W506–W511 *Nucleic Acids Research, 2005, Vol. 33, Web Server issue*
doi:10.1093/nar/gki453

AMOD: a morpholino oligonucleotide selection tool

Eric W. Klee[1,4], Kyong Jin Shim[3,4], Michael A. Pickart[2,4], Stephen C. Ekker[2,4] and Lynda B. M. Ellis[1,4,*]

[1]Department of Laboratory Medicine and Pathology, [2]Department of Genetics, Cell Biology and Development, [3]Department of Computer Science and Engineering and [4]Arnold and Mabel Beckman Center for Transposon Research, University of Minnesota, Minneapolis, MN, USA

Received February 11, 2005; Revised and Accepted April 1, 2005

ABSTRACT

AMOD is a web-based program that aids in the functional evaluation of nucleotide sequences through sequence characterization and antisense morpholino oligonucleotide (target site) selection. Submitted sequences are analyzed by translation initiation site prediction algorithms and sequence-to-sequence comparisons; results are used to characterize sequence features required for morpholino design. Within a defined subsequence, base composition and homodimerization values are computed for all putative morpholino oligonucleotides. Using these properties, morpholino candidates are selected and compared with genomic and transcriptome databases with the goal to identify target-specific enriched morpholinos. AMOD has been used at the University of Minnesota to design ~200 morpholinos for a functional genomics screen in zebrafish. The AMOD web server and a tutorial are freely available to both academic and commercial users at http://www.secretomes.umn.edu/AMOD/.

INTRODUCTION

Vertebrate genomes contain an estimated 20–30K genes involved in diverse processes; many encoding proteins with unknown function. The annotation of these genes remains a major step in understanding the vertebrate genome. The development of morpholino-based gene 'knockdown' technology provides a method for identifying function from primary sequence, on a genome-wide scale in many vertebrates (1–7). Sequence-driven screens for functional annotation of genomic data are being developed in systems that lack the high cost, significant time and infrastructure commitments associated with traditional model vertebrates. RNAi-based screening in nematode (8) and fly tissue culture cells (9) have applied 'knockdown' strategies to sequence-specific annotation, and siRNA has been effectively applied in mammalian tissue culture models (10). However, these approaches remain impractical for large-scale *in vivo* work.

Phosphorodiamidate morpholino oligonucleotides (morpholinos) are non-classical antisense reagents that modulate gene expression by inhibiting protein translation or inducing alternative splicing events. A synthetic DNA analog that contains a six-member morpholino ring and a neutral charge phosphorodiamidate backbone, morpholinos are resistant to nuclease digestion (8) and are freely water-soluble (9). Morpholinos overcome many limitations associated with traditional antisense reagents (11) and have been effectively used in many eukaryotes (1–7,11,12). The effect morpholinos cause on the expression is determined by the position targeted within a nucleotide sequence. Morpholinos targeting the 5′-untranslated regions (5′-UTRs) in proximity to the translational initiation site (TIS) disrupt ribosomal complex formation and inhibit protein translation of mRNA, while morpholinos targeting pre-mRNA splice sites can induce alternative splicing events (12–14). Consequently, effective morpholino design requires a clear understanding of nucleotide sequence characteristics in addition to the biochemical properties of the morpholino oligonucleotides. Since significant sequence analyses are required for informed morpholino design, the application of this technology to large-scale screens (5,15) necessitates a software tool capable of efficient and accurate target sequence selection and morpholino design.

Programs for siRNA design include some, but not all, of the necessary processes required for morpholino design (16–19). Both siRNA and morpholino design require computation of biochemical properties of short oligonucleotides, including base composition and homogeneous nucleotide run calculations. However, siRNA does not require a detailed analysis of oligonucleotide binding position relative to target nucleotide

*To whom correspondence should be addressed at Mayo Mail Code 609, 420 SE Delaware Street, Minneapolis, MN 55455, USA. Tel: +1 612 625 9122; Fax: +1 612 624 6404; Email: lynda@umn.edu
Present address:
Michael A. Pickart, Department of Biology, University of Wisconsin–Stout, Menomonie, Wisconsin, WI, USA

sequence features. Similar programs for morpholino-specific design are not currently available, although the sole commercial supplier of morpholinos, Gene Tools, offers a free, proprietary design service (http://www.gene-tools.com/) that requires prior knowledge of the translational start codon in the mRNA and provides very limited sequence design and analysis options to the user. AMOD implements morpholino design guidelines similar to those recommended by Gene Tools, such as avoidance of nucleotide motifs that form stable localized secondary structures or decrease water solubility. In addition, AMOD includes an integrated multiple-species sequence comparison and host-specific genomic sequence validation and uniqueness assessment capabilities. The resulting output provides the user with a range of potential oligonucleotide designs suitable for use in a variety of biological applications, including the most common use as inhibitors of mRNA translation or for the alteration of pre-mRNA splicing. AMOD is a transparent, versatile and effective tool for short oligonucleotide and primer design.

