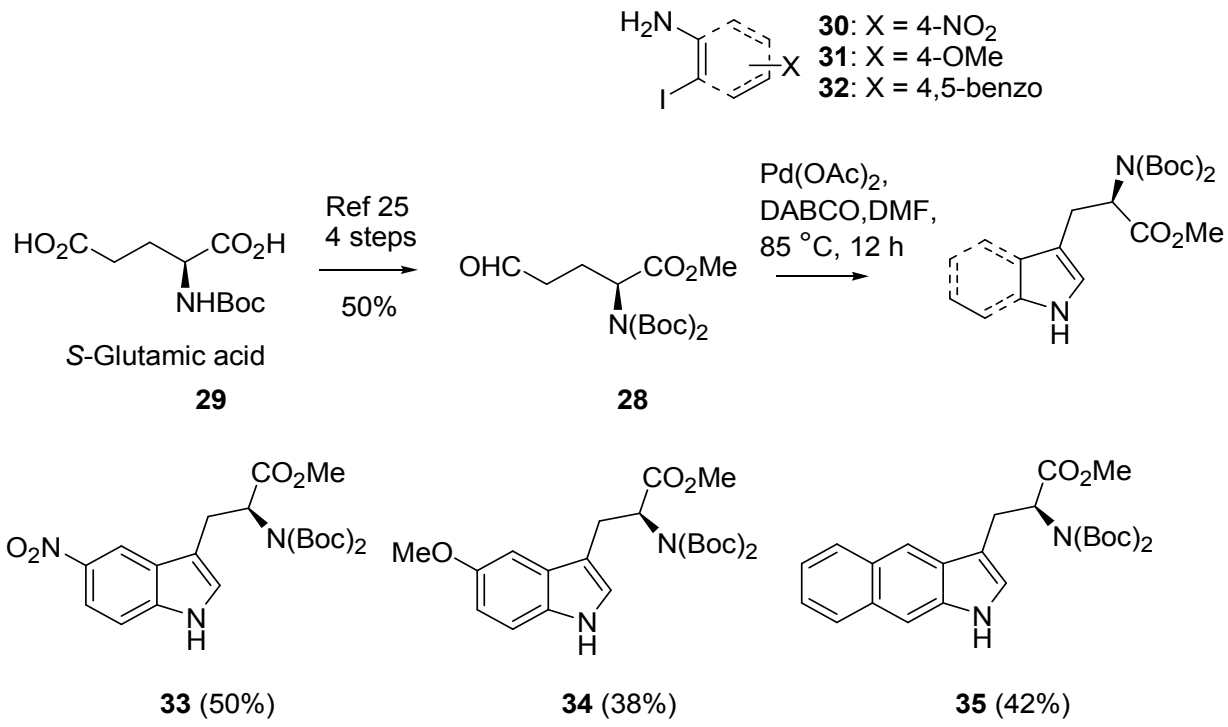


Reagents and conditions: a) Ir[(cod)OMe]$_2$ (5 mol%), phenanthroline (10 mol%), HBPin (0.25 equiv), B_2Pin_2 (4.0 equiv), hexane, 80 °C, 70%; b) NaN_3, $Cu(OAc)_2 \cdot H_2O$, MeOH, 70%; c) Phenyl acetylene, CuI, DIPA, HOAc, DCM, 72%; d) 4 N aq HCl, EtOAc, rt, 4 h; TBAF, THF, rt, 2 h, 85% (2 Steps); e) CuTC, Togni reagent, 1,10-Phen, DCM, LiOH, H_2O, rt, 60%; f) $Cu(OTf)_2$, KF; g) AgOTf, NaOH, Selectflour.

Scheme 3. C6 Functionalization of tryptophan using iridium catalyzed borylation chemistry

Functionalization of C6 of tryptophan was achieved by following a recently developed borylation of tryptophan **21** by Baran et al.,[17] which provided an inseparable 4:1 mixture of C6and C5 borylated Trp **22**[17] (Scheme 3). Treatment of the borylated tryptophan **22** with sodium azide in the presence of Cu(II) acetate as the catalyst produced 6-azido tryptophan **24** in good yield, which under Cu(I) catalyzed click

chemistry conditions provided the triazolyl-phenyltryptophan **25** in 72% yield. After removal[17] of the *tert*-butyl carbamate and triisopropyl silyl protection, the triazolyl-phenyltryptophan free amine **23** was obtained in excellent yield. Trifluoromethylation of **22** using the Togni reagent successfully yielded an

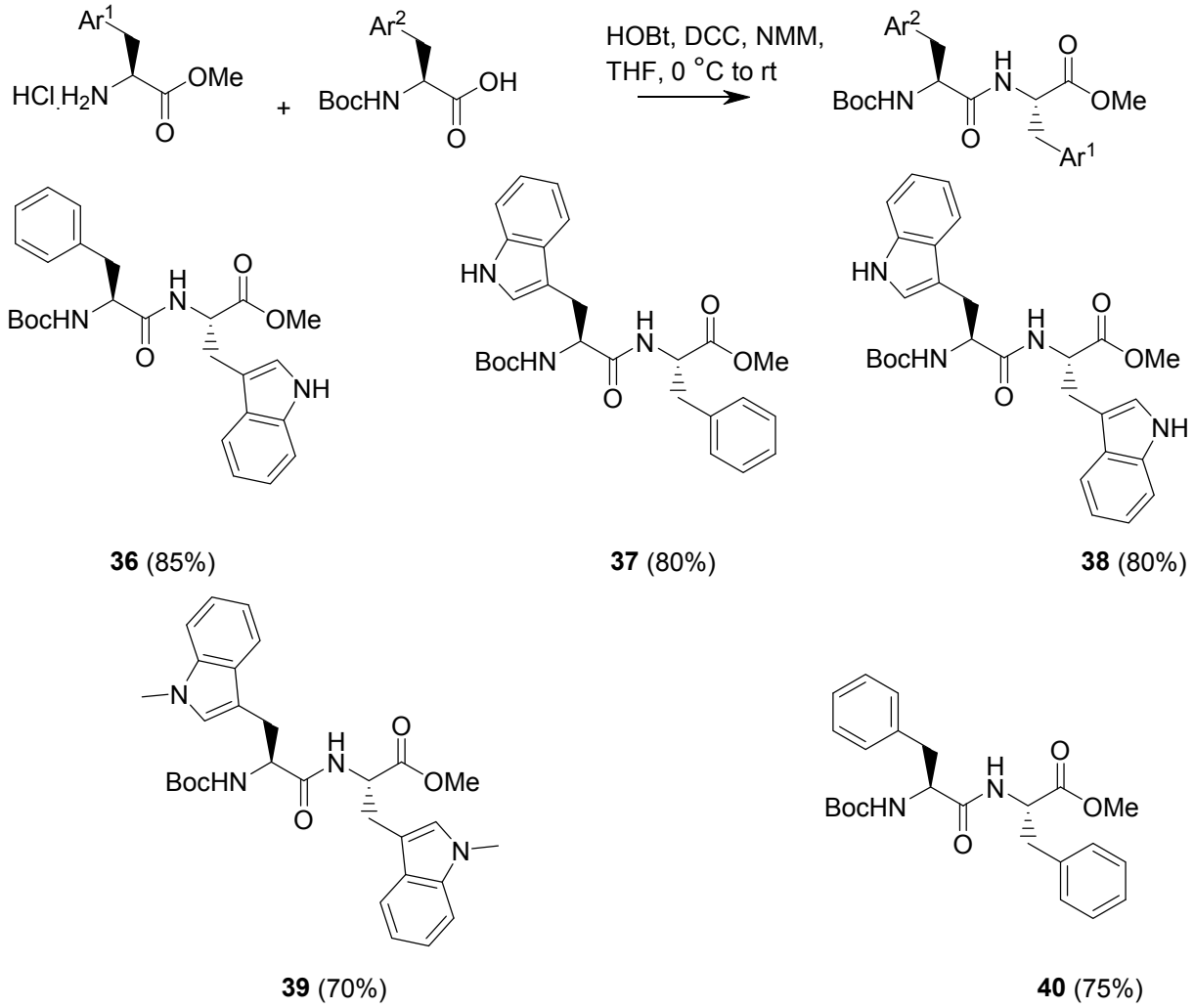


Scheme 4. Palladium catalyzed tryptophan synthesis

inseparable mixture of C5 and C6 trifluoromethlytated tryptophan **26** in 60% combined yield, complementing the requirement of the indole-NH protection. However, fluorination of **22** remained inaccessible under both nucleophilic and electrophilic conditions, owing to the labile nature of the triisopropyl silyl group (Scheme 3).

The palladium catalyzed tryptophan synthesis[18] methodology was utilized to access electron rich and electron deficient tryptophan residues. The required aldehyde **28**[19] was prepared from *S*-glutamic acid in 4 steps, subjected to the palladium catalyzed intramolecular cross-coupling reaction with three 2-iodoaniline derivatives(**30** to **32**) in the presence of palladium acetate as the catalyst and DABCO as the

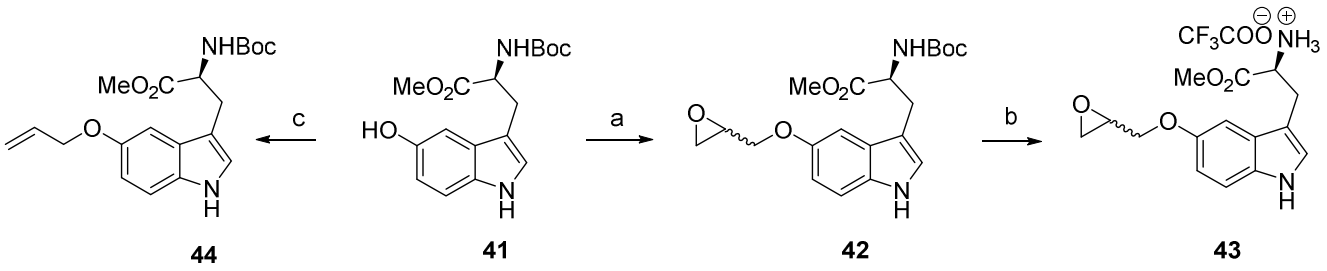
base. The reaction yielded 4-nitrotryptophan **33**[18] (50%), 4-methoxytryptophan **34**[18] (38%) and benzotryptophan **35**[18] (42%). NMR data of all the tryptophan derivatives (**33** - **35**) are in good agreement with previously reported data (Scheme 4).[18]



Scheme 5. Synthesis of select dipeptides by HOBt-DCC coupling

Five dipeptides (**36 - 40**)[20-22] were prepared following an FBDD approach to combine the two potent Phe and Trp structural elements. Using DCC-HOBt coupling reaction of corresponding NBoc protected amino acids and methyl ester hydrochlorides in the presence of *N*-methylmorpholine (NMM) as the base, the desired dipeptides were obtained in good yields (Scheme 5).

Treatment of **41**[23] with (±)-epibromohydrin in the presence of cesium carbonate as base afforded the 1:1 diastereomeric mixture of epoxy-tethered tryptophan **42**. Deprotection of **42** by TFA-DCM afforded the trifluoroacetate salt **43** in 80% yield. Similarly, reaction with allyl bromide yielded 5-*O*-allyl tryptophan derivative **44** in 82% yield. Both the *O*-allyl group and the epoxy residue could serve as reactive handles for further derivatization (Scheme 6).

