

AOBase: a database for antisense oligonucleotides selection and design

Xiaochen Bo, Shaoke Lou, Daochun Sun, Jing Yang and Shengqi Wang*

Beijing Institute of Radiation Medicine, Beijing 100850, China

Received August 14, 2005; Revised and Accepted October 8, 2005

ABSTRACT

Antisense oligonucleotides (ODNs) technology is one of the important approaches for the sequence-specific knockdown of gene expression. ODNs have been used as research tools in the post-genome era, as well as new types of therapeutic agents. Since finding effective target sites within RNA is a hard work for antisense ODNs design, various experimental methods and computational approaches have been proposed. For better sharing of the experimented and published ODNs, valid and invalid ODNs reported in literatures are screened, collected and stored in AOBase. Till now, ~700 ODNs against 46 target mRNAs are contained in AOBase. Entries can be explored via TargetSearch and AOSearch web retrieval interfaces. AOBase can not only be useful in ODNs selection for gene function exploration, but also contribute to mining rules and developing algorithms for rational ODNs design. AOBase is freely accessible via <http://www.bioit.org.cn/ao/aobase>.

INTRODUCTION

With the ability to selectively down-regulate the expression of genes, antisense oligonucleotides (ODNs) have been widely used in gene function determination, drug targets validation and pathways discovery (1–3). Recently, ODNs also serve as specific and efficient inhibitors for systematic loss-of-function analysis of miRNA (4,5). On the other hand, ODNs can be effective therapeutic agents. Several antisense compounds for disease treatment have been evaluated in clinical trials with promising results (6,7). However, the successful use of ODNs is somewhat limited since only a small number among all the possible antisense ODNs against a given target RNA show effective suppression of the target gene in living cells (8,9). It is commonly accepted that the selection of sensitive sites in

target RNA is of great importance for ODNs efficiencies. Various experimental approaches to identify promising local target sites have been presented in recent years (9–11). There has also been much interest in computational approaches to select target sites of ODNs, which get prominent advantages over experimental protocols in throughput, cost and efficiency (12–15).

In fact, for the researchers who use ODNs as gene expression modulation tools to explore gene functions or molecular networks, it is not necessary to screen ODNs targeting specific mRNA if they could find some with enough activity in literatures or database, considering that experimental ODNs screening methods are time consuming and expensive. However, for the researchers whose efforts focus on the development of *in silico* antisense ODNs design methods, information about both valid and invalid ODNs are of same value. Rules for rational target site selection can be mined from these positive and negative cases. Therefore, if the related data for ODNs are collected together, there would be obvious benefit for ODNs users and designers.

Three ODNs resources have been reported till now. The first public ODNs database named ODNBase was developed five years ago by Giddings *et al.* (16). Unfortunately, it cannot be accessed at present. Another database is a non-public database maintained by Isis Pharmaceuticals (17), which stores the data from the experiments performed at the corporation and is not publicly available. There is also a data list named AOdb, which is freely available for ODNs prediction algorithm research (13), containing basic descriptions of 315 ODNs. The AOBase we describe here is a database developed for both ODNs selection and ODNs design. It can be freely accessed at <http://www.bioit.org.cn/ao/aobase>. A more comprehensive dataset has been elaborately screened and constructed, and two retrieval tools have been developed. AOBase can be used to select effective ODNs for gene expression modulation, and can also contribute to mining rules and developing algorithms for rational ODNs design.