MATERIALS AND METHODS

AMOD is written in PERL (http://www.perl.org/) and uses HTML and JavaScript for the user interface. BioPerl modules (20) are used for BLAST parsing and nucleotide-to-protein sequence translation. TIS predictions are made using the ATGpr web server (21). Sequence-to-sequence comparisons are performed using a local installation of NCBI BLAST version 2.1.2. Sequence comparisons may be made against vertebrate RefSeq proteins and the Ensembl zebrafish genomic sequence set, housed in the Vertebrate Secretome and CTTome database (http://www.secretomes.umn.edu/).

RESULTS AND DISCUSSION

AMOD

The AMOD design process consists of six steps separated into two phases, as shown in Figure 1. Phase one includes steps to characterize the nucleotide sequence and aid users in identifying key sequence features, including the TIS and intron–exon splice sites. To facilitate the design of morpholinos for translational inhibition, TIS's are predicted using ATGpr, a linear discriminate analysis program (21). Nucleotide sequences are automatically submitted to the ATGpr web server and predicted TISs ranked by prediction reliability scores, indices defining the resulting open reading frame, and agreement with Kozak's consensus sequence (22) is obtained. This program is used in AMOD because of its superior performance when analyzing expressed sequence tag (EST) sequences (23). To compliment this analysis, AMOD also performs an internal scan for ATG start codons. The scan identifies all occurrences of the 'ATG' codon, the corresponding reading frame, assesses agreement to Kozak's consensus sequence and identifies the closest in-frame, upstream stop codon. Putative translational start sites identified by either method can be selected, and the TIS position highlighted in a display of the target gene sequence (Figure 2B).

Homology modeling is used to further characterize the nucleotide sequences and confirm the location of the TIS.

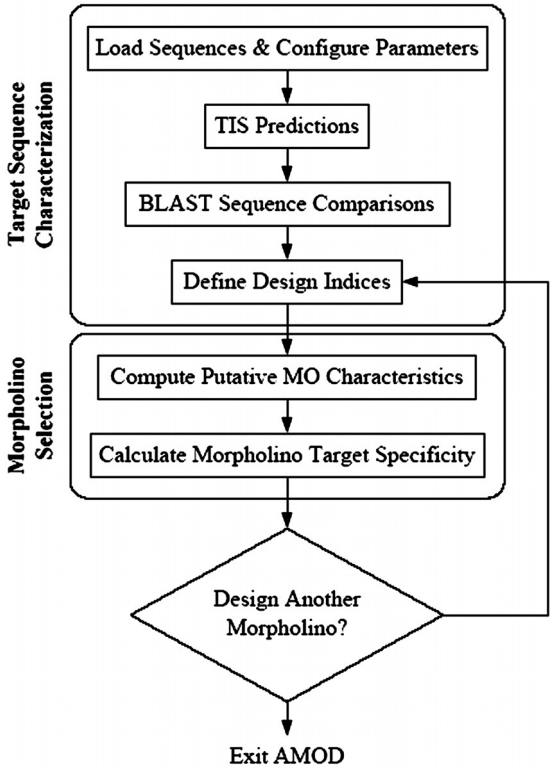

Figure 1. Process flow chart for AMOD. The program consists of six steps, encompassing two phases: target sequence characterization and morpholino selection.

Homology modeling is also useful for the identification of splice junctions for splice site morpholinos. Provided that 25 bases on either side of the splice junction can be identified, other parameters for the design of splice site morpholinos are either similar to TIS-targeted morpholinos or yet to be determined. Users may execute sequence comparisons against RefSeq protein sequences or against a user supplied reference sequence set on the AMOD server. Users can also choose to upload an output file from a previously executed BLAST comparison, provided the output is in text format and the target nucleotide sequences were from the BLAST query sequences. The 10 highest scoring alignments for a gene sequence are presented in a BLAST report and alignment summary (Figure 2A). For blastx comparisons, N-terminal alignment extensions are dynamically constructed to aid users in assessing the target gene's 5′ completeness, using methods described previously (24).

