ELSEVIER
DRUG DISCOVERY TODAY THERAPEUTIC STRATEGIES

Drug Discovery Today: Therapeutic Strategies
Vol. 9, No. 2–3 2012

Epigenetics

Future strategies in epigenetic drug discovery

Gerard Drewes

Cellzome AG, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

Most current research aimed at the discovery of epigenetic therapies adheres to the paradigm of target-based drug discovery, focusing on the modulation of single enzymes involved in DNA methylation and histone modifications. The recent discovery of promising small molecule inhibitors for a class of nonenzymatic chromatin regulators, the BET bromodomains, suggests that future drug discovery for epigenetic therapy will involve the modulation of protein–protein interactions and multiprotein complexes. Also, it is expected that target-based discovery strategies will be increasingly complemented by approaches based on chemical probes generated by phenotypic or mechanistic cell-based screening.

Introduction

Target-based drug discovery continues to be a dominating paradigm in industrial research, and current strategies for epigenetic therapy are no exception. Most activity is focused on the enzymes involved in DNA methylation and chromatin modification, including DNA methyltransferases, histone acetyl- and methyltransferases, and histone deacetylases and -demethylases [1,2]. The prevailing type of biochemical assays for small molecule screening and lead optimization chemistry typically measure the enzymatic activity of the purified target protein in isolation. Although enzymatic assays are usually preferred in screening projects, binding assays are used in some cases. The proteins utilized in these assays are typically purified from recombinant sources like bacteria or insect cell cultures, and in many cases, in particular in the case of large proteins, truncated constructs comprising the relevant catalytic or ligand-binding domains, but often lacking large portions of the sequence, are used. Hence, these recombinant proteins may lack regulatory sequences or domains, and because they are typically not isolated from mammalian systems, they may also lack the correct posttranslational modifications, such as phosphorylation or acetylation. Moreover, most proteins exist and are regulated by other proteins in protein complexes, and many proteins involved in chromatin structure and transcriptional regulation tend to reside in large multiprotein complexes [3–5]. Therefore, a recombinant protein or protein fragment in isolation may not reflect the conformation and activity of the target in its physiological context, owing to incorrect protein folding, the lack of regulatory domains and interacting proteins, or incorrect or absent post-translational modifications. The data generated in such assays do not reflect the complexity of the cellular context and hence may not correctly predict the efficacy of a compound or drug in cell-based or *in vivo* models (Fig. 1). In this review, we will discuss novel experimental strategies to the discovery of epigenetic drug targets based on assays which more closely reflect physiological complexity. In the first section, we outline different approaches towards the unbiased discovery of small molecule probes for targets that are relevant in the modulation of disease-relevant

E-mail address: G. Drewes (gerard.drewes@cellzome.com)

