EMBO reports

viewpoint

The case for open-access chemical biology

A strategy for pre-competitive medicinal chemistry to promote drug discovery

Johan Weigelt

Policy-makers and scientists are increasingly worried about the declining productivity of the pharmaceutical industry. Despite growing levels of investment from governments and industry, the number of new medicines that are approved each year has declined slowly during the past three decades (Betz, 2005; Booth & Zemmel, 2004; Cuatrecasas, 2006). The increased costs of drug discovery and development can no longer be supported by healthcare systems, and regulators are implementing control mechanisms such as price control schemes and increased regulatory demands, which could ultimately affect the drug discovery process by altering the financial and regulatory landscapes for new research programmes (Ess *et al*, 2003; Miller & Henderson, 2007).

The increased costs of drug discovery and development can no longer be supported by healthcare systems…

Over the years, the industry has relied on various technologies to improve drug discovery. However, despite the use of structure-based drug design, high-throughput screening, combinatorial chemistry and the various '-omics' technologies, productivity continues to decline. The various explanations for this lack of success include increasing regulatory requirements, the depletion of therapeutic opportunities and 'druggable' targets, competition and organizational inefficiencies (Booth & Zemmel, 2004; Cuatrecasas, 2006; Dickson & Gagnon, 2004; Garnier, 2008). Although each of these explanations rings true, at the heart of the problem is a lack of proper scientific understanding of human biology and disease mechanisms. This is ultimately manifested in the high fraction of clinical programmes that fail owing to a lack of efficacy or safety, often in spite of excellent pre-clinical data (Fingleton, 2008; Kola & Landis, 2004).

Without a doubt, improved drug-target validation would improve the success rate of drug discovery programmes, but this necessitates a better understanding of fundamental biological processes and the biological role of the drug target—usually a protein—under investigation. Paradoxically, the consolidation within the pharmaceutical and biotechnology industries during the past 10–15 years has not sufficiently advanced this understanding. The industry is retrenching from early-stage drug discovery and the fundamental exploration of novel therapeutic opportunities, and instead putting most of its efforts into the later stages of the drug discovery process and on clinically validated therapeutic opportunities. This paradox has not gone unnoticed; many authors and organizations are advocating the need to increase collaboration between industrial and academic researchers to allow new discoveries (Brewer, 2006; Garnier, 2008; Hughes, 2008).

Modulating the function of proteins through the use of chemical compounds is a powerful method to explore their physiological relevance. It also helps to validate drug targets with an experimental set-up that resembles more closely the clinical setting, as compared with, for example, genetic validation tools. Many universities have established screening centres with the aim of providing their researchers with access to chemical inhibitors to study biological processes. In many cases, these centres have also been advertised as a contribution to early-stage drug discovery by providing starting points to develop medicines using new biological targets.

Following the traditions of industry, universities now also tend to protect their intellectual property rights to small molecules to allow future commercial exploitation. However, mimicking pharmaceutical and biotechnology companies in both their commercial focus and choice of technologies might be futile; 'industry-styled' academic drug discovery programmes are just as likely to fail as industrial ones, ultimately with the same costs. Academic drug discovery is also less likely to provide rapid access to the best chemical tools, as many of the compounds will be subject to patent protection and encumbered by restrictive material transfer agreements.

…at the heart of the problem is a lack of proper scientific understanding of human biology and disease mechanisms

In the light of this, I therefore argue that academic approaches to chemical biology would have a greater impact on drug discovery if they focused predominantly on increasing our understanding of human biology—with more emphasis on open-access target validation and less emphasis on the commercial potential of any discovery. At the same time, if academia turns its attention to target validation and operates in an open manner, then industry should, in turn, facilitate the process by providing access to its resources, experience and expertise. In this regard, early-stage drug discovery research provides an ideal opportunity for closer collaboration in basic research.