*To whom correspondence should be addressed. Tel/Fax: +86 10 68210077 932211; Email: sqwang@bmi.com

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors

© The Author 2006. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org

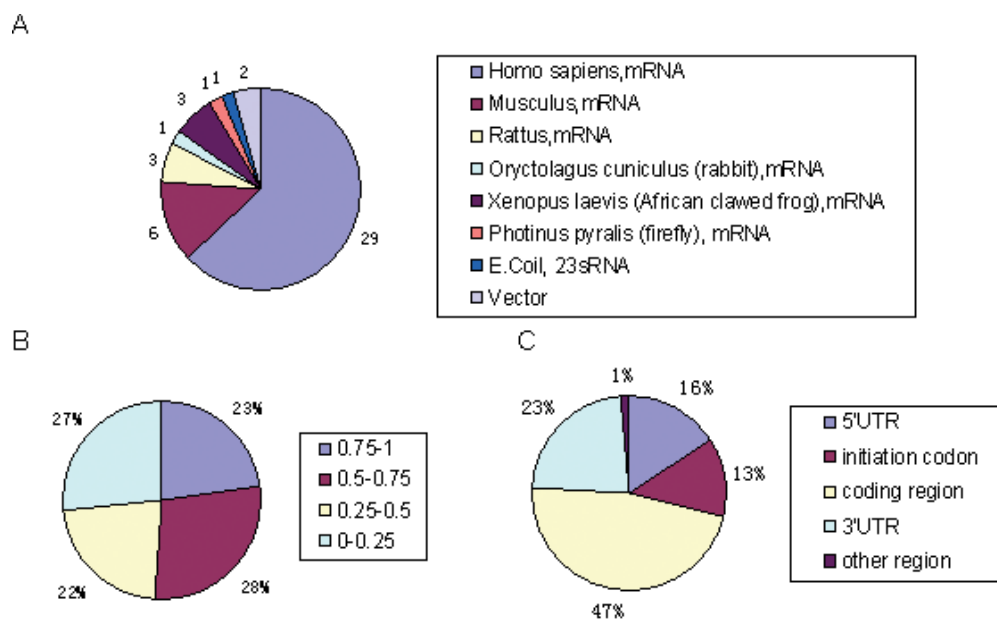


Figure 1. Overview of target and ODNs in AOBase. (A) Species distribution of target RNA molecular. (B) Efficacy of ODNs. (C) Target regions of ODNs.

DATABASE CONTENT AND IMPLEMENTATION

Criteria for data collection

Antisense ODNs stored in AOBase are all collected from published literatures. There have been a lot of techniques used in antisense studies, such as various chemical modification methods and different types of efficacy assay. Since the database is not specifically constructed for some ODNs effectiveness prediction system, the data selection criteria is not as strict as that used in efficacy prediction research (12,15). The reference selection criteria are (i) at least one scrambled or mismatched sequence was used as control; (ii) efficacy of ODNs was carried out against controls via *in vitro* assay at RNA or protein level; and (iii) efficacy was presented as a percentage of the control level of the target expression. The species of target RNAs are not restricted (Figure 1A). Presently, the database maintains ~700 ODNs against 46 different RNA molecules.

To keep in line with most of the researches on drug design, the ODNs efficacy in AOBase is transformed into $(1 - [\text{control expression}])$. The distribution of ODNs efficacy in the database is relatively uniform (Figure 1B). Considering that the selected ODNs were tested under different experimental conditions, some supplementary descriptions were also included in the database, such as chemical modifications employed for ODN synthesis, assay type used to measure the activity, concentration applied in test, etc.

Target region of ODNs

Target region selection is usually considered in ODNs design. Regions surrounding translation initiation codon are often chosen as target sites, since they are essential for gene expression and generally free from secondary structure (9). In the opinion that cleavage in 3'-untranslated regions (3'-UTRs) will

lead to rapid degradation of mRNA, the 3'-UTR of mRNA is also targeted frequently (9).

Target region of each ODN is annotated in AOBase. Bases of target RNA at different regions are marked with different colors shown in a detailed description page of each ODN (Figure 2G). The distribution of target regions in AOBase is shown in Figure 1C.

