Chimia 59 (2005) 803–807

ISSN 0009–4293

Application of RNAi in Biomedical Research

Jonathan Hall*

Abstract: Antisense techniques are powerful tools for the inhibition of gene expression in the target discovery and validation phases of pharmaceutical research. In particular, RNA interference has become a method of choice to investigate the roles of putative target genes in models of disease states. This technique is technically demanding, particularly when it is applied in complex experiments such as high-throughput genome analysis or in animal models of disease. We describe here two areas of current research in our group, dedicated to advancing these techniques: the use of neural networks to predict the biological activity of short interfering RNAs, and the application of siRNAs in animal models of disease.

Keywords: Antisense · Neural network · Oligonucleotide · RNAi · Target validation

Introduction

Pharmaceutical research often begins with the identification of a gene product (protein) which is dysfunctional, either through mis-regulated expression or mutation, in a disease state. It is however difficult to establish whether the protein represents a good pharmaceutical ('validated') target upon which to begin a drug discovery program. For example, the protein may be related to the result of the disease state and not related to the cause of the disease state, and thus its inhibition may be ineffective as a disease treatment. Alternatively, the protein may perform a critical function in tissues unaffected by the disease, and therefore its inhibition may result in unacceptable levels of toxicity. Target validation is most efficiently performed using low molecular weight (LMW) compounds, in fact, with the intended drug itself. However, the identification of said compounds, which inhibit specifically and potently the function of the protein in appropriate disease models is a lengthy, resource-intensive process. Ironically, the decision as to whether to initiate a medicinal chemistry program on a given target is best made using the very compounds produced by such a program.

The pharmacological consequences of inhibition of a candidate protein target can be simulated using antisense techniques in which oligonucleotides are applied to reduce the levels of the corresponding messenger RNA (mRNA), in cells or animal models (Fig. 1). We have described a number of recent examples where both antisense oligonucleotides (ASOs) and short interfering RNAs (siRNAs) have been used to simulate the knock-down of potentially important drug targets including *CD40* [1], *Hus1* [2], *Bcl-2/BclXl* [3], *Cdc34* [4] and *p21* [5]. Today, the RNA interference (RNAi) mechanism has replaced that of antisense as the most commonly used gene-knockdown technique.

The optimal use of oligonucleotides as validation tools requires an understanding of their *general* properties (*i.e.* as a class of molecule), as opposed to particular *sequence-specific* properties (*i.e.* properties of a single siRNA molecule). This is of particular importance when trying to select oligonucleotides which are both specific and potent inhibitors, but also, which do not cause non-sequence specific toxic effects. Due to the very large number of possible sequence motifs in a set of siRNAs – a consequence of the large number of sequence permutations – the general properties of oligonucleotides are best revealed by looking at sufficiently large data-sets [6]. This has been a common theme in our research of late.

RNA Interference

Mechanism and the Variable Potency of siRNAs

RNA interference (RNAi) is a powerful genetic technique for the validation of pharmaceutical targets. It is a natural mechanism which was first recognized in plants and model organisms as a defence against viral RNA [7]. Key components of this complex mechanism are siRNAs, 21-nucleotide (nt) duplexes bearing a dinucleotide 3′-overhang on each strand. The duplex is composed of sense and antisense strands, where the latter is complementary to the targeted mRNA and, when correctly designed, bears mismatches to all other transcripts in the cell. siRNAs are delivered into cells *in vitro* by one of several possible documented transfection methods, where they are processed. The duplex siRNA is unwound and the antisense strand is bound by the RNA-interference-silencing-complex (RISC), a protein complex which targets the complementary mRNA only for endonuclolytic cleavage in the centre of the duplex region [8]. After cleavage, the mRNA dissociates and the complex is available for further mRNA binding. Protein translation is therefore interrupted and residual

*Correspondence: Dr. J. Hall
Novartis Institutes for BioMedical Research
Genome and Proteome Sciences
WSJ-88.4.31
Postfach
CH-4002 Basel
Tel.: +41 61 324 55 49
E-Mail: jonathan.hall@novartis.com

