Discovery of an MLLT1/3 YEATS Domain Chemical Probe

Moses Moustakim*, Thomas Christott*, Octovia P. Monteiro, James Bennett, Charline Giroud, Jennifer Ward, Catherine M. Rogers, Paul Smith, Joanna Panagakou, Laura Díaz-Sáez, Suet Ling Felce, Vicki Gamble, Carina Gileadi, Nadia Halidi, David Heidenreich, Apirat Chaikuad, Stefan Knapp, Kilian V. M. Huber, Gillian Farnie, Jag Heer, Nenad Manevski, Gennady Poda, Rima Al-awar, Darren J. Dixon, Paul E. Brennan,* and Oleg Fedorov*

Abstract: YEATS domain (YD) containing proteins are an emerging class of epigenetic targets in drug discovery. Dysregulation of these modified lysine-binding proteins has been linked to the onset and progression of cancers. We herein report the discovery and characterisation of the first small-molecule chemical probe, **SGC-iMLLT**, for the YD of MLLT1 (ENL/YEATS1) and MLLT3 (AF9/YEATS3). **SGC-iMLLT** is a potent and selective inhibitor of MLLT1/3–histone interactions. Excellent selectivity over other human YD proteins (YEATS2/4) and bromodomains was observed. Furthermore, our probe displays cellular target engagement of MLLT1 and MLLT3. The first small-molecule X-ray co-crystal structures with the MLLT1 YD are also reported. This first-in-class probe molecule can be used to understand MLLT1/3-associated biology and the therapeutic potential of small-molecule YD inhibitors.

Keywords: CHEMICAL PROBE, CHEMICAL BIOLOGY, YEATS DIMENsION, YEATS DOMAIN, CARCINOGENESIS, CANCER THERAPEUTICS

Lyase residues bearing acetyl (Kac) or crotonyl (Kcr) marks are fundamental components of the epigenetic code. As such, YEATS (YAF9, ENL, AF9, TAF14, SAS5) domain containing proteins bind acetyl and crotonyl marks on histone tails. Recent reports have suggested that the dysregulation of YEATS domain (YD) containing proteins correlates with the onset and progression of cancers. There are four YD-containing genes in humans (MLLT1, YEATS2, MLLT3, and YEATS4). Despite high sequence homology (88% YD alignment) and similar roles in complex formation, both MLLT1 and MLLT3 appear to have independent roles in cancers such as acute myeloid leukaemia (AML). MLLT1, which associates with DOT1L, has been implicated in Wilms’ tumour progression when mutations occur in the YD. MLLT1/3 and associates with the AF4 subcomponent in certain rearranged leukaemia types. MLLT3 is a component of the super elongation complex (SEC). Unlike MLLT1, it is the most common fusion partner with MLL (mixed lineage leukaemia protein) in AML (ca. 30% of cases). Recent reports suggested that MLLT1/3 share a “KILK” motif interaction, also present in the extra-terminal (ET) domain of BRD3, which is responsible for the recruitment of chromatin-remodelling complexes, for example, NuRD, BAF, and INO80. Despite the YD-containing proteins...
being correlated with other diseases,[7] no small-molecule inhibitors have been reported to further the understanding of YD-associated biology.

Building on our own and others understanding of developing acetyllysine reading domain inhibitors (bromodomains) exemplified by collaborative efforts targeting the BET bromodomains,[14] p300/CBP,[15] PCAF/GCN5,[16] and others,[17–20] we sought to apply synthetic and medicinal chemistry efforts towards the identification of the first YD-containing protein chemical probe.[21] Owing to the high sequence homology of MLLT1 and MLLT3, particularly in the YD, achieving selective inhibitors would pose a challenge and indeed may not in fact be desired as dual inhibition may mitigate any functional redundancy in MLLT1/3.