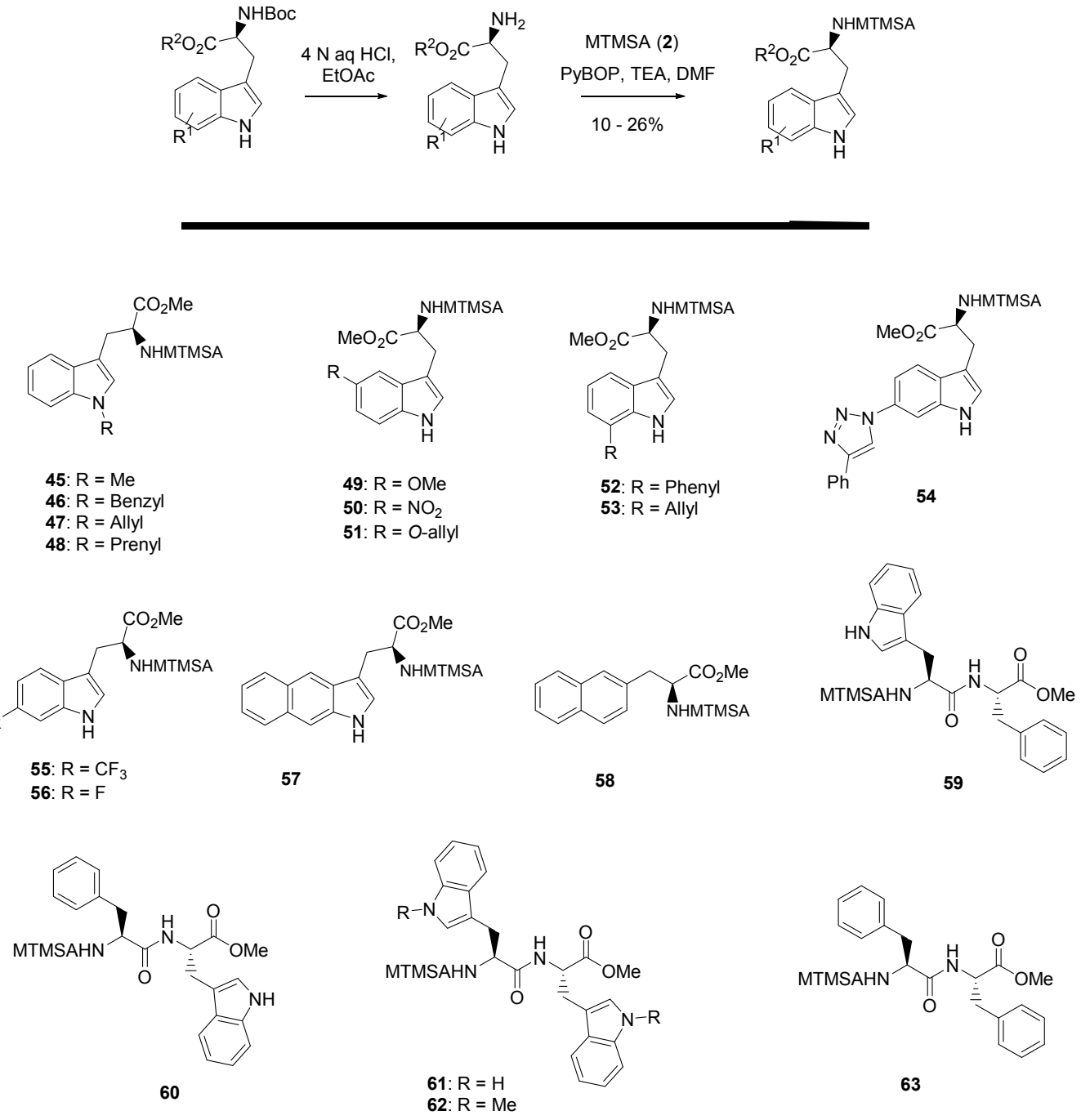


Reagents and conditions: a) Cs_2CO_3, DMF, (±)-Epibromohydrin, 80 °C, 12 h, 68%; b) TFA, DCM, rt, 6 h, 80%; c) Cs_2CO_3, DMF, Allyl bromide, 80 °C, 12 h, 82%.

Scheme 6. Alkylation of 5-hydroxy tryptophan

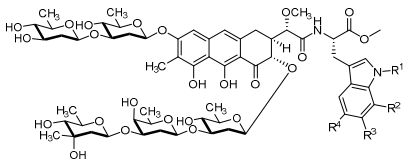
After deprotection of the *tert*-butyl carbamate of tryptophan derivatives by 4 N aq HCl in ethyl acetate, all of the free amines of the tryptophan derivatives were coupled with MTMSA (**2**) using PyBop as reagent and DIPEA as the base.[10] In each occasion, the corresponding MTMSA (**2**) analogues (**45** - **63**) were obtained in 10 - 26% yields (Scheme 7, vide SI for HPLC profile, HRMS data and individual yields of the reaction). The reaction of **43** with MTMSA (**2**) produced an inseparable mixture of diastereomeric MTMSA (**2**) coupled 1,2-diols, due to the opening of the epoxide ring under the reaction conditions.

To evaluate the anti-proliferative properties and selectivity of MTMSA (**2**) derivatives towards aberrant ETS transcription factors, such as EWS-FLI1, the following screening was performed: In the initial screen, MTM (**1**) analogues were tested for 72 h growth inhibition (GI_{50}) in TC-32 cells, a commonly



Scheme 7. Synthesis of MTMSA (**2**) analogues

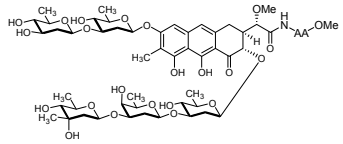
Table 1. Initial cytotoxicity (GI_{50}) screen in TC-32 (Ewing sarcoma) and PC-3 (non-Ewing sarcoma) cell line of MTMSA-Trp (3) analogues

		TC-32 Ewing sarcoma		PC-3 Prostate cancer		GI_{50} ratio PC-3 : TC-32
Entry	MTMSA (**2**) analogues	EWS-FLI1 Type 1		No ETS Translocation		
		GI_{50} (nM)	CI (95%)	GI_{50} (nM)	CI (95%)	
1	**45**, R^1 = Me, R^2 = R^3 = R^4 = H	139	113 - 174	1557	1165 - 2252	11.2
2	**46**, R^1 = Bn, R^2 = R^3 = R^4 = H	2297	1011 - 5884	> 10000	NE*	NE*
3	**47**, R^1= Allyl, R^2 = R^3 = R^4 = H	2715	1509 - 5216	3139	2553 - 9763	1.2
4	**48a**, R^1 = Prenyl, R^2 = R^3 = R^4 = H	> 10000	NE*	> 10000	NE*	NE*
5	**48b**, R^1 = Prenyl, R^2 = R^3 = R^4 = H	> 10000	NE*	> 10000	NE*	NE*
6	**52**, R^2 =Phenyl, R^1 = R^3 = R^4 = H	2402	2160 - 2652	> 10000	NE*	NE*
7	**53**, R^2 =Allyl, R^1 = R^3 = R^4 = H	794	535 - 1201	2989	2662 - 3687	3.8
8	**54**, R^3 = Trizolyl, R^1 = R^2 = R^4 = H	1030	441 - 2496	5878	> 2652	5.7
9	**55**, R^3 = CF_3, R^1 = R^2 = R^4 = H	2339	1958 - 3017	2395	> 2046	1.0
10	**56**, R^3 = F, R^1 = R^2 = R^4 = H	27	25 - 30	187	145 - 242	6.9
11	**49**, R^4 = OMe, R^1 = R^2 = R^3 = H	675	577 - 790	617	571 - 667	0.9
12	**50**, R^4 = NO_2, R^1 = R^2 = R^3 = H	353	313 - 408	646	174 - 2051	1.8
13	**51**, R^4 = *O*-Allyl, R^1 = R^2 = R^3 = H	53	37 - 75	454	275 - 763	8.6
14	**57**	3505	2974-5602	> 10000	NE*	NE*
15	**58**	541	353 – 834	1605	1106 - 2714	3.0
16	MTMSA-Trp (**3**)	16	13 - 20	76	66 - 88	4.8
17	MTMSA-Phe (**4**)	32	25 - 42	910	524 - 1606	28.4
18	MTM (**1**)	32	26 - 38	83	62 - 112	2.6
19	MTMSA (**2**)	> 10000	NE*	> 10000	NE*	NE*

*Regression not estimable

used Ewing sarcoma cell line expressing EWS-FLI1 sensitive to MTM (**1**). Analogues with a GI_{50} < 250 nM in the TC-32 cell line were then further tested against PC-3 cells, a prostate cancer cell line lacking EWS-FLI1 expression, also for 72 h growth inhibition (Table 1 and 2). In this assay, MTM (**1**) displayed 2.6-fold lower GI_{50} in TC-32 cells as compared to PC-3 cells (Table 1). Thus, analogues with >3-fold selectivity towards TC-32 cell line were considered selective and were further investigated in a broader panel of cell lines to confirm and validate the selectivity in the context of multiple ETS fusion or aberrant ETS expression and across multiple cancer types that do not depend on the ETS fusions. In this secondary screen, analogues were tested in panel of seven additional Ewing sarcoma cell lines that express the majority of the EWS-ETS fusions and on the only available prostate cancer cell line (VCaP)

Table 2. Initial cytotoxicity (GI_{50}) screen against TC-32 (Ewing sarcoma) and PC-3 (non- Ewing sarcoma) cell line of MTMSA (2)-dipeptide analogues

		TC-32 Ewing sarcoma EWS-FLI1 Type 1		PC-3 Prostate cancer No ETS Translocation		GI_{50} ratio PC-3 : TC-32
Entry	Analogues	GI_{50} (nM)	CI (95%)	GI_{50} (nM)	CI (95%)	
1	**59**, AA = Trp-Phe	37	26 - 55	75	51 - 108	2.0
2	**60**, AA = Phe-Trp	47	39 - 56	1128	578 - 2225	24.0
3	**61**, AA = Trp-Trp	41	25 - 68	568	376 - 862	13.9
4	**62**, AA = NMeTrpNMeTrp	7834	NE*	> 10000	NE*	NE*
5	**63**, AA = Phe-Phe	232	121 - 448	1132	504 - 2586	4.9
6	MTMSA-Trp (**3**)	16	13 - 20	76	66 - 88	4.8
7	MTMSA-Phe (**4**)	32	25 - 42	910	524 - 1606	28.4
8	MTM (**1**)	32	26 - 38	83	62 - 112	2.6

Regression not estimable

that expresses aberrant ERG. Results were compared to a panel of an additional 8 cancer cell lines that lack aberrant expression of ETS fusions (Table 3, Figure 2) to identify the most desired analogue. MTM (**1**), MTMSA-Trp (**3**) and MTMSA-Phe (**4**) were used as controls.

Analogue **45** was designed to evaluate the potential role of the indole-NH. *N*-methylation will cut off the potential hydrogen-bonding donation and will increase the hydrophobicity of the indole ring. The *N*-methyl analogue **45** was found to be less active than MTMSA-Trp (**3**) against the TC-32 and PC-3 cell lines (Table 1; Entry 1), but more selective towards TC-32. The increasing hydrophobicity of **45** probably makes it more target-specific than MTMSA-Trp (**3**). Inspired by these results, we synthesized *N*-benzylated analogue **46** to evaluate the effect of an additional aromatic ring. Surprisingly, *N*-benzylated analogue **46** completely lost its activity, against both the TC-32 and the PC-3 cell lines, probably due to steric hindrance at the binding site (Table 1;

Table 3. Ewing sarcoma selectivity index of select MTMSA (2) analogues determined by median cytotoxicity (GI_{50}) in Ewing sarcoma cell lines compared to non-Ewing sarcoma cell lines

Entry	MTMSA (**2**) Analogues	Median GI_{50} in Ewing sarcoma cell lines (nM)	Median GI_{50} in non-Ewing sarcoma cell lines (nM)	Ratio of Median GI_{50} non-Ewing sarcoma : Ewing sarcoma (Selectivity Index)	Selectivity Index ratio Analogues : MTM (1)
1	**60**	52	991	19.1	12.7
2	**61**	55	856	15.6	10.4
3	**59**	64	152	2.4	1.6
4	**45**	684	1644	2.4	1.6
5	**51**	485	532	1.1	0.7
6	**56**	466	419	0.9	0.6
7	**63**	561	1956	3.5	2.3
8	MTM (**1**)	46	71	1.5	1
9	MTMSA-Phe (**4**)	117	545	4.7	3.1
10	MTMSA-Trp (**3**)	47	109	2.3	1.5

Entry 2). Further modification by *N*-allylation (cf **47**) and *N*-prenylation (cf **48a** - **b**) to achieve a secondary interaction with the transcription factor also failed, resulting in complete loss of activity against both cell lines (Table 1; Entry 3, 4, and 5).

Analogues **52** and **53** were synthesized to understand the steric requirement of the C7 position of the tryptophan ring. Both, the C7 phenyl analogue **52** and C7 allyl analogue **53** lost their activity, suggesting that the C7 position has to remain non-substituted to avoid steric hindrance (Table 1; Entry 6, 7). To explore a click chemistry handle and to add a more distant (from the indole) phenyl ring, we designed the C6 functionalized analogue **54**, but it was found to be inactive (Table 1; Entry 8), which discouraged us to further follow this approach for FBDD expansion.

It is well documented that incorporation of CF_3 or F in a drug molecule can strongly affect the binding affinity, pharmacokinetic properties, and bioavailability, mostly by inserting strong H-bond acceptor sites. The presence of a CF_3 or F substituent in a drug also increases the hydrophilicity and the electronic environment of the molecule, and significantly slows down the oxidative metabolism of the molecule without altering its size drastically. We found that analogue **55** with C-6-CF_3 lost its activity completely in both cell lines (Table 1; Entry 9). However, the sterically less demanding C-6-F analogue **56** was found to be active in TC-32 cells with a selectivity of 6.9 and thus stood out as one of the potential candidates to be further studied and improved (Table 1; Entry 10). The increase in hydrophilicity while maintaining similar size as the unsubstituted tryptophan (cf **3**) could be the reason behind its activity with improved selectivity, in comparison to **55**. Electron rich or deficient tryptophans, **49** and **50**, respectively, were both found to be less active than the parent MTMSA-Trp (**3**), suggesting the importance of the steric factors over the electronic nature of the tryptophan ring (Table 1; Entry 11 and 12). Incorporation of an allylic residue at indole-*N* (cf. **47**) and C7 of the tryptophan ring (cf. **53**) resulted in loss of activity in both the TC-32 and PC-3 cell lines, hence we synthesized analogue **51** with a 5-*O*-

allyl tryptophan residue to explore substitution at the 5-position. Analogue **51** was found to be active in the TC-32 cell line with a selectivity of 8.6 (Table 1; Entry 13). These results indicate the accessibility of the C5 over the C7 position to incorporate an additional residue and to further expand the Trp residue by FBDD, with the ultimate goal to avail a secondary interaction with EWS-FLI1. Analogue **57**, with an additional fused benzene ring added to tryptophan, documented a ~ 34 fold decrease in activity in the TC-32 cell line, once more restricting the steric requirements for interaction with EWS-FLI1, and thus wiped out the possibility of adding π-donation directly to the tryptophan ring (Table 1; Entry 14). Similarly, analogue **58**, where the indole ring was replaced by a naphthalene ring, lost activity in the TC-32 cell line, which again showed the importance of the indole ring for interactions with EWS-FLI1 (Table 1; Entry 15).