CHIMIA **2005**, *59*, No. 11

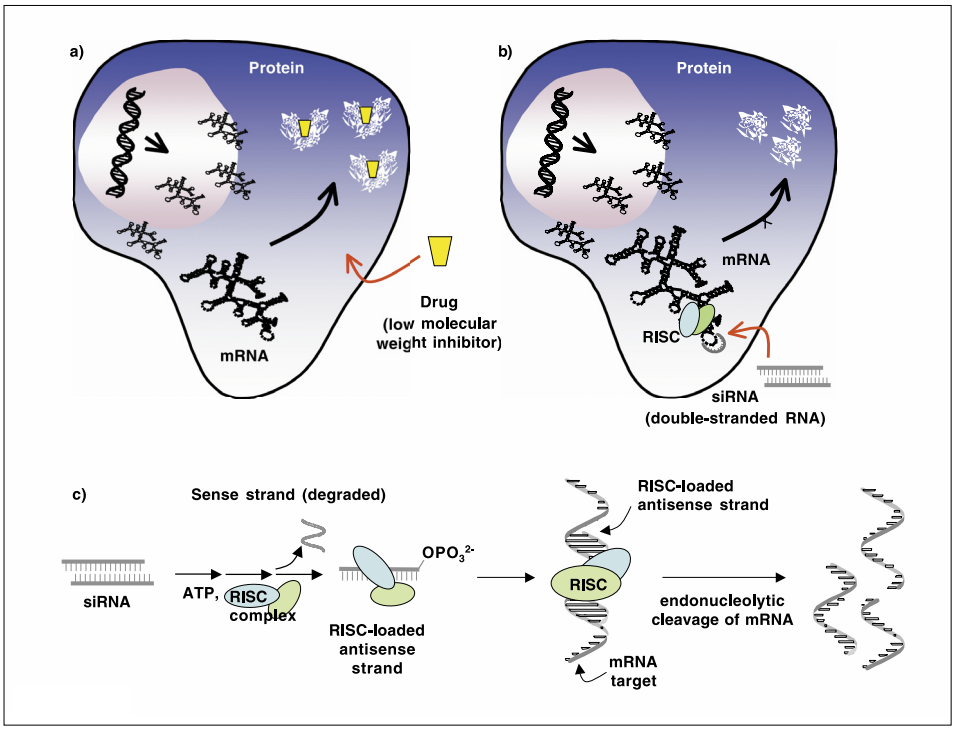

Fig. 1. A schematic comparison of two mechanisms of inhibiting the function of a protein, including the mechanism of action of siRNAs: a) A low-molecular weight (LMW) compound binds in the active site of the protein *e.g.* an enzyme, in order to block enzymatic function; b) A short interfering RNA (siRNA) binds to the messenger RNA (mRNA) responsible for synthesis of the corresponding protein, causing a gradual reduction in protein levels, and consequent loss of function; c) The double stranded RNA is bound and processed by the RISC complex during which the duplex is dissociated, the sense strand is discarded and degraded, and the antisense strand is phosphorylated at its 5′-terminus. This complex binds the mRNA and effects cleavage of the mRNA in the centre of the duplex region.

protein levels in the cell decrease as it is metabolized with its natural half-life. As protein levels fall so its effective function therefore is reduced. Therefore, this mechanism mimics the action of a LMW inhibitor of the protein itself, albeit with a temporal delay caused by kinetics of target protein loss (Fig. 1). RNAi can be used *in vitro* and *in vivo*, although there are significant technical hurdles to targeting oligonucleotides to a particular tissue *in vivo*.

Each mRNA is thousands of nucleotides in length. Therefore, there are thousands of possible complementary short oligonucleotide sequences which can be selected as either ASOs or siRNAs. Unfortunately, for a variety and combination of reasons, many of these target-specific oligonucleotides are biologically inactive (Fig. 2).

