

BioFocus

www.biofocus.com

Epigenetics - the social network of our genes

BioFocus combines 15 years' experience with expert teams to provide a drug discovery offering that has the ability to deliver success to its clients' therapeutic programs. Stand alone or integrated, our expertise in epigenetics includes:

Phenotypic assays. Cellular assays to study the role of epigenetic targets and co-factors in complex biological systems, including high content readouts, using compounds or shRNA/cDNA.

Biochemical assays and high throughput screening. Delivery of novel chemotypes in high throughput biochemical screening campaigns against epigenetic targets utilizing a screening deck specifically selected from our 900,000 compound library.

Fragment-based drug discovery. Cloning, protein expression, purification, and screening of in-house fragment library by surface plasmon resonance (Biacore 4000), crystallization and X-ray analysis with integrated medicinal chemistry.

Medicinal chemistry. Current programs for the development of Class II selective histone deacetylase (HDAC) inhibitors for the treatment of Huntington's disease and the investigation of other lysine deacetylases.

Computer-assisted molecular design. Design concepts and tools for protein class-focused libraries directed to histone acetyl transferases (HATs), histone methyl transferases (HMTs) and ligases.

According to Wikipedia, a social network is a social structure made up of individuals (or organizations) called "nodes", which are connected by one or more specific types of interdependency, such as friendship, kinship or common interest.

Understanding how the human genetic code is controlled to create the immense set of cellular phenotypes that form an organism is a fundamental challenge in biology and medicine. Almost every cell (or, in social-network terms, "node") in an organism contains an identical genotype (or "common interest") and yet different cells express a unique set of genes at different times. This control of gene expression has been studied for many years and is known to be mediated by many regulatory pathways including epigenetic mechanisms (the "social network"). Epigenetic mechanisms are governed by information that is not inherently present in the genetic code; these include covalent modification of DNA and histones, leading to remodelling of the chromatin structure, thus enabling access to essential promoter elements within the DNA. Figure 1 shows the types of histone modifications that have been identified to affect chromatin structure and therefore impact the regulation of gene expression.

Over the past two decades, our understanding of epigenetic control of cellular development and differentiation has grown immensely. Along with this, the role of dysregulated epigenetics in certain diseases has clearly been recognized.^{1,3} There is a growing body of evidence indicating that epigenetic mechanisms are implicated in cancer and in inflammatory, neurodegenerative and metabolic diseases. This poses the question of how epigenetic processes can be modulated and which molecular targets would be most suitable for therapeutic interference. Research in this direction has revealed new classes of enzymes for drug discovery research with the goal of creating small-molecule therapeutics effective in diseases.

Promisingly, the first inhibitors of histone deacetylases (HDACs) and DNA methylases are in clinical use for certain oncology indications. However, by and large, the field of epigenetic drug discovery is at an early stage and key questions address the basic understanding of pathways, optimal subtype selectivity within protein families, suitable assay conditions, and the search for novel chemical pharmacophores. In conclusion, the status of the epigenetic field may be compared to that of research in the kinase field two decades ago.

Prior to testing how modulation of an epigenetic target will translate into efficacy and safety in humans, the efficient development of the appropriate *in vitro* assays and selective/potent tool compounds is essential. At BioFocus, we support projects from early stage (target validation), assay development and compound design through to late-stage preclinical development involving integrated medicinal chemistry, structural biology and

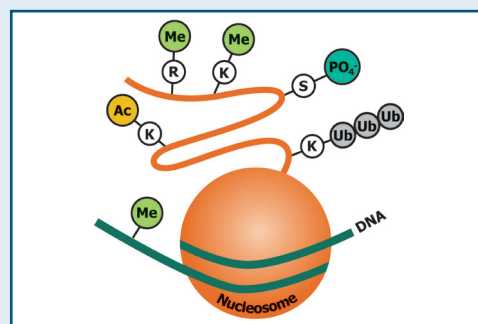


Figure 1: Post-translational modification of histone protein that can influence gene transcription. Abbreviations: Me, methyl; Ac, acetyl; Ub, ubiquitin; PO₄⁻, phosphate.

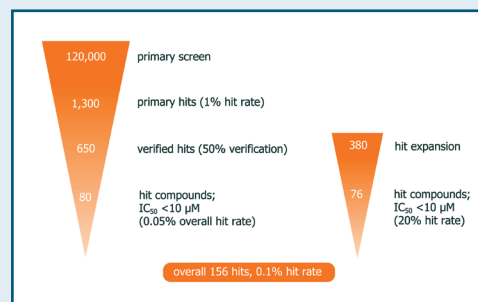


Figure 2: Hit finding campaign for a histone methyltransferase (HMT) – compound attrition for the primary screen and hit expansion.

ADME (absorption, distribution, metabolism and excretion) support.

Assay development and hit finding

Epigenetic targets present several challenges for the identification of small-molecule inhibitors. Many of these targets are in enzyme classes that are poorly covered by existing assay technologies, such as the histone methyltransferases (HMTs). As an example, an HMT assay platform BioFocus has used is the detection of ³H-methionine (³H-Me) incorporated into histone protein from ³H-S-adenosyl methionine (³H-SAM) using FlashPlates or filter plates. Similar to the early days in kinase drug discovery, this radiometric assay format is now the "workhorse" for hit finding campaigns in the epigenetics arena. For further characterization of hits, orthogonal assay technologies, such as the Caliper lab-on-a-chip, FLT (fluorescence lifetime technology) and binding assays have successfully been used. Detailed enzyme inhibitor kinetic studies can also be performed using a biophysical binding assay platform such as surface plasmon resonance.