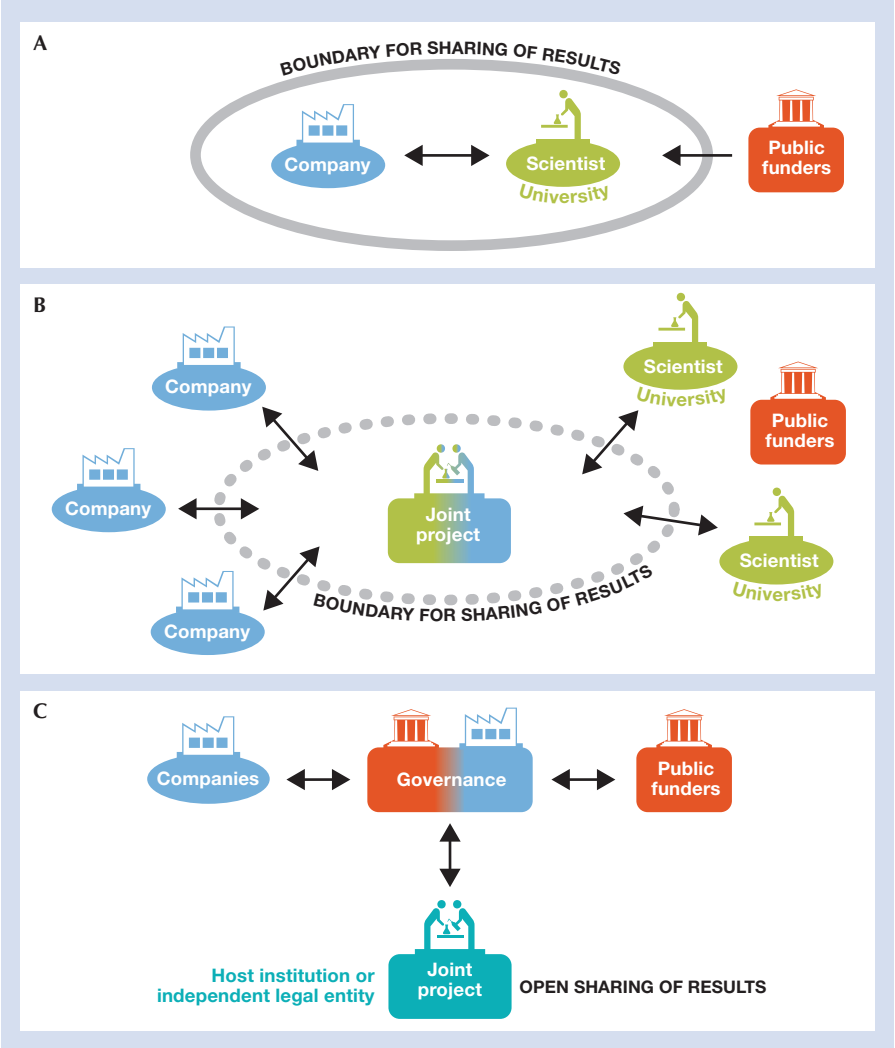

Fig 1 | Three models of the interaction between pharmaceutical companies and academia. (**A**) The traditional model comprises a closed collaboration between a company and an academic researcher. (**B**) The closed consortium model in which several parties contribute to a joint project. Results and/or intellectual property rights are normally retained by consortium members. (**C**) The open consortium model in the form of a public–private partnership. A joint project, governed by its funders, is carried out at a host institution or as a separate legal entity. Results are made freely available to allow research in the public domain.

Academics and partners from the pharmaceutical industry already have a long history of cooperation on various projects using a range of collaborative models. That said, new collaborative schemes and methods of interacting are constantly being developed and it is important to consider which of the current models is best suited to early-stage drug discovery. Generally, one can distinguish among three types of partnership between industry and academia (Fig 1).

The traditional model comprises a partnership between a single company and a single academic researcher and/or host institution. This type of collaboration is often highly focused on a specific problem in which the academic partner has significant expertise. It is a closed model in which the company usually retains the intellectual property rights, and the academic partner is compensated for services and sometimes given rights to license payments and royalties from future income. In essence, the closed model can be regarded as a mere outsourcing of research activities to a third party, and the research community as a whole does not generally benefit. Examples of this type of public–private collaboration are numerous and still represent the main model for interactions between industry and academia.

In a closed consortium model, one or more companies enter a joint project or collaboration with one or several academic partners. Such projects often have a broader focus or relate to an area of technology in which academic and industrial researchers have mutual interests. Data sharing within such a project is normally open to all participants but not to outsiders. Intellectual property rights are often tightly regulated and first-right-of-refusal mechanisms are often put in place to protect the interests of the industrial partners. Results are usually published, which benefits the research community, but some findings are withheld by consortium members. Examples of such joint ventures include the Dundee Division of Signal Transduction Therapy Consortium (www.lifesci.dundee.ac.uk/dstt) and the Biomarkers Consortium (www.biomarkersconsortium.org), which is managed by the US National Institutes of Health (NIH; Bethesda, MA, USA).

