

# Chemical Genetics and Orphan Genetic Diseases

## Review

Mitchell R. Lunn<sup>1,3</sup> and Brent R. Stockwell<sup>1,2,\*</sup>

<sup>1</sup>Department of Biological Sciences

<sup>2</sup>Department of Chemistry

Sherman Fairchild Center for the Life Sciences

Columbia University

1212 Amsterdam Avenue, MC 2406

New York, New York 10027

Many orphan diseases have been identified that individually affect small numbers of patients but cumulatively affect ~6%–10% of the European and United States populations. Human genetics has become increasingly effective at identifying genetic defects underlying such orphan genetic diseases, but little progress has been made toward understanding the causal molecular pathologies and creating targeted therapies. Chemical genetics, positioned at the interface of chemistry and genetics, can be used for elucidation of molecular mechanisms underlying diseases and for drug discovery. This review discusses recent advances in chemical genetics and how small-molecule tools can be used to study and ultimately treat orphan genetic diseases. We focus here on a case study involving spinal muscular atrophy, a pediatric neurodegenerative disease caused by homozygous deletion of the *SMN1* (survival of motor neuron 1) gene.

### Introduction

There are ~6,000 diseases that individually affect fewer than 200,000 patients in the United States, but that cumulatively have profound effects on the health of many people ([http://rarediseases.info.nih.gov/html/reports/fy2003/FY2003\\_index.html](http://rarediseases.info.nih.gov/html/reports/fy2003/FY2003_index.html)). The pharmaceutical industry, faced with the large costs of drug discovery and development, generally neglects these diseases due to the paucity of patients. In 1983, the Orphan Drug Act (ODA) was enacted to promote the development of therapeutics for these orphan diseases by offering incentives to the pharmaceutical industry such as fee waivers, tax credits, and market exclusivity for 7 years [1]. Congress reevaluated the rare disease research environment in 2002 and discovered that, in spite of the positive outcomes from the ODA, “rare diseases and disorders [deserve] greater emphasis in the national biomedical research enterprise” [2].

In response to these conclusions, Congress passed the Rare Diseases Act (RDA) of 2002. The act increased government investment in the development of diagnostic tools and therapeutics for rare diseases by legislatively establishing the Office of Rare Diseases at the National Institutes of Health (NIH) and specifically designating appropriations for the study of rare diseases [2].

Many countries have followed the United States model for promoting the discovery and development of therapeutics for orphan diseases. However, the definition of a rare disease varies (Table 1).

With both the ODA and the RDA enacted, academia and industry were financially primed to identify, analyze, optimize, formulate, and market targeted therapeutics for rare diseases. Unfortunately, the scientific community possessed little knowledge about nearly all of these diseases. Recent scientific advances in many fields have, however, improved our knowledge of these diseases and our approach to understanding them. The advent of modern human genetics has revealed that many of the ~6,000 rare diseases are genetic in origin [3]. These diseases include Huntington’s disease, spinal muscular atrophy (SMA), amyotrophic lateral sclerosis (ALS, or Lou Gehrig’s disease), Tourette syndrome, Crohn disease, cystic fibrosis, cystinosis, Duchenne muscular dystrophy, thalassemias, hemophilia, and rare cancers, such as sarcomas.

One approach to understanding genetic diseases involves using high-throughput cellular assays to find organic compounds that interact functionally, not physically, with disease alleles. This approach, using small molecules to perturb protein networks of biological systems, is referred to as chemical genetics by analogy to model organism genetics (classical genetics). Chemical genetics is a potential route to elucidating the cellular and molecular mechanisms of disease pathogenesis. Here, we discuss how this approach has been applied to the pediatric autosomal recessive disease spinal muscular atrophy, and how small molecules may aid in understanding and treating orphan genetic diseases.

### Advances in Biology

In the late eighteenth century, plant extracts were shown to exhibit various effects on animals. Shortly after, Frederick Serturmer isolated morphine from the opium plant in 1804 [4], indicating that biologically active molecules do indeed exist in nature. When Paul Ehrlich discovered that methylene blue selectively stains neuronal cells, he revealed that small molecules have specific biological targets [5, 6]. He won the 1908 Nobel Prize in Medicine for his work, and the term he coined, receptor, has remained valuable ever since. Even though these concepts—which represent the theoretical keystones of chemical genetics—were first articulated more than 200 years ago, chemical genetics required many subsequent advances to reach its current status as a powerful, alternative approach to classical genetics.

Modern human genetics flourished in the latter half of the twentieth century with the discovery of the structure of DNA (1953) [7], the breaking of the genetic code (1967) [8, 9], the discovery of reverse transcriptase (RT) (1970) [10, 11], the advent of recombinant DNA technology (1973) [12], and the introduction of polymerase chain reaction (PCR) (1985) [13]. These advances, coupled with linkage analysis techniques as well as with functional and positional cloning, have permitted

\*Correspondence: [stockwell@biology.columbia.edu](mailto:stockwell@biology.columbia.edu)

<sup>3</sup>Present address: Stanford University School of Medicine, 251 Campus Drive, Medical School Office Building, Room 323, Stanford, California 94305.

Table 1. Rare Diseases Definitions

Country/Organization	Definition of Rare Disease
Australia	<2,000 Australians
European Community (EC)	<5:10,000 inhabitants
Japan	<4:10,000 Japanese
United States	<200,000 Americans
World Health Organization	0.65–1:1,000 people

the identification and mapping of novel genes involved in disease pathogenesis [14, 15]. Using this approach, the first disease-causing gene—the X-linked gene for chronic granulomatous disease—was positionally cloned in 1986 [16]. This analysis also permitted the discovery of common polymorphisms such as the  $\epsilon 4$  allele of the *APOE* gene (ApoE4 protein) on chromosome 19 [17], which increases susceptibility to late-onset Alzheimer disease [18].

Knowledge of the genome increased in the twenty-first century, when 99% of the euchromatic human genome was sequenced in 2004 [19]. With simple database operations, this blueprint now permits rapid identification of genes and their mutated alleles, including detection of single nucleotide polymorphisms (SNPs). Modern gene cloning and knockout techniques allow for the generation of human genetic disease models in lower organisms such as fruit flies (*Drosophila melanogaster*), zebrafish (*Danio rerio*), or mice (*Mus musculus*). Phenotypic effects of gene modulation (e.g., overexpression, silencing) can easily be obtained from these models.

Unfortunately, these models in lower organisms rarely recapitulate the full spectrum of disease in humans. Human cell-based models of genetic diseases are complementary and can be engineered using molecular biology; these genetic-like models permit subsequent study of disease on a molecular and cellular level. However, correcting a cellular disease phenotype is difficult, as implicated proteins are often either novel or do not have a known small-molecule modulator.

First applied in a systematic fashion in the 1990s, chemical genetics is a potential approach for the study of genetic diseases [20–22]. Like classical genetics, this approach seeks to uncover the specific macromolecules (usually proteins) that act as critical regulators of cellular processes, whose functions are subsequently defined using protein biochemistry, molecular cell biology, and a significant amount of chemical synthesis. Chemical genetics offers advantages over the classical genetic approach: small molecules can permit researchers to temporally control modification of protein function by adding and washing away the small molecule and to specifically affect only one function of a multifunctional protein. Used in human cellular disease models, this approach offers a genetic-like method for studying human diseases and, in some cases, for uncovering novel therapeutic candidates. Chemical genetic approaches have indeed uncovered small molecules that ameliorate or rescue genetic diseases, including the rescue of mutant V2 vasopressin receptors (V2R) associated with nephrogenic diabetes insipidus [23], the chaperoning of proper  $\beta$ -glucosidase folding for Gaucher disease patients [24], and the restoration of signaling via mutant vitamin D

receptors (VDR) in patients with hereditary vitamin D-resistant rickets (HVDRR) [25].

