Bioorganic & Medicinal Chemistry Letters 24 (2014) 21–32

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

BMCL Digest

Design of small molecule epigenetic modulators ☆

Boobalan Pachaiyappan, Patrick M. Woster *

Department of Drug Discovery and Biomedical Sciences, Medical University of South Carolina, 70 President St., Charleston, SC 29425, United States

ARTICLE INFO

Article history:
Received 6 February 2013
Revised 28 October 2013
Accepted 4 November 2013
Available online 13 November 2013

Keywords:
Histone
Epigenetic
Methyllysine
Methylarginine
Acetyllysine

ABSTRACT

The field of epigenetics has expanded rapidly to reveal multiple new targets for drug discovery. The functional elements of the epigenomic machinery can be categorized as writers, erasers and readers, and together these elements control cellular gene expression and homeostasis. It is increasingly clear that aberrations in the epigenome can underly a variety of diseases, and thus discovery of small molecules that modulate the epigenome in a specific manner is a viable approach to the discovery of new therapeutic agents. In this Digest, the components of epigenetic control of gene expression will be briefly summarized, and efforts to identify small molecules that modulate epigenetic processes will be described.

The term epigenetics refers to the study of heritable changes in gene expression and/or phenotype that are mediated by processes that do not involve alterations in the primary sequence of DNA. Epigenetic changes occur either by covalent modification of histone proteins or by methylation of DNA, usually at CpG island sites near specific gene promoters. Specific epigenetic modifications allow the cell to solve the fundamental problem of storing approximately 2 m of DNA in a cell nucleus with an average diameter of 6 μm, while retaining the ability to control the expression of specific gene products. Prior to 2004, the field of epigenetics was focused on two primary processes, histone acetylation/deacetylation[1,2] and DNA methylation.[3] Histone proteins occur as octamers that consist of one H_3–H_4 tetramer and two H_{2A}–H_{2B} dimers,[4] and these structures are directed to DNA because of their predominating positive charges.[5] Histone protein tails contain residues that interact with the negative charges on the DNA backbone, and interact with double stranded DNA in such a way that a section of DNA is wrapped around the histone octamer, forming a structure known as a nucleosome. As part of this structure, histone lysine tails protrude through the DNA strand and become sites for post-translational modifications (PTMs) of chromatin, allowing alteration of higher order nucleosome structure.[1] The acetylation status of histones is controlled by a balance between two enzymes that functionalize specific histone lysines: histone acetyltransferase (HAT), which promotes histone hyperacetylation, and the histone deacetylases (HDACs), which catalyze acetyl group cleavage.[2,6] Normal mammalian cells efficiently control chromatin architecture in part by maintaining a balance between HAT and HDAC activity. Both the zinc-dependent[7,8] and sirtuin[9,10] classes of histone deacetylases have been regarded as validated drug targets for a number of years, and two HDAC inhibitors have reached the market, with several other drug candidates in Phase II or III trials. Aberrant methylation of DNA is also an important event in epigenetic silencing.[11] In many forms of cancer, cells exhibit a global loss of methylcytosine (hypomethylation) in the gene body, while simultaneously exhibiting hypermethylation at CpG islands in the promoter region.[12] Like HDAC inhibitors, DNA methyltransferase inhibitors are well known epigenetic modulators that have been used extensively in cancer chemotherapy.[13,14] Because these topics have been thoroughly reviewed, they will not be considered in this digest. This article will focus on druggable epigenetic targets that have come to light in the last decade, with a particular emphasis on recent developments.

Components of the epigenome: Conceptually, elements of the epigenome machinery can be divided into three categories that are commonly termed writers, erasers and readers.[1,15] In addition to DNA methylation, the epigenetic code is set by the formation and breakdown of specific PTMs at histone lysine (acetylation, methylation, ubiquitination, sumoylation, crotonylation, butyrylation and propionylation) and arginine (methylation, citrullination and ADP-ribosylation) residues, as well as phosphorylation of serine, tyrosine and threonine.[16–18] Taken together, modifications at these residues make up an epigenetic 'landscape' whereby cellular gene expression can be tightly and accurately controlled. From a drug discovery standpoint, the most important of these PTMs to date are lysine acetylation/deacetylation and lysine or arginine

* Corresponding author.
E-mail address: woster@musc.edu (P.M. Woster).

0960-894X/$ - see front matter © 2013 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmcl.2013.11.001

methylation/demethylation. The discussion below will focus on drug development efforts targeting epigenetic writers, erasers and readers; a deeper discussion of the biology underlying their function can be found in recent comprehensive reviews.[16,19] To limit the scope of this review, emphasis was given to novel compounds described in primary journal references. Information concerning epigenetic modulators disclosed in the patent literature has been presented elsewhere.[7,13,20,21]

Epigenetic writers: Inhibitors of histone acetyltransferases. As stated above, histone lysines can be post-translationally modified by the addition of acetyl groups to lysine, or by the addition of methyl groups to both lysine and arginine. These reactions are carried out by HATs or one of several histone methyltransferases, respectively. In the case of lysine acetylation, formation of an amide with the terminal amino group serves two functions. Acetylation of the ε-amino groups of lysine residues cancels the positive charge, promoting the formation of a relaxed form of DNA (euchromatin).[1,2] This allows greater access to DNA by transcription factors and RNA polymerase, and promotes the expression of various transcriptional products. It follows that removal of the acetyl group from histone lysine residues by HDACs restores the positive charge at ε-amino group, resulting in a densely packed form of chromatin (heterochromatin) that is transcriptionally inactive. As described below, acetylated lysine residues also serve as recognition and binding sites for specific histone readers such as bromodomain-containing proteins.[22,23] The association of these reader proteins with an acetylated lysine on the histone tail also promotes transcription.