We recently conducted a study in which we compared the potency of siRNAs with 2′-O-methoxyethyl (MOE)-modified ASOs, under identical cellular delivery conditions. Oligonucleotides to target recombinant rat pain-related cation-channel P2X3 expressed in CHO-K1 and the rat brain tumour derived cell line, 33B [10], were delivered by electroporation and results were evaluated on the mRNA, protein, and functional levels. The siRNA was found to be more potent than the sequence-identical ASO. We also detected and characterized a complementation effect of ASOs and siRNAs when siRNA was combined with non-homologous ASO targeting different regions on the common $P2X_3$ mRNA. Combination of both reagents yielded a more efficient down-regulation of $P2X_3$ mRNA than if they were administered separately. Possible reasons for variations in oligonucleotide potencies include: strong secondary and tertiary structure of the mRNA which prevents effective oligonucleotide binding, preferences of the RISC for certain nucleotide motifs in the mRNA or siRNA sequence, sequence-dependent transfection efficiency, sequence-dependent siRNA stability *etc.*, all of which are determined to some extent by nucleotide sequence motifs. It is usually not possible to recognize such motifs by visual inspection of active siRNAs because of the vast number of possible permutations, and therefore it is difficult to predict active siRNAs based upon nucleotide sequence alone. This became an acute problem in a genomics project where it was necessary to construct a large, genome-wide library of siRNAs. The purpose of this library is to examine the effects of inhibition of each individual gene of a genome in a selected pathway through the use of oligonucleotides in suitable cellular assays [11]. Where a given gene plays an important role in the pathway, inhibition of its mRNA will reveal a change in phenotype in the assay readout. Until recently, the largest gene knock-down experiments performed to date in human cells have employed multiple si/shRNAs per gene and have only addressed a subset of the human genome [12–14]. One of the principal problems in this expensive area of research is that many siRNAs are biologically inactive and it is not possible to experimentally identify these prior to inclusion in a library. Thus multiple reagents

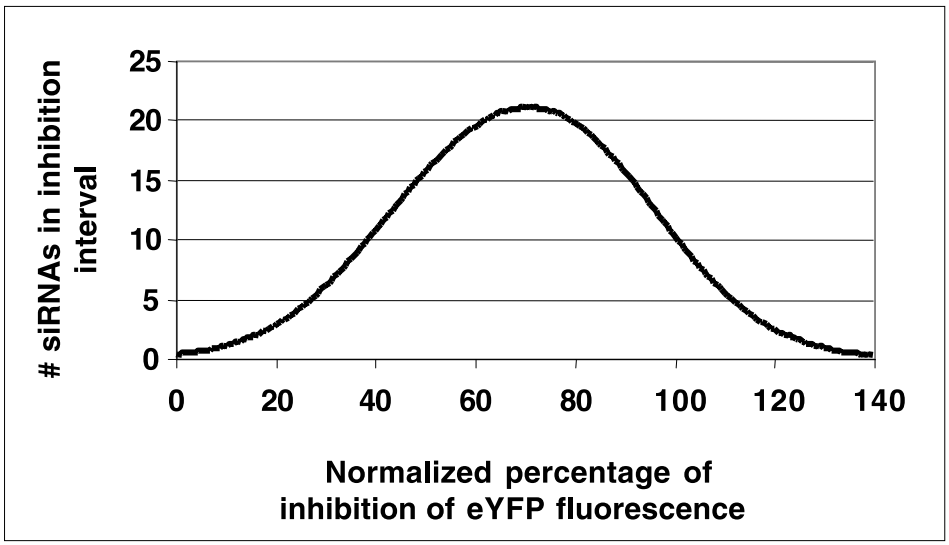

Fig. 2. 2431 Randomly-selected siRNAs complementary to 34 reporter plasmids coding for yellow fluorescent protein (YFP), differing in the sequence of their 3′-untranslated regions (UTRs), were tested in H1299 cancer cells for their ability to down regulate YFP. Residual levels of YFP fluorescence were measured after 48 h and normalized to unrelated negative and positive controls. The Y-axis is the number of siRNAs, normalized for correspondence to a 256 bin histogram. The X-axis is YFP inhibition normalized with the positive control set to 90.0% inhibition and the negative control set to 35.4% inhibition, such that the least active siRNA becomes 0%. siRNAs show an approximate Gaussian distribution of potencies relative to controls (taken from [9]).

need to be employed per gene to ensure that gene-knock-down in the assay can be guaranteed for fidelity of cellular readout. In order to minimize this burden, we turned to an artificial neural network (ANN) in order to try and predict potent siRNAs based upon their nucleotide sequence.