Dipeptide analogues **59 - 63** were tested in TC-32 cell lines to validate the FBDD concept in the most straightforward way, combining Phe and Trp. The fact that the *N*-methyl tryptophan analogue **45** showed a ~ 4 fold increase in selectivity (Table 1; Entry 1) impelled us to include Trp-*N*-methylation into this series of analogues. Analogue MTMSA-Trp-Phe (**59**) was found to have similar activity as MTMSA-Trp (**3**), but was less selective, while MTMSA-Phe-Trp (**60**), with the opposite arrangement of the amino acid residues, showed significantly increased selectivity compared to **3** (Table 2; Entry 1 and 2). Likewise, analogue MTMSA-Trp-Trp (**61**), which contains two consecutive tryptophan moieties, was found to have much better selectivity than MTMSA-Trp (**3**) (Table 2; Entry 3). Both **60** and **61** are about equally active in the TC-32 cell line (as **3**). This shows the importance of the Trp residue to be in a more distant position from the DNA-interacting MTM (**1**) core, to interact with the transcription factor and thus may account for the observed improved selectivity towards the EWS-FLI1 expressing TC-32 cell line over the PC-3 cell line. It is noteworthy that the di-*N*-methyl analogue of **61** (cf **62**) lost its activity completely, which indicates the importance of the beta indole-NH for the secondary interaction with EWS-FLI1 (Table 2; Entry 4). The loss of activity against TC-32 of analogue **63** indicates that a

second phenylalanine residue is not advantageous (Table 2; Entry 5). These studies concluded that a tryptophan moiety in the more distant, second position of the analogues is crucial, and increases drastically the selectivity towards EWS-FLI1 while maintaining reasonable cytotoxic activity.

Overall, the above described initial screen identified 5 novel MTMSA (**2**) analogues with cytotoxicity (GI_{50}) in TC-32 cells of less than 250 nM, and selectivity against TC-32 cells greater than 3 times of that of PC-3, determined by PC-3 : TC-32 GI_{50} ratio (Table 1 and 2, highlighted in red). To deepen these findings and to further investigate the most promising analogues, we expanded the array of cancer cell lines to include a panel of 8 Ewing sarcoma cell lines (expressing aberrant ETS transcription factors) versus 9 non-Ewing sarcoma cell lines (lacking aberrant ETS transcription factors) (Figure 2, vide SI).

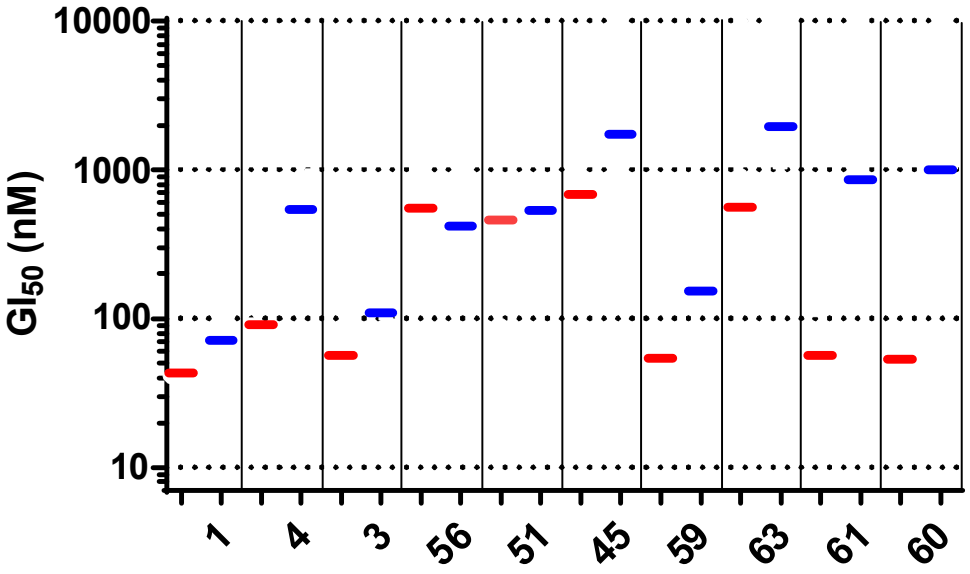


Figure 2. Median GI_{50} of select MTMSA (**2**) analogues in a panel of 8 Ewing sarcoma (TC-32, 5838, RD-ES, TC-71, A-673, ES-2, ES-7, ES-8) as well as 9 non-Ewing sarcoma (PC-3, DU 145, A549, LNCaP, U-118 MG, HeLa, HCT116, DMS 114, PANC-1) cell lines.

Despite of initial lack in selectivity and cytotoxicity, we included analogues **59** and **63**, respectively, to get a complete overview of dipeptide analogues. MTM (**1**), MTMSA-Trp (**3**) and MTMSA-Phe (**4**) were used as controls. The final selectivity ratio was calculated by taking the median GI_{50} of the non-Ewing sarcoma panel over the Ewing sarcoma panel (Table 3, Figure 2). The median is presented as we were not able to estimate the GI_{50} (exceeded 10μM), which precludes the estimation of a mean value. Of the 7 novel MTMSA (**2**) analogues that were tested on this secondary screen, only 3 were found to maintain selectivity greater than 3 against the panel of Ewing sarcoma cell lines versus non-Ewing sarcoma cell lines, namely **60**, **61** and **63**. The resulting Ewing sarcoma selectivity index ranked as follows: **60** (19.1) > **61** (15.6) > > MTMSA-Phe (**4**) (4.7) > **63** (3.5) > **59** (2.4) = **45** (2.4) > MTMSA-Trp (**3**) (2.3) > MTM (**1**) (1.5) > **51** (1.1) > **56** (0.9) (Table 3, Figure 2).

Taking both Ewing sarcoma cytotoxicity (GI_{50}) and selectivity index into account, a potency order of MTMSA (**2**) analogues was determined to be **60** > **61** > > MTMSA-Trp (**3**) = MTMSA-Phe (**4**) > MTM (**1**). Several initially promising MTMSA (**2**) analogues (cf **45**, **51**, **56**, and **63**) were eliminated in the secondary screen when tested against the panel of Ewing sarcoma cell lines because of their poor cytotoxicity (GI_{50} > 250 nM). In contrast, analogue **59** was cytotoxic (median GI_{50} of 64), but eliminated due to poor selectivity (selectivity index < 3) (Table 3, Figure 2). Interestingly and somewhat surprisingly, MTMSA-Phe (**4**) had a three times better ETS selectivity index of 4.7 than MTMSA-Trp (**3**, selectivity index 1.5), which initially was the lead molecule in this study (Table 3, Figure 2). Based on these cytotoxicity screens, the overall best candidates for further development are analogues **60** and **61**, with a median ETS cytotoxicity of (GI_{50}) 52 nM and 55 nM, respectively, and a drastically improved selectivity of > 10-fold towards ETS depended cell lines, in comparison to MTM (**1**) (Table 3, Figure 2).

Previous modification of MTM (**1**) leading to the identification of EC-8105 with improved suppression of EWS-FLI1 by almost 10-fold, focused on the introduction of an allyl carbonate residue in the 3B position of the disaccharide residue of MTM (**1**) (Figure 1).[24] However, from our previous study of

DNA-MTMSA-Trp (**3**) and DNA-MTMSA-Phe (**4**) crystal structures, it was anticipated that the 3-side chain can interact with FLI1 DNA binding domain.[12] Therefore, the identification of MTMSA-Phe-Trp (**60**) and MTMSA-Trp-Trp (**61**) with improved selectivity of 19.1 and 15.6, respectively, cemented our initial hypothesis that adding an additional tryptophan residue to MTMSA-Phe (**4**) and MTMSA-Trp (**3**) at the 3-side chain position would enhance their interaction with EWS-FLI1.

Additionally, we tested all 7 MTMSA (**2**) analogues in VCaP cells, which overexpress the TMPRSS2-ERG gene fusion, a common genomic alteration harbored by prostate cancer cells. We initially expected that they would follow a similar selectivity trend to 5838 cells, which overexpress the more rare EWS-ERG gene alteration in Ewing sarcoma. Our results indeed show that analogues **60** and **61** are the most selective for VCaP cells over other prostate cancer cell lines, with a selectivity ratio of 25.4 and 18.3, respectively (Table 4), consistent with their Ewing sarcoma selectivity index (cf Table 3, Figure 2). Moreover, these analogues were more selective than MTM (**1**) for 5838 cells, however, not in the exact same rank order (Table 4).

Table 4. GI_{50} and selectivity index of select MTMSA (2) analogues in cancer cells expressing aberrant ERG transcription factors.

Entry	MTMSA (**2**) Analogues	VCaP Prostate Cancer TMPRSS2-ERG		Median GI_{50} in prostate cell lines, lacking TMPRSS2-ERG	Ratio of median GI_{50} Prostate cells lacking TMPRSS2-ERG : VCaP	5838 Ewing Sarcoma EWS-ERG		Ratio of median GI_{50} Non-Ewing sarcoma cell lines (cf Table 3) : 5838
		GI_{50} (nM)	CI (95%)			GI_{50} (nM)	CI (95%)	
1	**60**	39	34 - 45	991	25.4	149	121 - 186	6.7
2	**61**	31	25 - 37	568	18.3	53	40 - 70	16.2
3	**59**	21	12 - 34	75	3.6	39	25 - 60	3.9
4	**45**	307	87 - 1430	1557	5.1	121	98 - 151	13.3
5	**51**	485	371 - 645	454	0.9	24	20 - 29	22.2
6	**56**	319	235 - 457	187	0.6	466	229 - 929	0.9
7	**63**	442	377 - 513	1132	2.6	579	350 - 966	3.4
8	MTM (**1**)	41	32 - 52	48	1.2	43	30 - 61	1.7
9	MTMSA-Phe (**4**)	150	70 - 321	732	4.9	38	30 - 49	14.3
10	MTMSA-Trp (**3**)	60	45 - 80	76	1.3	6	4 - 10	18.2

Since it is recognized that MTM displaces Sp1 from DNA and likely affects the expression of Sp1 target genes (e.g., BCL-2), we performed qRT-PCR to determine the effect of MTM (**1**) and MTMSA analogue treatment on the expression of those genes as well as on EWS-FLI1 and its target gene NR0B1. The expression of EWS-FLI1 was reduced to approximately 70 - 75% after 6 h and then further reduced after 12 h treatment (Figure 3A). The NR0B1 promoter contains a microsatellite region of GGAA repeats of DNA that binds to EWS-FLI1 for transcriptional regulation, through the conserved ETS binding domain of FLI1.[25] NR0B1 mRNA expression was not affected by MTM (**1**) at the GI_{50} at 6 and 12 h, but was reduced to ~75% after 6 h of treatment with MTMSA analogues, and it was further reduced after 12

h (Figure 3B). In contrast both Sp1 and BCL-2 mRNA expression was reduced with MTM (**1**) and MTMSA analogues treatments after 6 and 12 h (Figure 3C and 3D). Previously it was reported that 6 h treatment with 100 nM MTM (**1**) had no effect on mRNA expression of CCK.[26] Therefore, we used this gene as a negative control. Our results also showed no change in mRNA expression of CCK after 6 h treatment at the GI_{50} (32 nM, ref. Table 1) for MTM (**1**), and minimal effect with MTMSA analogues. However, after 12 h, CCK mRNA expression is reduced to ~60% for both MTM (**1**) and MTMSA analogues (Figure 3E).