### Advances in Chemistry, Engineering, and Computing

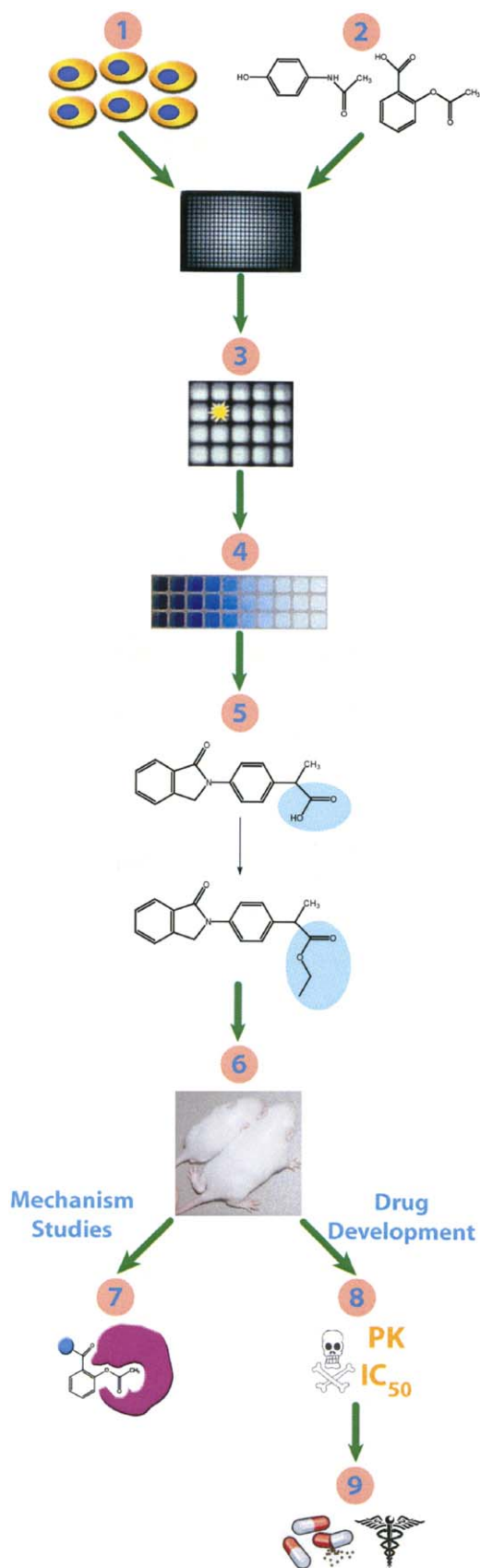
In traditional genetic screens, an effective mutagen (e.g., ultraviolet light) creates random mutations throughout the genome. Similarly, chemical genetics requires an effective collection of small molecules to perturb protein networks. Useful molecules are specific for a single target, have high intrinsic activity (i.e., completely inhibit or activate a target), and are potent. Ideal libraries contain small molecules that are (1) structurally complex (i.e., rigid like most natural products with known biological targets) to decrease the entropic cost of binding a target and (2) diverse (i.e., contain many chemical scaffolds) to facilitate the presentation of functional groups for interactions with targets.

Although rule sets have facilitated the selection of small molecules for screening, two difficulties remain: (1) creating sufficient structural diversity to ensure compounds interact with many different target proteins and (2) synthesizing large numbers of compounds for screening. The advent of diversity-orientated synthesis (DOS) and combinatorial syntheses offer solutions to these problems. DOS profits from the already existing practicality, efficiency, and scale of split-pool synthesis [26] but incorporates reactants to produce a compound collection that is both diverse and structurally and isomerically complex [27]. With careful planning, large numbers of both diverse and complex molecules have been synthesized using few synthetic steps with DOS methodology in which products of one reaction are the substrates of the subsequent reaction [28, 29].

Have optimal regions of chemical space already been defined by nature? Or has nature simply synthesized the best small molecules possible using chemical functionalities that were readily available? If the latter is true, then the ability to create diverse natural product-like compounds that have not been evolutionarily selected for is perhaps one of the most interesting prospects of modern synthetic organic chemistry. This process lends itself to discovering novel regulators of gene products for which no naturally occurring small molecule exists.

Chemical screens necessitate multiple tests to maintain quality control. High-throughput automated screening pipelines incorporate innovations in robotics, materials engineering, and computer science to help accommodate this need. Advanced liquid handling (ALH) robotic systems and microtiter platereaders, both in 384-well format, have significantly increased screening throughput while lowering the cost per experiment. Miniaturization (to 96-well, 384-well, and microarray format [30]) has, however, created new problems, including media evaporation, accumulation of toxins, nutrient depletion, and the need for advanced detection systems for assay execution.

Conventional testing of small-molecule libraries in replicate creates hundreds of thousands of data points and presents a remarkable challenge in terms of both organizational bookkeeping and data analysis. Using plate barcodes for tracking each compound and its linked data, data analysis software facilitates the organization and analysis of compound structures, physical



locations, annotation information, and raw screening data [31]. Following normalization, the systematic spatial errors that are common with modern robotics can be automatically detected and corrected [32]. Computational analyses are performed, and active compounds are highlighted for subsequent study. Hit compounds undergo secondary screening, which typically includes confirmation in a different assay and a counter screen to eliminate possible artifacts (e.g., fluorescent compounds) that presented in the primary screen.

#### Assay Development and Target Identification

With compound libraries and an automated screening pipeline available, laboratories must develop an effective, sensitive assay that models the disease of interest. Although, in theory, many biological questions and diseases can be studied using a chemical approach, developing a specific assay that is amenable to high-throughput screening may be problematic. Financial constraints can additionally limit the types of assays that can be used (e.g., high-content, image-based assays require an automated microscope and image analysis software). Two broad categories of screens are possible: target-based or phenotype-based.

Target-based screens involve the same concept as reverse genetics. Commonly used by pharmaceutical companies, this type of screen involves the identification, isolation, and purification of the specific protein (i.e., small-molecule target) to be studied. With the availability of entire genome DNA sequences, it may be possible in some cases to perform this type of screen against all members of a protein family. Target-based assays have discovered small-molecule partners for numerous proteins, such as helicases [33], E3 ubiquitin ligases [34], steroid receptors [35], transglutaminases [36], coagulation factors [37], and T cell costimulatory receptors [38].

Phenotype-based screens, on the other hand, are comparable to forward genetic screens (Figure 1). Whole cells are exposed to small molecules until a phenotype of interest is found. Phenotypes may be detected through functional assays that measure various cellular activities, such as apoptosis, cell division, or adhesion to a substrate. Other assays, such as reporter gene assays or immunofluorescence, involve the detection of a specific marker indicating the phenotypic change. Recently, phenotype-based assays have discovered small molecules that modulate cell cycle arrest [39, 40], enhance melanin production [41], prevent cell migration [42], prevent IL-6 secretion [43], prevent progression in

Figure 1. Chemical Genetic Pipeline

In forward phenotype-based chemical genetic screens, cells (1) and small organic molecules (2) are dispensed into 384-well microtiter plates. After incubation, an appropriate assay (e.g., cell viability) is executed, and the phenotype of interest is identified (3). Hit compounds enter secondary testing, including testing in dilution series (4) and other confirmatory methods. Medicinal chemistry optimization and testing of analogs (5) identifies the most active structure. Disease model organisms (6) provide a final confirmation of desired phenotype and permit initial observation of compound-induced systemic effects. Affinity purification techniques can be used to identify the macromolecule target of the small molecule (7) providing mechanistic information. Drug development includes studying and optimizing toxicity, pharmacokinetics, and pharmacodynamics (8) before moving to formulation and clinical testing (9).

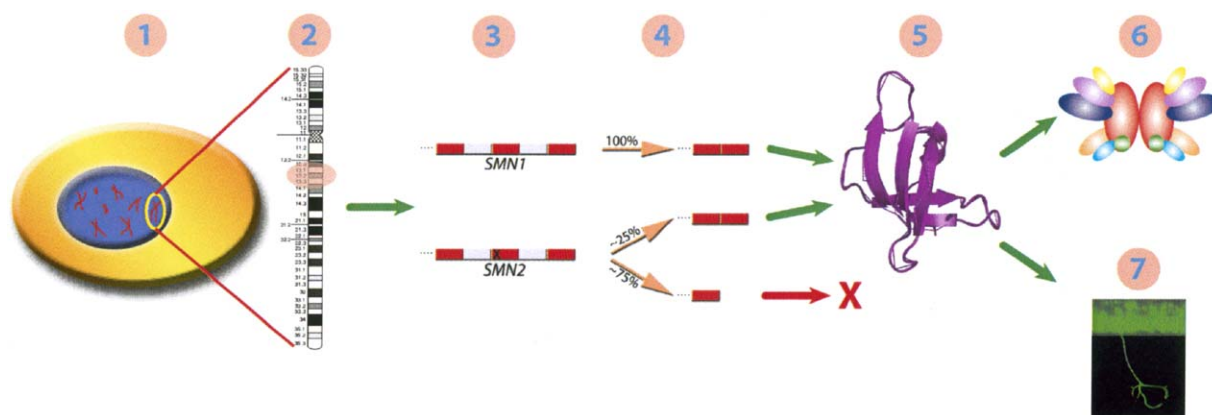


Figure 2. Spinal Muscular Atrophy Molecular Pathology

The survival of motor neuron (SMN) protein is a ubiquitously expressed protein found in all somatic tissues (1). Unaffected humans possess two nearly identical copies of the *SMN* (survival of motor neuron) gene at locus 5q13 (2). A single nucleotide difference inside exon 7 of *SMN2* (3) causes alternative pre-mRNA splicing that results in truncation (i.e., lacking exon 7) of ~75% of the final *SMN2* mRNA transcripts (compared to *SMN1*) (4). Most patients with spinal muscular atrophy have a homozygous deletion of *SMN1*. *SMN2* transcripts may not be functional and their protein products may be rapidly degraded (red X). Full-length transcripts are translated into functional SMN protein (5) (SMN tudor domain is shown). SMN protein, in conjunction with other proteins, has been implicated in the biogenesis of small nuclear ribonucleoproteins (snRNPs) (6) and growth cone localization with axon pathfinding (7). (Number 7 is reproduced from *The Journal of Cell Biology*, 2003, volume 162(5), p. 923, by copyright permission from The Rockefeller University Press.)