1740-6773/$ DOI: 10.1016/j.ddstr.2011.12.001

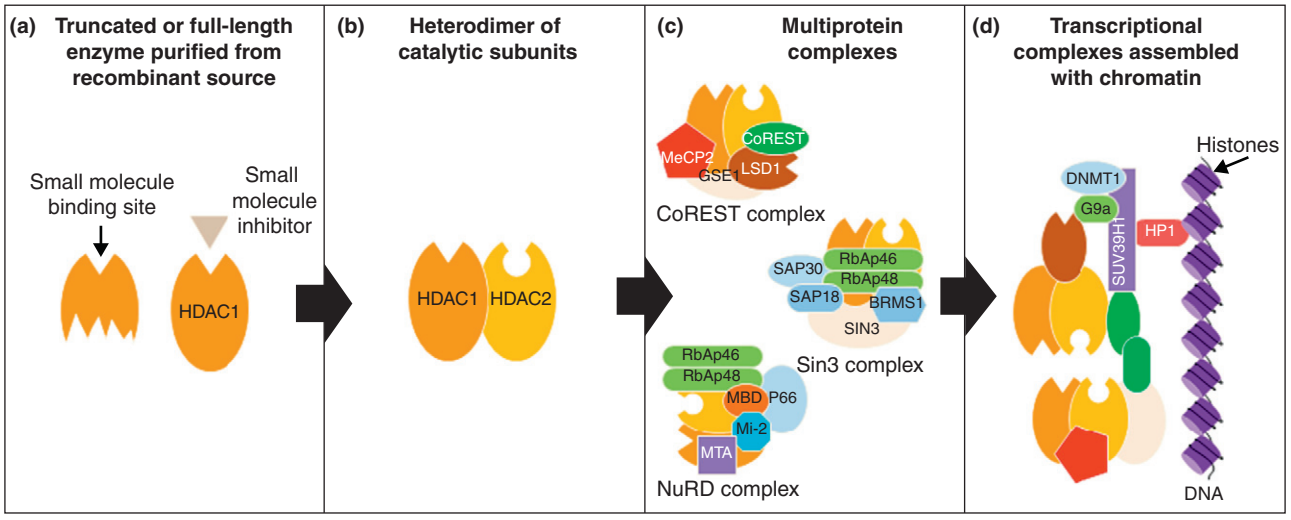

Fig. 1. Measuring drug-target interactions at multiple levels of complexity. **(a)** Industry standard assay utilizing purified protein, for example a histone deacetylase (HDAC1) from recombinant source, or a truncated protein fragment containing the small molecule binding domain. **(b)** Heterodimer of two catalytic subunits. **(c)** Defined multiprotein complexes present in cell extract, as used for drug profiling in chemoproteomic techniques **(d)** Representation of transcriptional complexes assessed in cell-based approaches.

cellular and epigenetic processes. In the second part of this review, recent progress in the application of chemical probes in drug discovery utilizing native cellular proteins and protein complexes is discussed.

Target discovery and target validation based on small molecule probes

As discussed in several of the articles in this volume, many epigenetic enzymes and nonenzymatic regulatory proteins are members of large families characterized by functional domains with a considerable degree of sequence homology. This may present scientific challenges with respect to preclinical target validation and potentially achieving inhibitor selectivity. In this context, a target is regarded as validated if it has been demonstrated to mediate a pathophysiological process such that modulation of its activity reverses a disease-relevant parameter, which is measured in cell-based or animal models and expected to be predictive of human disease. Early target discovery and validation strategies are frequently based on protein knockdown screens using RNA interference by siRNA or shRNA [6–8] which can be combined with high-throughput screening of compound libraries in cell-based assays or with biochemical assays for the targets identified by siRNA. Active compounds derived from such screens are then used as chemical probes to confirm the knockdown data and further validate the target. A recent elegant study by Zuber *et al.* explored chromatin regulators as drug targets by assembling a library of 1094 shRNAs customized to target 243 selected chromatin regulators. The library was transduced as a single pool into a genetic mouse model of acute myeloid leukemia (AML) and changes in library representation were monitored over time. The screen identified the bromodomain adaptor protein Brd4 as a candidate target, which was subsequently validated by a small molecule inhibitor of this protein [9]. The main challenges with RNA interference-based screens arise from the substantial off-target activities of siRNAs [10], and the potential problems with the selectivity of small molecule inhibitors used for target validation, which are more pronounced for target classes with a high degree of structural conservation. In addition targets may be masked in siRNA screens due to functional redundancy on the protein level with respect to the cellular readout. Despite these challenges, the industry has been successful in generating selective small molecule inhibitors directed against the conserved co-substrate binding site of targets from highly conserved large families, as demonstrated by the successes in developing protein kinases inhibitors [11]. For epigenetic enzymes like the histone methyltransferases [12] and demethylases [13], similar challenges as with ATP-competitive kinase inhibitors may be expected, because inhibitors are likewise designed based on the structures of the natural cosubstrates. The industry has dealt with these challenges by assessing inhibitor selectivity in large biochemical assay panels which in the case of kinases comprise hundreds of enzymes [14]. Both the Structural Genomics Consortium (http://www.thesgc.org) and the Center for Protein Research (http://www.cpr.ku.dk) have initiated programs for the purification of proteins and the development of chemical probes which will eventually become available to the scientific community. However, extensive resources and time will be required to purify and