A medium-throughput screen of the Ontario Institute of Cancer Research (OICR) library (40000 compounds)[22] revealed compound 1 as a micromolar inhibitor of the MLLT1 YD in an AlphaScreen (AS) assay (MLLT1: IC_{50} = 2.1 μM; Figure 1). Compound 1 posed as an attractive chemical starting point for analogue generation owing to intuitive retrosynthetic disconnections and a lack of structural alerts after PAINS[23] filtering. Guided by flexible docking with ICM[24] using a model extracted from a previously reported co-crystal structure of MLLT1:H3Kac27 peptide (Figure 2A, PDB ID 5J9S), docked poses of compound 1 overlay with the amide bond present in the Kac residue in a flipped conformation (Figure 2B). YDs demonstrate higher affinities for crotonylated lysine peptides over acetylated ones,[25,26] which is thought to be related to the presence of a π–π–π network between the crotonyl double bond and residues F28, Y78, and F59 in MLLT1.[7] Docking studies revealed a potential interaction when compound 1 bound to MLLT1 YD between the amide bond CO and backbone NH of Y78, in addition to the amide N-H interacting with the side chain of S58. The piperidine ring in compound 1 was likely protonated in its bound form (predicted pK_a 8.1 using ACD/Percepta pK_a),[27] and because of a number of polar residues close to the entrances of the YD binding channel such as E75, any additional substituents or modifications that decrease the piperidine ring basicity would likely cause a drop off in binding affinity. If analogues did indeed overlay with Kac/ε hinged on the amide bond depicted in Figure 2B, substituents capable of improving π–π–π stacking would be favoured. Although the Kac/Kcr channel in MLLT1 is linear and narrow, both ends contain adequate space for ligand elaboration (Figure 2B).

Ligand development efforts focused on the conversion of the potentially labile methyl ester of compound 1 into another substituent that would be tolerated and offer a suitable vector to identify new binding interactions. Utilising a “poised” approach,[28] compound 1 was disconnected into synthons for rapid diversification (Figure 3). Structure–activity relationship (SAR) studies were carried out on compound 1, culminating in the synthesis of > 200 analogues (selected example compounds shown Table 1, others in the Supporting Information).

Figure 1. Identified hit MLLT1/3 inhibitor compound 1.

Figure 2. A) Co-crystal structure of MLLT1 YD with H3Kac27 peptide bound (PDB ID 5J9S). B) Docking studies of compound 1 bound to MLLT1 YD from the co-crystal structure with H3Kac27 forms similar hydrogen bonds with the backbone NH moieties of Y78 and A79, structural water HOH312, the S58 side chain, and π–π–π stacking interactions between the side chains of F28, F59, Y78, and H56.

Figure 3. Synthetic disconnection of hit compound 1 into simple building blocks for the rapid construction of potential MLLT1/3 YD inhibitors.

Derivatives of compound 1 were synthesised from 4-nitrobenzene-1,2-diamine 2, which was treated with chloroalkyl esters under acidic conditions to form condensed chloromethyl and 2-chloroethyl benzimidazoles 3–5. Benzimidazoles 3–5 were then substituted with amines, providing compounds 6–36. Reduction of nitro compounds 6–36 furnished the corresponding anilines 37–67, which were converted into amidines and sulfonamides, namely compounds 68–183. Compounds 68–183 were screened by AS, and selected examples were further validated by isothermal titration calorimetry (ITC; Scheme 1).

Scheme 1. Synthetic of MLLT1/3 benzimidazole based inhibitors. Reagents and conditions: a) ethyl 2-chloroacetate, ethyl 2-chloropropionate (1.2 equiv), 4 N HCl (0.6 M), 16 h, 100°C, quant.; b) amine (1.2–1.5 equiv), Na_2CO_3 (1.5 equiv), 23 °C, 3–82%; c) H_2, Pd/C (10%), MeOH, RT, 16 h, 17–88%; d) sulfonfyl/acid chloride (1.2 equiv), PS-DIPEA (2 equiv), CH_2Cl_2 (0.2 M), 16 h, 11–100%; e) acid (1.2 equiv), PS-DIPEA (2 equiv), 16 h, 6–100%.