Local structures of target sites

Among the factors that influence the activity of ODNs, the local secondary structures of target RNA are well known to play a very significant role in determining ODNs efficacy *in vitro* (1,9,18). Although it is commonly accepted that computation-based structure models of long RNA molecules are not yet reliable enough to represent the RNA structure in living cells (19), the computational predicted structure of target RNA, especially the minimum free energy (MFE) structure, is still of particular importance in ODNs design strategies (20–22). Considering that single-stranded regions of RNA molecules play many important roles in RNA–RNA, RNA–DNA and RNA–protein interactions, the single-stranded probability profile, which is a kind of probabilistic representation of RNA structure, is also used in computer-aided ODN selection (23,24).

Both these two structural representations of ODN target region addressed above can be found in AOBase to help the structure–efficacy relationship exploration. For each target RNA, all the secondary structures within 5% of the computed MFE were predicted by MFold (25). The upper limit on the number of computed structures is 50. The corresponding single-stranded probability profile was estimated from the MFold ss-count file. In the detailed description page of ODN, the target site is highlighted in MFE structure illustration, while the single-stranded probability profile of target is plotted as waveform (Figure 2G).

ACKNOWLEDGEMENTS

This work is supported by a grant from the National Nature Science Foundation of China (No. 30171111), a grant from the National High Technology Research and Development Program of China (863 Program) (No. 2004AA234031) and a grant from the Special Funds for Major State Basic Research Program of China (973 Program) (No. 2004CB518904). Funding to pay the Open Access publication charges for this article was provided by Ministry of Science and Technology of China.

Conflict of interest statement. None declared.

