Design, Cloning and *In silico* Analysis of Efficient siRNA-inducing Cassette for Silencing Wheat γ-gliadins

Bahram Baghban Kohnehrouz* and Shahnoush Nayeri

Department of Plant Breeding & Biotechnology, University of Tabriz, Tabriz-51666, Iran

Received: September 13, 2015       Revised: December 14, 2015       Accepted: December 25, 2015

Abstract

Gluten is an important protein source for human beings and a wide diversity of foods has been developed to take advantage of this protein in wheat flour. However, some individuals, suffering from Celiac Disease (CD), cannot tolerate proline and glutamine-rich gluten peptides including γ-gliadins. A life-long gluten free diet is the only known effective treatment for such patients. However, the sensitivity intensity in CD patients can be controlled by RNA interference (RNAi) technology. The main aim of the present study is to develop an efficient and specific siRNA-inducing cassette, as a first and critical step for an effective targeting of mRNAs of wheat γ-gliadins. To achieve this aim, we have followed the strategy based on 200 bp in sense and antisense orientation with the ~160bp sequence of none potent siRNA-containing region from γ-gliadin gene as a spacer in between. The endosperm-specific γ-gliadin promoter was sub-cloned into upstream the cassette. The nucleotide alignment results validated the sequence data of γ-gliadin promoter and the direct inserts with high homology identities of 99% and 99.97%, respectively. Here, six potential and consecutively arranged-siRNA sites were predicted using computational approaches. All of these sites covered the inverted repeats region with high efficacy and performance values for triggering RNAi.

Keywords: Celiac disease, Common wheat, γ-gliadins, *In silico* analysis, siRNA-inducing cassette.

1. Introduction

Gluten is a complex mixture of protein components comprising the gliadins and glutelins in wheat and equivalent proteins in barley and rye (Rossell et al., 2014). Gliadins can be subdivided into four sub-fractions: α/β, γ and α-gliadins, whereas glutelins consist of low and high molecular weight glutelins (LMW, HMW) (Colomba and Gregorini, 2012). The γ-type gliadins that contribute to the visco-elastic properties of the dough are mainly heterogeneous collection of 30-78 kDa monomeric proteins with poor solubility in dilute salt solutions and good solubility in 70% ethanol (Guo et al., 2012; Shewry and Tatham, 1990; Wieser, 2007). These peptides account for approximately 12% of the flour proteins of the hexaploid wheat (Dupont et al., 2011). They are encoded by clusters of linked genes, present at the Gli-1 loci on the short arms of the respective homologous group 1 chromosomes of hexaploid bread wheat (*Triticum aestivum* L.) and are tightly linked to the Glu-3 (LMW glutenins) and Gli-3 (α-gliadin) loci (Dupont et al., 2011; Gao et al., 2007). The number of different γ-gliadin genes in the genome of bread wheat was estimated at 15-40 copies and these can be clustered into four up to eleven groups (Payne, 1987; Qi et al., 2009). γ-gliadins show a strong association with Celiac Disease (CD), a widely prevalent (0.5-1% of the general populations) chronic inflammatory condition of small intestine triggered by the ingestion of gluten fractions derived from wheat, barley and rye (Colomba and Gregorini, 2012; Ferretti et al., 2012; Van den Broeck et al., 2009). Since the several sets of CD-epitope cores (9-mer peptides) are located in the first variable domain R1 (domain II) region of γ-gliadins, they could be great initiators of CD (Anderson et al., 2012; Salentijn et al., 2012). These epitopes, generally with their highly proline (P) and glutamine (Q) content, are perfect substrate for transglutaminase reaction of tissue transglutaminase 2 (tTG2) enzyme, which are critical for the creation of active T-cell epitopes...