Phase two of the AMOD process entails the selection of a specific morpholino oligonucleotide, within the defined gene subsequence possessing biochemical properties to improve morpholino water solubility and efficacy as pioneered by Gene Tools. These considerations are the same for morpholinos targeting both TIS as well as splice junctions. AMOD creates an exhaustive list of putative morpholino oligonucleotides 25 bases long; the oligonucleotide length may be modified but is set to the size considered to result in optimal morpholino performance (25,26). For each morpholino, the G%, C%, A%, GC%, maximum number of consecutive adenines and maximum number of consecutive guanines in the oligonucleotide are calculated. Morpholino intra-strand

Figure 2. Target sequence evaluation phase. (**A**) The 10 highest scoring alignments for a gene sequence are presented in a BLAST report and alignment summary. (**B**) Putative start sites identified by either method can be selected, and the TIS position highlighted in gray text in a display of the target gene sequence.

binding is estimated by adding the number of hydrogen bonds occurring in all possible self-to-self, ungapped alignments. For example, Figure 3 illustrates a morpholino self-to-self alignment with an intra-strand binding score of 30, 18 from six G–C bonds and 12 from six A–T bonds. All intra-strand binding alignments are reported, and the alignments yielding the maximum global intra-strand binding and the maximum intra-strand binding constrained to a four-base window are noted (Figure 4A). The list of putative morpholino oligonucleotides can be filtered using base composition, or using the intra-strand binding score to allow users to identify sequences with the least intra-strand binding. Default thresholds for base composition and intra-strand binding properties are set to existing morpholino design guidelines; however, users may modify these values to fit custom design criteria.

A critical aspect of antisense reagent design is the evaluation of the reagent's target specificity (11,13). AMOD provides a second set of BLAST comparisons between select morpholinos and genomic sequences or the transcriptome to identify all sequences complementary to the selected morpholino oligonucleotides (Figure 4B). These comparisons enable a user to select a morpholino that maximizes binding to the target gene with high specificity while minimizing binding to other gene targets that may induce effects unrelated to the specific knockdown of the target gene.

User interaction

A design session on the AMOD web server begins with the entry of a user-selected target nucleotide sequence in the

FASTA format. This sequence may be either pasted in or a sequence file may be uploaded. AMOD then performs TIS calculations on all entered sequences and asks users whether BLAST comparisons are desired. Following the computation or bypass of BLAST comparisons, users are prompted to select a single input sequence against which morpholino design will proceed. Once a user has defined a subsequence, reviewed the putative morpholinos, and identified an acceptable oligonucleotide, the morpholino sequence characteristics are written to an output file. The user may then analyze another sequence from the input sequence set or exit the program.

```
-----------------------------------------------------
Maximum Binding Score for global alignment is 30
-----------------------------------------------------
CGTCCTTAGTAATAAGACTCCGTTC  :A
 |  |  |||  ||  |||  |  |
 CTTGCCTCAGAATAATGATTCCTGC :B
-----------------------------------------------------
```

Figure 3. Illustrates intra-strand binding approximation from a morpholino self-to-self alignment. In this figure, the total binding score is 30: 18 from six G–C bonds and 12 from six A–T bonds.

AMOD application

AMOD was used to design morpholinos for the University of Minnesota's functional genomics screen in zebrafish. This screen utilizes comparative genomic data mining strategies to identify conservation of higher-order biological processes among vertebrates. Use of morpholinos allows high-throughput rates of genetic analysis without the inherent redundancy of forward-genetic approaches that typically result in multiple mutations within the same gene. The screen has analyzed 255 gene sequences to date, using 189 morpholinos targeting 157 unique genes were designed. During AMOD's use, the program's functionality and user interface were evaluated and improved.

The positive results of the ongoing functional genomics morpholino screen validate the use of morpholinos for this approach (27) and demonstrate the capacity of AMOD to

Figure 4. Oligonucleotide selection phase. (**A**) All intra-strand binding alignments are reported, and the alignments yielding the maximum global intra-strand binding and the maximum intra-strand binding constrained to a four-base window are noted. (**B**) AMOD uses a second set of BLAST comparisons between select morpholinos and genomic sequences or the transcriptome to identify genes that could be targeted by the selected morpholino oligonucleotides.

guide the target sequence selection and morpholino design. In spite of this success, the analysis of relationships between phenotypic outcomes and sequence characteristics of morpholinos used in this screen have only identified a modest correlation between morpholino effectiveness and a minimum base composition threshold that was subsequently integrated into AMOD. No other improvement in parameters for morpholino design based on target gene sequence was strongly suggested by these outcomes. In addition, multiple factors arising from the use of a continually evolving zebrafish genomic sequence database and incomplete EST sequence data limit the ability to discern specific guidelines for morpholino design or to evaluate the capacity of AMOD to improve morpholino design from this initial functional genomics screen. The exhaustive, robust evaluation of morpholino design parameters and programs for morpholino design await the identification of a controlled reference set, including both efficacious and non-efficacious morpholinos. Until then, AMOD addresses the need to streamline target sequence selection and morpholino design, e.g. in support of high-throughput approaches for morpholino screening.