Nuclear-localized HATs are more correctly referred to as lysine acetyltransferases, since they are also known to acetylate a variety of non-histone lysine residues.[24] These enzymes are classified into 5 families based on structural homology and catalytic mechanism,[25,26] but three of these constitute the major classes that are involved in chromatin remodeling: the Gcn5-related N-acetyltransferases or GNAT family (tGcn5, PCAF and ELP3), the p300/CBP family (p300 and cyclic AMP-responsive element (CREB binding protein) and the MYST family (Tip60 and MYST 1–4).[27–29] The effects of these acetyltransferases can be gene-specific or global, depending on their structure and they may participate in a multisubunit protein complex that determines substrate specificity. All of these enzymes require acetyl CoA as a co-substrate, and despite having low sequence homology, they all exhibit a conserved acetyl CoA binding pocket.[25] The p300/CBP and GNAT family acetyltransferases appear to play a role in cancer, and elevated HAT activity has been observed in asthma and chronic obstructive pulmonary disease (COPD).[30] The levels of the acetyltransferase p300 have been suggested as biomarkers for prognosis in both lung and prostate cancer.[31,32]

Only a handful of HAT inhibitors have been identified to date (Fig. 1), and these compounds were useful in delineating the enzyme mechanism and identifying the role of HAT in cancer. These inhibitors are classified as bisubstrate analogues, natural products or synthetic small molecules.[33] Lau et al. described the synthesis and application of compound **1**, an adduct between lysine and acetyl CoA, a potent (IC_{50} = 500 nM) and selective (~200-fold) bisubstrate inhibitor of p300 that displayed an unusual hit-and-run (Theorell-Chance) catalytic mechanism.[34] Interestingly, addition of a 20-mer peptide analogous to the sequence of the H3 lysine tail produces a compound that is selective for tGcn5 (IC_{50} = 300 nM) and PCAF over p300.[35] Olhava et al. performed a virtual screen of about 500,000 small molecules against the p300 HAT bisubstrate inhibitor binding site and identified compound **2**, which displayed an IC_{50} of 400 nM in an in vitro assay.[36] Treatment with **2** reduces histone H3 and H4 acetylation levels and abrogates TSA-induced acetylation in cultured cells. In addition, a dose-dependent growth

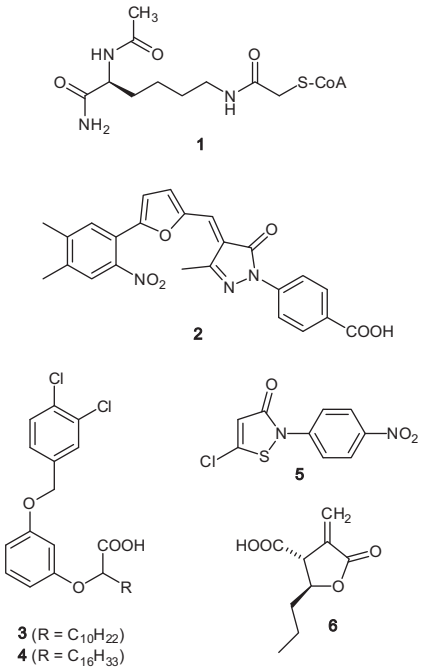

Figure 1. Known inhibitors of histone acetyltransferases.

inhibitory effects were observed when tested against melanoma, lung and brain cancer cells.[36] A number of natural products such as curcumin, garcinol, plumbagin, EGCG, gambogic acid and anacardic acid act as micromolar inhibitors of HATs,[30] and these findings led to the design of small molecule synthetic analogues. The most notable derivatives of anacardic acid, compounds **3** and **4**, reported by Eliseeva et al. produced HAT inhibition that correlates with its antiproliferative effects, and caused H4 hypoacetylation in MCF7 cells in vitro.[37] In addition, isothiazolone derivatives such as **5** reported by Stimson et al. have been shown to inhibit the GNAT family acetyltransferase PCAF,[38] while α-methylene-γ-butyrolactone **6** described by Biel et al. have been shown to inhibit GNC5, a GNAT family HAT.[39]

Inhibitors of histone methylation: N-terminal histone tails undergo methylation at specific histone lysine and arginine chromatin marks.[2,40] To date, 17 lysine residues and 7 arginine residues on histone proteins have been shown to undergo methylation[38] catalyzed by 28 lysine methyltransferases (KMTs)[41,42] and 9 protein arginine methyltransferases (PRMTs).[40,43] All of these enzymes require S-adenosylmethionine (SAM) as methyl donor, and produce the by-product S-adenosylhomocysteine (SAH), which is recycled through salvage pathways. Histone KMTs generally target a specific lysine residue, and can promote or repress transcription, depending on the lysine residue involved.[42,44,45] There are 8 classes of histone KMTs: 7 SET1 domain KMTs, 5 SET2 domain KMTs, 6 KMTs in the SUV39 group, EZH1 and 2, 3 KMTs in the PRDM group, 4 KMTs in the other-SET group and 1 non-SET domain KMT.[46] Except for the non-SET KMT hDOT1L, all histone KMTs contain a conserved SET (SU(VAR)3–9, enhancer-of-Zeste, Trihorax) domain that is responsible for the methylation activity.[47] As shown in Figure 2A, histone KMTs catalyze the stepwise conversion of unmethylated histone lysines to the mono-, di- and trimethylated forms. There are also 2 separate groups of PRMTs, 5 Type I PRMTs and 4 Type II PRMTs. Type I PRMTs methylate arginine to produce *N*-monomethyl arginine (Figure 2B). From there, a second methyl group can be added to the same carbon to form gem-*N,N*-dimethylarginine, or a Type II PRMT can methylate the second arginine