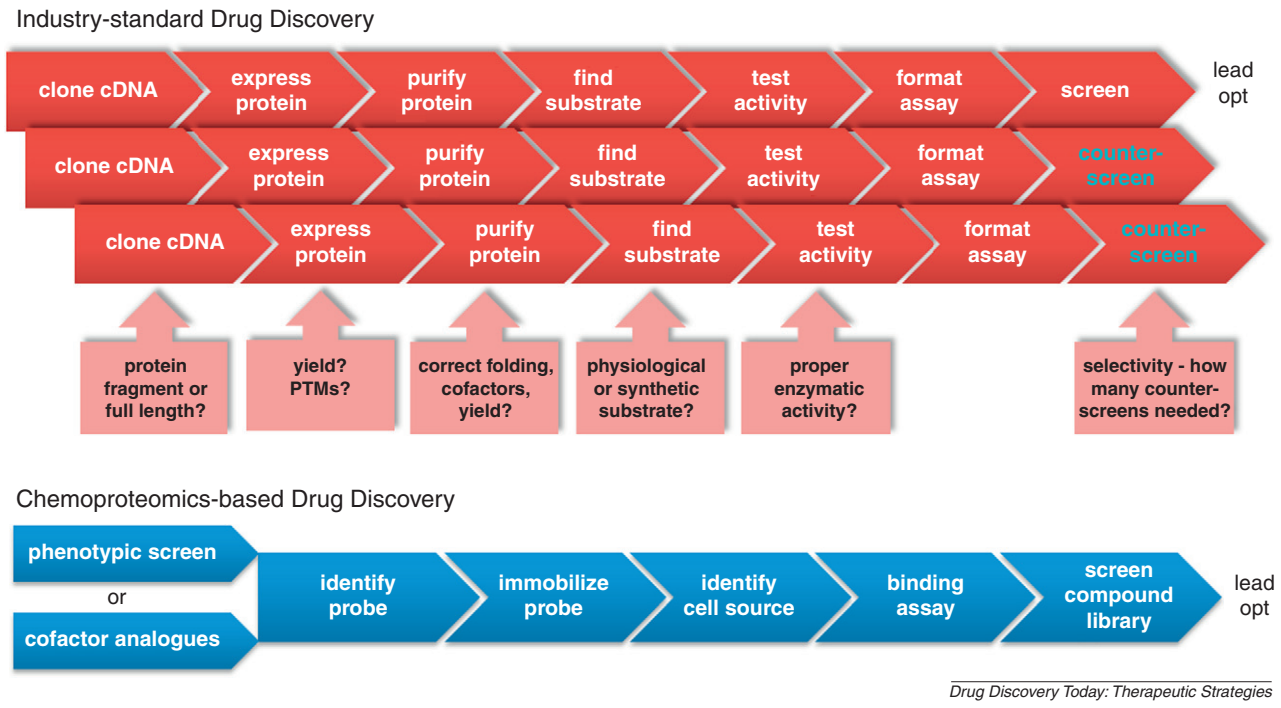

Fig. 2. Industry-standard and chemoproteomics drug discovery. Industry standard assay cascades for small molecule screening consist of multiple target-based assays. For each target, a sequence of steps is carried out involving the production of the target protein and the formatting of the assays supporting screening and lead optimization. In each step, the researcher may encounter particular technical hurdles as indicated. Chemoproteomics drug discovery is conceptually different as it is based on probe compounds which are used in binding assays together with native proteins from whole cell extracts or cell fractions. The sequence of steps shown is sufficient to provide assays for all proteins binding to the probe. For many target classes, probes can be based on cofactors or cosubstrates, allowing the use of a single probe for many different targets.

screen the large number of proteins needed to cover the major epigenetic protein classes, and it can be expected that many proteins will present problems with respect to recombinant expression and purification in an active form. Therefore, alternative and complementary strategies for the discovery of chemical probes and drug leads should be considered, and recently powerful strategies based on Chemical Genomics and Proteomics are emerging with the potential of discovering chemical probes, and drug candidates, with or without a bias for a defined target class (Fig. 2).