Comparison with other morpholino design services

The sole commercial supplier of morpholinos, Gene Tools, offers a free, proprietary design service (http://www.gene-tools.com/). To our knowledge, this is presently the only other morpholino design service, requires prior user knowledge of the translational start codon in the selected target mRNA, and provides very limited sequence design and analysis options. Based on the success of this proprietary method for morpholino design, AMOD was developed to implement similar morpholino design guidelines, such as avoidance of nucleotide motifs that form stable localized secondary structures or decrease water solubility. In an effort to improve on these initial design parameters and to incorporate the advancing genomic and transcriptome databases into morpholino design, AMOD includes an integrated multiple-species sequence comparison and host-specific genomic sequence validation and uniqueness assessment capabilities. The resulting output provides the user with a range of potential oligonucleotide designs suitable for use in a variety of biological applications, including their most common use as inhibitors of mRNA translation or for the alteration of pre-mRNA splicing. AMOD thus provides the user with a transparent design tool offering an increased number of oligonucleotide design options while implementing genomics-wide sequence criteria not found in the other available morpholino design option.

AVAILABILITY

AMOD web server is freely accessible to academic and commercial users at http://www.secretomes.umn.edu/AMOD/. A graphical step-by-step tutorial is available. For questions, comments and suggestions pertaining to AMOD, there is a link to an online 'Contact Us' form on every web page.

FUTURE DEVELOPMENTS AND CONCLUSIONS

The AMOD web server is a tool designed for selecting oligonucleotides for the construction of antisense morpholinos against target nucleotide sequences. Planned additions to the server include intron–exon and exon–intron splice site predictions, entropy calculations and a ranking system for putative morpholino designs.

ACKNOWLEDGEMENTS

The authors thank Michelle Knowlton, Aidas Nasevicius and Eric Mendenhall, for providing valuable feedback on AMOD's function and user interface. Vanessa Louis, 2004 University of Minnesota undergraduate Bioinformatics Summer Intern (NSF/NIBIB EEC-0234112), assisted with analysis of relationships between phenotypic outcomes and sequence characteristics. This work was supported in part by NIH R01-GM63904, a predoctoral traineeship to E.K. (NLM TG-0704l), and a postdoctoral traineeship to M.A.P. (NIDCR: T32 DE07288-8). Funding to pay the Open Access publication charges for this article was provided by NIH R01-GM 63904.

Conflict of interest statement. None declared.