mitosis [44], and kill tumor cells in a genotype-specific fashion [45]. These assays can be used to screen thousands of molecules, which may be used as therapeutic candidates for disease treatment and as chemical tools to perturb disease protein networks.

After finding a small molecule that affects a cellular process and produces the desired phenotype, the challenge remains of identifying the small-molecule target or targets. Fifteen years after chemical genetic studies began in earnest, target identification is still a challenge. At its onset, scientists using chemical screening approaches had success with this arduous task (e.g., purification of calcineurin as the target of FK506 [46], histone deacetylase (HDAC) as the target of trapoxin [47], and tubulin as the target of colchicine [48]). The same affinity purification techniques applied then continue to be used today; this usually entails immobilizing the active small molecule on a solid resin (e.g., plastic bead) for subsequent incubation with cellular lysate, elution of nonspecific proteins, purification, and sequencing of bound target proteins. Several opinions have been offered to explain the infrequent success of this approach: attachment of the small molecule to a solid support either requires a vital functional group or affects the required molecular confirmation, steric hindrance of the compound-target interaction is involved, or low binding affinity prohibits target isolation [49]. With some successes using this classic pull-down experiment, new modifications are under development, including creating chimeric ligands [50, 51], developing new binding assays that properly compete with the target [52], comparing proteins bound to a true target to those bound to a "false" molecule (e.g., structural analog, ideally an enantiomer) [53, 54], and using a photo-inducible linker to covalently link the molecule to its target in vivo [55].

### Spinal Muscular Atrophy

SMA is a fatal, autosomal recessive, pediatric disease that results in rapid degeneration and death of  $\alpha$ -motor

neurons in the anterior horn of the spinal cord. SMA is the leading genetic cause of infant (under 2 years of age) mortality in the United States and Western Europe, with an incidence of ~1:6,000 live births and a carrier frequency of 1:40 [56, 57]. Certain populations (e.g., Saudi), however, have higher rates due to consanguineous marriages [58]. No cure or treatment exists for SMA. The small numbers of patients with this genetic disease have resulted in a lack of interest from the pharmaceutical industry despite the scientific tractability of SMA relative to most other orphan genetic diseases. Studies of the underlying molecular pathology of SMA have the potential to reveal essential aspects of motor neuron function and functions of the gene product, and also to suggest various therapeutic strategies for this disease [20, 21, 59].

The *SMN1* (survival of motor neuron 1) gene codes for a 294 amino acid, 38 kD, essential, ubiquitously expressed protein (Figure 2). Humans possess two nearly identical copies of this gene (*SMN1* and *SMN2*) located on a 500 kb inverted repeat in a relatively unstable region of the genome at locus 5q13 [60]. These two genes differ only in five nucleotides within their coding regions, none of which alter the predicted amino acid sequence but do alter the pre-mRNA splicing [61].

In humans, the archetypal primary mRNA (pre-mRNA) transcript contains many introns (i.e., non-protein coding sequences) and exons (i.e., protein coding sequences). The spliceosome is composed of several small nuclear ribonucleoproteins (snRNPs) that form the complex required to catalyze splicing, a post-transcriptional process that removes introns from the primary transcript and ligates exons to form a final mRNA transcript suitable for translation.

Nucleotide polymorphisms and/or mutations in both intronic and exonic sequences can affect splicing. These sequence changes may create *cis*-regulatory elements, such as an exonic splicing enhancer (ESE) or an intronic splicing enhancer (ISE). ESEs stimulate

the inclusion of the exon in the final transcript while ISEs promote removal of the intron. In fact, alternative forms of pre-mRNA splicing exist for ~74% of multiexon genes [62]. Alternative splicing is the major source of proteomic diversity in eukaryotic organisms [63]. Expectedly, alternative splicing can adversely affect the gene product (e.g., truncation, poor folding). This can be, in certain instances, the sole cause of disease, or it can modulate a disease phenotype. Uncovering methods of regulating aberrant splicing represents a largely unexplored avenue for therapeutic development.

The *MAPT* gene, which encodes the microtubule-associated protein TAU, can be spliced into six different isoforms in the human brain. A proper ratio of isoforms is required to prevent development of chromosome 17's frontotemporal dementia with parkinsonism (FTDP-17) [64, 65]. Similarly, various isoforms of the *CFTR* (cystic fibrosis transmembrane conductance regulator) gene cause the typical cystic fibrosis pulmonary disease, as well as male sterility and pancreatic deficiency [66, 67]. Alternative splicing also plays a role in spinal muscular atrophy patients.

One hundred percent of the pre-mRNA transcript from the telomeric gene *SMN1* (or *SMN<sup>T</sup>*) produces full-length mRNA with eight exons. A single C-to-T transition at position six in exon 7 of *SMN2* (or *SMN<sup>C</sup>*) causes alternative pre-mRNA splicing that is 75%–80% effective, resulting in a truncated mRNA (usually lacking exon 7) and protein product. *SMN2* transcripts may not be functional and their protein products may be rapidly degraded, although this is disputed by some investigators. There is also current debate as to whether this nucleotide difference disrupts an exonic splice enhancer (ESE) [68, 69] or creates an exonic splice suppressor (ESS) [70]. Translation of *SMN2*, therefore, yields four to five times less full-length SMN protein as is produced from *SMN1*. SMA patients have either a homozygous deletion of, or a mutation in, the *SMN1* gene and are therefore left with a level of full-length SMN protein that is not sufficient to preserve function and life of  $\alpha$ -motor neurons [71, 72]. Synthetic activators directed to specific pre-mRNA targets have been proposed as emerging technologies with therapeutic potential for SMA [73].

Mice, on the other hand, have a single *SMN* gene (denoted *Smn*), which produces a gene product that is 82% identical in sequence to human SMN protein [74, 75]. This gene is expressed as early as embryonic day 7 [74]. Homozygous deletion of this gene causes early embryonic lethality [76]. Addition of a human *SMN2* transgene into an *Smn*<sup>-/-</sup> background rescues this lethality, but the pups rapidly develop SMA pathology after birth [77, 78], indicating that SMN protein is essential through at least postnatal day 1.

In humans, SMA presents phenotypically with proximal, symmetrical, lower limb and trunk muscle weakness, muscular atrophy, paraplegia, and, frequently, early childhood death. The International SMA Consortium has defined three types of juvenile-onset SMA, based on severity and time of onset [79]. Type I (Werdnig-Hoffman disease; OMIM #253300) is the most severe and the most common (~50% of SMA patients); it is characterized by disease onset before 6 months of age and death before 2 years. Children with

type I SMA are unable to sit without assistance. The second chronic form of SMA, type II (OMIM #253550), is of intermediate severity, with onset before 18 months of age and death after two years. Type II patients are able to sit but cannot walk or stand without help. Type III (Kugelberg-Welander disease; OMIM #253400), the least severe, is distinguished by disease onset after 18 months of age. Type III patients reach all landmarks in motor development and are able to walk until muscles atrophy; they live into adulthood. Pearn et al. have characterized a very rare, type IV SMA (OMIM #271150) that presents during adulthood, between 20 and 50 years of age. Type IV patients experience some difficulty walking but have a normal life expectancy [80, 81]. Due to the rareness of this SMA variant, it has not been subject to significant study, although gene conversions [82] or deletions of exons seven and eight [83, 84] have been reported with this variant.