In the open consortium model, one or more companies collaborate with public partners to explore jointly a scientific question. Projects are funded collectively and research activities are normally carried out at a host institution and/or as an independent legal entity. Data are shared openly among consortium members and made publicly available with no restriction on use. Such models are rare, however: some examples are the Single Nucleotide Polymorphism Consortium (Holden, 2002), the Structural Genomics Consortium (SGC; Cottingham, 2008; Gileadi *et al*, 2007) and the international Severe Adverse Events Consortium (www.saeconsortium.org).

Each model has its advantages and disadvantages, and the choice of which to use will depend on the scientific question at hand. Specific questions are probably best dealt with by outsourcing research to a single investigator, whereas larger projects might be better pursued through a consortium. If we postulate that an increased focus on the basic science of drug discovery will be of great benefit to the pharmaceutical industry, there is a strong case for

declaring that chemical biology is pre-competitive research and thus opting for an open consortium model.

Although the creation of new chemical entities has always been considered the realm of patents, I think that it is time for change. Novel chemical tools, most of which will not have drug-like properties, are too valuable to be restricted; they will be of far greater benefit to research if freely available without restrictions on their use. Chemical biologists would benefit from the many advantages that the open consortium model brings: rapid access to research tools; less bureaucratic workload to enter legal agreements; the ability to work with the best people through collaborations focused on the publication of results; and freedom to operate for companies, harnessing the synergies between academic freedom and industrial approaches to systematically tackle a scientific challenge. My call for open-access chemistry public–private partnerships might sound impractical, but pilot projects are already underway.

…'industry-styled' academic drug discovery programmes are just as likely to fail as industrial ones, ultimately with the same costs

The SGC is a one example of an open public–private partnership. It was created as a legal charity in 2004 to determine the three-dimensional high-resolution structures of medically important proteins. As an open consortium, the resulting structures are placed in the public domain without restriction on their use. The SGC was conceived nearly ten years ago, based on the conviction that high-quality structural information is of tremendous value in promoting drug discovery and a belief that patenting protein structures could limit the freedom to operate for academic and industrial organizations. The consortium includes laboratories at three institutions—the University of Oxford (UK), the University of Toronto (Canada) and Karolinska Institutet (Stockholm, Sweden)—and has an annual budget of approximately US$25 million, more than 15% of which comes from industry. Three pharmaceutical companies—GlaxoSmithKline (Brentford, UK), Merck (Whitehouse Station, NJ, USA) and Novartis (Basel, Switzerland)—are supporting the project. The funders set the general direction of the project by providing the research objectives—in practice, by listing protein targets—and by setting quantitative milestones. The operations of the SGC are overseen by a Board of Directors with an independent Chairman, currently Wayne Hendrickson of Columbia University (New York, NY, USA), and a Scientific Committee. Both meet quarterly and each funder has a representative or a delegate on each committee. The mandate of the SGC prescribes that all protein structures be placed promptly in the public domain, and not even the project sponsors receive any prior rights or exclusive access to data and results.

As both funders and non-funders have contemporaneous access to the results generated by the SGC and are free to use those results as they wish, one might ask what motivates a commercial organization to join the consortium? There are at least five easily identifiable reasons: consortium members can influence the direction of the research, which allows them to maximize synergies with internal research programmes; the investment of each corporation is leveraged by public funding; consortium members are more likely to rapidly pick up on important findings; the support from the public sector crucially depends on contributions from industry and, without industrial sponsors, the likelihood of public funding for the project would decrease significantly; and consortium members gain direct insights into the value of various technologies and can more efficiently exploit this knowledge.

On a quantitative level, the model is a success: the SGC is now a leading contributor of human protein structures to the Protein Data Bank (www.pdb.org). The open-access mandate has also facilitated collaborations with leading academics; the project has published more than 150 scientific articles, most of which have resulted from collaborations with outside researchers (Avvakumov *et al*, 2008; Barr *et al*, 2009; Baumli *et al*, 2008; Filippakopoulos *et al*, 2008; Lunin *et al*, 2006; Ng *et al*, 2007; Schuetz *et al*, 2006). In essence, the SGC represents an organizational paradigm of how academia and industry can collaborate to pursue basic research with commercial relevance.