A Neural Net Algorithm for the Prediction of Potent siRNAs

ANNs identify and work with highly complex patterns in data-sets, using automated and unbiased learning. The numerous, complex parameters that they derive are combined and weighted for accurate prediction. Often, the prediction quality of an ANN depends on a sufficiently large, unbiased training set, preferably where data points are directly comparable.

We reasoned that a large homogeneous data set of siRNA potency for training ANNs could be best generated using a high-throughput (HT) reporter screening assay which we had previously developed for ASOs [15]. The high-throughput screening assay is based upon a mRNA fusion transcript comprising a full length reporter gene (luciferase or fluorescent proteins) with sequences inserted into its 3′-UTR. The insert can be from partial or full length genes, can be from genes from alternate species, and from any region of an mRNA (Fig. 3).

A large number (58) of 'gapmer' 18-nt MOE-oligonucleotides were screened for antisense activity against eight such luciferase reporter fusion constructs in a CHO cell line and the inhibitory profiles obtained were compared and quantified to those observed through screening the oligonucleotides against the corresponding endogenous genes, assayed at the mRNA level in H1299. A high degree of similarity in the profiles was obtained indicating that the fusion constructs are suitable surrogates for the endogenously-expressed mRNAs for characterization of ASOs. We then ensured that this assay would also function with siRNAs, and designed a series of assays suitable for raising ANNs. A very large number (2431) of siRNAs were selected randomly across 34 different plasmids, and screened in H1299 cells [9]. The data was randomly divided into a training set of 2182 data points and an independent test set of 249 data points. The ANNs were tested by plotting predicted biological activity against experimental activity (Fig. 4), and quantified using Pearson correlation coefficients.

The highest performing algorithm, named BIOPREDsi, yielded a Pearson correlation coefficient r = 0.66 at P = 2.2exp-16, proving that artificial intelligence alone can be used to predict activity. BIOPREDsi was then used to select the top two predicted siRNAs for six genes which were then assayed for inhibition of target mRNA by Q-PCR (normalized to control siRNA). All siRNAs were found to be active. BIOPREDsi was used to design a genome-wide set of 50,000 siRNAs targeted to 22,473 predicted human genes. In this process, BIOPREDsi consistently selects siRNAs which commence with a U at the 5′-terminus of the sequence. This curious observation suggested a very dominant effect of the 5′-U in contributing to the potency of siRNAs. This effect was subsequently explained in publications as contributing to a preference of the RISC complex to load the correct antisense siRNA strand [16–18]. This library of reagents is currently being used to screen the human genome for genes involved in a wide variety of genetic pathways related to human disease. BIOPREDsi is available for