Full-length SMN protein level has been shown to inversely correlate with SMA severity (patient life span, time of onset, symptoms). Moreover, Feldkötter and colleagues recently reported a striking correlation between SMA type and *SMN2* copy number in humans [85], which can be greater than two due to gene duplication and conversion events [86–88]. Of type I SMA patients, 80% possessed one or two copies of *SMN2*; 82% of type II patients had three *SMN2* copies; and 96% of patients with type III SMA carried three or four copies of *SMN2*. This correlation was further shown in *Smn*<sup>-/-</sup> mice with eight copies of a human *SMN2* transgene; these were phenotypically normal and lived normal lives [77, 78]. However, other modifier genes may exist as well [89].

Since the isolation of the *SMN* gene product in 1995 [61], SMN protein has been subject to intense scrutiny in hopes of understanding disease pathogenesis through protein-protein interactions, as well as the structure and localization of the protein itself. SMN protein, which is expressed at particularly high levels in neurons [90], is diffusely localized in the cytoplasm and in small, SMN-containing, nuclear punctate structures called gems. SMN protein has been implicated in the biogenesis of snRNPs involved in spliceosome formation [91, 92] and in axon pathfinding with localization in the axon growth cones [93, 94].

#### SMA and Phenotype-Based Screens

In hopes of elucidating the molecular mechanism of SMA disease pathogenesis and uncovering therapeutic options for SMA patients, research groups have recently pursued various methods of upregulating the *SMN2* promoter [95], correcting alternative *SMN2* pre-mRNA splicing [96–99], preventing the degradation of SMN protein [100], and modulating *SMN2* translation [101]. These laboratories have chosen a partially hypothesis-based approach, in which tested compounds were chosen for a previously reported effect (e.g., HDAC inhibition) that would putatively have a similar effect on SMN protein regulation. Specifically, hydroxyurea was recently reported as a candidate SMA therapeutic. Chosen for its strong safety profile, high pediatric bioavailability, and known gene upregulation, hydroxyurea was found to restore exon 7 inclusion in *SMN2* transcripts, as well as increasing both SMN

protein level and the number of nuclear gems in SMA patient lymphoblastoid cell lines [96]. Knowing that the SMN protein carboxyl terminus is critical for proper localization, Lorson and colleagues discovered several aminoglycosides, known to suppress stop codon identification, that cause the SMN $\Delta$ 7 protein to acquire five additional amino acids [101]. This addition increases SMN protein expression and the number of nuclear gems in SMA patient cells.

Several laboratories, including ours, have embarked on high-throughput screening of diverse, small-molecule libraries using phenotype-based screens to discover regulators of SMN protein production. This approach has yielded novel regulators of SMN protein that may serve as both drug leads and as new chemical tools to probe the regulation system of SMN protein. It should be noted that, although decreases in SMN protein level result in lower gem counts, small molecules may alter SMN protein distribution between the nucleus and the cytoplasm or affect the time that SMN protein resides in gems. An increase in residency time offers a possible explanation why some compounds exhibit large gem count increases with modest SMN protein upregulation (see below).

In 2001, Andreassi and colleagues reported that aclarubicin (aclacinomycin A), an old topoisomerase-based oncology drug, promotes the retention of exon 7 in the *SMN2* transcript in type I SMA patient fibroblasts (3813) and in a mouse motor neuron-like cell line (NSC34) [102]. Using a semiquantitative RT-PCR assay [71], 960 compounds of the Microsource Plus library had been screened. While aclarubicin is a known chemotherapy agent and possesses a toxicity profile incompatible with a treatment regimen for SMA patients, its discovery indicates that high-throughput assays can be used to study splicing modifications, and it may be valuable in studying other diseases with alternative splicing. In addition, this compound may serve as a probe to identify mechanisms regulating *SMN2* splicing.

Using an assay designed to detect exon 7 inclusion in the *SMN2* transcript, we reported a screen of 47,062 compounds that could affect the production of SMN protein [103]. The small molecules screened were from four libraries: a combinatorial chemistry library (ComGenex), a National Institute for Neurological Disorders and Stroke (NINDS) library, our Annotated Compound Library of known bioactive molecules [104], and our TIC library of compounds selected for specific properties including stereocomplexity, saturated heterocycles, and other natural product qualities. This screen resulted in two selective compounds: aclarubicin (discovered previously, as described above) and indoprofen, a non-steroidal anti-inflammatory drug (NSAID) and cyclooxygenase (COX) inhibitor.

Indoprofen was the only NSAID that exhibited this effect on SMN protein production, indicating that its action does not implicate COXs. Surprisingly, real time RT-PCR (qRT-PCR) revealed no change in either the absolute level or the ratio of full-length to truncated *SMN2* mRNA transcript from indoprofen-treated type I SMA patient fibroblasts (3813). Indoprofen treatment did, however, modestly increase endogenous SMN protein level in 3813 cells and caused an increase in the number

of nuclear gems. To confirm indoprofen's effect, the ability of indoprofen to rescue embryonic lethality in an SMA mouse model was examined: an increase in mean litter size and a trend showing increased viability of SMA model mice resulted. Indoprofen may serve as a possible therapeutic candidate for SMA patients and may be developed into a tool for probing mechanisms governing SMN protein production.

Recently, Jarecki and colleagues used a  $\beta$ -lactamase reporter gene assay to detect increases in *SMN2* transcription [105]. Driven by a 3.4 kb fragment of the endogenous promoter, fluorescence produced by cleavage of a  $\beta$ -lactamase substrate was read from NSC34 cells treated with ~550,000 synthetic compounds. Two chemical scaffolds (an indole and a quinazoline) exhibited significant increases in exon 7 inclusion in the *SMN2* transcript and increased production of SMN protein as quantified by Western blot. Only the quinazoline, however, was amenable to medicinal chemistry (i.e., retained activity after modification). Both scaffolds are undergoing medicinal chemistry optimization for drug-like characteristics before entering animal testing and subsequent clinical trials.

Several compounds have been found to ameliorate SMA pathogenesis by upregulating *SMN2* transcription [106], modulating pre-mRNA splicing of the *SMN2* transcript [96, 97, 99, 102, 105], or upregulating SMN protein through an unknown mechanism [103]. All of the potential therapeutics for SMA induce a modest SMN protein or gem count increase in SMA patient cells and have shortcomings that range from gastrointestinal bleeding to short half-life in human serum, and from high toxicity to upregulation of hundreds of genes (Figure 3). There is no additional evidence to indicate that these compounds will improve existing SMA pathology or prevent disease pathogenesis in vivo. While the initial data are promising [107, 108], the effects of these small molecules in SMA clinical trials remain to be determined. Additionally, determining how much SMN protein is needed, along with proper spatial and temporal data, to rescue the SMA phenotype are critically important outstanding questions that need attention.

A complementary approach to SMA drug discovery involves searching for small molecules that replace or enhance the function of the SMN protein. For example, Dreyfuss and colleagues have pioneered this approach and developed an assay to find such compounds [109]. Such SMN-replacing compounds would be complementary to those that enhance production of SMN protein from the *SMN2* gene, and together these two types of compounds might yield a powerful combination therapy.

Undoubtedly, additional small molecules with therapeutic potential in SMA remain undiscovered. The future of chemical genetics, as it relates to spinal muscular atrophy, is 2-fold: first, the mechanism of action of reported candidate therapeutics should be elucidated by using the small molecule itself as a chemical tool in target-identification assays, such as affinity purification and mRNA binding assays. In addition, putative HDAC inhibitors, such as valproic acid, can be profiled for their specificity against all HDAC family members. Identified target(s) will influence the design of future target-based assays to uncover more effective (i.e., more potent,