Angew. Chem. Int. Ed. 2018, 57, 16302–16307 © 2018 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim www.angewandte.org 16303
Structural modifications of the 2- or 3-positions in the benzoyl motif of compound 1 were unfavoured (compounds 68–70, Table 1). The introduction of para substitutions or 3,4-disubstitutions was better tolerated although the binding activity still decreased relative to compound 1 (compounds 71–74). Replacement of the amide bond of ester 1 with a sulfonamide also ablated activity (compound 75). The addition of electron-poor heteroaromatic moieties at the benzoyl position to mimic the methyl ester of compound 1 (compounds 76–79) improved binding activity. With suitable methyl ester replacements identified in compounds 77–79, we focused our attention on further improving potency by modification of the basic amine. Introduction of a chiral centre on the piperidine ring of ester 1 would potentially increase preference for a particular conformation for a salt bridge (e.g., E75). Modifications to the basic amine involving substituted piperidine rings or homologation of the benzylic centre ablated or had no effect (compounds 80–84) on binding affinity compared to unsubstituted compound 73. Rearrangement of the piperidyl motif to a 2-methylpyrrolidine or a fused cyclopropyl-pyrrolidine in combination with the 1-methyl-1H-indazol-5-yl motif to give compounds 85 and 86 resulted in a net increase in potency (MLLT1 YD: IC_{50} 0.33 \mu M and 1.4 \mu M). Elaboration of nanomolar inhibitor 85 into 2-ethyl- and 2-methoxymethylpyrrolidines 87 and 88 gave similarly or less active compounds.

In efforts to optimise substitution of the pyrrolidine core of compound 85, all stereoisomers of both 2-methyl- and 3-methylpyrrolidine derivatives were synthesised (89–92, Table 1). Interestingly, the 2-methyl-substituted pyrrolidines 91 and 92 display a larger difference in binding activity between both enantiomers, with compound 92 displaying the most potent activity ((S)-92 MLLT1 YD IC_{50} 0.26 \mu M, (R)-91 MLLT1 YD IC_{50} 2.0 \mu M). Binding of compound 92 to MLLT1 YD was validated by ITC (MLLT1 YD IC_{50} 0.83 \mu M), which allows it to be used as a chemically similar, but less active control compound. Compound 92 is predicted to be slightly more conformationally restricted about the aliphatic pyrrolidine ring due to substitution compared with original hit piperidine 1, which may confer stabilised electrostatic interactions with charged side chain residues in MLLT1 YD. More potent binding observed from the introduction of heteroaromatics in place of the methyl benzoyl motif as with compound 92 may be attributed to a more complementary π-π-π “sandwich” stack in the binding site. Compounds 91

---

**Table 1:** Selected binding affinity characterisation data of the benzimidazole series for MLLT1 YD.

| R | MLLT1 IC_{50} [\mu M] | R | MLLT1 IC_{50} [\mu M] |
|---|----------------|---|----------------|
| 1 | 2.1 ± 0.78 | 80 | > 20 |
| 68 | > 20 | 81 | > 20 |
| 69 | > 20 | 82 | 5.7 ± 2.4 |
| 70 | > 20 | 83 | > 20 |
| 71 | 4.5 ± 0.96 | 84 | > 20 |
| 72 | 5.5 ± 0.06 | 85 | 0.33 ± 0.11 |
| 73 | 4.7 ± 1.1 | 86 | 1.4 ± 0.86 |
| 74 | 6.4 ± 0.65 | 87 | 0.3 ± 0.06 |
| 75 | > 20 | 88 | 0.97 ± 0.47 |
| 76 | 1.2 ± 0.62 | 89 | 0.6 ± 0.46 |
| 77 | 1.0 ± 0.59 | 90 | 0.73 ± 0.59 |
| 78 | 2.1 ± 1.0 | 91 | 2.0 ± 1.1 |
| 79 | 1.6 ± 0.54 | 92 | 0.26 ± 0.09 |

[a] X = C unless otherwise stated. [b] X = SO.
and 92 were found to be approximately equipotent against highly homologous MLLT3 YD by ITC (compound 91 $K_d$ 0.54 μM; compound 92 $K_d$ 0.077 μM).

Profiling of compound 92 against YEATS2 and YEATS4 revealed excellent selectivity for MLLT1/3 with no activity observed (YEATS2/4 $IC_{50} > 10$ μM). Profiling of compound 92 against a selection of bromodomains showed complete selectivity: There was no inhibition of BRD4 (I), CBP, TAF1, CECD2, and FALZ (10 μM using AS). This was validated in a thermal shift assay where both the original hit compound 1 and compound 92 showed no activity against 48 bromodomains (50 μM compound concentrations; see the Supporting Information).