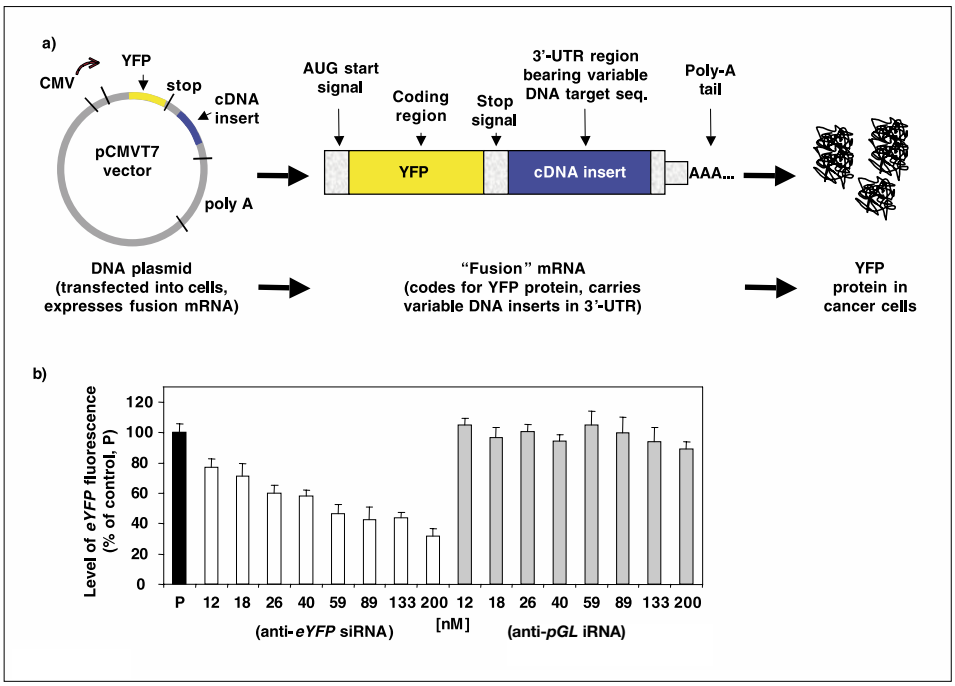

Fig. 3. A high-throughput reporter assay to quantify potency of siRNAs. a) Transfection into cells of the DNA plasmid results in transcription of a 'fusion' mRNA, the sequence of which contains a variable sequence insert, which serves as target for complementary siRNAs. The mRNA is translated into YFP protein, the sequence of which does not contain the target insert; b) Co-transfection of the plasmid with an siRNA (YFP siRNA) targeted to the mRNA causes a dose-dependent cleavage of the mRNA and consequently loss of YFP fluorescence, whereas an unrelated siRNA (anti-pGL3) has no effect, in H1299 cells at 48 h. Normalized relative intensity of YFP is plotted (n = 3) (taken from [9]).

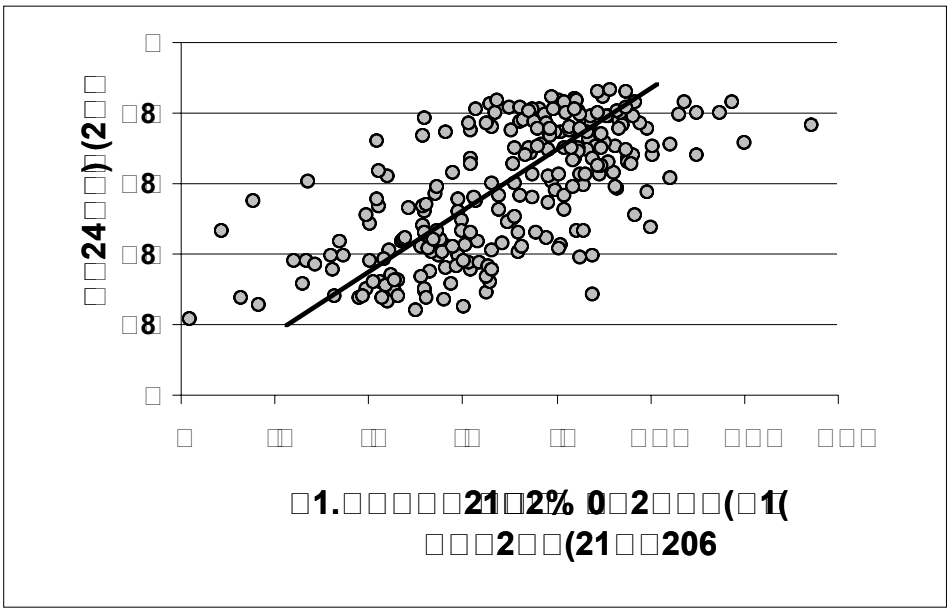

Fig. 4. Plot of predicted potency (Biopred score) of 249 siRNAs by the artificial neural network (ANN) BIOPREDsi against the observed inhibitory activity obtained in the HT YFP reporter assay. For a Biopred score ≥0.75, the top 10% most potent siRNAs yield a specificity of 79% and a sensitivity of 53% (taken from [9]).

public use on the internet (*www.biopredsi.org*).