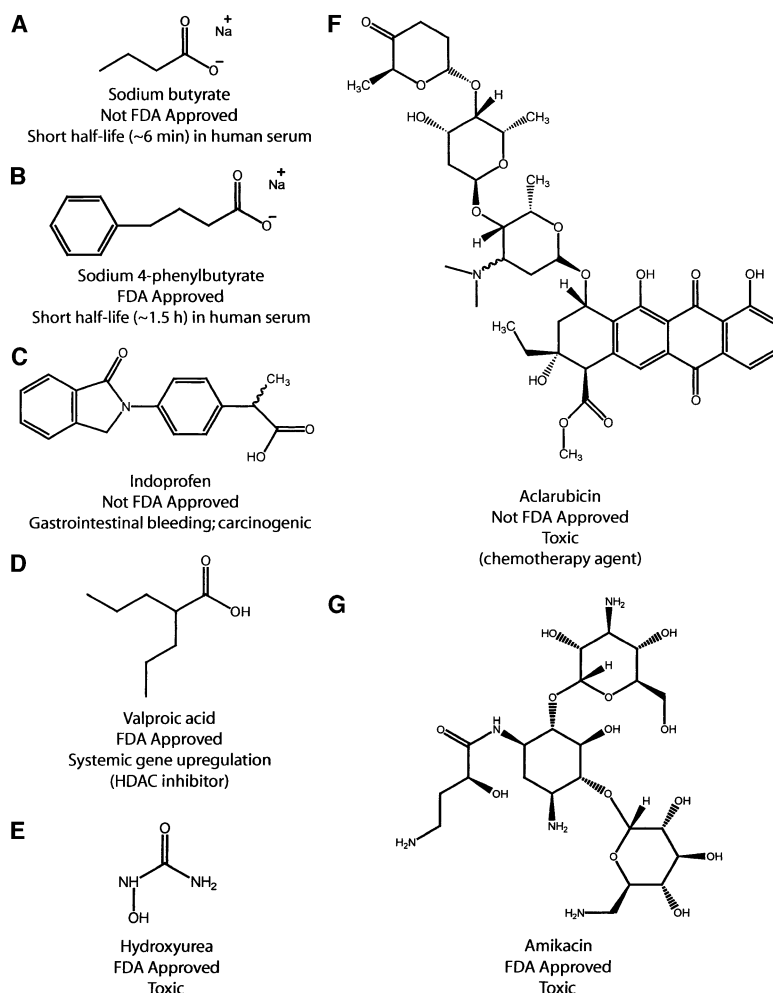


Figure 3. Candidate Therapeutics for SMA Patients

(A) Sodium butyrate [99], (B) sodium 4-phenylbutyrate [106], (C) indoprofen [103], (D) valproic acid [97, 98], (E) hydroxyurea [96], (F) aclarubicin [102], and (G) amikacin (a representative aminoglycoside) [101] have all been reported to ameliorate SMA pathogenesis through known and unknown mechanisms. FDA approval status and possible shortcomings of each candidate small molecule are presented.

higher intrinsic activity) small molecules. Virtual screening (the *in silico* counterpart of high-throughput screening) promises to be a popular approach to determining functionalities predicted to bind to such protein targets [110–112]. Second, new phenotype-based assays will be developed to detect increases in SMN protein level or gem count, or compounds that otherwise overcome the effects of *SMN1* deletion, which are the sought-after activities of candidate therapeutics. Such assays have been developed for other proteins [39, 71, 113–115]. Coupled with libraries composed of small molecules selected for disease-specific applications (e.g., ability to cross the blood-brain barrier for neurological diseases), these assays will identify new molecules to use as chemical tools for probing the SMN protein regulation system in addition to being candidate drugs.

#### Future Improvements and Outlook

Small molecules play vital roles at the cellular level, but the downstream effects manifest at all levels of biological organization (i.e., cell, tissue, organ, system, organism). Despite the presence of small molecules in essentially every aspect of life, small molecules have yet to be included in the ‘central dogma’ (DNA to RNA to protein) [116]. Recent advances have outfitted the scientific community with the technologies needed to explore biological processes using chemical tools.

Chemical genetics applies these advances to studying human genetic diseases and simultaneously developing therapeutic candidates. However, important challenges remain to be solved.

The throughput of screening assays must be improved to allow for the testing of larger numbers of compounds, which should lead to the discovery of more potent compounds. Several possibilities exist for increasing screening throughput: (1) eliminating the required spatial separation of compounds and (2) retaining the spatial separation but designing more efficient detection methods. New nonspatial approaches involve encoded self-assembling chemical (ESAC) libraries [117], which enable novel small molecules to be uncovered in target-based screens. Using this methodology, small molecules linked to oligonucleotide sequences undergo self-assembly *in vitro* to form a library in which rapid compound identification is done via standard PCR amplification. On the other hand, spatial separation can be preserved if detection methods are developed with sufficient sensitivity to be quantifiable. Recently, a novel microarray technique for phenotype-based screens was created that embeds small molecules into polymer spots (~200 μm in diameter) on a standard microscope slide [30]. Using a biodegradable polymer, the small molecule is slowly released, and the phenotypic effects of cells growing over the slide can

be readily assessed by visual inspection. Automated microscopy, coupled with computational algorithms, will allow less compound (and fewer cells) to be used in screens.

The medical community frequently treats diseases with several drugs. Correspondingly, this combination therapy approach has already proven to be valuable in high-throughput screens [118, 119]. This multicomponent, systems-based method can uncover synergistic therapeutics while providing information about pathway crosstalk. However, determining which small molecules to screen in combination and interpreting the results of such screens in terms of underlying mechanism remain largely untapped areas of research.

As a long-range outlook, several "audacious goals" for chemical biology were recently proposed [116]. These include cataloging all naturally occurring small molecules, discovering a small molecule that modulates every individual function of human proteins, and linking the genotype of cells to the phenotypic response(s) elicited by small molecules. Over the coming years, both 'biological-activity space' and 'chemical-activity space' will be further mapped and linked in pursuit of these goals [120]. RNA interference (RNAi) [121, 122], antibodies [123], and gene expression analysis [124] will complement study with small molecules. The complete realization of these goals will renovate and rejuvenate science, medicine, and orphan disease research.

## Conclusion

The chemical genetic approach to studying human disease pathogenesis has evolved over the last 15 years. Despite this progress, many challenges remain. The fatal pediatric disease spinal muscular atrophy is an example in which successful phenotype-based screens using high-throughput screening technologies have been applied to study of a rare disease. Several upregulators of SMN protein have been discovered, and additional compounds will likely emerge in the near future.

With many rare genetic diseases in human populations, both the political and scientific communities need to embrace the challenge to identify and define therapeutic candidates for disease treatment. With a concerted effort, academics, nonprofit organizations, and industrial groups can work together to develop the equipment, technologies, and assays needed for investigating these devastating and neglected human diseases.

## Acknowledgments

We thank Reka Letso for valuable comments and critical reading of this manuscript. B.R.S. is supported in part by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund and by the SMA Foundation.

Received: July 7, 2005

Revised: September 8, 2005

Accepted: September 8, 2005

Published: October 21, 2005

## References

1. Orphan Drug Act. (1983). Public Law 97-414, as amended. U.S. Code 21, § 360ee.
2. Rare Diseases Act. (2002). Public Law 107-280. U.S. Code 21, § 360aa-ee.
3. McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University, and NCBI (2000). OMIM: Online Mendelian Inheritance in Man (<http://www.ncbi.nlm.nih.gov/omim/>).
4. Schmitz, R. (1985). Friedrich Wilhelm Sertürner and the discovery of morphine. *Pharm. Hist.* 27, 61-74.
5. Ehrlich, P. (1913). Chemotherapeutics: scientific principles, methods and results. *Lancet* 2, 445-451.
6. Travis, A.S. (1989). Science as receptor of technology: Paul Ehrlich and the synthetic dyestuffs industry. *Sci. Context* 3, 383-408.
7. Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171, 737-738.
8. Crick, F.H. (1963). On the genetic code. *Science* 139, 461-464.
9. Crick, F.H., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961). General nature of the genetic code for proteins. *Nature* 192, 1227-1232.
10. Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209-1211.
11. Temin, H.M., and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226, 1211-1213.
12. Cohen, S.N., Chang, A.C., Boyer, H.W., and Helling, R.B. (1973). Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. USA* 70, 3240-3244.
13. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350-1354.
14. Collins, F.S. (1992). Positional cloning: let's not call it reverse anymore. *Nat. Genet.* 1, 3-6.
15. Collins, F.S. (1995). Positional cloning moves from perditional to traditional. *Nat. Genet.* 9, 347-350.
16. Royer-Pokora, B., Kunkel, L.M., Monaco, A.P., Goff, S.C., Newburger, P.E., Baehner, R.L., Cole, F.S., Curnutte, J.T., and Orkin, S.H. (1986). Cloning the gene for an inherited human disorder—chronic granulomatous disease—on the basis of its chromosomal location. *Nature* 322, 32-38.
17. Pericak-Vance, M.A., Bebout, J.L., Gaskell, P.C., Jr., Yamaoka, L.H., Hung, W.Y., Alberts, M.J., Walker, A.P., Bartlett, R.J., Haynes, C.A., Welsh, K.A., et al. (1991). Linkage studies in familial Alzheimer disease: evidence for chromosome 19 linkage. *Am. J. Hum. Genet.* 48, 1034-1050.
18. Strittmatter, W.J., Saunders, A.M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G.S., and Roses, A.D. (1993). Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl. Acad. Sci. USA* 90, 1977-1981.
19. International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431, 931-945.
20. Stockwell, B.R. (2000). Chemical genetics: ligand-based discovery of gene function. *Nat. Rev. Genet.* 1, 116-125.
21. Stockwell, B.R. (2004). Exploring biology with small organic molecules. *Nature* 432, 846-854.
22. Schreiber, S.L. (2003). The small-molecule approach to biology: chemical genetics and diversity-oriented organic synthesis make possible the systematic exploration of biology. *Chem. Eng. News* 81, 51-61.
23. Morello, J.P., Salahpour, A., Laperriere, A., Bernier, V., Arthus, M.F., Lonergan, M., Petaja-Repo, U., Angers, S., Morin, D., Bichet, D.G., et al. (2000). Pharmacological chaperones rescue cell-surface expression and function of misfolded V2 vasopressin receptor mutants. *J. Clin. Invest.* 105, 887-895.
24. Sawkar, A.R., Cheng, W.C., Beutler, E., Wong, C.H., Balch, W.E., and Kelly, J.W. (2002). Chemical chaperones increase the cellular activity of N370S beta -glucosidase: a therapeutic strategy for Gaucher disease. *Proc. Natl. Acad. Sci. USA* 99, 15428-15433.
25. Gardezi, S.A., Nguyen, C., Malloy, P.J., Posner, G.H., Feldman, D., and Peleg, S. (2001). A rationale for treatment of hereditary