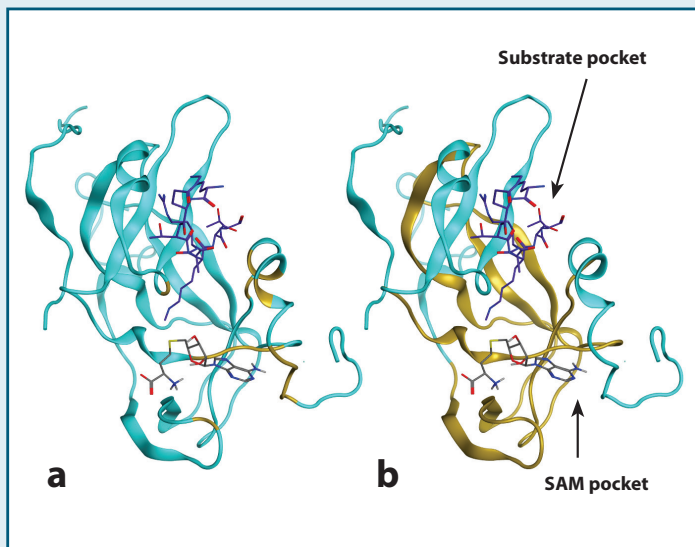


Figure 3: Primary (a) and secondary (b) structure conservation in HMTs using example EHMT1 (PDB:3HNA) as an example. Blue represents variability across the family whilst gold represents homology.

In addition to the lack of assay reagents, focused screening libraries directed at these proteins are not yet available; compound collections that are heavily based around chemotypes synthesized or selected for previous projects are unlikely to provide useful starting points. The combination of a highly diverse compound library (greater than 900,000 compounds) and a sophisticated hit expansion process with a well established assay platform has enabled BioFocus to generate a variety of novel chemical scaffolds that served as starting points for inhibitors of epigenetic targets. Figure 2 shows a typical example of a hit finding campaign for an HMT target where the screening of only 120,380 compounds has yielded 156 hit compounds covering a novel chemical inhibitor space.

Computer-assisted molecular design

BioFocus has successfully employed computer assisted molecular design strategies for over 10 years in the design of their SoftFocus family-targeted compound libraries. Protein family structural superposition followed by sequence alignment underpins the chemogenomic models used to support the design process. Such models have been developed into protein family “roadmaps”, where key residues are considered and selected, aiding the understanding of pocket-based family homology and selectivity. Using the information of such models enables the understanding of protein targets, for which structural information is not necessarily available. We use this ensemble of information to design new target-family focused libraries, to prioritize compounds for specific assays (e.g. hit expansion) and/or to predict potential selectivity targets for a counter-screen. Protein structures used to generate the initial superposition may either be available in the public domain or generated in-house by the structural biology group at BioFocus.

As an illustration, we have analyzed the histone methyl transferase (HMT) family. HMTs are split into protein lysine methyltransferases (PKMTs) and protein arginine methyltransferases (PRMTs). The PKMTs transfer methyl groups to a lysine residue on the histone whilst the PRMTs methylate an arginine. We generated a chemogenomic model for the PKMT subfamily. Structural superposition revealed high conservation in the primary and secondary structure around the SAM pocket (Figure 3).

This high conservation is likely to be an evolutionary consequence of the fact that the same cofactor is being utilized throughout the family. Analysis of the crystallized cofactor conformations of SAM or S-adenosyl-homocysteine (SAH) bound into the SAM pocket reveals, in general, excellent preservation in conformation and interactions with the protein. This observation can help prioritize compounds for screening; for example, adenosine is present in both ATP and SAM, therefore screening minimally functionalized compounds designed for the ATP site in kinases may provide a useful component of a screening deck for hit identification against this novel protein class. From a drug discovery perspective, the conserved homology may pose a challenge in that molecules that bind exclusively to the SAM pocket could lack selectivity over other family members. Hence, adjacent pockets and less conserved residues will need to be taken into account in the compound design.

We have generated a phylogenetic tree around the SAM pocket of the PKMT subfamily. Using the structural superposition, all amino acid residues within 8 Å of the cofactor were selected. Translation of this information to a sequence alignment allowed the extension of this selection to incorporate proteins for which no structural information was available. This tree provides insights as to how the family members are structurally related to each other, which may help in the choice of the optimal selectivity assay. Further, application of structural alignment or homology models may help develop a strategy to address selectivity by rational design. Interestingly, whilst the tree is focused on the SAM pocket, the distribution of targets also reflects the substrates known to bind.

Notably, DOT1L was a clear outlier in this analysis. Structural alignment to the PKMT subfamily was inconsistent revealing notable differences in structure and sequence. Key differences were observed around the SAM pocket, where recognition elements differed from other family members and resulted in a distinct SAM conformation. Our conclusion from this sequence and SAM recognition assessment is that DOT1L has a greater homology with the PRMT subfamily than with the PKMTs.

At BioFocus we have the skills and experience to help you succeed in epigenetic drug discovery.

References:

1. Pray, L. At the flick of a switch: epigenetic drugs. *Chem. Biol.* 2008, **15**:640-641.
2. Copeland, R. A., Solomon, M. E. and Richon, V. M. : Protein methyltransferase as a target class for drug discovery . *Nat. Rev. Drug. Discov.* 2009, **8**: 724-732.
3. Best, J.D. and Carey, N. Epigenetic opportunities and challenges in cancer 2010, *Drug Discov. Today* **15**, 65-70.



CONTACT DETAILS:

Dr Kate Hilyard
 BioFocus
 Tel: +44 (0) 1799 533 500
 Email: info@glpg.com
 www.biofocus.com