In vivo *Target Validation*

The use of oligonucleotides *in vivo* is fraught with technical difficulties: large quantities of these highly expensive reagents are required, penetration into cells in the desired target tissue is usually poor, the reagents are rapidly degraded by nucleases in blood and other body fluids, and have unfavourable pharmacokinetics due to high molecular weights (7–15 kD) and high negative charge. These difficulties together probably account for the general lack of success of oligonucleotides in the clinic to date. Despite this, there are increasing numbers of reports of successful use of ASOs and siRNAs *in vivo*, particularly where it is possible to deliver the oligonucleotide locally.

The ability to achieve a rapid, inexpensive and specific knockdown of neuronal genes, for *in vivo* target validation, would be a major step forward in pharmaceutical research in neuroscience. We have used siRNAs extensively in rodent models of chronic neuropathic pain [19]. The purpose of the studies was to evaluate the roles of putative pain genes in this indication. The $P2X_3$ gene was selected as a target in a proof of concept study to evaluate technical feasibility. It is a ATP-gated cation channel, restricted in expression to cell bodies of sensory neurons and known to be involved in pain from $P2X_3$ mRNA expression in models of chronic pain, and altered pain-related responses in knock-out mice. Furthermore, we had previously described the successful use of ASOs in these models to validate the target [20]. Characterized siRNAs were delivered intrathecally directly into the spinal fluid using osmotic mini-pumps. Two of the aforementioned technical hurdles were circumvented as the use of the mini-pumps guaranteed an effective local delivery of high concentration, while requiring small absolute amounts of reagents. Furthermore, a similar method is commonly used for drug delivery into patients with severe pain. The experiments yielded reduced $P2X_3$ mRNA in dorsal root ganglia and reduced $P2X_3$ protein translocated in the dorsal horn of spinal chord of rats compared to untreated and mismatched controls (Fig. 5).

The oligoribonucleotides carried 3′-overhangs of MOE-containing ribonucleotides to ensure protection against nuclease-degradation. We showed in earlier studies *in vitro* that the modified oligonucleotides were as potent as those carrying deoxythymidine overhangs, and that that the modified termini did not inhibit the RNAi mechanism [10]. Although the degree of target knock-down was partial, probably reflecting an incomplete cellular uptake, a strong phenotypic effect was observed compared to mismatched and untreated controls with doses of 400 microgram/day. Partial sciatic ligation models in rats were used with readouts of allodynia and mechanical hyperalgesia. The double-stranded siRNA was more potent than the single-stranded ASO, confirming the results obtained *in vitro*. This was the first report of an endogenous disease-related neuronal gene targeted in the nervous system with a measurable pathophysiological response in a relevant and well-defined disease model. The phenotypic results caused by a genetic inhibition of the $P2X_3$ gene were reproduced using low molecular weight compounds against the channel [21], confirming the original postulate in this example that oligonucleotides can function as mimics of compounds in target validation studies.