- vitamin D-resistant rickets with analogs of 1 alpha,25-dihydroxyvitamin D(3). *J. Biol. Chem.* **276**, 29148–29156.
26. Furka, A., Sebestyen, F., Asgedom, M., and Dibo, G. (1991). General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **37**, 487–493.
  27. Schreiber, S.L. (2000). Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* **287**, 1964–1969.
  28. Tan, D.S., Foley, M.A., Shair, M.D., and Schreiber, S.L. (1998). Stereoselective synthesis of over two million compounds having structural features both reminiscent of natural products and compatible with miniaturized cell-based assays. *J. Am. Chem. Soc.* **120**, 8565–8566.
  29. Krishnan, S., and Schreiber, S.L. (2004). Syntheses of stereochemically diverse nine-membered ring-containing biaryls. *Org. Lett.* **6**, 4021–4024.
  30. Bailey, S.N., Sabatini, D.M., and Stockwell, B.R. (2004). Microarrays of small molecules embedded in biodegradable polymers for use in mammalian cell-based screens. *Proc. Natl. Acad. Sci. USA* **101**, 16144–16149.
  31. Kelley, B.P., Lunn, M.R., Root, D.E., Flaherty, S.P., Martino, A.M., and Stockwell, B.R. (2004). A flexible data analysis tool for chemical genetic screens. *Chem. Biol.* **11**, 1495–1503.
  32. Root, D.E., Kelley, B.P., and Stockwell, B.R. (2003). Detecting spatial patterns in biological array experiments. *J. Biomol. Screen.* **8**, 393–398.
  33. Tanner, J.A., Zheng, B.J., Zhou, J., Watt, R.M., Jiang, J.Q., Wong, K.L., Lin, Y.P., Lu, L.Y., He, M.L., Kung, H.F., et al. (2005). The adamantane-derived bananins are potent inhibitors of the helicase activities and replication of SARS coronavirus. *Chem. Biol.* **12**, 303–311.
  34. Bottger, V., Bottger, A., Howard, S.F., Picksley, S.M., Chene, P., Garcia-Echeverria, C., Hochkeppel, H.K., and Lane, D.P. (1996). Identification of novel mdm2 binding peptides by phage display. *Oncogene* **13**, 2141–2147.
  35. Norris, J.D., Paige, L.A., Christensen, D.J., Chang, C.Y., Huacani, M.R., Fan, D., Hamilton, P.T., Fowlkes, D.M., and McDonnell, D.P. (1999). Peptide antagonists of the human estrogen receptor. *Science* **285**, 744–746.
  36. Choi, K., Siegel, M., Piper, J.L., Yuan, L., Cho, E., Strnad, P., Omary, B., Rich, K.M., and Khosla, C. (2005). Chemistry and biology of dihydroisoxazole derivatives: selective inhibitors of human transglutaminase 2. *Chem. Biol.* **12**, 469–475.
  37. Spiegel, P.C., Kaiser, S.M., Simon, J.A., and Stoddard, B.L. (2004). Disruption of protein-membrane binding and identification of small-molecule inhibitors of coagulation factor VIII. *Chem. Biol.* **11**, 1413–1422.
  38. Huxley, P., Sutton, D.H., Debnam, P., Matthews, I.R., Brewer, J.E., Rose, J., Trickett, M., Williams, D.D., Andersen, T.B., and Classon, B.J. (2004). High-affinity small molecule inhibitors of T cell costimulation: compounds for immunotherapy. *Chem. Biol.* **11**, 1651–1658.
  39. Stockwell, B.R., Haggarty, S.J., and Schreiber, S.L. (1999). High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. *Chem. Biol.* **6**, 71–83.
  40. Stockwell, B.R., Hardwick, J.S., Tong, J.K., and Schreiber, S.L. (1999). Chemical genetic and genomic approaches reveal a role for copper in specific gene activation. *J. Am. Chem. Soc.* **121**, 10662–10663.
  41. Snyder, J.R., Hall, A., Ni-Komatsu, L., Khersonsky, S.M., Chang, Y.T., and Orlow, S.J. (2005). Dissection of melanogenesis with small molecules identifies prohibitin as a regulator. *Chem. Biol.* **12**, 477–484.
  42. Yarrow, J.C., Totsukawa, G., Charras, G.T., and Mitchison, T.J. (2005). Screening for cell migration inhibitors via automated microscopy reveals a rho-kinase inhibitor. *Chem. Biol.* **12**, 385–395.
  43. Yamaguchi, K., Yada, M., Tsuji, T., Hatanaka, Y., Goda, K., and Kobori, T. (1999). 4-Phenylthiazole derivatives inhibit IL-6 secretion in osteoblastic cells and suppress bone weight loss in ovariectomized mice. *Bioorg. Med. Chem. Lett.* **9**, 957–960.
  44. Haggarty, S.J., Mayer, T.U., Miyamoto, D.T., Fathi, R., King, R.W., Mitchison, T.J., and Schreiber, S.L. (2000). Dissecting cellular processes using small molecules: identification of colchicine-like, taxol-like and other small molecules that perturb mitosis. *Chem. Biol.* **7**, 275–286.
  45. Dolma, S., Lessnick, S.L., Hahn, W.C., and Stockwell, B.R. (2003). Identification of genotype-selective antitumor agents using synthetic lethal chemical screening in engineered human tumor cells. *Cancer Cell* **3**, 285–296.
  46. Liu, J., Farmer, J.D., Lane, W.S., Friedman, J., Weissman, I., and Schreiber, S.L. (1991). Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66**, 807–815.
  47. Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408–411.
  48. Correia, J.J. (1991). Effects of antimetabolic agents on tubulin-nucleotide interactions. *Pharmacol. Ther.* **52**, 127–147.
  49. Burdine, L., and Kodadek, T. (2004). Target identification in chemical genetics: the (often) missing link. *Chem. Biol.* **11**, 593–597.
  50. Profit, A.A., Lee, T.R., and Lawrence, D.S. (1999). Bivalent inhibitors of protein tyrosine kinases. *J. Am. Chem. Soc.* **125**, 13404–13414.
  51. Maly, D.J., Choong, I.C., and Ellman, J.A. (2000). Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors. *Proc. Natl. Acad. Sci. USA* **97**, 2419–2424.
  52. Alluri, P.G., Reddy, M.M., Bachhawat-Sikder, K., Olivos, H.J., and Kodadek, T. (2003). Isolation of protein ligands from large peptoid libraries. *J. Am. Chem. Soc.* **125**, 13995–14004.
  53. Ding, S., Wu, T.Y., Brinker, A., Peters, E.C., Hur, W., Gray, N.S., and Schultz, P.G. (2003). Synthetic small molecules that control stem cell fate. *Proc. Natl. Acad. Sci. USA* **100**, 7632–7637.
  54. Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999.
  55. Dorman, G., and Prestwich, G.D. (2000). Using photolabile ligands in drug discovery and development. *Trends Biotechnol.* **18**, 64–77.
  56. Melki, J., Lefebvre, S., Burglen, L., Burlet, P., Clermont, O., Millasseau, P., Reboullet, S., Benichou, B., Zeviani, M., Le Paslier, D., et al. (1994). De novo and inherited deletions of the 5q13 region in spinal muscular atrophies. *Science* **264**, 1474–1477.
  57. Pearn, J. (1980). Classification of spinal muscular atrophies. *Lancet* **1**, 919–922.
  58. Al-Jumah, M., Majumdar, R., Al-Rajeh, S., Awada, A., Chaves-Carbello, E., Saih, M., Al-Shahwan, S., Al-Subiey, K., and Al-Uthaim, S. (2003). Molecular analysis of the spinal muscular atrophy and neuronal apoptosis inhibitory protein genes in Saudi patients with spinal muscular atrophy. *Saudi Med. J.* **24**, 1052–1054.
  59. Stockwell, B.R. (2002). Chemical genetic screening approaches to neurobiology. *Neuron* **36**, 559–562.
  60. Burglen, L., Lefebvre, S., Clermont, O., Burlet, P., Viollet, L., Cruaud, C., Munnich, A., and Melki, J. (1996). Structure and organization of the human survival motor neurone (SMN) gene. *Genomics* **32**, 479–482.
  61. Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., et al. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**, 155–165.
  62. Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R., and Shoemaker, D.D. (2003). Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* **302**, 2141–2144.
  63. Maniatis, T., and Tasic, B. (2002). Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* **418**, 236–243.
  64. Stanford, P.M., Shepherd, C.E., Halliday, G.M., Brooks, W.S., Schofield, P.W., Brodaty, H., Martins, R.N., Kwok, J.B., and Schofield, P.R. (2003). Mutations in the tau gene that cause an increase in three repeat tau and frontotemporal dementia. *Brain* **126**, 814–826.