One could argue that protein structural information is clearly 'pre-competitive' and as long as chemical tool compounds have potential commercial value, the SGC model is irrelevant. It is therefore important to note that when the SGC was formed, the merits of releasing structural information into the public domain were not widely accepted owing to its potential use in developing commercial products. Over time, however, both academia and industry realized that the benefits of having more structures freely available outweighed any potential commercial disadvantage. Now, protein structures are widely regarded as pre-competitive information. I would like to make a similar argument for early-stage chemistry and the development of tool compounds for drug targets.

Tool compounds—'chemical probes'—are valuable for studying basic human biology and are complementary to other methods such as transgenic/knockout animals or RNA interference experiments. The advantage of tool compounds is their ability to permeate cells, their systemic distribution and their immediate action, which make them attractive reagents for cellular and *in vivo* pharmacology. Moreover, if the probes are well characterized and specific for a given target, they might also act as starting points for drug discovery programmes. Unfortunately, high-quality chemical tool compounds are scarce in the academic sector; most chemicals that are available to academic researchers through commercial vendors are poorly characterized in terms of specificity and cell permeability. Recognizing that a major strength of the pharmaceutical industry is its ability to generate potent, selective and bio-available inhibitors or agonists of protein function, it would seem natural to carry out chemical probe generation in collaboration with industry. Realistically, the best strategy to increase knowledge is to engage the wider academic community. This can only be achieved by releasing the chemical probes freely into the realm of creative commons.

As a pilot project for 'open-access chemical biology', GlaxoSmithKline, the NIH Chemical Genomics Center, the SGC, the Universities of Oxford and Toronto, and a group of academic chemists with funding from the Wellcome Trust and the Ontario government have established a project to develop chemical probes with cellular activity for targets implicated in epigenetic signalling (www.thesgc.org/epigenetics; Fig 2). Eight medicinal chemists from GlaxoSmithKline and several experienced medicinal chemists from academia are collaborating to provide the community with high-quality reagents that can be used without restriction and purchased through commercial vendors. GlaxoSmithKline, which

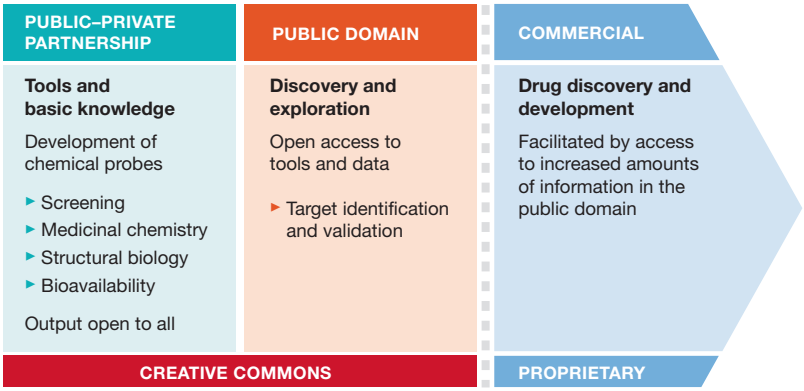

Fig 2 | Model of the open-access chemical biology consortium. The public–private partnership (PPP) is mandated to create chemical probes that target proteins involved in epigenetic signalling. Results—tools and data—are shared freely to facilitate further exploration and new discoveries. The increased knowledge will allow commercial projects at a later stage with an increased chance of success. The model can also be viewed as a general scheme for pre-competitive public–private collaboration. The Structural Genomics Consortium was originally created as a structural genomics PPP following the same principles.

agreed to generate and release new chemical matter without restriction, is championing the concept that the resulting knowledge will potentially lead to new concepts and/or targets for therapeutic intervention.

Within the field of epigenetics, our understanding of the molecular mechanisms has grown rapidly over the past years, and it is clearly an area of potential therapeutic relevance. The prospects for inhibitor-based therapeutic intervention seem promising; DNA methyltransferase inhibitors and a histone deacetylase inhibitor have been approved for use in the treatment of certain cancers (Gore *et al*, 2006; Kaminskas *et al*, 2005; Mann *et al*, 2007). This suggests that compounds that modulate other proteins and enzymes that read, write or erase epigenetic marks might also be of pharmaceutical interest. However, our current level of knowledge is not sufficient to determine which of the hundreds of epigenetic signalling proteins are suitable targets for therapeutic intervention. Chemical probes that specifically target epigenetic proteins will therefore help to elucidate their roles in human physiology and disease, and to identify the most promising targets for pharmacological modulation of disease states. Clearly, GlaxoSmithKline will not benefit exclusively from this expansion of knowledge, but the alternative—to rely on internal resources or exclusive collaborations with only a few academics—is perceived as a less effective approach to the same goal.