After extensive crystal soaking experiments, an X-ray co-crystal structure of compound 92 in complex with MLLT1 YD was obtained (Figure 4A). Compound 92 occupies the Kac

![Figure 4. A) Detailed interactions of compound 92 (orange sticks) with MLLT1 YD (loop 1 in cyan, loop 4 in magenta, loop 6 in yellow; PDB ID 6HT1). B) Detailed interactions of compound 92 (orange sticks) with MLLT1 YD (cyan sticks) overlaid with a co-crystal structure of Kcr (yellow sticks):MLLT3 YD (yellow sticks) (PDB ID 5HJB).](image)

Kcr binding site of MLLT1 YD, making a number of interactions with loop 1, loop 4, and loop 6 (Figure 4A; cyan, magenta, and yellow, respectively) adjacent to a structural water molecule. Interestingly, the binding mode of compound 92 matches docked predictions for compound 1. Y78 adopts two conformations, namely an “in” pose where a π stacking interaction with the amide of 92 can take place, similar to the Kcr:MLLT3 YD crystal structure (PDB ID 6HT1), along with an “out” pose in which the Y78 side chain is now displayed edge-to-face with the adjacent side chain of F28.

As demonstration of target engagement is a key criteria in chemical probe qualification, we sought to demonstrate cell activity of compound 92 through multiple methods. We demonstrated MLLT1 target engagement using a cellular thermal shift assay (CETSA) with endogenous MLLT1 in MV4;11 cells. Compound 92 showed stabilisation of MLLT1 (Figure 5A–B) whereas the less active control compound JQ1 showed no activity against 48 bromodomains (50 μM compound concentrations; see the Supporting Information).

![Figure 5. A) Western blot analysis of dose-dependent thermal stabilisation of endogenous MLLT1 in MV4;11 cells with compounds 92 and 91. B) Stabilisation of MLLT1 induced by 92 and 91 normalised to DMSO. C) FRAP assay showing the half-life recovery time in U2OS cells transfected with GFP-tagged MLLT1 wild-type (WT) and mutant (MUT) and D) MLLT3 WT after treatment with compound 92 or 91 (24 h) in the presence of 2.5 μM SAHA. Graphs represent $n=3$ biological replicates, with $n = 10$ cells per treatment group. Mean ± SD, one-way ANOVA with Tukey-Kramer correction for multiple comparisons **** $P<0.0001$, *** $P<0.001$, ** $P<0.01$, * $P<0.05$. E) NanoBRET dose response for 91 and 92 after 24 h treatment using N-terminal nanoLuc-MLLT3 and C-terminal HaloTag-H3.3 in HEK293 cells, in the presence of 2.5 μM SAHA. ■ compound 92, ▼ compound 91, ● DMSO, ○ DMSO + SAHA. Graphs represent $n=6$ biological replicates, with $n > 4$ technical replicates. Mean ± SEM, mBRET units. F) qPCR for MYC, MPO, DNN, CTSG, and CD86 expression in MV4;11 cells transfected with GFP-tagged MLLT1 wild-type (WT) and mutant (MUT) and D) MLLT3 WT after treatment with compound 92 or 91 (24 h) in the presence of 2.5 μM SAHA. Graphs represent $n=6$ biological replicates, with $n = 4$ technical replicates. Mean ± SEM, one-way ANOVA with Tukey-Kramer correction for multiple comparisons **** $P<0.0001$, *** $P<0.001$, ** $P<0.01$, * $P<0.05$.](image)
times to study the effect of compound 92. Therefore, we preincubated the cells with HDAC inhibitor suberoylanilide hydroxamic acid (SAHA, 2.5 μM), preserving global histone acetylation and thus increasing binding of the wild-type MLLT1 and MLLT3 but not mutant MLLT1 (Figure 5C,D). Preincubation with SAHA increased $t_{1/2}$ to 0.7767 ± 0.09 s and 1.32 ± 0.45 s for wild-type MLLT1 and MLLT3, respectively. Incubation of cells with compound 92 in presence of SAHA significantly decreased $t_{1/2}$ to 0.46 ± 0.05 s and 0.56 ± 0.02 s for MLLT1 and MLLT3, respectively (MLLT1: $P < 0.0001$, MLLT3: $P = 0.0185$). On the other hand, $t_{1/2}$ of cells incubated in compound 91 was not significantly different from that of wild-type MLLT1 or MLLT3 in the presence of SAHA (Figure 5C,D). We also developed a full length MLLT1-Histone 3.3 (H3.3) NanoBRET assay to test compound 92. Although the MLLT1 NanoBRET assay was responsive to SAHA treatment, showing a significant increase in BRET activity (mBU) relative to the DSMO control, there was no reduction in mBU in response to MLLT1/3 inhibitor treatment. For MLLT3, the NanoBRET assay showed clear dose-dependent displacement of full-length MLLT3-NanoLuc from histone H3.3-HaloTag (average IC$_{50}$ 0.4 ± 0.08 μM) in HEK293 cells (Figure 5E and Figure S5).