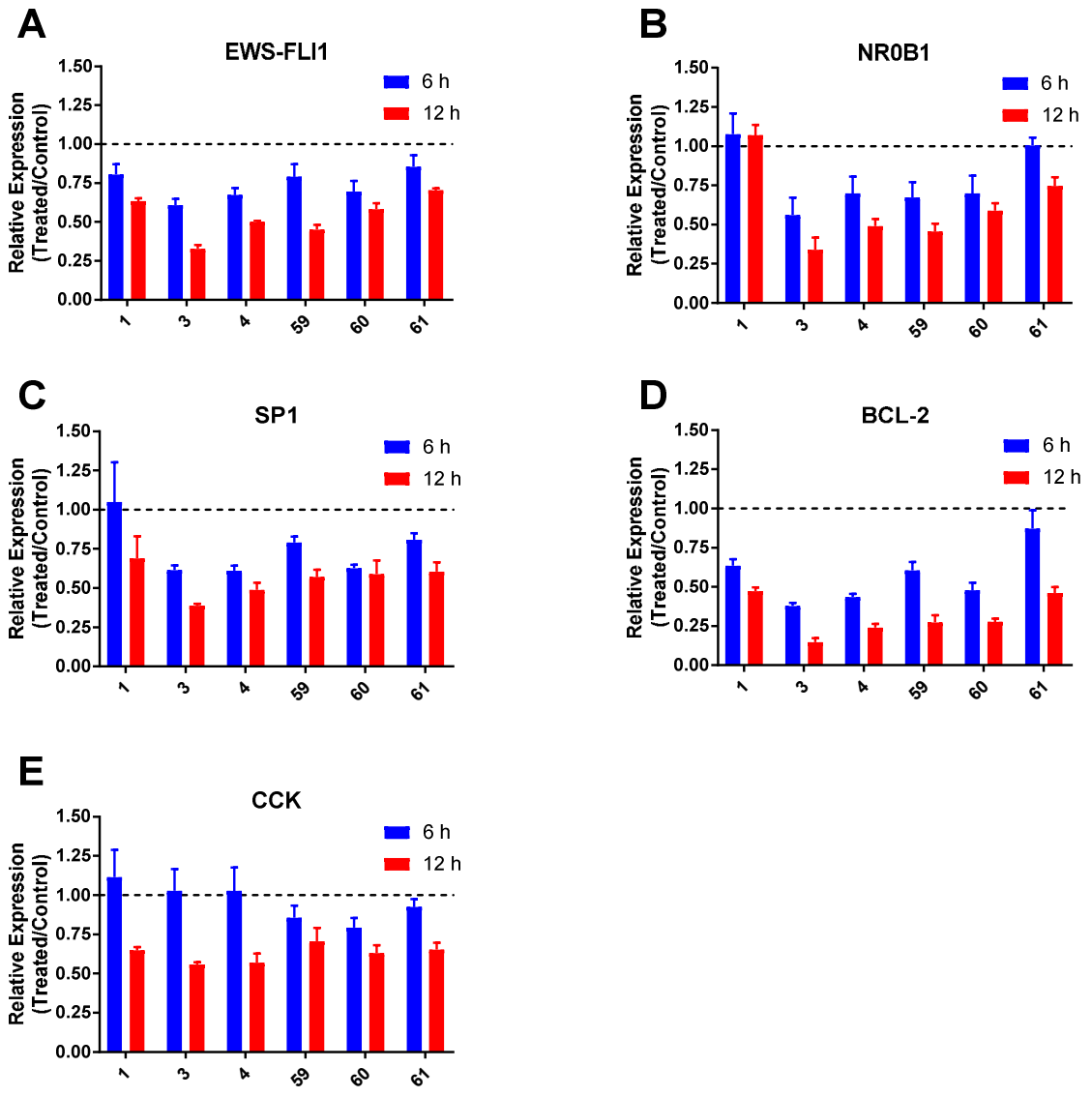


Figure 3. mRNA expression of (A) EWS-FLI1 and associated gene, (B) NR0B1, are decreased after 6 and 12 h treatments with MTMSA analogues at respective GI_{50}, analyzed by qRT-PCR. mRNA expres-

sion of (C) Sp1, a well-known downstream target of MTM (1), and associated gene, (D) BCL-2, were also analyzed. mRNA expression of (E) CCK was analyzed as a negative control, previously reported as unaffected after 6 h of treatment with MTM at 100 nM.[26] Relative expressions were calculated using GAPDH expression.

To further investigate the interference of MTM (**1**) and MTMSA analogues with the binding of ETS transcription factors, we developed cell lines that express luciferase under the control of EWS-FLI1. TC-32 cells, which express EWS-FLI1, were stably transfected with a luciferase reporter vector und the control of the full length NR0B1 promoter. Cells were treated for 12 h with multiple concentrations between 0 and 10 µM. All MTMSA analogues decrease luciferase expression to 50% (Figure 4A). Additionally, a luciferase reporter vector under the control of a CMV promoter was tested as a non-specific control. None of the analogues reduced CMV driven luciferase signal at the range of their GI_{50} concentrations. MTM (**1**) and analogue **59** decrease CMV driven luciferase expression down to ~50% and ~75%, respectively, at concentrations between 1 and 10 µM, while analogues **60** and **61** had no effect at all, even when treated up to 10 µM (Figure 4B).

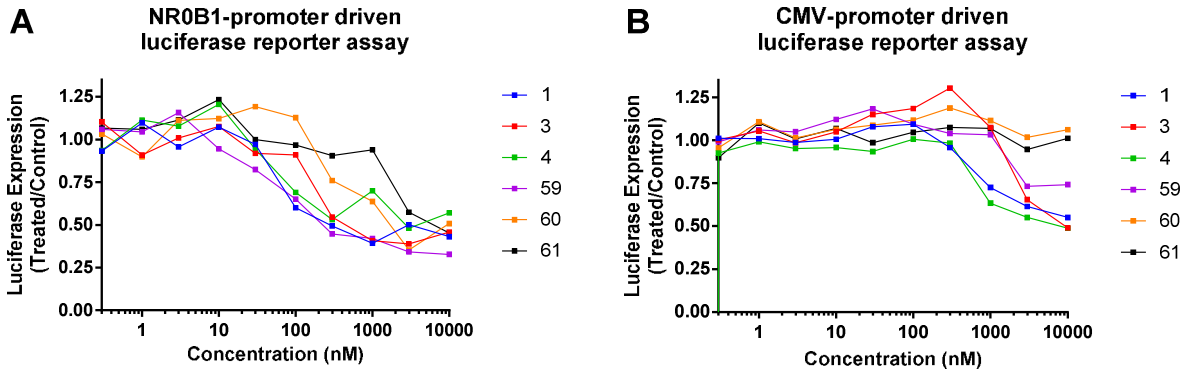


Figure 4. (A) MTMSA analogues decrease the expression of luciferase controlled by NR0B1 promoter, a validated binding promoter of EWS-FLI1. TC-32 cells, expressing EWS-FLI1, were stably transfected with luciferase reporter vectors and treated for 12 h. (B) MTMSA analogues **60** and **61** did not decrease luciferase expression controlled by a non-specific CMV promoter.

Lack of effect of analogue MTMSA-Phe-Trp (**60**) and MTMSA-Trp-Trp (**61**) on CMV promoter driven transcription while maintaining activities against EWS-FLI1 mediated transcription support the observed selectivity in our cytotoxicity assays (cf. Table 3, Figure 2). Furthermore, MTMSA-Trp-Phe (**59**), which reduces CMV promoter driven transcription, lacks selectivity in our cytotoxicity assays (cf. Table 3, Figure 2). This supports the conclusions made after the initial screening (Table 2), namely that a Trp residue in a second, more distant position from the MTM (**1**)-DNA binding core, is crucial for an interaction with the EWS-FLI1 transcription factor. The importance of the Phe or the Trp moiety in the first position of the MTMSA-3-side chain and the impact of an additional Trp residue (cf MTMSA (**2**) tripeptide analogues) remains to be investigated.

CONCLUSIONS

This study was aimed on refinement of two amino acid derivatives of MTMSA (**2**), namely, MTMSA-Trp (**3**) and MTMSA-Phe (**4**), which both showed promising activity and increased selectivity in a preliminary cytotoxicity assay looking at effects on cell lines overexpressing aberrant ETS transcription factors. This was expected, since crystallographic studies investigating DNA-FLI1 interactions and the DNA binding modalities of the MTMSA-Phe (**4**) and MTMSA-Trp (**3**) analogues resulted in a new mode-of-action hypothesis of MTM (**1**) and these derivatives: a ternary complex of MTM (**1**)-DNA-FLI1 (similarly MTM (**1**)-DNA-ERG in the context of prostate cancer), in which the MTM (**1**) analogue binds to the minor groove of certain DNA microsatellites, with its core and trisaccharide side chain, and simultaneously to the FLI1 portion of EWS-FLI1 within the major groove of DNA. Aromatic residues of MTMSA (**2**) analogues are well poised to increase protein binding. Initially, the focus was on refining the Trp residue, since MTMSA-Trp (**3**) showed better cytotoxicity than MTMSA-Phe (**4**) (Table 3). The array of MTMSA-Trp (**3**) derivatives was achieved through iridium-catalyzed borylation and palladium catalyzed tryptophan syntheses. The borylated tryptophans were further used to diversify the indol ring

of tryptophan, before the PyBop coupling reaction with MTMSA (**2**) to generate sterically and electronically different MTMSA-Trp (**3**) analogues.

Later, the studies were expanded to a fragment based drug development (FBDD) approach, to combine the protein interactive aromatic rings of Trp and Phe, which both appeared to be advantageous elements to increase selectivity towards ETS fusion expressing cell lines. These studies started with phenyl-expanded Trp derivatives, in which the indole of Trp was expanded by an aromatic ring, or through phenyl-Trp derivatives, and ended with dipeptide analogues of MTMSA (**2**), with all combinations of dipeptide side chains, namely MTMSA-Trp-Phe (**59**), MTMSA-Phe-Trp (**60**), MTMSA-Phe-Phe (**63**) and MTMSA-Trp-Trp (**61**).

Initially all analogues were tested in the Ewing sarcoma TC-32 cell line that expresses EWS-FLI1 type I, the genotype in the majority of Ewing sarcoma patients. For target specificity analysis, these analogues were also tested in prostate PC-3 cell line that lacks dependence on aberrant ETS transcription factors. These comparative tests resulted in the identification of 7 MTMSA analogues with increased specificity towards the Ewing sarcoma cell line (Table 1 and 2). As a result they were further evaluated against multiple Ewing sarcoma cell lines, as well as VCaP prostate cancer cell line, all of whom express various aberrant ETS transcription factors. This was compared to a panel of cell lines lacking expression of aberrant ETS transcription factors. Analogues **60** and **61** were found to have cytotoxic activities comparable to MTM (**1**), but showed > 10-fold increased selectivity towards aberrant ETS dependent cell lines (Table 3 and 4, Figure 2). To further investigate if this selectivity is a result of activity against aberrant ETS transcription factors we looked at the effects on EWS-FLI1 and target gene NR0B1 mRNA expression with analogue treatment. Furthermore, we tested a luciferase reporter vector controlled by NR0B1 promoter, which binds to EWS-FLI1. We found that analogues **60** and **61** decreased EWS-FLI1 and NR0B1 mRNA expression, as well as NR0B1 promoter driven luciferase signal (Figure 3A and 3B, Figure 4A). However, analogue **60** and **61** did not decrease non-specific CMV promoter driven luciferase signal (Figure 4B), correlating well with the observed selectivity in cytotoxicity

assays (Table 3, Figure 2). This is in comparison to MTM (**1**) and analogue **59**, which decreased non-specific CMV promoter driven luciferase signal (Figure 4B), and therefore had no observed selectivity in cytotoxicity assays (Table 3, Figure 2).

This expanded study confirmed that the tryptophan ring in second position of the 3-side chains of analogue **60** and **61** may play a significant role to afford better selectivity against ETS dependent cell lines.

EXPERIMENTAL SECTION

Chemistry

General

All commercial reagents were used without further purification. The required amine for the synthesis of **56** and **58** were prepared by protecting the commercially available corresponding tryptophan in the presence of thionyl chloride in methanol. Solvents were dried and distilled following the standard procedures. TLC was carried out on pre-coated plates (Merck silica gel 60, GF254), and the spots were visualized with UV, fluorescent light or by charring with phosphomolybdic acid hydrate (PMA). Column chromatography was performed on silica gel (230-400 mesh). ^{1}H and ^{13}C NMR spectra for the compounds were recorded with Varian 400 or 600 MHz spectrometers. ^{1}H and ^{13}C chemical shifts are reported in ppm downfield of tetramethylsilane and referenced to residual solvent peaks ($CHCl_3$; δ_H = 7.26 and δ_C = 77.23, d_4-MeOH; δ_H= 3.31 and δ_C = 49.1). Multiplicities are reported using the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad resonance, ap = apparent. The phrase ‘usual work up’ or ‘worked up in usual manner’ refers to washing of the organic phase with water (2 × 1/4 the volume of organic phase) and brine (1 × 1/4 the volume of organic phase), and drying (anhydrous Na_2SO_4), filtration, and concentration under reduced pressure. Yields referred to isolated yields after purification. Analytical LC/MS was performed on a Waters 2965 (Kinetex 5U EVO C18 100A , 250 × 4.6 mm, a linear gradient from A/B 75:25 to 30:70 (20 min), 30:70 to 0:100 (2 min), 0:100 (2 min,); 0:100 to 75:25 (2 min), 75:25 (4 min) (A = H_2O + 0.1% formic acid, B = MeCN + 0.1% formic acid), flow rate 0.5 mL/min) equipped with an Waters ZQ 2000 mass spectrometer and Waters

2996 photodiode array detector. The purity of all analogues used in the bioassays was determined by this method to be >95%. Mass spectra were taken on ABSciex QTOF mass spectrometer.