65. Kowalska, A., Hasegawa, M., Miyamoto, K., Akiguchi, I., Ike-moto, A., Takahashi, K., Araki, W., and Tabira, T. (2002). A novel mutation at position +11 in the intron following exon 10 of the tau gene in FTDP-17. *J. Appl. Genet.* **43**, 535–543.
66. Noone, P.G., and Knowles, M.R. (2001). 'CFTR-opathies': disease phenotypes associated with cystic fibrosis transmembrane regulator gene mutations. *Respir. Res.* **2**, 328–332.
67. Costes, B., Girodon, E., Ghanem, N., Flori, E., Jardin, A., Soufir, J.C., and Goossens, M. (1995). Frequent occurrence of the CFTR intron 8 (TG)<sub>n</sub> 5T allele in men with congenital bilateral absence of the vas deferens. *Eur. J. Hum. Genet.* **3**, 285–293.
68. Lorson, C.L., Hahnen, E., Androphy, E.J., and Wirth, B. (1999). A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc. Natl. Acad. Sci. USA* **96**, 6307–6311.
69. Cartegni, L., and Krainer, A.R. (2002). Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nat. Genet.* **30**, 377–384.
70. Kashima, T., and Manley, J.L. (2003). A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nat. Genet.* **34**, 460–463.
71. Covert, D.D., Le, T.T., McAndrew, P.E., Strasswimmer, J., Crawford, T.O., Mendell, J.R., Coulson, S.E., Androphy, E.J., Prior, T.W., and Burghes, A.H. (1997). The survival motor neuron protein in spinal muscular atrophy. *Hum. Mol. Genet.* **6**, 1205–1214.
72. Lefebvre, S., Burlet, P., Liu, Q., Bertrand, S., Clermont, O., Munnich, A., Dreyfuss, G., and Melki, J. (1997). Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat. Genet.* **16**, 265–269.
73. Cartegni, L., and Krainer, A.R. (2003). Correction of disease-associated exon skipping by synthetic exon-specific activators. *Nat. Struct. Biol.* **10**, 120–125.
74. DiDonato, C.J., Chen, X.N., Noya, D., Korenberg, J.R., Nadeau, J.H., and Simard, L.R. (1997). Cloning, characterization, and copy number of the murine survival motor neuron gene: homolog of the spinal muscular atrophy-determining gene. *Genome Res.* **7**, 339–352.
75. Bergin, A., Kim, G., Price, D.L., Sisodia, S.S., Lee, M.K., and Rabin, B.A. (1997). Identification and characterization of a mouse homologue of the spinal muscular atrophy-determining gene, survival motor neuron. *Gene* **204**, 47–53.
76. Schrank, B., Gotz, R., Gunnarsen, J.M., Ure, J.M., Toyka, K.V., Smith, A.G., and Sendtner, M. (1997). Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 9920–9925.
77. Monani, U.R., Sendtner, M., Covert, D.D., Parsons, D.W., Andreassi, C., Le, T.T., Jablonka, S., Schrank, B., Rossol, W., Prior, T.W., et al. (2000). The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in *Smn*(<sup>-/-</sup>) mice and results in a mouse with spinal muscular atrophy. *Hum. Mol. Genet.* **9**, 333–339.
78. Hsieh-Li, H.M., Chang, J.G., Jong, Y.J., Wu, M.H., Wang, N.M., Tsai, C.H., and Li, H. (2000). A mouse model for spinal muscular atrophy. *Nat. Genet.* **24**, 66–70.
79. Munsat, T.L., and Davies, K.E. (1992). International SMA consortium meeting. (26–28 June 1992, Bonn, Germany). *Neuro-muscul. Disord.* **2**, 423–428.
80. Pearn, J. (1978). Autosomal dominant spinal muscular atrophy: a clinical and genetic study. *J. Neurol. Sci.* **38**, 263–275.
81. Pearn, J.H., Hudgson, P., and Walton, J.N. (1978). A clinical and genetic study of spinal muscular atrophy of adult onset: the autosomal recessive form as a discrete disease entity. *Brain* **101**, 591–606.
82. Mazzei, R., Gambardella, A., Conforti, F.L., Magariello, A., Patitucci, A., Gabriele, A.L., Sprovieri, T., Labate, A., Valentino, P., Bono, F., et al. (2004). Gene conversion events in adult-onset spinal muscular atrophy. *Acta Neurol. Scand.* **109**, 151–154.
83. Brahe, C., Servidei, S., Zappata, S., Ricci, E., Tonali, P., and Neri, G. (1995). Genetic homogeneity between childhood-onset and adult-onset autosomal recessive spinal muscular atrophy. *Lancet* **346**, 741–742.
84. Clermont, O., Burlet, P., Lefebvre, S., Burglen, L., Munnich, A., and Melki, J. (1995). SMN gene deletions in adult-onset spinal muscular atrophy. *Lancet* **346**, 1712–1713.
85. Feldkötter, M., Schwarzer, V., Wirth, R., Wienker, T.F., and Wirth, B. (2002). Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am. J. Hum. Genet.* **70**, 358–368.
86. Burghes, A.H. (1997). When is a deletion not a deletion? When it is converted. *Am. J. Hum. Genet.* **61**, 9–15.
87. McAndrew, P.E., Parsons, D.W., Simard, L.R., Rochette, C., Ray, P.N., Mendell, J.R., Prior, T.W., and Burghes, A.H. (1997). Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMN1 and SMN2 gene copy number. *Am. J. Hum. Genet.* **60**, 1411–1422.
88. van der Steege, G., Grootsholten, P.M., Cobben, J.M., Zap-pata, S., Scheffer, H., den Dunnen, J.T., van Ommen, G.J., Brahe, C., and Buys, C.H. (1996). Apparent gene conversions involving the SMN gene in the region of the spinal muscular atrophy locus on chromosome 5. *Am. J. Hum. Genet.* **59**, 834–838.
89. Scharf, J.M., Endrizzi, M.G., Wetter, A., Huang, S., Thompson, T.G., Zerres, K., Dietrich, W.F., Wirth, B., and Kunkel, L.M. (1998). Identification of a candidate modifying gene for spinal muscular atrophy by comparative genomics. *Nat. Genet.* **20**, 83–86.
90. Battaglia, G., Princivalle, A., Forti, F., Lizier, C., and Zeviani, M. (1997). Expression of the SMN gene, the spinal muscular atrophy determining gene, in the mammalian central nervous system. *Hum. Mol. Genet.* **6**, 1961–1971.
91. Pellizzoni, L., Yong, J., and Dreyfuss, G. (2002). Essential role for the SMN complex in the specificity of snRNP assembly. *Science* **298**, 1775–1779.
92. Meister, G., and Fischer, U. (2002). Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *EMBO J.* **21**, 5853–5863.
93. McWhorter, M.L., Monani, U.R., Burghes, A.H., and Beattie, C.E. (2003). Knockdown of the survival motor neuron (*Smn*) protein in zebrafish causes defects in motor axon outgrowth and pathfinding. *J. Cell Biol.* **162**, 919–931.
94. Rossol, W., Jablonka, S., Andreassi, C., Kroning, A.K., Karle, K., Monani, U.R., and Sendtner, M. (2003). *Smn*, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of beta-actin mRNA in growth cones of motoneurons. *J. Cell Biol.* **163**, 801–812.
95. Majumder, S., Varadharaj, S., Ghoshal, K., Monani, U., Burghes, A.H., and Jacob, S.T. (2004). Identification of a novel cyclic AMP-response element (CRE-II) and the role of CREB-1 in the cAMP-induced expression of the survival motor neuron (SMN) gene. *J. Biol. Chem.* **279**, 14803–14811.
96. Grzeschik, S.M., Ganta, M., Prior, T.W., Heavlin, W.D., and Wang, C.H. (2005). Hydroxyurea enhances SMN2 gene expression in spinal muscular atrophy cells. *Ann. Neurol.* **58**, 194–202.
97. Sumner, C.J., Huynh, T.N., Markowitz, J.A., Perhac, J.S., Hill, B., Covert, D.D., Schussler, K., Chen, X., Jarecki, J., Burghes, A.H., et al. (2003). Valproic acid increases SMN levels in spinal muscular atrophy patient cells. *Ann. Neurol.* **54**, 647–654.
98. Brichta, L., Hofmann, Y., Hahnen, E., Siebzehrubli, F.A., Raschke, H., Blumcke, I., Eyupoglu, I.Y., and Wirth, B. (2003). Valproic acid increases the SMN2 protein level: a well-known drug as a potential therapy for spinal muscular atrophy. *Hum. Mol. Genet.* **12**, 2481–2489.
99. Chang, J.G., Hsieh-Li, H.M., Jong, Y.J., Wang, N.M., Tsai, C.H., and Li, H. (2001). Treatment of spinal muscular atrophy by sodium butyrate. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 9808–9813.
100. Chang, H.C., Hung, W.C., Chuang, Y.J., and Jong, Y.J. (2004). Degradation of survival motor neuron (SMN) protein is mediated via the ubiquitin/proteasome pathway. *Neurochem. Int.* **45**, 1107–1112.
101. Wolstencroft, E.C., Mattis, V., Bajer, A.A., Young, P.J., and Lorson, C.L. (2005). A non-sequence-specific requirement for SMN protein activity: the role of aminoglycosides in inducing elevated SMN protein levels. *Hum. Mol. Genet.* **14**, 1199–1210.