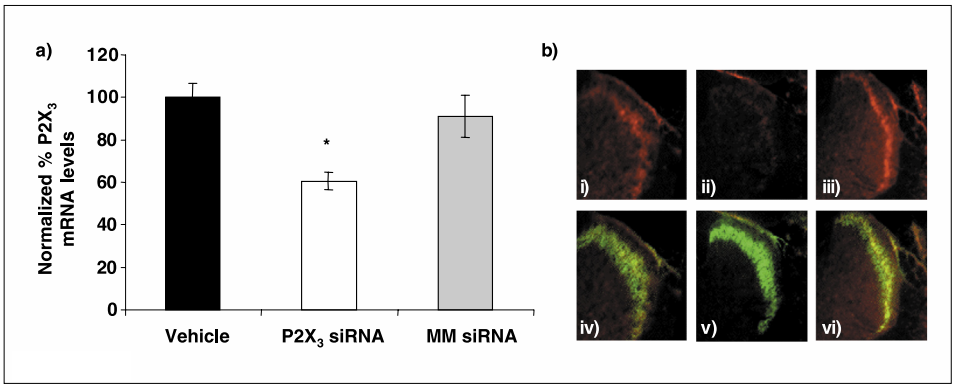

Fig. 5. Molecular down regulation of $P2X_3$ RNA and protein in the nervous system of a rat upon treatment with siRNA: a) $P2X_3$ mRNA levels in dorsal root ganglia (DRG), normalized to β-actin, were significantly reduced in $P2X_3$ siRNA-treated compared to mismatched control siRNA (MM)-treated and vehicle treated rats (n/group = 3); b) $P2X_3$ immunoreactivity in the lamina II of the spinal cord was visibly reduced in siRNA-treated (ii) animals compared to vehicle-treated (i) and mismatched (iii) negative controls. Co-immunostaining was performed with neuron-specific isolectin B4 (iv–vi) and showed that equivalent regions of each spinal cord were analyzed (n/group = 3) (taken from [19]; reproduced by permission of Oxford University Press).

The power of using siRNAs to inhibit genes in the nervous system prompted us to attempt a new approach with infusion of siRNAs into the ventricular system of the rat brain, using a similar delivery protocol with the mini-pumps [22]. The extent of RNAi-induced downregulation was analyzed by targeting enhanced green fluorescent protein (eGFP), in transgenic mice over-expressing eGFP. Extensive mRNA target inhibition was observed, in regions adjacent or dorso-ventrally and medio-laterally distant to the infusion site. Lower levels of knockdown were observed in more distal regions. This technical success permitted us to generate specific knockdown of the endogenous gene encoding the dopamine transporter (DAT). DAT controls the activity of dopamine by facilitating a rapid uptake of the neurotransmitter into pre-synpatic terminals. DAT is a key regulator of dopamine actions on locomotion, emotion and cognition, and implicated in the etiology of several disorders. Infusion of DAT-siRNA in adult mice produced a significant down-regulation of DAT mRNA and protein in the brain, and elicited a temporal hyperlocomotor response in comparison to mismatch and untreated controls.

Once again, the effects with the siRNA mirrored those seen with a pharmacologically-selective LMW DAT inhibitor, but were delayed by 24–48 h, as expected for a kinetically slower-acting drug.

Outlook

RNAi is a very powerful means to study gene function and to simulate the effects of LMW-compound inhibition of a target protein. We have extended and exploited the power of RNAi-mediated gene knockdown technologies to evaluate the roles of genes in models of disease, both *in vitro* and *in vivo*.

The research performed as part of the genomics project has allowed us to advance to another major milestone in functional genomics, that of systematic independent inhibition of all human genes in model systems. The studies performed *in vivo* are particularly exciting as they promise for the first time the possibility to apply siRNAs as a therapeutic treatment for a number of disorders in the central nervous system.

Acknowledgements

I thank sincerely the entire Nucleic Acids group of NIBR for their past and continuing skillful contributions in this area.

Received: September 9, 2005

[1] S.A. Rushworth, C.A. Bravery, J. Hall, F. Natt, N.J. Parsons, J. Weiler, R. Häner, S. Thompson, *Transplantation* **2002**, *73*, 635.

[2] B. Kinzel, J. Hall, F. Natt, J. Weiler, D. Cohen, *Cancer* **2002**, *94*, 1808.