Academic institutions have been keen to be involved in chemical biology because they appreciate the value of tool compounds in basic science. To this end, many universities are building research capacity by establishing high-throughput, high-content screening centres and traditional drug discovery operations. Their stated purpose is to harness the capabilities of chemistry to advance biological understanding—in keeping with their academic mandate.

> **The paradigm of keeping secret potentially valuable tool compounds for poorly validated targets […] does not stimulate scientific discovery and is ultimately detrimental to the public good**

As examples, several chemical screening centres have been established in the USA as part of the NIH Molecular Libraries Initiative (http://nihroadmap.nih.gov/molecularlibraries). Primary screening data are made publicly available to allow further activity in the public and private sectors. In Europe, discussions are underway to launch an OpenScreen initiative to create the infrastructure needed for high-throughput screening and the development of biologically active compounds (www.fmp-berlin.de/eu-openscreen.html). Many universities have also established core facilities for chemical biology that combine screening platforms and medicinal chemistry.

These initiatives are laudable, yet they may not achieve the maximum benefit. Although the efforts were designed to enable academic research, many of the operations were 'sold' to the institutions, the funding bodies or legislatures based on potential returns on investment through the development of clinical drug candidates. To our knowledge, in all cases, follow-up chemistry is being kept secret to protect intellectual property rights. As such, the centres face a potential inherent dilemma: they seek to maximize global knowledge, which, in my view, is best achieved by making high-quality chemical matter freely available; conversely, they seek financial gains by protecting high-quality chemical matter and restricting access to it. This duality of purpose sends mixed messages to the funders and scientists, delays the release of information, and often encumbers the distribution of any patented reagent with restrictive and time-consuming legal agreements.

Drug development requires the ability to generate potent, selective, non-toxic, biologically active compounds with acceptable pharmacokinetic and pharmacodynamic properties. Indeed, once a protein target has been validated, the creativity of the industry to develop new and/or improved drugs seems almost limitless—in this regard, chemical tool compounds, which do not usually have drug-like properties and occupy only a small region in the vast space of chemistry, have little commercial value. On the contrary, one of the main hurdles in drug discovery is to determine which protein targets are the best points of intervention. With this in mind, it seems appropriate to raise concerns about current developments within the academic sector and its apparent desire to develop drugs or clinical candidates. These academic initiatives should not try to copy what is done in industry because even the most experienced commercial entities encounter difficulties in developing new drugs. The paradigm of keeping secret potentially valuable tool compounds for poorly validated targets to retain the potential—and usually imaginary—commercial benefits does not stimulate scientific discovery and is ultimately detrimental to the public good.

Although it will be nigh on impossible to turn around the 'academic drug discovery'

supertanker, I suggest that there is an opportunity to launch a different ship. If academic medicinal chemistry initiatives were truly interested in increasing the number of new medicines, they should move away from the main model in which potential commercialization opportunities have a crucial role. Instead, the focus should be on allowing target validation by taking an open-access route. High-quality, well-characterized reagents that are made available to the scientific community could be expected to markedly change the process of target validation by engaging the wider community.

Although it is clear that open-access chemistry is in the best interests of society, the challenge is the cost. My arguments can be defended on the macroeconomic level, but costs for assay development and for chemical screening and synthesis are incurred locally, by the institutions and from the public purse. Free release of chemical probes by academia would ultimately benefit the pharmaceutical industry and society, but the possibilities for royalty and license payments for universities would decrease. One solution is to explore models in which both the public and private sectors contribute up-front in return for unrestricted access to the results and compounds, as in the SGC. It should also be noted that an open-access model is not in conflict with the aim to commercialize, at least not in the long term. It could be argued that experience built around specific biological systems would allow commercial development at a later stage if findings by the community indicate that a particular protein or pathway is a valid target. A chemical biology centre with such experience would be in an ideal position to develop new chemistry and launch a proprietary programme.

CONFLICT OF INTEREST

The author declares no conflict of interest beyond his affiliation to the Structural Genomics Consortium as noted below.