Following recent links of MLLT1 knockdown and CRISPR knockout reducing AML tumorigenesis,[5,6] we profiled genes that were shown to be down- (MYC, myeloperoxidase (MPO), dendrin (DDN), cathepsin G (CTSG)) and up (CD86)-regulated in MV4;11 AML cells after MLLT1 knockdown/out.[35] Compound 92 changed gene expression of 3/5 genes, showing a reduction in MYC, DDN and an increase in CD86 compared to DSMO. The weaker enantiomer, compound 91, displayed diminished cell activity in the MLLT1 CETSA (Figure 5A,B), FRAP (Figure 5C,D), NanoBRET (Figure 5E, average IC$_{50}$ 5.8 ± 0.07 μM), and gene expression (Figure 5F).

We also investigated the metabolic stability of compound 92 in primary human hepatocytes. Compound 92 shows moderate metabolic resistance ($t_{1/2}$ 53 min, 48% remaining after 60 min), with a primary process for metabolism being N demethylation. An N-cyclopropyl indazole analogue, compound 94, designed to improve pharmacokinetics, largely mitigates N dealkylation observed and retains potent binding activity (MLLT1 $K_a$ 0.058 μM), which was also rationalised in a co-crystal structure with MLLT1 YD (see the Supporting Information, PDB ID 6HT0), but the overall half-life of 94 was inferior to 92, with pyrrolidine oxidation occurring more rapidly ($t_{1/2}$ ~30 min). These analogues provide information on how to develop MLLT1/3 probes with good PK properties.

We have reported the discovery of the first small-molecule inhibitors and a cell-potent, selective chemical probe for MLLT1 and MLLT3 YDs. Compound 1 was discovered as an initial hit in a medium-throughput biochemical screen. Simple synthesis allowed rapid generation of >200 analogues and the chemical probe 92 (SGC-iMLLT) with its less active control 91 (SGC-4iMLLT-N). Selectivity was demonstrated over acylsine binding modules, YEATS2/4, and 48 bromodomain.$^{[2]}$. Using orthogonal combinations of cell target engagement studies (NanoBRET, FRAP, and CETSA), submicromolar cellular activity was confirmed. SGC-iMLLT will enable researchers to design the first biological experiments exploring MLLT1/3 YD inhibition.

**Acknowledgements**

We gratefully acknowledge N. Manevski (UCB Pharma) for carrying out PK studies. The SGC is a registered charity number (1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, the Canada Foundation for Innovation, the Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766]. Janssen, Merck KGaA Darmstadt Germany, MSD, Novartis Pharma AG, the Ontario Ministry of Economic Development and Innovation, Pfizer, the São Paulo Research Foundation-FAPESP, Takeda, and Wellcome [106169/ZZ14/Z]. M.M. is grateful to the EPSRC Centre for Doctoral Training in Synthesis for Biology and Medicine (EP/L015838/1) for a studentship, generously supported by AstraZeneca, Diamond Light Source, Defence Science and Technology Laboratory, Evotec, GlaxoSmithKline, Janssen, Novartis, Pfizer, Syngenta, Takeda, UCB, and Vertex. Base funding for OICR is provided by the Government of Ontario.