General Procedure A: *N*-alkylation of **6**

To a stirred solution of **6** (1 equiv) in dry DMF was added 60% NaH (1.5 equiv) in portion at 0 °C and stirred for 30 min after which respective halide (1.5 equiv) was added. The reaction mixture was stirred at rt for 12 h, cooled to 0 °C and quenched with methanol. It was diluted with ethyl acetate, worked up in usual manner, and subjected to column chromatography with silica gel using 20% ethyl acetate in hexane as eluent.

General Procedure for the deprotection of phthalimide:

To a stirred solution of phtalimide (cf. **7** to **10**) (1 equiv) in MeOH-DCM (1 : 1; 10 mL) was added hydrazine hydrate (1.5 equiv) and the reaction mixture was stirred at rt for 24 h. It was then filtered through celite, washed with ethyl acetate and concentrated under reduced vacuum. It was purified in reverse phase silica using 20% acetonitrile in water as mobile phase to obtain the free amine (cf. **11** to **14**) in 45 - 60% yield. The free amine was only characterized by LCMS and directly used without further purification in the PyBop coupling reaction with MTMSA (**2**).

General Procedure B: Procedure of synthesis of tryptophan from aldehyde **28**

A mixture of *o*-iodoaniline (1.1 equiv), aldehyde **28** (1.0 equiv), and DABCO (3 equiv) in dry DMF was degassed for 30 min using argon. $Pd(OAc)_2$ (5 mol%) was added to the reaction, and the resulting reaction mixture was heated at 85 °C in a pressure tube for 24 h. The reaction mixture was cooled to room temperature, diluted with water, extracted with ethyl acetate and worked up in usual manner. The crude product was purified by flash column chromatography to obtain the corresponding tryptophan derivative.

General Procedure C: Procedure for HOBt-DCC coupling reaction[20]

To a solution of NBocAA-OH (1 equiv) in dry THF were added HOBt (1.2 equiv) and DCC (1.2 equiv) and stirred at 0 °C for 1 h. Then HCl·AA-OMe (1.1 equiv) was added to the reaction mixture and

pH was adjusted in between 8 - 9 by adding excess NMM. The reaction mixture was stirred for 12 h at room temperature, filtered and concentrated under reduced pressure. The residue was dissolved in ethyl acetate, washed with saturated aq $NaHCO_3$ and worked up in usual manner. The crude product was purified by flash column chromatography to obtain the corresponding protected dipeptide.

General Procedure D: Procedure for PyBop coupling reaction

To a stirred solution of MTMSA (**2**) (1 equiv) in dry DMF were added PyBop (1.5 equiv), free amine (3.0 equiv) and excess triethyl amine (adjusted to pH 8). The reaction mixture was stirred at room temperature under argon atmosphere until disappearance of MTMSA (**2**). It was quenched by adding saturated NaCl solution and extracted with *n*-BuOH. The organic fraction was collected and concentrated under reduced pressure and purified by HPLC to obtain pure MTMSA (**2**) analogues.

Methyl (*S*)-3-(1-benzyl-1*H*-indol-3-yl)-2-(1,3-dioxoisoindolin-2-yl)propanoate (**8**)

Compound **8** (512 mg, 70%) was prepared as a yellow semisolid by *N*-benzylation of **6** (581 mg, 1.67 mmol) with benzyl bromide (0.3 ml, 2.5 mmol) using 60% NaH (100 mg, 2.5 mmol) as base, following the general procedure A. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.75 – 7.73 (m, 2H), 7.67 – 7.62 (m, 3H), 7.17 – 7.05 (m, 6H), 6.93 (s, 1H), 6.89 – 6.86 (m, 2H), 5.30 (dd, *J* = 9.9, 6.4 Hz, 1H), 5.16 (ABq, *J* = 14.8 Hz, 2H), 3.81 (s, 3H), 3.79 – 3.75 (m, 2H).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.8, 167.6, 137.6, 136.6, 134.1, 131.8, 128.7, 128.0, 127.5, 127.0, 126.6, 123.5, 122.1, 119.5, 118.9, 110.3, 109.8, 53.0, 52.8, 49.9, 24.9.; HRMS (TOF MS ES+) m/z calcd for $C_{27}H_{23}N_2O_4$ $[M+H]^+$ 439.1659, found 439.1638.

Methyl (*S*)-3-(1-allyl-1*H*-indol-3-yl)-2-(1,3-dioxoisoindolin-2-yl)propanoate (**9**)

Compound **9** (400 mg, 85%) was prepared as a yellow semisolid by *N*-allylation of **6** (425 mg, 1.22 mmol) with allyl bromide (0.16 ml, 1.83 mmol) using 60% NaH (75 mg, 1.83 mmol) as base, following the general procedure A. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 – 7.73 (m, 2H), 7.66 – 7.60 (m, 2H), 7.61 (d, *J* = 7.9 Hz, 1H), 7.19 (d, *J* = 8.4 Hz, 1H), 7.13 (t, *J* = 8.0, 1H), 7.05 (t, *J* = 8.4, 1H), 6.88 (s, 1H), 5.83 – 5.74 (m, 1H), 5.27 (t, *J* = 7.6 Hz, 1H), 4.97 – 4.93 (m, 1H), 4.78 – 4.73 (m, 1H), 4.62 – 4.50 (m, 2H), 3.80 (s, 3H), 3.75 (d, *J* = 8.1 Hz, 2H).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.8, 167.6, 136.4, 134.1,

133.5, 131.9, 127.9, 126.6, 123.5, 121.9, 119.3, 118.8, 116.7, 110.1, 109.7, 53.0, 52.8, 48.5, 24.9.; HRMS (TOF MS ES+) m/z calcd for $C_{23}H_{21}N_2O_4$ $[M+H]^+$ 389.1501, found 389.1488.

Methyl (*R*)-2-(*tert*-butoxycarbonyl)-3-(7-iodo-1*H*-indol-3-yl)propanoate (**18**)

To a solution of **15** (740 mg, 1.67 mmol) in methanol (15 mL) in a pressure tube, were added CuI (35 mg, 0.183 mmol, 10.0 mol%), 1,10-phen (60.0 mg, 0.34 mmol, 20.0 mol%) and KI (420 mg, 2.53 mmol, 1.50 equiv). The mixture was stirred at room temperature, water (3.5 mL) was added and it was sealed under air. The mixture was heated at 80 °C for 2 h, cooled to room temperature and diluted with water (40 ml). It was extracted with Et_2O (3 × 35 mL) and worked up in usual manner. The crude compound was subjected to column chromatography with silica gel using 30% ethyl acetate in hexane as eluent to obtain **18** as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.19 (s, 1H), 7.54 (apt, *J* = 8.4 Hz, 2H), 7.06 (s, 1H), 6.88 (t, *J* = 7.7 Hz, 1H), 5.08 (d, *J* = 8.2 Hz, 1H), 4.64 (dd, *J* = 12.8, 6.0 Hz, 1H), 3.67 (s, 3H), 3.30 – 3.20 (m, 2H), 1.42 (s, 9H).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.7, 155.3, 138.1, 131.0, 128.0, 123.3, 121.5, 119.1, 112.1, 80.1, 76.1, 54.3, 52.5, 28.52, 28.49.; HRMS (TOF MS ES+) m/z calcd for $C_{17}H_{22}IN_2O_4$ $[M+H]^+$ 445.0625, found 445.0620.

Methyl (*R*)-3-(7-allyl-1*H*-indol-3-yl)-2-(*tert*-butoxycarbonyl)propanoate (**17**)

To a stirred solution of **18** (115 mg, 0.29 mmol) in dry toluene (5 ml), were added allyl tributyltin (0.13 ml, 0.4 mmol) and tetrakis(triphenylphosphine)palladium(0) (70 mg, 0.06 mmol, 20 mol%). The reaction mixture was refluxed under argon atmosphere for 24 h, cooled to room temperature and diluted with ethyl acetate (50 ml). After usual work up followed by column chromatography with silica gel using 25% ethyl acetate in hexane as eluent obtained 17 (56 mg, 60%) as a yellow liquid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.17 (s, 1H), 7.44 (d, *J* = 7.8 Hz, 1H), 7.07 (t, *J* = 7.5 Hz, 1H), 7.03 – 6.97 (m, 2H), 6.11 – 6.01 (m, 1H), 5.28 – 5.13 (m, 2H), 5.07 (d, *J* = 8.2 Hz, 1H), 4.64 (q, *J* = 6.3 Hz, 1H), 3.68 (s, 3H), 3.62 (d, *J* = 6.5 Hz, 2H), 3.28 (dd, *J* = 5.6, 2.6 Hz, 2H), 1.42 (s, 9H).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.0, 155.4, 136.9, 135.7, 122.7, 122.5, 122.3, 120.1, 117.4, 116.6, 110.7, 80.0, 54.4, 52.4, 37.0, 28.5, 28.2.; HRMS (TOF MS ES+) m/z calcd for $C_{20}H_{27}N_2O_4$ $[M+H]^+$ 359.1971, found 359.1965.

Methyl (*R*)-2-(*tert*-butoxycarbonyl)-3-(7-phenyl-1*H*-indol-3-yl)propanoate (**20**)

To a stirred solution of **18** (160 mg, 0.36 mmol) in dry toluene (5 ml), were added iodobenzene (70 µL, 0.6 mmol), K_3PO_4 (180 mg, 0.84 mmol), Sphos (18 mg, 0.044 mmol, 12 mol%), and tris(dibenzylideneacetone)dipalladium(0) (20 mg, 0.02 mmol, 5 mol%). The reaction mixture was heated under argon atmosphere at 80 °C for 16 h, cooled to room temperature and diluted with ethyl acetate (50 ml). Work up in usual manner followed by column chromatography with silica gel using 20% ethyl acetate in hexane as eluent obtained 17 (100 mg, 70%) as a yellow liquid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.34 (s, 1H), 7.62 (d, *J* = 7.2 Hz, 2H), 7.56 – 7.48 (m, 3H), 7.39 (t, *J* = 7.2 Hz, 1H), 7.22 – 7.19 (m, 2H), 7.02 (s, 1H), 5.11 (d, *J* = 8.2 Hz, 1H), 4.66 (q, *J* = 6.0 Hz, 1H), 3.70 (s, 3H), 3.37 – 3.27 (m, 2H), 1.44 (s, 9H).;^{13}C NMR (100 MHz, $CDCl_3$) δ 172.9, 155.4, 139.2, 134.2, 129.3, 128.4, 128.3 127.7, 125.9, 123.2, 122.3, 120.4, 118.3, 110.9, 80.0, 54.4, 52.5, 28.5, 28.3.; HRMS (TOF MS ES+) m/z calcd for $C_{25}H_{26}N_2O_4$ $[M+H]^+$ 395.1971, found 395.1965.

Methyl (*R*)-3-(6-azido-1-(triisopropylsilyl)-1*H*-indol-3-yl)-2-(*tert*-butoxycarbonyl)propanoate (**24**, mixture of C6 and C5 isomers, major C6 isomer)

To a solution of **22** (244 mg, 0.41 mmol) in methanol (5 ml), were added sodium azide (40 mg, 0.61 mmol) and copper(II)acetate monohydrate (8.2 mg, 0.041 mmol, 10 mol%). The reaction mixture was heated at 55 °C under air for 12 h, diluted with ethyl acetate and worked up in usual manner. Purification by column chromatography with silica gel using 25% ethyl acetate in hexane as eluent obtained **24** (147 mg, 70%) as a yellow liquid as an inseparable mixture of two isomers. ^{1}H NMR (400 MHz, $CDCl_3$, Major C6-azido isomer) δ 7.49 (d, *J* = 8.5 Hz, 1H), 7.08 (d, *J* = 1.9 Hz, 1H), 7.00 (s, 1H), 6.86 (dd, *J* = 8.5, 1.9 Hz, 1H), 5.09 (d, *J* = 8.2 Hz, 1H), 4.64 (q, *J* = 5.6 Hz, 1H), 3.63 (s, 3H), 3.26- 3.22 (m, 2H), 1.68 - 161 (m, 3H), 1.43 (s, 9H), 1.13 (d, *J* = 6.0 Hz, 18H).; HRMS (TOF MS ES+) m/z calcd for $C_{26}H_{42}N_5O_4Si$ $[M+H]^+$ 516.3006, found 516.3001.