ACKNOWLEDGEMENTS

The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and the Wellcome Trust. The author is grateful to Aled M. Edwards for critical review of the manuscript and productive discussions.

REFERENCES

Avvakumov GV, Walker JR, Xue S, Li YJ, Duan SL, Bronner C, Arrowsmith CH, Dhe-Paganon S (2008) Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* **455:** 822–825

Barr AJ, Ugochukwu E, Lee WH, King ONF, Filippakopoulos P, Alfano I, Savitsky P, Burgess-Brown NA, Muller S, Knapp S (2009) Large-scale structural analysis of the classical human protein tyrosine phosphatome. *Cell* **136:** 352–363

Baumli S, Lolli G, Lowe ED, Troiani S, Rusconi L, Bullock AN, Debreczeni JE, Knapp S, Johnson LN (2008) The structure of P-TEFb (CDK9/cyclin T1), its complex with flavopiridol and regulation by phosphorylation. *EMBO J* **27:** 1907–1918

Betz UA (2005) How many genomics targets can a portfolio afford? *Drug Discov Today* **10:** 1057–1063

Booth B, Zemmel R (2004) Prospects for productivity. *Nat Rev Drug Discov* **3:** 451–456

Brewer GJ (2006) Fundamental problems lie ahead in the drug discovery and commercialization process: restructuring of the pharmaceutical industry and an improved partnership with academia are required. *J Investig Med* **54:** 291–302

Cottingham K (2008) The structural genomics consortium makes its presence known. *J Proteome Res* **7:** 5073

Cuatrecasas P (2006) Drug discovery in jeopardy. *J Clin Invest* **116:** 2837–2842

Dickson M, Gagnon JP (2004) Key factors in the rising cost of new drug discovery and development. *Nat Rev Drug Discov* **3:** 417–429

Ess SM, Schneeweiss S, Szucs TD (2003) European healthcare policies for controlling drug expenditure. *Pharmacoeconomics* **21:** 89–103

Filippakopoulos P *et al* (2008) Structural coupling of SH2-kinase domains links fes and Abl substrate recognition and kinase activation. *Cell* **134:** 793–803

Fingleton B (2008) MMPs as therapeutic targets—still a viable option? *Semin Cell Dev Biol* **19:** 61–68

Garnier JP (2008) Rebuilding the R&D engine in big pharma. *Harv Bus Rev* **86:** 68–70, 72–76, 128

Gileadi O *et al* (2007) The scientific impact of the Structural Genomics Consortium: a protein family and ligand-centered approach to medically-relevant human proteins. *J Struct Funct Genomics* **8:** 107–119

Gore SD, Jones C, Kirkpatrick P (2006) Decitabine. *Nat Rev Drug Discov* **5:** 891–892

Holden AL (2002) The SNP consortium: summary of a private consortium effort to develop an applied map of the human genome. *Biotechniques* **June (Suppl):** 22–24, 26

Hughes B (2008) Pharma pursues novel models for academic collaboration. *Nat Rev Drug Discov* **7:** 631–632

Kaminskas E, Farrell AT, Wang YC, Sridhara R, Pazdur R (2005) FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. *Oncologist* **10:** 176–182

Kola I, Landis J (2004) Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov* **3:** 711–715

Lunin VV, Dobrovetsky E, Khutoreskaya G, Zhang RG, Joachimiak A, Doyle DA, Bochkarev A, Maguire ME, Edwards AM, Koth CM (2006) Crystal structure of the CorA Mg^{2+} transporter. *Nature* **440:** 833–837

Mann BS, Johnson JR, Cohen MH, Justice R, Pazdur R (2007) FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *Oncologist* **12:** 1247–1252

Miller HI, Henderson DR (2007) Governmental influences on drug development: striking a better balance. *Nat Rev Drug Discov* **6:** 532–539

Ng SS *et al* (2007) Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448:** 87–91

Schuetz A, Allali-Hassani A, Martin F, Loppnau P, Vedadi M, Bochkarev A, Plotnikov AN, Arrowsmith CH, Min JR (2006) Structural basis for molecular recognition and presentation of histone H3 By WDR5. *EMBO J* **25:** 4245–4252

Johan Weigelt is the Associate Director of the Structural Genomics Consortium and the Chief Scientist of the Structural Genomics Consortium laboratory at the Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden.
E-mail: johan.weigelt@ki.se

doi:10.1038/embor.2009.193