REFERENCES

- Scherer, L.J. and Ross, J.J. (2003) Approaches for the sequence-specific knockdown of mRNA. *Nat. Biotechnol.*, **21**, 1457–1465.
- Taylor, M.F., Wiederholt, K. and Sveldrup, F. (1999) Antisense oligonucleotides: a systematic high-throughput approach to target validation and gene function determination. *Drug Discov. Today*, **4**, 562–567.
- Flaherty, K.T., Stevenson, J.P. and O'Dwyer, P.J. (2001) Antisense therapeutics: lessons from early clinical trials. *Curr. Opin. Oncol.*, **13**, 499–505.
- Leaman, D., Chen, P.Y., Fak, J., Yalcin, A., Pearce, M., Unnerstall, U., Marks, D.S., Sander, C., Tuschl, T. and Gaul, U. (2005) Antisense-mediated depletion reveals essential and specific functions of microRNAs in drosophila development. *Cell*, **121**, 1097–1108.
- Cheng, A.M., Byrom, M.W., Shelton, J. and Ford, L.P. (2005) Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. *Nucleic Acids Res.*, **33**, 1290–1297.
- Jansen, B. and Zangemeister-Wittke, U. (2002) Antisense therapy for cancer—the time of truth. *Lancet Oncol.*, **3**, 672–683.
- Crooke, S.T. (1998) An overview of progress in antisense therapeutics. *Antisense Nucleic Acid Drug Dev.*, **8**, 115–122.
- Far, R.K., Nedbal, W. and Sczakiel, G. (2001) Concepts to automate the theoretical design of effective antisense oligonucleotides. *Bioinformatics*, **17**, 1058–1061.
- Sohail, M. and Southern, E.M. (2000) Selecting optimal antisense reagents. *Adv. Drug Deliv. Rev.*, **44**, 23–34.
- Allawi, H.T., Dong, F., Ip, H.S., Neri, B.P. and Lyamichev, V.I. (2001) Mapping of RNA accessible sites by extension of random oligonucleotide libraries with reverse transcriptase. *RNA*, **7**, 314–327.
- Zhang, H.Y., Modn, J., Zhou, D., Xu, Y., Thonberg, H., Liang, Z. and Wahlestedt, C. (2003) mRNA accessibility site tagging (MAST): a novel high throughput method for selecting effective antisense oligonucleotides. *Nucleic Acid Res.*, **31**, e72.
- Camps-Valls, G., Chalk, A.M., Serrano-Lopez, A., Martin-Guerrero, J.D. and Sonnhammer, E.L.L. (2004) Profiled support vector machine for antisense oligonucleotide efficacy prediction. *BMC Bioinformatics*, **5**, 135.
- Chalk, A.M. and Sonnhammer, E.L.L. (2002) Computational antisense oligo prediction with a neural network model. *Bioinformatics*, **18**, 1567–1575.
- Giddings, M.C., Shah, A.A., Freier, S., Atkins, J.F., Gesteland, R.F. and Matveeva, O.V. (2002) Artificial neural network prediction of antisense oligodeoxynucleotide activity. *Nucleic Acids Res.*, **30**, 4295–4304.
- Sætrom, P. (2004) Predicting the efficacy of short oligonucleotides in antisense and RNAi experiments with boosted genetic programming. *Bioinformatics*, **20**, 3055–3063.
- Giddings, M.C., Matveeva, O.V., Atkins, J.F. and Gesteland, R.F. (2000) ODNBase—a web database for antisense oligonucleotide effectiveness studies. *Bioinformatics*, **16**, 843–844.
- Matveeva, O.V., Mathews, D.H., Tsodikov, A.D., Shabalina, S.A., Gesteland, R.F., Atkins, J.F. and Freier, S.M. (2003) Thermodynamic criteria for high hit rate antisense oligonucleotide design. *Nucleic Acids Res.*, **31**, 4989–4994.
- Vickers, T.A., Wyatt, J.R. and Freier, S.M. (2000) Effects of RNA secondary structure on cellular antisense activity. *Nucleic Acids Res.*, **28**, 1340–1347.
- Kurreck, J. (2003) Antisense technologies improvement through novel chemical modifications. *Eur. J. Biochem.*, **270**, 1628–1644.
- Lima, W.F., Monia, B.P., Ecker, D.J. and Freier, S.M. (1992) Implication of RNA structure on antisense oligonucleotide hybridization kinetics. *Biochemistry*, **31**, 12055–12061.
- Thierry, A.R., Rahman, A. and Dritschilo, A. (1993) Overcoming multi drug resistance in human tumor cells using free and liposomally encapsulated antisense oligodeoxynucleotides. *Biochem. Biophys. Res. Commun.*, **190**, 952–960.
- Laptev, A.V., Lu, Z., Colige, A. and Prockop, D.J. (1994) Specific inhibition of expression of a human collagen gene (COL1A1) with modified antisense oligonucleotides. *Biochemistry*, **33**, 11033–11039.
- Ding, Y. and Lawrence, C.E. (2001) Statistical prediction of single-stranded regions in RNA secondary structure and application to predicting effective antisense target sites and beyond. *Nucleic Acids Res.*, **29**, 1034–1046.
- Bo, X. and Wang, S. (2005) TargetFinder: a software for antisense oligonucleotide target site selection based on MAST and secondary structures of target mRNA. *Bioinformatics*, **21**, 1401–1402.
- Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization. *Nucleic Acids Res.*, **31**, 3406–3415.
- Smetsers, T.F., Boezeman, J.B. and Mensink, E.J. (1996) Bias in nucleotide composition of antisense oligonucleotides. *Antisense Nucleic Acid Drug Dev.*, **6**, 63–67.
- Tu, G.C., Cao, Q.N., Zhou, F. and Israel, Y. (1998) Tetranucleotide GGGA motif in primary RNA transcripts. Novel target site for antisense design. *J. Biochem.*, **273**, 25125–25131.
- Matveeva, O.V., Tsodikov, A.D., Giddings, M., Freier, S.M., Wyatt, J.R., Spiridinov, A.M., Shabalina, S.A., Gesteland, R.F. and Atkins, J.F. (2000) Identification of sequence motifs in oligonucleotides whose presence is correlated with antisense activity. *Nucleic Acids Res.*, **28**, 2862–2865.