Discovery of chemical probes by cell-based screening and target deconvolution

Cell-based screens with phenotypic or mechanistic readouts are a useful complement to target-based strategies, and because they are agnostic with respect to the type or class of target, they may simultaneously enable the discovery of both chemical probes and new drug targets [15]. Notably, many approved drugs were in fact discovered by phenotypic approaches [16]. The major drawback of this approach is that the optimization and development of a lead compound is difficult if the molecular target is not established [17]. The retrospective identification of the target proteins responsible for the observed phenotypic response is often termed target deconvolution. Knowledge of a compound's target is not only important for understanding the biological or disease mechanisms, including potential target-based toxicities, but also enables Medicinal Chemists to understand the structure-activity relationship (SAR) necessary for rational drug design. Most direct target deconvolution approaches rely on some variation of affinity chromatography combined with protein mass spectrometry. This approach was pioneered in the seminal studies on the targets of immunosuppressants [18,19] and inhibitors of histone acetylation [20] by Stuart Schreiber and colleagues in the 1990s. Notably, Paul Marks and Ronald Breslow discovered one of the first approved epigenetic drugs, SAHA (vorinostat), when they studied compounds which induced differentiation and growth arrest of erythroleukemia cells [21]. It took several years before it was realized that these compounds act by inhibiting histone deacetylases (HDACs) [22,23].

Phenotypic screens can be carried out with smaller focused compound sets biased towards particular structures or targets, or without any target bias using highly diverse large compound libraries [24]. The use of large libraries has the potential to enable the identification of new and often unexpected target classes. Recent studies have successfully employed Chemical Biology-based strategies combining the screening of structurally diverse compound collections in cell-based transcriptional reporter assays with subsequent