* Corresponding author. e-mail: bahramrouz@yahoo.com

**Abbreviations**: CD Celiac Disease, CTAB Cetyl Triethyl Ammonium Bromide, dNTPs deoxy nucleoside5´-triphosphates, dsRNA double strand RNA, hpiRNA hairpin RNA, HMW High Molecular Weight, IPTG Isopropyl β-D-1-thiogalactopyranoside, kD kilo Dalton, LB Luria Bertani, LMW Low molecular weight, MFE Minimum Free Energy, mRNA messenger RNA, NCBI National Center for Biotechnological Information, PCR Polymerase Chain Reaction, psRNAi plant short small RNA interfering tool, PTGS Post Transcriptional Gene Silencing, RISC RNA-Induced Silencing Complex, RNAi RNA interference, siRNA small interfering RNA, UPE Unpaired Energy, X-gal 5-bromo-4-chloro-3-indolylβ-D-galactopyranoside,
(Anderson et al., 2012; Kim et al., 2004; Meresse et al., 2012; Salentijn et al., 2012). Although a lifelong gluten-exclusion diet is the only available and effective treatment for CD patients, complying with a gluten-free diet is tedious and affects the patients’ quality of life (Gil-Humanes et al., 2014). According to some reports, most of down-regulation of gluten proteins in wheat and other cereals has been based on RNAi technology (Rosell et al., 2014). RNA silencing of specific individual gliadins or groups of prolamins may therefore, be of interest in relation to determining their role in both grain-processing properties and in triggering celiac disease (Gil-Humanes et al., 2008).

Since RNAi is a sequence-specific RNA degradation system that is conserved in a wide range of organisms, it was rapidly favored as a powerful research tool for the Post-Transcriptional Gene Silencing (PTGS) in plants. The mechanism of RNAi is similar in all eukaryotes and is triggered by double strand RNA (dsRNA) molecules through a two-step mechanism (Gil-Humanes et al., 2010). RNAi gene silencing protects the organism’s genome from transposons and viruses as well as a part of the defense system in plants (Kemp et al., 2013; Nicolas et al., 2013). In recent decade, with realizing the importance of RNAi technology for effective down-regulating in the expression of a plant gene of interest, the various specific RNAi-based vectors for the construction and the expression of hairpin-like RNA constructs in plant cells have been developed. The expression of an RNAi-inducing cassette results in a dsRNA molecules composed of two distinct regions: a single-stranded loop, encoded by the spacer region and a double strand stem encoded by an inverted repeats (Edman and Waterhouse, 2011).

The main aim of the present work was to generate the tissue specific and effective siRNA-inducing cassette, as a first and critical step for effective targeting of γ-gliadin transcripts from T. aestivum. Therefore, an effort has been exerted to identify potential siRNA candidates for silencing the target messenger RNA (mRNA) using a computational approach. Furthermore, we attempted to predict the secondary structure formed between siRNA and the target mRNAs.

2. Materials and Methods

2.1. Sequence retrieval and bioinformatics analysis

The total number of 7 γ-gliadin promoter and 140 mRNA sequences was identified as orthologous of wheat γ-gliadins genes from GenBank at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST/). CLUSTALW program (accurate) was used to carry out nucleotide sequence alignments separately among the sequences of γ-gliadin promoter and the total mRNA sequences of wheat γ-gliadin genes (McWilliam et al., 2013). The accession numbers of the γ-gliadin promoter, used for sequence alignment, were as follows: Wheat transformation vector pGhp-omega/alpha/beta, HM352558; T. aestivum, FJ231103, FJ234648, FJ234649, EF426565, and AF234647; Wheat γ-gliadin, M36999. The accession numbers of γ-gliadins mRNA from T. aestivum, used for sequence alignment analysis, were as follows: A total 103 complete gene sequences, including HQ631424, JN849087-96, KF412602-614, KC715955-67, KC715971-91, KC715996, KC716000-6004, JQ943400-406, JN849087, JN849090-93, AY338386-390 and FJ006589-623; a total of 37 pseudogene sequences, including FJ006678-83, KF412615-21, KC715968-70, KC715992-95, KC715997-999, KC716005-011, JN849083-86, JN849088, JN849089, and JN849094.HM352558 and FJ006593 accession numbers were selected for designing the primer sets based on alignment results respectively belonged to γ-gliadin promoter and γ-gliadin mRNA sequences.