102. Andreassi, C., Jarecki, J., Zhou, J., Coover, D.D., Monani, U.R., Chen, X., Whitney, M., Pollok, B., Zhang, M., Androphy, E., et al. (2001). Aclarubicin treatment restores SMN levels to cells derived from type I spinal muscular atrophy patients. *Hum. Mol. Genet.* *10*, 2841–2849.
103. Lunn, M.R., Root, D.E., Martino, A.M., Flaherty, S.P., Kelley, B.P., Coover, D.D., Burghes, A.H., Man, N.T., Morris, G.E., Zhou, J., et al. (2004). Indoprofen upregulates the survival motor neuron protein through a cyclooxygenase-independent mechanism. *Chem. Biol.* *11*, 1489–1493.
104. Root, D.E., Flaherty, S.P., Kelley, B.P., and Stockwell, B.R. (2003). Biological mechanism profiling using an annotated compound library. *Chem. Biol.* *10*, 881–892.
105. Jarecki, J., Chen, X., Bernardino, A., Coover, D.D., Whitney, M., Burghes, A., Stack, J., and Pollok, B.A. (2005). Diverse small-molecule modulators of SMN expression found by high-throughput compound screening: early leads towards a therapeutic for spinal muscular atrophy. *Hum. Mol. Genet.* *14*, 2003–2018.
106. Andreassi, C., Angelozzi, C., Tiziano, F.D., Vitali, T., De Vincenzi, E., Boninsegna, A., Villanova, M., Bertini, E., Pini, A., Neri, G., et al. (2004). Phenylbutyrate increases SMN expression in vitro: relevance for treatment of spinal muscular atrophy. *Eur. J. Hum. Genet.* *12*, 59–65.
107. Mercuri, E., Bertini, E., Messina, S., Pelliccioni, M., D'Amico, A., Colitto, F., Mirabella, M., Tiziano, F.D., Vitali, T., Angelozzi, C., et al. (2004). Pilot trial of phenylbutyrate in spinal muscular atrophy. *Neuromuscul. Disord.* *14*, 130–135.
108. Russman, B.S., Iannaccone, S.T., and Samaha, F.J. (2003). A phase 1 trial of riluzole in spinal muscular atrophy. *Arch. Neurol.* *60*, 1601–1603.
109. Wang, J., and Dreyfuss, G. (2001). A cell system with targeted disruption of the SMN gene: functional conservation of the SMN protein and dependence of Gemin2 on SMN. *J. Biol. Chem.* *276*, 9599–9605.
110. Taylor, R.D., Jewsbury, P.J., and Essex, J.W. (2002). A review of protein-small molecule docking methods. *J. Comput. Aided Mol. Des.* *16*, 151–166.
111. Lengauer, T., Lemmen, C., Rarey, M., and Zimmermann, M. (2004). Novel technologies for virtual screening. *Drug Discov. Today* *9*, 27–34.
112. Schneider, G., and Bohm, H.J. (2002). Virtual screening and fast automated docking methods. *Drug Discov. Today* *7*, 64–70.
113. Chen, H., Kovar, J., Sissons, S., Cox, K., Matter, W., Chadwell, F., Luan, P., Vlahos, C.J., Schutz-Geschwender, A., and Olive, D.M. (2005). A cell-based immunocytochemical assay for monitoring kinase signaling pathways and drug efficacy. *Anal. Biochem.* *338*, 136–142.
114. Blaskovich, M.A., Sun, J., Cantor, A., Turkson, J., Jove, R., and Sebti, S.M. (2003). Discovery of JSI-124 (cucurbitacin I), a selective Janus kinase/signal transducer and activator of transcription 3 signaling pathway inhibitor with potent antitumor activity against human and murine cancer cells in mice. *Cancer Res.* *63*, 1270–1279.
115. Koeller, K.M., Haggarty, S.J., Perkins, B.D., Leykin, I., Wong, J.C., Kao, M.C., and Schreiber, S.L. (2003). Chemical genetic modifier screens. Small molecule trichostatin suppressors as probes of intracellular histone and tubulin acetylation. *Chem. Biol.* *10*, 397–410.
116. Schreiber, S. (2005). Small molecules: the missing link in the central dogma. *Nat Chem Biol* *1*, 64–66.
117. Melkko, S., Scheuermann, J., Dumelin, C.E., and Neri, D. (2004). Encoded self-assembling chemical libraries. *Nat. Biotechnol.* *22*, 568–574.
118. Borisy, A.A., Elliott, P.J., Hurst, N.W., Lee, M.S., Lehar, J., Price, E.R., Serbedzija, G., Zimmermann, G.R., Foley, M.A., Stockwell, B.R., et al. (2003). Systematic discovery of multicomponent therapeutics. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 7977–7982.
119. Keith, C.T., Borisy, A.A., and Stockwell, B.R. (2005). Multicomponent therapeutics for networked systems. *Nat. Rev. Drug Discov.* *4*, 71–78.
120. Kim, Y.-k., Arai, M.A., Arai, T., Lamenza, J.O., Dean, E.F., III, Patterson, N., Clemons, P.A., and Schreiber, S.L. (2004). Relationship of stereochemical and skeletal diversity of small molecules to cellular measurement space. *J. Am. Chem. Soc.* *126*, 14740–14745.
121. Hannon, G.J., and Rossi, J.J. (2004). Unlocking the potential of the human genome with RNA interference. *Nature* *431*, 371–378.
122. Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature* *431*, 343–349.
123. Moore, P., and Clayton, J. (2003). To affinity and beyond. *Nature* *426*, 725–731.
124. Stegmaier, K., Ross, K.N., Colavito, S.A., O'Malley, S., Stockwell, B.R., and Golub, T.R. (2004). Gene expression-based high-throughput screening (GE-HTS) and application to leukemia differentiation. *Nat. Genet.* *36*, 257–263.