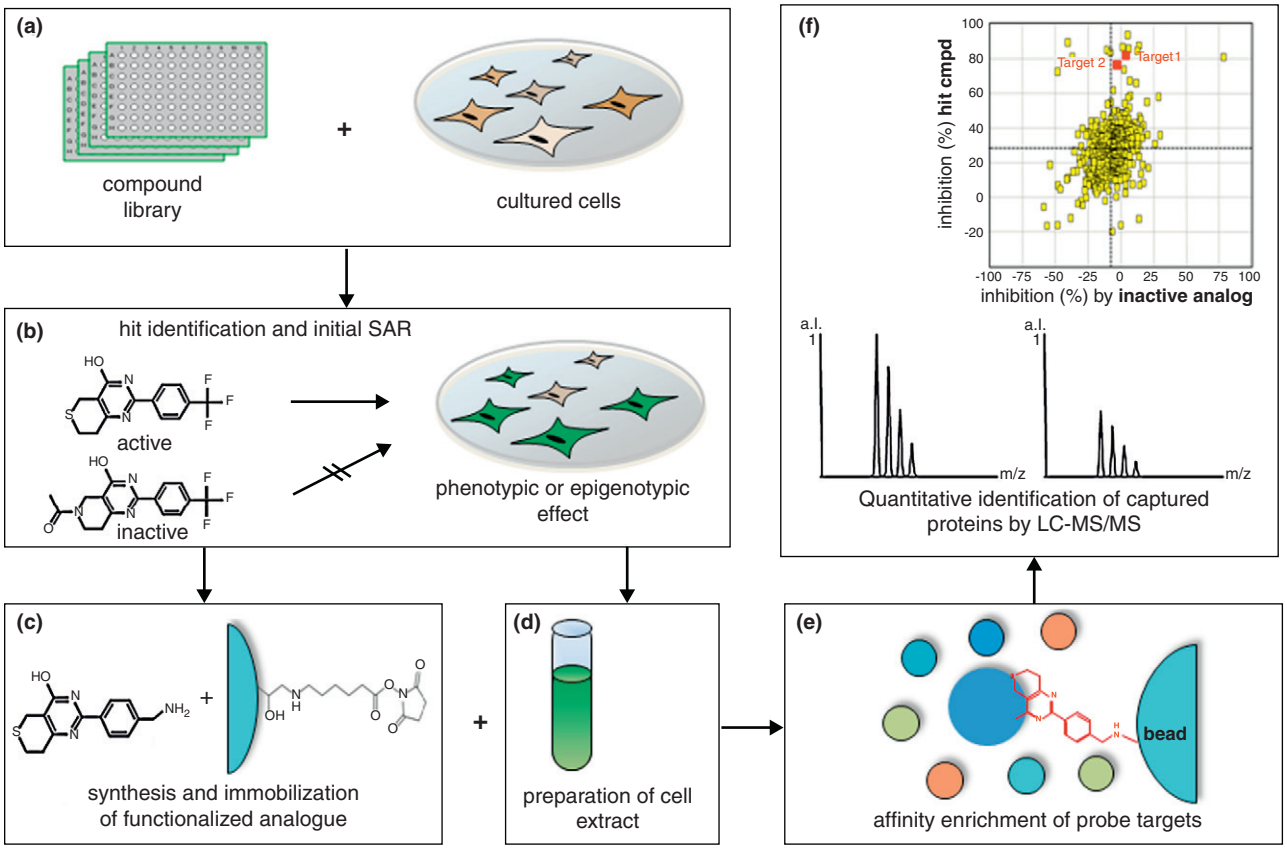

Fig. 3. Discovery of chemical probes by cell-based screening and target deconvolution. Cell-based screens with phenotypic or mechanistic readouts are agnostic with respect to the type or class of target and provide a useful complement to target-based strategies. Combined with target deconvolution, they may simultaneously enable the discovery of both chemical probes and new drug targets. **(a)** A compound library is screened in a cell-based assay, for instance for the induction of a particular gene. **(b)** The active structures obtained can often be used for a first analysis of the structure-activity relationship (example taken from [26]), which may guide the design of a probe compound functionalized for linker attachment **(c)**. The immobilized or biotin-tagged probe is then combined with the cell extract **(d)** allowing capturing of target proteins by the immobilized probe **(e)**. This step is typically conducted in the absence and presence of excess free active and inactive compounds. **(f)** Quantitative mass spectrometry is ideally suited to determine the identity and quantity of the captured proteins. Targets are identified by the fact that they are competed by excess of active compound but remain unaffected by excess inactive analogue.

chemoproteomics-based target identification (Fig. 2). Emami *et al.* screened colon carcinoma cells with a secondary structure-templated compound library for inhibitors of β-catenin/TCF dependent transcription, leading to the identification of a compound that selectively induced apoptosis in transformed cells but not in normal cells. Subsequent affinity capture with a biotinylated derivative of the active compound led to the identification of the transcriptional coactivator cAMP response element-binding protein (CBP) as the target [25]. A derivative of this compound has entered clinical development for colon cancer. Huang *et al.* reported the screening of a large diverse library for inhibitors of a cell-based β-catenin-dependent transcriptional reporter, combined with affinity-capture and mass spectrometry-based target deconvolution, and identified the Poly (ADP-ribose) polymerase 5 (PARP5, tankyrase) as the target of a novel small molecule inhibitor of the Wnt pathway, which plays an important role in colorectal cancers and is characterized by a lack of tractable targets [26]. Using a comparable strategy, Nicodeme *et al.* performed a screen for compounds with the ability to upregulate Apolipoprotein A-I, which is associated with cholesterol clearance, protection from atherosclerosis and anti-inflammatory effects, using a hepatocyte line with an APOA1 luciferase reporter. Target deconvolution of hit compounds led to the discovery of BET bromodomain proteins as targets for a new class of small molecules which can be viewed as histone mimetics, because they block the protein–protein interaction formed between acetylated histones and the bromodomain [27,28]. These recent studies demonstrate the general value of a generic strategy combining a cell-based phenotypic or transcriptional screen with target deconvolution (Fig. 3). It appears conceivable that also cell-based screens with mechanistic or epigenetic readouts, for example DNA or histone modification, could be