2.2. Plant, plasmid, enzymes, and chemicals

Wheat (T. aestivum L.) plants were grown in a pod in a greenhouse condition. The pTG19-T cloning vector (Vivantis, USA) derived from pTZ19-R vector (Accession no. Y14835.1) was used as a plasmid backbone for all cloning purposes. Escherichia coli strain DH5α (Invitrogen, USA) was used as a host strain for molecular cloning. Restriction enzymes, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and Isopropyl β-D-1-thiogalactopyranoside (IPTG) were obtained from Thermo Fisher Scientific Inc. (USA). All the other chemicals were molecular biology grade (Merck, Germany). The sequence of the γ-gliadin siRNA-inducing cassette is available at Gene Bank under accession number KT732419.

2.3. Genomic DNA extraction

After a 24-hour darkness treatment for a full break-down of the cellular starch content, the total genomic DNA was isolated from 2-3 leaves growth stage using Cetyl Trimethyl Ammonium Bromide (CTAB) reagent method (Doyle and Doyle, 1987). The concentration of genomic DNA was measured spectrophotometrically and its quality was checked by 0.8% agarose gel electrophoresis.

2.4. Primer design and PCR amplification

The specific primer sets for γ-gliadin promoter were synthesized from accession number HM352558, direct and inverted fragments from FJ006593. All primers were designed using Primer3 software at NCBI and analyzed by Oligo analyzer software (version 7.56). The restriction sites, arbitrarily defined (underlined letters), were added at the 5'-ends of forward and reverse primer sets (Table 1).
Table 1. Primers used for the synthesis of the siRNA-inducing cassette

<table>
<thead>
<tr>
<th>Primer + The enzyme added at the 5'-end</th>
<th>Description</th>
<th>Sequence 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ-GliPromoter_F+ HindIII</td>
<td>Forward primer for the synthesis of γ-gliadin promoter</td>
<td>ATAACGTTTCCAGAAAAAAGTTTCTA</td>
</tr>
<tr>
<td>γ-GliPromoter_R+ PstI</td>
<td>Reverse primer for the synthesis of γ-gliadin promoter</td>
<td>ATCTGACAGGGTGATTTGGTAAAACATAC</td>
</tr>
<tr>
<td>γ-GliDirect_F+ KpnI</td>
<td>For the synthesis of the 360bp γ-gliadin direct fragment</td>
<td>ATGGTACGGCCACACAAACAAACAGTCC</td>
</tr>
<tr>
<td>γ-GliDirect_R+ NheI</td>
<td>Reverse primer for the synthesis of the 360bp γ-gliadin direct fragment</td>
<td>ATGCTAAGCTTTCAGCGGTTCTACGTT</td>
</tr>
<tr>
<td>γ-GliInverted_F+ XbaI</td>
<td>For the synthesis of the 200bp γ-gliadin inverted fragment</td>
<td>ATTCTAGAAACAAACATTCCCCAACGA</td>
</tr>
<tr>
<td>γ-GliInverted_R+ PstI</td>
<td>Reverse primer for the synthesis of the 200bp γ-gliadin inverted fragment</td>
<td>ATCTGACAGTGACTGAATCGCCGTTTG</td>
</tr>
</tbody>
</table>

2.5. Molecular cloning

First PCR-amplified promoter and direct fragment were purified by gel purification kit (Bioneer, South Korea; cat. no. k-3035) according to the manufacturer’s instructions. Then, they were TA-cloned into linear KpnI (peptone 1% (w/v), yeast extract 0.5% (w/v), NaCl 1% (w/v), agar 1.2% (w/v)) containing ampicillin (100μg/ml), X-gal (100μg/ml) and IPTG (1mM). The positive clones for direct and promoter inserts were sequenced. At the second step, the inverted fragment was amplified from plasmid containing direct fragment. At the final step, the inverted and promoter fragments were sub-cloned at upstream of the direct insert and the cassette, respectively. The cassette was validated by HindIII/NheI double restriction digestion and verified by sequencing. T7 promoter primer (TAATACGACTCACATTAGGG), presented in Figure 2, was used for sequencing.