Methyl (2*R*)-2-(*tert*-butoxycarbonyl)-3-(6-(4-phenyl-1*H*-1,2,3-triazol-1-yl)-(triisopropylsilyl)-1*H*-indol-3-yl)propanoate (**25**, mixture of C6 and C5 isomers, major C6 isomer)

To a solution of **24** (150 mg, 0.3 mmol) in dry DCM (2 mL), were added phenyl acetylene (32 µL, 0.28 mmol), copper(I) iodide (1.2 mg, 0.02 equiv), DIPEA (0.04 equiv), and acetic acid (1 µL, 0.06 equiv). The reaction mixture was stirred at room temperature for 2 h, diluted with DCM (3 mL) and purified by column chromatography with silica gel using 30% ethyl acetate in hexane as eluent obtained **25** (130 mg, 72%) as a yellow liquid as an inseparable mixture of two isomers. ^{1}H NMR (400 MHz, $CDCl_3$, major C6 isomer) δ 8.20 (s, 1H), 8.01 (d, *J* = 1.8 Hz, 1H), 7.95 – 7.91 (m, 2H), 7.63 (d, *J* = 8.5 Hz, 1H), 7.47 – 7.39 (m, 3H), 7.37 – 7.30 (m, 1H), 7.14 (s, 1H), 5.13 (d, *J* = 8.4 Hz, 1H), 4.67 (q, *J* = 6.3 Hz, 1H), 3.63 (s, 3H), 3.29 (apt, *J* = 6.2 Hz, 2H), 1.73 – 166 (m, 3H), 1.43 (s, 9H), 1.14 (d, *J* = 7.5 Hz, 18H).; HRMS (TOF MS ES+) m/z calcd for $C_{34}H_{48}N_5O_4Si$ $[M+H]^+$ 618.3476, found 618.3447.

Methyl 2-*tert*-butoxycarbonylamino-3-(6-trifluoromethyl-1-triisopropylsilanyl-1*H*-indol-3-yl)propanoate (**26**)

A mixture of 22 (150 mg, 0.25 mmol), 1,10-Phen (9 mg, 0.05 mmol), $LiOH{\cdot}H_2O$ (21 mg, 0.5 mmol) CuTC (5 mg, 0.025 mmol) and 3,3-Dimethyl-1-(trifluoromethyl)-1,2-benziodoxol (Togni's reagent) (91 mg, 0.28 mmol) in DCM (2 mL) were stirred under argon atmosphere at rt for 12 h. The reaction mixture was diluted with DCM (10 mL), filtered through a cilite pad and concentrated under reduced pressure. The crude compound was purified by column chromatography with silica gel using 30% ethyl acetate in hexane as eluent obtained **26** (58 mg, 60%) as a yellow liquid as an inseparable mixture of two isomers. ^{1}H NMR (400 MHz, $CDCl_3$, Mixture of C5 and C6 isomers) δ 7.77 (d, *J* = 1.3 Hz), 7.70 (s), 7.59 (d, *J* = 8.4 Hz), 7.35 (dd, *J* = 14.3, 8.4 Hz), 7.27 (d, *J* = 10.8 Hz), 7.15 (s), 6.93 (s), 5.05 (t, *J* = 8.6 Hz), 4.63 (dq, *J* = 13.5, 5.9 Hz), 3.61 (s), 3.27 – 3.05 (m), 1.71 – 1.56 (m), 1.42 (s), 1.21 – 1.06 (m). ^{19}F NMR (376 MHz, $CDCl_3$, Mixture of C5 and C6 isomers) δ - 54.2, - 60.5. HRMS (TOF MS ES+) m/z calcd for $C_{18}H_{22}F_3N_2O_4$ $[M+H]^+$ 386.1453, found 386.1462.

Methyl 2-[2-*tert*-butoxycarbonylamino-3-(1-methyl-1*H*-indol-3-yl)-propionylamino]-3-(1-methyl-1*H*-indol-3-yl)propanoate (**39**)

Compound **39** (58 g, 70%) was prepared following the general procedure A by HOBt-DCC coupling of BocNH-NMeTrp-OH (50 mg, 0.15 mmol) with NH_2-NMeTrp-OMe (40 mg, 0.17 mmol) in the presence of HOBt (20 mg, 0.15 mmol) and DCC (35 mg, 0.17 mmol) in THF (5 ml) as a white solid. Mp: 128 °C.; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.66 (d, *J* = 7.9 Hz, 1H), 7.29 (d, *J* = 8.1 Hz, 1H), 7.27 – 7.23 (m, 1H), 7.23 – 7.18 (m, 1H), 7.14 (t, *J* = 7.0 Hz, 3H), 6.94 (s, 1H), 6.85 (t, *J* = 7.3 Hz, 1H), 6.38 – 6.26 (m, 2H), 5.09 (d, *J* = 8.1 Hz, 1H), 4.79 (q, *J* = 6.0 Hz, 1H), 4.45 (d, *J* = 7.5 Hz, 1H), 3.67 (s, 3H), 3.59 (s, 3H), 3.55 (s, 3H), 3.36 (q, *J* = 4.7 Hz, 1H), 3.15 (td, *J* = 13.8, 12.8, 6.3 Hz, 2H), 3.06 (dd, *J* = 14.7, 5.4 Hz, 1H), 1.38 (s, 9H).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.6, 171.1,155.3, 136.9, 136.7, 128.2, 128.0, 127.7, 127.5, 121.8, 121.7, 119.3, 119.1, 119.0, 118.3, 109.3, 108.9, 107.8, 52.8, 52.2, 32.6, 32.5, 28.2, 28.1, 27.5. HRMS (TOF MS ES+) m/z calcd for $C_{30}H_{37}N_4O_5$ $[M+H]^+$ 533.2764, found 533.2835.

Methyl 2-(2-*tert*-butoxycarbonylamino-3-phenyl-propionylamino)-3-phenylpropanoate (**40**)

Compound **40** (1.8 g, 75%)) was prepared following the general procedure A by HOBt-DCC coupling of BocNH-Phe-OH (2.0 g, 5.65 mmol) with HCl·NH_2-Phe-OMe (1.12, 6.21 mmol) in the presence of HOBt (900 mg, 6.67 mmol) and DCC (1.38 g, 6.67 mmol) in THF (60 ml) as a white solid. Mp: 117 °C.; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 – 7.26 (m, 1H), 7.24 (td, *J* = 3.8, 1.4 Hz, 2H), 7.21 (dt, *J* = 5.8, 1.6 Hz, 3H), 7.17 (d, *J* = 7.4 Hz, 2H), 6.95 (dd, *J* = 6.4, 2.3 Hz, 2H), 6.24 (br s, 1H), 4.91 (br s, 1H), 4.76 (d, *J* = 6.7 Hz, 1H), 4.31 (br s, 1H), 3.65 (s, 3H), 3.06 – 2.95 (m, 4H), 1.38 (s, 9H).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.3, 170.7, 155.2 136.5, 135.6, 129.3(3C), 129.2(2C), 128.6, 128.5 (2C), 127.1, 126.9, 53.2, 52.3 (2C), 37.9 (2C), 28.2(3C).; HRMS (TOF MS ES+) m/z calcd for $C_{24}H_{30}N_2NaO_5$ $[M+Na]^+$ 449.2052, found 449.2027.

Methyl 2-*tert*-butoxycarbonylamino-3-(5-oxiranylmethoxy-1*H*-indol-3-yl)propanoate (**42**)

To a stirred solution of **41** (200 mg, 0.60 mmol) in dry DMF (6 mL) were added cesium carbonate (390 mg, 1.2 mmol) and (±)-epibromohydrin (77 µL, 0.90 mmol). The reaction mixture was stirred at 60 °C for 12 h, cooled to rt and concentrated under reduced pressure. It was diluted with ethyl acetate and worked up in usual manner and subjected to column chromatography with silica gel using 30% ethyl

acetate in hexane as eluent to afford **42** (150 mg, 68%, dr = 1:1) as a yellow semi soild. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.40 (s, 1H), 7.20 (d, *J* = 8.7 Hz, 1H), 7.02 (d, *J* = 2.4 Hz, 1H), 6.94 (s, 1H), 6.86 (dd, *J* = 8.8, 2.4 Hz, 1H), 5.13 (dt, *J* = 6.7, 3.4 Hz, 1H), 4.61 (dt, *J* = 14.7, 5.9 Hz, 1H), 4.25 (dt, *J* = 11.0, 3.3 Hz, 1H), 3.99 (dt, *J* = 10.7, 5.2 Hz, 1H), 3.65 (s, 3H), 3.38 (dq, *J* = 7.1, 3.1 Hz, 1H), 3.21 (d, *J* = 5.7 Hz, 2H), 2.90 (t, *J* = 4.6 Hz, 1H), 2.77 (dd, *J* = 5.0, 2.7 Hz, 1H), 1.41 (s, 9H).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.0, 155.5, 153.1, 131.8, 128.1, 124.0, 113.0, 112.96, 112.2, 109.9, 102.3, 80.1, 70.0, 69.9, 54.4, 52.4, 50.6, 44.9, 28.5, 28.3. HRMS (TOF MS ES+) m/z calcd for $C_{20}H_{27}N_2O_6$ $[M+H]^+$ 391.1869, found 391.1874.

Methyl 3-(5-allyloxy-1*H*-indol-3-yl)-2-*tert*-butoxycarbonylaminopropanoate (**44**)

To a stirred solution of **41** (200 mg, 0.60 mmol) in dry DMF (6 mL) were added cesium carbonate (390 mg, 1.2 mmol) and allyl bromide (0.12 mL, 1.40 mmol). The reaction mixture was stirred at 60 °C for 12 h, cooled to rt and concentrated under reduced pressure. It was diluted with ethyl acetate and worked up in usual manner and subjected to column chromatography with silica gel using 30% ethyl acetate in hexane as eluent to afford **44** (182 mg, 82%) as a yellow semi soild. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.11 (s, 1H), 7.22 (d, *J* = 8.8 Hz, 1H), 7.03 (s, 1H), 6.96 (d, *J* = 2.5 Hz, 1H), 6.88 (dd, *J* = 8.8, 2.3 Hz, 1H), 6.11 (ddt, *J* = 16.3, 10.6, 5.3 Hz, 1H), 5.45 (dd, *J* = 17.3, 1.9 Hz, 1H), 5.36 – 5.23 (m, 1H), 5.10 (d, *J* = 8.3 Hz, 1H), 4.64 (d, *J* = 7.2 Hz, 1H), 4.58 (d, *J* = 5.4 Hz, 2H), 3.68 (s, 3H), 3.23 (d, *J* = 5.5 Hz, 2H), 1.42 (s, 9H).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.0, 155.5, 153.3, 134.1, 131.6, 128.2, 123.8, 117.6, 113.3, 112.1, 110.1, 102.3, 80.1, 70.0, 54.4, 52.5, 28.5, 28.3.; HRMS (TOF MS ES+) m/z calcd for $C_{20}H_{27}N_2O_5$ $[M+H]^+$ 375.1920, found 375.1916.

Analogue **60**

Analogue **60** (12 mg, 10%) was prepared as a yellow solid from MTMSA (**2**) (90 mg, 0.09 mmol) using PyBop (72 mg, 0.14 mmol), NH_2-Phe-Trp-OMe (100 mg, 0.27 mmol) and DMF (5 mL), following the general procedure D. Mp: 154 °C.; ^{1}H NMR (600 MHz, CD_3OD) δ 7.51 (d, J = 7.8 Hz, 1H), 7.37 – 7.25 (m, 5H), 7.14 – 7.10 (m, 2H), 7.07 -7.04 (m, 1H), 6.69 (t, *J* = 7.8 Hz, 1H), 6.41 (s, 1H), 6.26 (s,

1H), 5.07 – 5.00 (m, 2H), 4.84 – 4.79 (m, 2H), 4.77 – 4.67 (m, 2H), 4.40 (brs, 1H), 4.08 (dd, *J* = 8.7, 5.3 Hz, 1H), 3.99 (brs, 1H), 3.91 – 3.86 (m, 2H), 3.82 – 3.78 (m, 2H), 3.76 – 3.70 (m, 4H), 3.69 (s, 3H), 3.67 – 3.59 (m, 3H), 3.55 – 3.51 (m, 1H), 3.45 – 3.40 (m, 2H), 3.38 (s, 3H), 3.29 – 3.21 (m, 6H), 3.17 – 3.10 (m, 3H), 3.03 – 2.97 (m, 3H), 2.65 – 2.60 (m, 1H), 2.47 – 2.45 (m, 1H), 2.39 – 2.34 (m, 2H), 2.28 – 2.10 (5H), 2.02 – 1.80 (m, 6H), 1.66 – 1.62 (m, 4H), 1.44 (d, *J* = 6.1 Hz, 3H), 1.40 (d, *J* = 5.6 Hz, 3H), 1.37 (d, *J* = 6.1 Hz, 3H), 1.33 (d, *J* = 7.3 Hz, 3H), 1.30 (d, *J* = 6.2 Hz, 3H), 1.28 (s, 3H).; ^{13}C NMR (100 MHz, CD_3OD) δ 204.3, 174.3, 173.7, 173.5, 169.6, 164.7, 160.3, 156.7, 139.4, 139.0, 138.1, 137.9, 136.3, 135.5, 130.6, 130.5, 130.2, 129.7, 128.9, 128.7, 128.6, 127.8, 124.8, 124.5, 122.6, 122.5, 119.9, 119.9, 119.2, 119.1, 118.3, 112.4, 112.4, 111.8, 110.4, 109.0, 108.6, 102.0, 100.1, 99.9, 98.9, 98.1, 80.9, 80.8, 78.1, 77.9, 77.4, 76.6, 76.4, 73.7, 73.4, 72.1, 72.0, 71.9, 71.8, 70.5, 60.4, 55.5, 55.0, 54.9, 52.9, 49.9, 47.9, 45.2, 44.8, 40.8, 38.6, 38.2, 33.2, 28.6, 28.5, 27.3, 18.8, 18.8, 18.2, 17.1. HRMS (TOF MS ES-) m/z calcd for $C_{70}H_{90}N_3O_{25}$ $[M-H]^-$ 1372.5863, found 1372.5890.