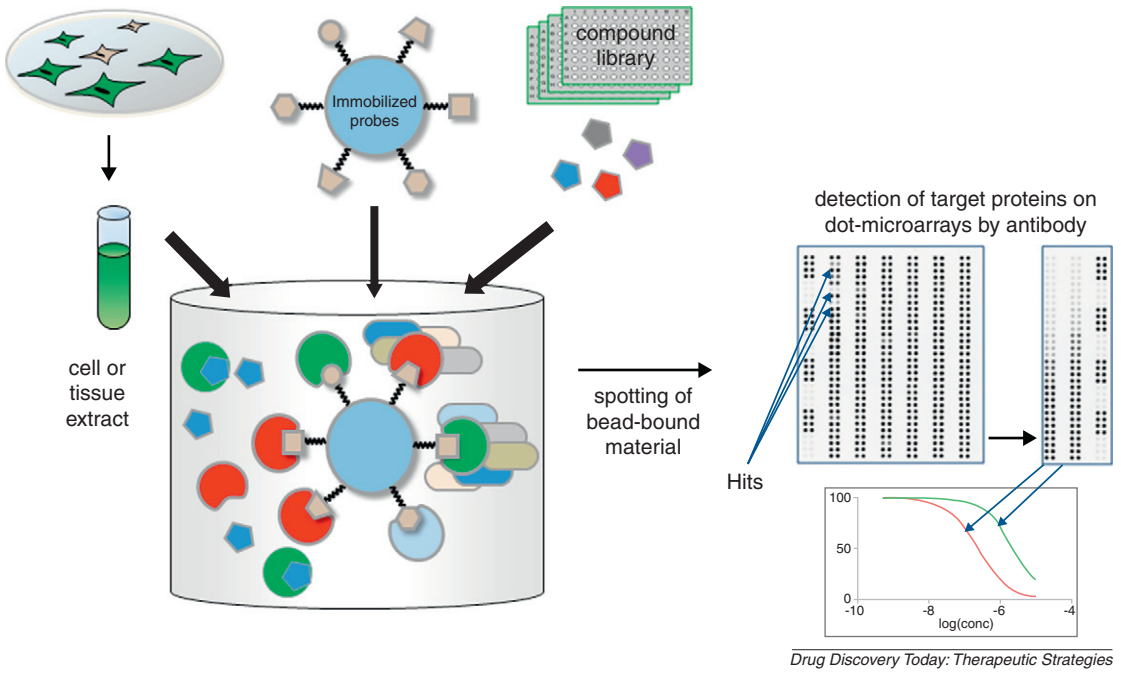

Fig. 4. Chemoproteomics-based screening of small molecule libraries. One or more chemical probes discovered by phenotypic or target-based approaches are immobilized on a bead matrix. The matrix is added to multiwell plates containing a suitable cell extract or cell fraction, and one compound each per well of a suitable chemical library. Beads are removed from the mixture, for example by filtration, captured proteins are eluted with detergent, and spotted as microarrays. The arrays are subsequently developed with fluorescently labeled antibodies for the target protein. Active compounds are evident by a reduction of signal (courtesy of Friedrich Reinhard and Gitte Neubauer, Cellzome).

employed. The recent technical advances in protein mass spectrometry providing ever greater sensitivity and more robust protein quantification can also be expected to contribute to a wider applicability of these strategies [29,30]. Because cell-based target deconvolution allows for the simultaneous discovery of targets and chemical probes, they can be seamlessly extended into target-based approaches by employing the binding assays based on the probe compounds for the discovery of drug candidates, as described below (Fig. 4).

Targeting multiprotein complexes and protein–protein interactions in epigenetic regulation

The chemical-probe-based proteomic approaches developed for target deconvolution can be turned into a generic methodology for the discovery of lead molecules which is independent from the use of purified proteins from recombinant sources. Because many epigenetic targets reside in large protein complexes, these techniques may be particularly advantageous in this research area. For instance, it has been reported that the methyltransferase activity of the EZH2 complex requires a minimum of three components, EED, EZH2 and SUZ12 [31], and that purified class I HDACs exhibit increased activity in the presence of interacting proteins, with HDAC3 essentially inactive in its monomeric form [32,33]. Recent proteomics studies based on stable isotope labeling with amino acids in cell culture (SILAC) and quantitative mass spectrometry have initiated the cell-wide characterization of chromatin-associated protein complexes [34]. In these studies, protein complexes were further characterized based on their preference for defined nucleosome modifications [35] and chromatin immunoprecipitation profiling was employed to map their downstream transcriptional targets [36]. Quantitative mass spectrometry-based strategies can also be applied to directly characterize the drug-binding properties of such protein complexes, and to render these complexes accessible to the screening and profiling of small molecule inhibitors. A recent study reported binding inhibition data for a set of 16 HDAC inhibitors against HDAC complexes, utilizing a test system comprising the inhibitor of choice, an immobilized hydroxamate probe, and a crude cell extract [37,38]. In this system, the test inhibitor competes with the nonspecific immobilized probe for the binding to the target HDACs, and the degree of competition was measured for each target in the cell extract by quantitative mass spectrometry. Because complexes were preserved in the cell extract, a set of IC_{50} or K_d (dissociation constant) values for each inhibitor against the catalytic and noncatalytic subunits of the HDAC complexes was provided. Remarkably, the selectivity profiles differed from profiles based on purified proteins, with some compounds actually exhibiting a higher degree of selectivity than previously reported. Unexpected