**Conflict of interest**

N. Manevski and J. Heer are employees and hold shares in UCB Pharma Ltd.

**Keywords:** chemical probes · epigenetics · MLLT1 · MLLT3 · YEATS

**How to cite:** Angew. Chem. Int. Ed. 2018, 57, 16302–16307

**Angew. Chem.** 2018, 130, 16540–16545

[1] P. Filippakopoulos, S. Knapp, Nat. Rev. Drug Discovery 2014, 13, 337 – 356.
[2] M. Lawrence, S. Daujat, R. Schneider, Trends Genet. 2016, 32, 42 – 56.
[3] M. Moustakim, P. G. K. Clark, D. A. Hay, D. J. Dixon, P. E. Brennan, MedChemComm 2016, 7, 2246–2264.
[4] S. G. Smith, M. M. Zhou, ACS Chem. Biol. 2016, 11, 598–608.
[5] L. Wan et al., Nature 2017, 543, 265–269.
[6] M. A. Erb et al., Nature 2017, 543, 270–274.
[7] D. Zhao, Y. Li, X. Xiong, Z. Chen, H. Li, J. Mol. Biol. 2017, 429, 1994–2002.
[8] A. Yokoyama, M. Lin, A. Naresh, I. Kitabayashi, M. L. Cleary, Cancer Cell 2010, 17, 198–212.
[9] D. Mueller et al., Blood 2007, 110, 4445–4454.
[10] N. He, C. K. Chan, B. Sobhian, S. Chou, Y. Xue, M. Liu, T. Alber, M. Benkirane, Q. Zhou, Proc. Natl. Acad. Sci. USA 2011, 108, E636–E645, SE636/1-SE636/2.
[11] E. J. Perlman et al., Nat. Commun. 2015, 6, 10013.
[12] D. T. Zeisig, C. B. Bittner, B. B. Zeisig, J. M. García-Cuéllar, J. L. Hess, R. K. Slany, Oncogene 2005, 24, 5525–5532.
[13] D. C. Wai et al., J. Biol. Chem. 2018, 293, 7160–7175.
[14] P. Filippakopoulos et al., Nature 2010, 468, 1067–1073.
[15] D. A. Hay et al., J. Am. Chem. Soc. 2014, 136, 9308–9319.
[16] M. Moustakim et al., Angew. Chem. Int. Ed. 2017, 56, 827–831; Angew. Chem. 2017, 129, 845–849.
[17] P. Chen et al., J. Med. Chem. 2016, 59, 1410–1424.
[18] L. Drouin et al., J. Med. Chem. 2015, 58, 2553–2559.
[19] A. E. Fernández-Montalván et al., ACS Chem. Biol. 2017, 12, 2730–2736.
[20] D. A. Hay, C. M. Rogers, O. Fedorov, C. Tallant, S. Martin, O. P. Monteiro, S. Muller, S. Knapp, C. J. Schosfield, P. E. Brennan, MedChemComm 2015, 6, 1381–1386.
[21] “SGC Chemical Probes” can be found under http://www.thesgc.org/chemical-probes, 2018.
[22] The “Ontario Institute of Cancer Research” can be found under https://oicr.on.ca.
[23] D. A. Erlanson, J. Med. Chem. 2015, 58, 2088–2090.
[24] M. A. C. Neves, M. Totrov, R. Abagyan, J. Comput.-Aided Mol. Des. 2012, 26, 675–686.
[25] Y. Li et al., Cell 2014, 159, 558–571.
[26] Y. Buganim et al., PLoS One 2010, 5, e9657.
[27] ACD/Labs, ACD/Labs Percepta Predictor, ACD/Labs, Toronto, ON, Canada, 2015.
[28] O. B. Cox et al., Chem. Sci. 2016, 7, 2322–2330.
[29] M. E. Bunnage, E. L. P. Chekler, L. H. Jones, Nat. Chem. Biol. 2013, 9, 195–199.
[30] R. Jafari, H. Almqvist, H. Axelsson, M. Ignatushchenko, T. Lundbäck, P. Nordlund, D. M. Molina, Nat. Protoc. 2014, 9, 2100–2122.
[31] T. Machleidt, C. C. Woodroofe, M. K. Schwinn, J. Méndez, M. B. Roberts, K. Zimmerman, P. Otto, D. L. Daniels, T. A. Kirkland, K. V. Wood, ACS Chem. Biol. 2015, 10, 1797–1804.

Manuscript received: September 14, 2018
Accepted manuscript online: October 5, 2018
Version of record online: November 16, 2018