2.6. Determination of multiple siRNA candidates

So far, a number of experimental rules on siRNA duplex features have been frequently reported. These include the asymmetry rules for siRNA duplex ends, high A/U content at the 5'-end of the antisense strand, high G/C content at the 5'-end of sense strand, 30-50% GC content, thermodynamic properties in terms of the secondary structures, accessibilities of siRNA and target mRNA of gene(s) of interest (Nur et al., 2014). Therefore, the plant short small RNA interfering tool (psRNAit web tool), which provides a functional and an off-target minimized siRNA design, was used to in-silico identify efficient siRNAs candidates in the target mRNAs. The sequence data of the direct insert were used to determine potential siRNA sites in the target repeat.

2.7. GC calculation and off-target alignment

The DNA/RNA GC content calculator (http://www.endmemo.com/Tools/Biology) was used to calculate the GC content of the predicted siRNA. For checking any off-target sequence resemblance in other non-targeted organism’s genome, the BLAST tool (http://www.ncbi.nlm.nih.gov/blast) was used against the whole GenBank database by applying expected thresholds value 10 and BLOSUM 62 matrix as parameter (Nur et al., 2014).

2.8. Prediction of the secondary structure

The UNAfold program module in mfold web server (http://mfold.rna.albany.edu) was used to predict the secondary structure formed between mRNA and RNAi guide strands (Zuker, 2003).

3. Results and Discussion

Since RNAi technology is being used for down-regulating wheat γ-gliadin genes, an efficient inducing cassette is needed (Gil-Humanes et al., 2010). To reach this aim, it is important to develop the efficient hpRNA-inducing cassette as a critical triggering to target mRNAs of wheat γ-gliadins. Therefore, we have designed and cloned the γ-gliadin-specific siRNA-inducing cassette to produce a ~200bp stem with six siRNAs sites harboring a ~360bp direct fragment, was determined using psRNAit server tool (Table 2). Twenty out of the total of thirty-two siRNA candidates were selected with 7-10 and 8-10 scores for off-target accessibility and efficiency values, respectively. For further validation of reduced off-targets, all siRNA are subjected to BLASTn. As it is important to ensure whether RNAi binds RISC or not, results presented, as in Table 2, prove the successful incorporation of siRNA antisense strand with RISC. All the siRNA holds GC content within 36-48% were determined by GC content calculator. Lower GC content leads to stronger inhibitory effects and 30-57% of GC content is considered sufficient for the execution of RNAi’s action (Nur et al., 2014; Bret et al., 2005; Liu et al., 2013). Consequently, after the analysis of all siRNA
candidates, six potential siRNAs, which were consecutively arranged in the target repeats, were considered as the efficient candidates for effective triggering γ-gliadin target mRNAs.

3.2. The secondary structures of siRNA-target mRNA duplexes

Secondary structures of RNAi do not only provide a convenient and computationally tractable estimation to structures but they also provide them to the thermodynamics of RNA-RNA interaction (Nur et al., 2014). Furthermore, the secondary structure of RNAi can be involved in the RISC cleavage. Therefore, the prediction of the siRNA-mRNA secondary structure is the most important factor to select the efficient guide strand of siRNA sites for the effective triggering of RNAi (Bret et al., 2005; Liu et al., 2013). UNAfold software predicts the most stable secondary structure of an oligonucleotide by minimizing folding free energy. The secondary structures of the selected efficient siRNAs are shown in Figure 1(A)-(F). The ΔG-binding of siRNA-mRNA duplexes with minimum free energy are shown in Table 2.