differences for some types of inhibitor with respect to the inhibition of HDAC1 and HDAC2 in different protein complexes were observed, such that a clustering of targets according to their inhibition across the set of 16 compounds yielded a co-clustering of proteins into complexes. This demonstration that small molecules can exhibit differences in selectivity for endogenous protein targets and may even discriminate between protein complexes suggests that the concept of an epigenetic drug target may need to be refined. It should be noted that the HDAC inhibitors in this study were developed without any notion of this complexity. Therefore, it appears attractive to employ chemoproteomic approaches directly in the screening of compound libraries [37,39]. This strategy addresses target potency and potential selectivity in a more physiological situation early in the discovery process, while at the same time reducing the requirement for the expression and purification of active protein in the milligrams to grams amounts usually required in high throughput screening (Fig. 2). Notably, there are many types of protein targets where the protein cannot be expressed at high levels, does not fold properly, or requires other endogenous factors or modifications for proper activity.

Any epigenetics discovery approach based on the proteomic characterization of protein complexes may also provide important clues to target biology and target validation. Recently, Dawson *et al.* characterized the target profile of a small molecule BET bromodomain inhibitor (I-BET), by using a combination of drug affinity capture and immunopurification of target complexes [40]. They found that I-BET inhibited the binding of several distinct BET complexes to histone peptides. In particular, I-BET inhibited the chromatin association of complexes comprising the BET proteins Brd3 and Brd4 and the so-called superelongation (SEC) and polymerase-associated (PAFc) complexes. These complexes contain components which are subject to proleukemic fusion events in mixed-lineage leukemia (MLL), a disease with currently limited treatment options. The data suggested that I-BET should inhibit the expression of genes involved in MLL pathogenesis and should have therapeutic potential, and the study provided data to support this hypothesis including demonstration of efficacy in patient-derived cells and in animal studies.

Finally, it is tempting to speculate that the availability of native protein complexes in assay formats compatible with small molecule screening and profiling may reinvigorate the discovery of inhibitors of protein–protein interactions, especially when combined with diversity-oriented synthesis (DOS) and other combinatorial small molecule strategies [41,42].

Conclusions

Recent advances in chemical biology and in proteomics have promoted new Drug Discovery strategies based on assays utilizing native proteins or protein complexes providing a better appreciation of the molecular context in which protein targets operate. In particular when combined with cell-based screening based on RNA interference or small molecules, these approaches should provide important complementary approaches to epigenetic drug target identification, selectivity profiling, and lead finding, and have the potential to substantially contribute to *in vivo* studies and clinical studies of drug-target interactions.

Acknowledgements

I would like to thank my Cellzome colleagues, in particular Marcus Bantscheff, Carsten Hopf, Gitte Neubauer and David Simmons, for many stimulating discussions, Alan Watt for comments on the manuscript, and Frank Weisbrodt for help with the figures.