Analogue **61**

Analogue **61** (17 mg, 14%) was prepared as a yellow solid from MTMSA (**2**) (90 mg, 0.09 mmol) using PyBop (72 mg, 0.14 mmol), NH_2-Trp-Trp-OMe (110 mg, 0.27 mmol) and DMF (5 mL), following the general procedure D. Mp: 165 °C.; ^{1}H NMR (600 MHz, CD_3OD) δ 7.72 (d, *J* = 7.9 Hz, 1H), 7.50 (d, *J* = 7.8 Hz, 1H), 7.46 (d, *J* = 8.2 Hz, 1H), 7.28 (d, *J* = 8.1 Hz, 1H), 7.23 (d, *J* = 7.1 Hz, 2H), 7.09 (d, *J* = 8.9 Hz, 3H), 7.01 (t, *J* = 7.7 Hz, 1H), 6.55 (s, 1H), 5.91 (s, 1H). 5.29 (brs, 1H), 5.06 (brs, 1H), 5.01 – 5.00 (m, 2H), 4.83 (t, *J* = 6.6 Hz, 1H), 4.79 (d, *J* = 9.8 Hz, 1H), 4.71 – 4.62 (m, 1H), 4.41 (d, *J* = 11.1 Hz, 1H), 4.00 (s, 1H), 3.90 – 3.80 (m, 2H), 3.76 – 3.69 (m, 2H), 3.68 (s, 3H), 3.63 -3.57 (m, 2H). 3.42 – 3.35 (m, 4H), 3.29 – 3.26 (m, 3H), 3.19 – 3.14 (m, 2H), 3.07 (t, *J* = 8.7 Hz, 1H), 3.02 – 2.93 (m, 2H), 2.57 – 2.47 (m, 1H), 2.34 – 2.13 (m, 5H), 1.97 – 1.93 (m, 2H), 1.90 – 1.85 (m, 1H), 1.80 (q, *J* = 11.3 Hz, 1H), 1.72 (d, *J* = 15.1 Hz, 1H), 1.64 – 1.53 (m, 3H), 1.46 (d, *J* = 6.1 Hz, 3H), 1.37 – 1.27 (m, 15H).; ^{13}C NMR (100 MHz, CD_3OD) δ 204.0, 174.2, 173.9, 173.7, 160.2, 139.6, 138.2, 137.9, 136.4, 128.9, 128.8, 125.0, 124.8, 122.7, 122.5, 120.1, 119.9, 119.8, 119.2, 112.6, 112.4, 111.5, 111.1, 110.4,

109.1, 101.8, 100.0, 100.0, 99.0, 97.9, 80.9, 78.2, 77.9, 77.6, 77.4, 76.6, 76.3, 73.7, 73.3, 72.1, 72.0, 71.9, 71.8, 70.4, 64.4, 60.4, 54.9, 54.5, 52.9, 49.9, 47.4, 47.4, 45.3, 40.7, 38.2, 38.1, 33.2, 28.5, 27.4, 27.3, 27.3, 18.8, 18.7, 18.2, 18.2, 17.1. HRMS (TOF MS ES-) m/z calcd for $C_{72}H_{91}N_4O_{25}$ $[M-H]^-$ 1411.5972, found 1411.5985.

Isolation of MTMSA (2) from *S. argillaceus* M7W1[9]

S. argillaceus M7W1 colonies were selected by multiple spore to spore passages over R5A agar plates supplied with 50 μg/ml apramycin. The visually darkest colony was cultured for 48 h (30 °C, 220 rpm) in TSB media supplemented with 50 μg/ml apramycin and subsequently used to inoculate a modified R5A media (100 g/L sucrose, 5 g/L glucose, 5 g/L soybean powder, 1 g/L yeast extract, 15 g/L MOPS, 5 g/L glycerol, 5 g/L $MgCl_2 \cdot 6H_2O$, 1 g/L $CaCO_3$, pH 7.5) for 8 days (30 °C, 240 rpm). The culture broth was centrifuged (3500 rpm, 30 min), the supernatant liquid was adjusted to pH 5.5 and extracted with *n*-BuOH (2 X equal volume). The butanol fraction was concentrated under reduced pressure and purified with silica gel chromatography (gradient: Chloroform : Methanol : Acetic acid = 15 : 1 : 0.1 to 10 : 1 : 0.1) to obtain 90% pure MTMSA (**2**), which was further purified by HPLC[15] to obtain pure MTMSA (**2**, yellow solid, 20 mg/L).

Cell culture media and materials

RPMI-1640 (Sigma, St. Louis, MO), DMEM (Sigma, St. Louis, MO), and F12K (Sigma, St. Louis, MO) media were prepared with 10% v/v heat-inactivated fetal bovine serum (FBS) (Atlanta Biologicals, Flowery Branch, GA) and 1% v/v 10,000 units/mL penicillin and 10,000 μg/mL streptomycin (PS) (Life Technologies, Carlsbad, CA). McCoy's 5A (Sigma, St. Louis, MO) medium was prepared with 15% v/v FBS and 1% v/v PS. All media were prepared with 1.5 – 2 g/L sodium bicarbonate (Sigma, St. Louis, MO), pH was adjusted to 7.2 and filtered with 0.2 μm filters (Corning, Corning, NY) prior to use. TC-32 (RPMI-1640), RD-ES (RPMI-1640), TC-71 (RPMI-1640), A-673 (DMEM), and 5838 (McCoy's 5A) cell lines (culture medium) were gifts from Dr. Timothy Cripe (Hospital Research Foundation, Columbus, OH). ES-8 (RPMI-1640), ES-2 (RPMI-1640), and ES-7 (RPMI-1640) were gifts from Dr. Peter

Houghton (Greehey Children's Cancer Research Institute, San Antonio, TX). VCaP (DMEM), PC-3 (RPMI-1640), DU 145 (RPMI-1640), PANC-1 (DMEM), U-118 MG (DMEM), HeLa (DMEM), A549 (F12K), and DMS 114 (RPMI-1640) cell lines were from ATCC (Manassas, VA). LNCaP (RPMI-1640) cell line was a gift from Dr. Vivek Rangnekar (University of Kentucky College of Medicine, Lexington, KY). HCT 116 (McCoy's 5A) cell line was a gift from the Genetic Resource Core Facility (Dr. Bert Vogelstein, John Hopkins School of Medicine, Baltimore, MD). All cell lines were grown at 37 °C under 5% CO_2 in a humid incubator and were tested regularly for mycoplasma using the MycoAlert mycoplasma detection kit (Lonza, Basel, Switzerland).

72 h growth inhibition (GI_{50}) assay

Cells were seeded in clear 96-well plates (VWR, Radnor, PA) at a cell density appropriate for exponential growth over 5 days. Following a 24 h attachment period, cells in duplicate wells were treated with half-log increments of respective compounds (0 nM and 0.3 nM - 10 μM). Working stocks were prepared from an initial 10 mM drug stock diluted in either 100% EtOH or DMSO. All wells contained a final concentration of 0.1% v/v respective organic solvent. Immediately following treatment, cell viability was measured for (Day 0) no-treatment control wells. Cell viability was measured in the remaining wells after 72 h of incubation with compound or vehicle control. For viability measurements, 0.1 mM resazurin (Sigma, St. Louis, MO) was added to wells and following 3 h of incubation at 37 °C, fluorescence readings (EM 560 nm, EX 590 nm) were recorded using a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA). Percent cell viability, relative to the initial seeding concentration, was calculated using the following formula:

$$Percent\ Viability\ (\%) = \frac{(Treatment_{Day\,3} - Vehicle\ Control_{Day\,0})}{(Vehicle\ Control_{Day\,3} - Vehicle\ Control_{Day\,0})} x\ 100$$

Percent cell viability was plotted against concentration (Log [M]) and regression software GraphPad Prism 7.0 (GraphPad Software, La Jolla, CA) was used to curve fit data and calculate a respective 72 h growth inhibition (GI_{50}).

In vitro selectivity screen using Ewing sarcoma GI_{50} model

The GI_{50} of MTMSA (**2**) analogues was estimated first in TC-32 cells. Analogues with GI_{50} values less than 250 nM in TC-32 cells were further tested in PC-3 cells. The ratio of GI_{50} values (PC-3 : TC-32) was then estimated and MTMSA (**2**) analogues with a ratio greater than 3 were selected for additional testing. A selectivity index for the compounds was then determined by taking the ratio of the median GI_{50} found in the non-Ewing sarcoma cell lines compared to the median GI_{50} found in ETS dependent cell lines.

Statistical analysis of GI_{50} results

The GI_{50} value was determined by pooling all available experiments and reported with 95% confidence interval. All compounds were tested at least once in multiwall replicates and each experiment included one or two control compounds (i.e., MTM (**1**) or MTMSA-Trp (**3**)) to ensure the stable response of the cell lines. The selectivity index was estimated by measuring the ratio of GI_{50} estimates in non-Ewing sarcoma/ETS dependent cell lines. Median values are reported since the GI_{50} was not estimable in, some cases.

Cloning of NR0B1 and CMV promoter-driven luciferase reporter vectors

Full-length NR0B1 promoter sequence was PCR amplified using Q5 high-fidelity DNA polymerase (New England Biolabs, Ipswich, MA) from genomic DNA of TC-32 cells and cloned into the pGLUC-Basic 2 vector (New England Biolabs, Ipswich, MA). XhoI and KpnI restriction sites were incorporated into the primers. After restriction enzyme (New England Biolabs, Ipswich, MA) digest and purification by gel electrophoresis, the target amplicon was ligated into the pGLUC-Basic 2 vector using T4 DNA ligase (Thermo Fisher Scientific, Waltham, MA). Chemically competent TOP10 *E.coli* cells were used to propagate vectors under ampicillin selection (100 μg/mL) (Thermo Fisher Scientific, Waltham, MA). The resulting vectors were extracted and purified using GeneJET plasmid prep kit (Thermo Fisher Scientific, Waltham, MA). The purified vector was transfected into TC-32 cells using lipofectamine 3000 (Thermo Fisher Scientific, Waltham, MA) with subsequent G418 selection (1 mg/mL) (VWR, Radnor,

PA). Separately, TC-32 cells were transfected with pCMV-Red Firefly Vector (Thermo Fisher Scientific, Waltham, MA) with subsequent puromycin selection (0.1 mg/mL) (Sigma, St. Louis, MO) as a control.

Luciferase reporter assay in stably transfected TC-32 cells

Selected TC-32 cells, expressing either NR0B1 or CMV luciferase reporter vectors, were seeded in clear 96-well plates at a density of 10,000 cells/well. Following a 24 h attachment period, cells were treated in duplicate with half-log increments of respective compounds (0 nM and 0.3 nM - 10 μM). After a 12 h treatment, media was removed and cells were washed 3 times with DPBS (Thermo Fisher Scientific, Waltham, MA). Cells were directly lysed on a plate shaker for 30 min at room temperature using 100 μL of passive membrane lysis solution. Lysates (80 μL) were transferred to a white luminescence plate. Luciferase substrate, either coelenterazine or D-luciferin (50 μL of 1X solution), for NR0B1 or CMV vectors, respectively, was added in a Glomax 96 microplate luminometer (Promega, Madison, WI) and luminescence was immediately measured. Delay before and after injections were set to the default of 0.4 s and a 10 s integration time was used. All reagents used in this assay were from the dual luciferase reporter assay system (Promega, Madison, WI). Concurrently, an additional 96 well plate was seeded and treated under the exact same conditions to determine cell viability using resazurin assay. Luminescence results were normalized to cell viability.