[3] R.A. Olie, C. Hafner, R. Kuttel, B. Sigrist, J. Willers, R. Dummer, J. Hall, R.A. Stahel, U. Zangermeister-Wittke, *J. Investigative Derm.* **2002**, *118*, 505.
[4] N. Butz, S. Ruetz, F. Natt, J. Hall, J. Weiler, J. Mestan, M. Ducarre, R. Grossenbacher, P. Hauser, D. Kempf, F. Hofmann, *Exp. Cell Res.* **2005**, *303*, 482.
[5] I. Beuvink, A. Boulay, S. Fumagalli, F. Zilbermann, S. Ruetz, T. O'Reilly, F. Natt, J. Hall, H.A. Lane, G. Thomas, *Cell* **2005**, *120*, 747.
[6] J. Hall, *Nature Reviews Genet.* **2004**, *5*, 552.
[7] A. Fire, S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, C.C. Mello, *Nature* **1998**, *391*, 806.
[8] W. Filipowicz, *Cell* **2005**, *122*, 17.
[9] D. Huesken, J. Lange, C. Mickanin, J. Weiler, F. Asselbergs, J. Warner, B. Meloon, S. Engel, A. Rosenberg, D. Cohen, M. Labow, M. Reinhardt, F. Natt, J. Hall, *Nature Biotechnol.* **2005**, *23*, 995.
[10] M. Hemmings-Mieszczak, G. Dorn, F. Natt, J. Hall, W. Wishart, *Nucleic Acids Res*. **2003**, *31*, 2117.
[11] M. Boutros, A.A. Kiger, S. Armknecht, K. Kerr, M. Hild, B. Koch, S.A. Haas, Consortium, Heidelberg Fly Array, R. Paro, N. Perrimon, *Science* **2004**, *303*, 832.
[12] P.J. Paddison. J.M. Silva, D.S. Conklin, *Nature* **2004**, *428*, 427.
[13] K. Berns, E.M. Hijmans, J. Mullenders, T.R. Brummelkamp, A. Velds, M. Heimerikx, R.M. Kerkhoven, M. Madiredjo, W. Nijkamp, B. Weigelt, R. Agami, W. Ge, G. Cavet, P.S. Linsley, R.L. Beijersbergen, R. Bernards, *Nature* **2004**, *428*, 431.
[14] R. Kittler. G. Putz, L. Pelletier, I. Poser, A.-K. Heninger, D. Drechsel, S. Fischer, I. Konstantinova, B. Habermann, H. Grabner, M.-L. Yaspo, H. Himmelbauer, B. Korn, K. Neugebauer, M.T. Pisabarro, F. Buchholz, *Nature* **2004**, *432*, 1036.
[15] D. Hüsken, F. Asselbergs, B. Kinzel, F. Natt, J. Weiler, P. Martin, R. Häner J. Hall, *Nucleic Acids Res.* **2003**, *31*, e102.
[16] A. Khvorova, A. Reynolds, S.D. Jayasena, *Cell* **2003**, *115*, 209.
[17] D. S. Schwarz, G. Hutvagner, T. Du, Z. Xu, N. Aronin, P.D. Zamore, *Cell* **2003**, *115*, 199.
[18] A. Reynolds, D. Leake, Q. Boese, S. Scaringe, W.S. Marshall, A. Khvorova, *Nat. Biotechnol.* **2004**, *22*, 326.
[19] G. Dorn, S. Patel, G. Wotherspoon, M. Hemmings-Mieszczak, J. Barclay, F. Natt, P. Martin, S. Bevan, A. Fox, P. Ganju, W. Wishart, J. Hall, *Nucleic Acids Res.* **2004**, *32*, e49.
[20] J. Barclay, S. Patel, G. Dorn, G. Wotherspoon, S. Moffatt, L. Eunson, S. Abdel'al, F. Natt, J. Hall, J. Winter, S. Bevan, W. Wishart, A. Fox, P. Ganju, *J. Neuroscience* **2002**, *22*, 8139.
[21] M.F. Jarvis, E.C. Burgard, S. McGaraughty, P. Honore, K. Lynch, T.J. Brennan, A. Subieta, T. Van Biesen, J. Cartmell, B. Bianchi, W. Niforatos, K. Kage, H. Yu, J. Mikusa, C.T. Wismer, C.Z. Zhu, K. Chu, C.-H. Lee, A.O. Stewart, J. Polakowski, B.F. Cox, E. Kowaluk, M. Williams, J. Sullivan, C. Faltynek, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 17179.
[22] D.R. Thakker, F. Natt, D. Huesken, R. Maier, M. Mueller, H van der Putten, D. Hoyer, J.F. Cryan, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17270.