References

1 Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395
2 Heightman, T.D. (2011) Therapeutic prospects for epigenetic modulation. *Expert. Opin. Ther. Targets* 15, 729–740
3 Alberts, B. (1998) The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* 92, 291–294
4 Malovannaya, A. *et al.* (2011) Analysis of the human endogenous coregulator complexome. *Cell* 145, 787–799
5 Margueron, R. and Reinberg, D. (2010) Chromatin structure and the inheritance of epigenetic information. *Nat. Rev. Genet.* 11, 285–296
6 Low, J. *et al.* (2008) Prioritizing hits from phenotypic high-content screens. *Curr. Opin. Drug Discov. Dev.* 11, 338–345
7 Quon, K. and Kassner, P.D. (2009) RNA interference screening for the discovery of oncology targets. *Expert Opin. Ther. Targets* 13, 1027–1035
8 Zuber, J. *et al.* (2011) Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat. Biotechnol.* 29, 79–83
9 Zuber, J. *et al.* (2011) RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* 478, 524–528
10 Jackson, A.L. and Linsley, P.S. (2010) Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat. Rev. Drug Discov.* 9, 57–67
11 Knight, Z.A. and Shokat, K.M. (2005) Features of selective kinase inhibitors. *Chem. Biol.* 12, 621–637
12 Copeland, R.A. *et al.* (2009) Protein methyltransferases as a target class for drug discovery. *Nat. Rev. Drug Discov.* 8, 724–732
13 Rose, N.R. *et al.* (2011) Inhibition of 2-oxoglutarate dependent oxygenases. *Chem. Soc. Rev.* 40, 4364–4397
14 Karaman, M.W. *et al.* (2008) A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* 26, 127–132
15 Feng, Y. *et al.* (2009) Multi-parameter phenotypic profiling: using cellular effects to characterize small-molecule compounds. *Nat. Rev. Drug Discov.* 8, 567–578
16 Swinney, D.C. and Anthony, J. (2011) How were new medicines discovered? *Nat. Rev. Drug Discov.* 10, 507–519
17 Terstappen, G.C. *et al.* (2007) Target deconvolution strategies in drug discovery. *Nat. Rev. Drug Discov.* 6, 891–903
18 Brown, E.J. *et al.* (1994) A mammalian protein targeted by G1-arresting rapamycin-receptor complex. *Nature* 369, 756–758
19 Harding, M.W. *et al.* (1989) A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase. *Nature* 341, 758–760
20 Taunton, J. *et al.* (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* 272, 408–411
21 Richon, V.M. *et al.* (1996) Second generation hybrid polar compounds are potent inducers of transformed cell differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 93, 5705–5708

22 Marks, P.A. (2010) The clinical development of histone deacetylase inhibitors as targeted anticancer drugs. *Expert Opin. Investig. Drugs* 19, 1049–1066
23 Marks, P.A. and Breslow, R. (2007) Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nat. Biotechnol.* 25, 84–90
24 Mayr, L.M. and Bojanic, D. (2009) Novel trends in high-throughput screening. *Curr. Opin. Pharmacol.* 9, 580–588
25 Emami, K.H. *et al.* (2004) A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. *Proc. Natl. Acad. Sci. U. S. A.* 101, 12682–12687
26 Huang, S.M. *et al.* (2009) Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* 461, 614–620
27 Chung, C.W. *et al.* (2011) Discovery and characterization of small molecule inhibitors of the BET family bromodomains. *J. Med. Chem.* 54, 3827–3838
28 Nicodeme, E. *et al.* (2010) Suppression of inflammation by a synthetic histone mimic. *Nature* 468, 1119–1123
29 Domon, B. and Aebersold, R. (2010) Options and considerations when selecting a quantitative proteomics strategy. *Nat. Biotechnol.* 28, 710–721
30 Mallick, P. and Kuster, B. (2010) Proteomics: a pragmatic perspective. *Nat. Biotechnol.* 28, 695–709
31 Pasini, D. *et al.* (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* 23, 4061–4071
32 Guenther, M.G. *et al.* (2001) The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. *Mol. Cell Biol.* 21, 6091–6101
33 Zhang, Y. *et al.* (1999) Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* 13, 1924–1935
34 Eberl, H.C. *et al.* (2011) Quantitative proteomics for epigenetics. *Chembiochem* 12, 224–234
35 Bartke, T. *et al.* (2010) Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143, 470–484
36 Vermeulen, M. *et al.* (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967–980
37 Bantscheff, M. *et al.* (2011) Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat. Biotechnol.* 29, 255–265
38 Holson, E.B. and Schreiber, S.L. (2011) Chemoproteomics quantifies complexity. *Nat. Biotechnol.* 29, 235–236
39 Ramsden, N. *et al.* (2011) Chemoproteomics-based design of potent LRRK2-selective lead compounds that attenuate Parkinson's disease-related toxicity in human neurons. *ACS Chem. Biol.* 6, 1021–1028
40 Dawson, M.A. *et al.* (2011) Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* 478, 529–533
41 Koehler, A.N. (2010) A complex task? Direct modulation of transcription factors with small molecules *Curr. Opin. Chem. Biol.* 14, 331–340
42 Wells, J.A. and McClendon, C.L. (2007) Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* 450, 1001–1009