Relative mRNA expression (qRT-PCR)

TC-32 cells were seeded in 6-well plates at a density of 300,000 cells/well. When 80% confluent, ~72 h later, cells were washed with DPBS and treated with respective compounds (0 nM and GI_{50} nM). After 6 and 12 h treatments media was removed. Cells were washed with DPBS and 500 μL of 0.05% Trypsin with 0.53 mM EDTA (Corning, Corning, NY) was added. After a 2-3 min incubation, cells detached from the plate and 1 mL of fresh RPMI media was immediately added. Cells were collected in 1.5 mL tubes on ice and centrifuged to a pellet at 1200 x g for 5 min. Supernatant was removed and cell pellets were lysed with 600 μL of RLT lysis buffer (Qiagen, Hilden, Germany). Lysates were centrifuged through QIAshredder inserts for complete cell disruption (Qiagen, Hilden, Germany). RNeasy mini spin

columns were used to isolate pure RNA (Qiagen, Hilden, Germany). RNA concentrations were measured using NanoDrop 2000 Spectrometer (Thermo Fisher Scientific, Waltham, MA) and a 100 ng/μL stock solution was prepared. RNA (1 μg) was used to prepare cDNA with a MultiScribe reverse transcriptase (Thermo Fisher Scientific, Waltham, MA). qRT-PCR reactions were conducted using Maxima SYBR green Taq DNA Polymerase (Thermo Fisher Scientific, Waltham, MA). EpMotion5070 (Eppendorf, Hamburg, Germany) robot was used to mix reactions in 384 well plates and thermocycling was completed on a QuantStudio 7 Flex (Thermo Fisher Scientific, Waltham, MA). GAPDH served as the housekeeping gene for comparing relative expression of target genes. All primers (IDT, Coralville, IA) were verified to amplify a single amplicon of appropriate size by gel electrophoresis and melting curve analysis.

AUTHOR INFORMATION

Corresponding Author

*Jürgen Rohr; Email: jrohr2@ uky.edu

*Markos Leggas; Email: mark.leggas@uky.edu

† Contributed equally

Author Contributions

PM and JR designed all the molecules; PM and AM synthesized all the analogues; SMS and AJ optimized the media for the production of MTMSA (**2**); PM, AJ and AM contributed to the isolation of MTMSA (**2**). JME and ML designed the cytotoxicity and molecular assays; JME performed all the assays. The manuscript was written through contributions of all authors. †These authors contributed equally.

Notes

The authors declare the following competing financial interest(s): n/a

ACKNOWLEDGMENT

This work was supported by the National Institutes of Health grants CA 091901 and GM 105977 to JR as well as the US Department of Defense grant PC150300 to JR and ML. We acknowledge Markey Cancer Center and Dance Blue Foundation for their financial support. We also acknowledge Dr. Oleg V. Tsodikov for his valuable thoughts in designing analogues. We thank Dr. Madan Kharel and Dr. Khaled Shabaan of the UK Center for Pharmaceutical Research Innovation (CPRI) for their help to prepare larger amounts of mithramycin SA and UK College of Pharmacy for NMR facility.

ASSOCIATED CONTENT

The supporting information associated with this manuscript is available free of charge on the ACS Publications website at DOI XXXXXXXX.

HPLC profile and HRMS data of all the MTMSA analogues.

LCMS data of all the free amines.

Detailed cytotoxicity data of select MTMSA analogues.

Primers used for qRT-PCR.

Molecular formula strings with associated biochemical and biological data (CVS).

ABBREVIATIONS USED

FLI1, friend leukemia integration 1 transcription factor; ETS, E26 transformation-specific; ERG, ETS-related gene; TMPRSS2, transmembrane protease, serine 2; Sp1, specificity protein 1; BCL-2, B-cell lymphoma-2; NR0B1, nuclear receptor subfamily 0, group B, member 1; CCK, cholecystokinin; CMV, Cytomegalovirus; DMF, dimethylformamide; DCM, dichloromethane; DCC, *N*,*N*′-dicyclohexylcarbodiimide; HOBt, hydroxybenzotriazole; Boc, *tert*-butyloxycarbonyl; TFA, trifluoroacetic acid; qRT-PCR, quantitative reverse transcription polymerase chain reaction; GADPH, glyceraldehyde 3-phosphate dehydrogenase.

REFERENCES

1. Wohlert, S.; Künzel, E.; Machinek, R.; Mendez, C.; Salas, J.; Rohr, J. The structure of mithramycin reinvestigated. *J. Nat. Prod.* **1999**, *62*, 119-121.

2. Rohr, J.; Méndez, C.; Salas, J. A. The biosynthesis of aureolic acid group antibiotics. *Bioorg. Chem.* **1999**, *27*, 41-54.

3. Kofman, S., Perlia, C. P, Economou, S. G. Mithramycin in the treatment of metastatic ewing's sarcoma. *Cancer* **1973**, *31*, 889-893.; b) Balamuth, N., Womer, R. B.: Ewing's sarcoma, *Lancet Oncol.* **2010**, *11*, 184-192.

4. (a) Kofman, S.; Eisenstein, R. Mithramycin in the treatment of disseminated cancer. *Cancer Chemother. Rep.* **1963**, *32*, 77-96.; (b) Kofman, S.; Medrek, T. J.; Alexander, R. W. Mithramycin in the treatment of embryonal cancer. *Cancer* **1964**, *17*, 938-948.

5. Delattre, O.; Zucman, J.; Plougastel, B.; Desmaze, C.; Melot, T.; Peter, M.; Kovar, H.; Joubert, I.; de Jong, P.; Rouleau, G. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **1992**, *359*, 162.

6. May, W. A.; Arvand, A.; Thompson, A. D.; Braun, B. S.; Wright, M.; Denny, C. T. EWS/FLI1-induced manic fringe renders NIH 3T3 cells tumorigenic. *Nat. Genet.* **1997**, *17*, 495-497.

7. Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R. and Lee, C. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*, **2005**, *310,*644-648.

8. Sastry, M.; Patel, D. J. Solution structure of the mithramycin dimer-DNA complex. *Biochemistry* **1993**, *32*, 6588-6604.

9. Remsing, L. L.; González, A. M.; Nur-e-Alam, M.; Fernández-Lozano, M. J.; Braña, A. F.; Rix, U.; Oliveira, M. A.; Méndez, C.; Salas, J. A.; Rohr, J. Mithramycin SK, a novel antitumor drug with improved therapeutic index, mithramycin SA, and demycarosyl-mithramycin SK: three new products generated in the mithramycin producer streptomyces argillaceus through combinatorial biosynthesis. *J. Am. Chem. Soc.* **2003**, *125*, 5745-5753.

10. Scott, D.; Chen, J. M.; Bae, Y.; Rohr, J. Semi□synthetic mithramycin SA derivatives with improved anti-cancer activity. *Chem. Biol. Drug. Des*. **2013**, *81*, 615-624.

11. Leggas, M.; Eckenrode, J.; Mitra, P.; Jha, J.; Salem, S.; Mandal, A.; Thorson, J.; Rohr, J. [abstract]. In: Proceedings of the AACR-NCI-EORTC international conference: molecular targets and cancer therapeutics; 2017 Oct 26-30; philadelphia, PA. philadelphia (PA): AACR; *Mol Cancer Ther***.** **2018**, *17* (1 Suppl):Abstract nr B043.

12. Hou, C.; Weidenbach, S.; Cano, K. E.; Wang, Z.; Mitra, P.; Ivanov, D. N.; Rohr, J.; Tsodikov, O. V. Structures of mithramycin analogues bound to DNA and implications for targeting transcription factor FLI1. *Nucleic Acids Res*. **2016**, *44*, 8990-9004.

13. Alqahtani, N.; Porwal, S. K.; James, E. D.; Bis, D. M.; Karty, J. A.; Lane, A. L.; Viswanathan, R. Synergism between genome sequencing, tandem mass spectrometry and bio-inspired synthesis reveals insights into nocardioazine B biogenesis. *Org. Biomol. Chem*. **2015**, *13*, 7177-7192.

14. Cardoso, A. S. P.; Marques, M. M. B.; Srinivasan, N.; Prabhakar, S.; Lobo, A. M.; Rzepa, H. S. Studies in sigmatropic rearrangements of *N*-prenylindole derivatives–a formal enantiomerically pure synthesis of tryprostatin B. *Org. Biomol. Chem*. **2006**, *4*, 3966-3972.

15. Loach, R. P.; Fenton, O. S.; Amaike, K.; Siegel, D. S.; Ozkal, E.; Movassaghi, M. Derivatization of C3-alkylindoles including tryptophans and tryptamines. *J. Org. Chem*. **2014**, *79*, 11254-11263.

16. Partridge, B. M.; Hartwig, J. F. Sterically controlled iodination of arenes via iridium-catalyzed C–H borylation. *Org. Lett*. **2012**, *15*, 140-143.

17. Feng, Y.; Holte, D.; Zoller, J.; Umemiya, S.; Simke, L. R.; Baran, P. S. Total synthesis of erruculogen and fumitremorgin a enabled by ligand-controlled CH borylation. *J. Am. Chem. Soc.* **2015**, *137*, 10160-10163.

18. Jia, Y.; Zhu, J. Palladium-catalyzed, modular synthesis of highly functionalized indoles and tryptophans by direct annulation of substituted *o*-haloanilines and aldehydes. *J. Org. Chem*. **2006**, *71*, 7826-7834.

19. Kokotos, G.; Padrón, J. M.; Martín, T.; Gibbons, W. A.; Martín, V. S. A general approach to the asymmetric synthesis of unsaturated lipidic α-amino acids. The first synthesis of α-aminoarachidonic acid. *J. Org. Chem*. **1998**, *63*, 3741-3744.

20. Bi, W.; Bi, Y.; Xue, P.; Zhang, Y.; Gao, X.; Wang, Z.; Li, M.; Baudy-Floc'h, M.; Ngerebara, N.; Li, X. Novel *β*-carboline-tripeptide conjugates attenuate mesenteric ischemia/reperfusion injury in the rat. *Eur. J. Med. Chem*. **2011**, *46*, 2441-2452.

21. Coste, A.; Toumi, M.; Wright, K.; Razafimahaléo, V.; Couty, F.; Marrot, J.; Evano, G. Copper-catalyzed cyclization of iodo-tryptophans: A straightforward synthesis of pyrroloindoles. *Org. Lett*. **2008**, *10*, 3841-3844.

22. Cozett, R. E.; Venter, G. A.; Gokada, M. R.; Hunter, R. Catalytic enantioselective acyl transfer: the case for 4-PPY with a C-3 carboxamide peptide auxiliary based on synthesis and modelling studies. *Org. Biomol. Chem.* **2016**, *14*, 10914-10925.

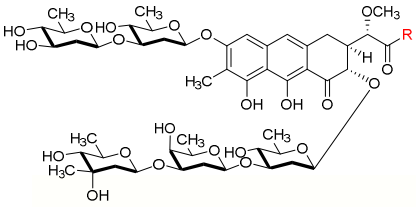
23. Choi, J. Y.; Calvet, C. M.; Gunatilleke, S. S.; Ruiz, C.; Cameron, M. D.; McKerrow, J. H.; Podust, L. M.; Roush, W. R. Rational development of 4-aminopyridyl-based inhibitors targeting trypanosoma cruzi CYP51 as anti-chagas agents. *J. Med. Chem.* **2013**, *56*, 7651-7668.

24. Osgood, C. L.; Maloney, N.; Kidd, C. G.; Kitchen-Goosen, S.; Segars, L.; Gebregiorgis, M.; Woldemichael, G. M.; He, M.; Sankar, S.; Lessnick, S. L.; Kang, M.; Smith, M.; Turner, L.; Madaj, Z. B.; Winn, M. E.; Núñez, L. E.; González-Sabín, Z.; Helman, L. J.; Morís, F.; Grohar, P. J. Identification of mithramycin analogues with improved targeting of the EWS-FLI1 transcription factor. *Clin. Cancer Res*. **2016**, *22*, 4105-4118.

25. Garcia-Aragoncillo, E., J. Carrillo, E. Lalli, N. Agra, G. Gomez-Lopez, A. Pestana, and J. Alonso. "DAX1, a direct target of EWS/FLI1 oncoprotein, is a principal regulator of cell-cycle progression in ewing's tumor cells." *Oncogene* **2008**, *27*, 6034-6043.

26. Grohar, P. J.; Woldemichael, G. M.; Griffin, L. B.; Mendoza, A.; Chen, Q.-R.; Yeung, C.; Currier, D. G.; Davis, S.; Khanna, C.; Khan, J. Identification of an inhibitor of the EWS-FLI1 oncogenic transcription factor by high-throughput screening. *J. Natl. Cancer Inst.* **2011**, *103*, 962-978.

Table of Contents graphic



R = NHTrp-TrpOMe; MTMSA-Trp-Trp (**61**);
GI_{50} = 55 nM, Selectivity Index = 15.6
R = NHPhe-TrpOMe; MTMSA-Phe-Trp (**60**);
GI_{50} = 52 nM, Selectivity Index = 19.2
R = $CH(OH)CH(OH)CH_3$; Mithramycin (MTM; **1**);
GI_{50} = 45 nM, Selectivity Index = 1.5