REFERENCES

1. Kos,R., Reedy,M.V., Johnson,R.L. and Erickson,C.A. (2001) The wingedhelix transcription factor FoxD3 is important for establishing the neural crest lineage and repressing melanogenesis in avian embryos. *Development*, **128**, 1467–1479.
2. Boonanuntanasarn,S., Yoshizaki,G., Takeuchi,Y., Morita,T. and Takeuchi,T. (2002) Gene knock-down in rainbow trout embryos using antisense morpholino phosphorodiamidate oligonucleotides. *Mar. Biotechnol.*, **4**, 256–266.
3. Nasevicius,A. and Ekker,S.C. (2000) Effective targeted gene 'knockdown' in zebrafish. *Nature Genet.*, **26**, 216–220.
4. Mellitzer,G., Hallonet,M., Chen,L. and Ang,S.L. (2002) Spatial and temporal 'knock down' of gene expression by electroporation of double-stranded RNA and morpholinos into early postimplantation mouse embryos. *Mech. Dev.*, **118**, 57–63.
5. Yamada,L., Shoguchi,E., Wada,S., Kobayashi,K., Mochizuki,Y., Satou,Y. and Satoh,N. (2003) Morpholino-based gene knockdown screen of novel genes with developmental function in *Ciona intestinalis*. *Development*, **130**, 6485–6495.
6. Satou,Y., Imai,K.S. and Satoh,N. (2001) Action of morpholinos in Ciona embryos. *Genesis*, **30**, 103–106.
7. Heasman,J., Kofron,M. and Wylie,C. (2000) b-Catenin signaling activity dissected in the early Xenopus embryo: a novel antisense approach. *Dev. Biol.*, **222**, 124–134.
8. Summerton,J. and Weller,D. (1997) Morpholino antisense oligomers: design, preparation and properties. *Antisense Nucleic Acid Drug Dev.*, **7**, 187–195.
9. Summerton,J., Stein,D., Huang,S.B., Matthews,P., Weller,D. and Partridge,M. (1997) Morpholino and phosphorothioate oligomers compared in cell-free and in-cell systems. *Antisense Nucleic Acid Drug Dev.*, **7**, 63–70.
10. Corey,D.R. and Abrams,J.M. (2001) Morpholino antisense oligonucleotides: tools for investigating vertebrate development. *Genome Biol.*, **2**, reviews 1015.1–1015.3.
11. Heasman,J. (2002) Morpholino oligos: making sense of antisense? *Dev. Biol.*, **243**, 209–214.
12. Gebski,B.L., Mann,C.J., Fletcher,S. and Wilton,S.D. (2003) Morpholino antisense oligonucleotide induced dystrophin exon 23 skipping in mdx mouse muscle. *Hum. Mol. Genet.*, **12**, 1801–1811.
13. Ekker,S.C. and Larson,J.D. (2001) Morphant technology in model developmental systems. *Genesis*, **30**, 89–93.
14. Draper,B.W., Morcos,P.A. and Kimmel,C.B. (2001) Inhibition of zebrafish *fgf8* pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown. *Genesis*, **30**, 154–156.
15. Kenwrick,S., Amaya,E. and Papalopulu,N. (2004) Pilot morpholino screen in *Xenopus tropicalis* identifies a novel gene involved in head development. *Dev. Dyn.*, **229**, 289–299.

16. Henschel,A., Buchholz,F. and Habermann,B. (2004) DEQOR: a web-based tool for the design and quality control of siRNAs. *Nucleic Acids Res.*, **32**, W113–W120.
17. Dudek,P. and Picard,D. (2004) TROD: T7 RNAi oligo designer. *Nucleic Acids Res.*, **32**, W121–W123.
18. Naito,Y., Yamada,T., Ui-Tei,K., Morishita,S. and Saigo,K. (2004) siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference. *Nucleic Acids Res.*, **32**, W124–W129.
19. Yuan,B., Latek,R., Hossbach,M., Tuschl,T. and Lewitter,F. (2004) siRNA Selection Server: an automated siRNA oligonucleotide prediction server. *Nucleic Acids Res.*, **32**, W130–W134.
20. Stajich,J.E., Block,D., Boulez,K., Brenner,S.E., Chervitz,S.A., Dagdigian,C., Fuellen,G., Gilbert,J.G., Korf,I., Lapp,H. *et al*. (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res.*, **12**, 1611–1618.
21. Salamov,A.A., Nishikawa,T. and Swindells,M.B. (1998) Assessing protein coding region integrity in cDNA sequencing projects. *Bioinformatics*, **14**, 384–390.
22. Kozak,M. (1989) The scanning model for translation: an update. *J. Cell Biol.*, **108**, 229–241.
23. Nadershahi,A., Fahrenkrug,S.C. and Ellis,L.B. (2004) Comparison of computational methods for identifying translation initiation sites in EST data. *BMC Bioinformatics*, **5**, 14.
24. Klee,E.W., Carlson,D.F., Fahrenkrug,S.C., Ekker,S.C. and Ellis,L.B. (2004) Identifying secretomes in people, pufferfish and pigs. *Nucleic Acids Res.*, **32**, 1414–1421.
25. Summerton,J. (1999) Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochem. Biophys. Acta*, **1489**, 141–158.
26. He,J., Liu,G., Zhang,S., Vanderheyden,J.L., Liu,N., Liu,C., Zhang,Y., Gupta,S., Rusckowski,M. and Hnatowich,D.J. (2003) A comparison of *in vitro* and *in vivo* stability in mice of two morpholino duplexes differing in chain length. *Bioconjug. Chem.*, **14**, 1018–1023.
27. Pickart,M.A., Sivasubbu,S., Nielsen,A.L., Shiram,S., King,R.A. and Ekker,S.C. (2004) Functional genomics tools for the analysis of zebrafish pigment. *Pigment Cell Res.*, **17**, 461–470.