![Figure 1](image)

**Figure 1.** (A)-(F). Schematic diagram of the secondary structures of duplexes between antisense strands of the selected six siRNAs (left strands) and the target mRNA of wheat γ-gliadins. The folded structures are predicted by minimizing folding free energy at the specified temperature (28 °C).

**Table 2.** Effective siRNA molecules with target accessibilities, efficiency, GC percentage, free energy of binding with target

<table>
<thead>
<tr>
<th>Target No.</th>
<th>Location of target within mRNA</th>
<th>Sequence of siRNA duplex</th>
<th>GC%</th>
<th>ΔG binding (k/Cal)</th>
<th>UPE</th>
<th>siRNA Efficiency</th>
<th>RISC binding antisense scores</th>
<th>RISC binding sense scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>256-276</td>
<td>GUUUGUUGAGCUCGGGAAAU</td>
<td>42.85</td>
<td>-38.47</td>
<td>7.08</td>
<td>9.2</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UACAAACACUCGGACCCCCUUU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>284-304</td>
<td>GGGUCUCGUACUGGGAAAUUGU</td>
<td>47.61</td>
<td>-38.49</td>
<td>8.43</td>
<td>8.18</td>
<td>0.4</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UUCCAACAAUGACCCCCUUAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03</td>
<td>165-185</td>
<td>GGGCUGCCGGAAUUGUGUUG</td>
<td>50.08</td>
<td>-38.86</td>
<td>8.37</td>
<td>7.97</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACCCGACGCCCUUUAACCAACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04</td>
<td>304-325</td>
<td>UGGGGAAGGGUUGUAUCUGG</td>
<td>47.60</td>
<td>-38.02</td>
<td>8.58</td>
<td>7.59</td>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CGACCCCCUUAACCCUUAGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05</td>
<td>219-239</td>
<td>GGGAAAUGUAGUUGGGGUGUCG</td>
<td>47.98</td>
<td>-30.76</td>
<td>9.34</td>
<td>7.41</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACCCUUACCACAAAACCCCCCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06</td>
<td>268-289</td>
<td>AAGGUGAUGUUCGGCUUGULGG</td>
<td>47.61</td>
<td>-30.42</td>
<td>10.2</td>
<td>7.01</td>
<td>0.33</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CUUACCACUAACCCCCAGCAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* Off-target accessibility or unpaired energy (UPE): the energy required to open mRNA secondary structure around target site is represented by UPE score in the range of 0-25 accessible scores. The less UPE score means the more possibility that siRNA is able to contact with target mRNA, which leads to the less off-target accessibility (Dai and Zhao, 2011).

*b* siRNA efficiency: Efficiency denotes the effectiveness of designed siRNA to silence transcripts. The efficiency range can vary from 0-10, higher the value greater silencing of transcript (Dai and Zhao, 2011).

*c* RISC binding score including sense and antisense strands indicate the binding abilities of each siRNA strands to RISC complex for triggering RNAi. The binding score range vary from -0.2 to 2 values and the higher score shows an efficient binding to RISC complex.
3.3. Cloning and screening of the specific siRNA-inducing cassette

To achieve such a high performance of siRNA-inducing molecules, the endosperm-specific γ-gliadin promoter was used for a specific and a strong expression in endosperm, representing approximately 80% total grain proteins (Rosell et al., 2014; Dupont et al., 2011; Gil-Humanes et al., 2010; Piston et al., 2013; Piston et al., 2009). The alignment result of the γ-gliadin promoter insert showed a maximum homology with the γ-prolamin promoter sequences from public databases and the previously reported γ-gliadin promoter with efficient performance of 98% and 100%, respectively. The nucleotide alignment of 360bp direct insert with 140 γ-gliadin mRNA sequence data were performed, and they showed average homology identities of 99.97% (accession numbers of data presented at materials and methods). According to the high affinity in binding RISC to antisense strand of double stranded siRNAs, the orientation of direct and inverted fragments was adjusted as in Figure 2. Consequently, the promoter fragment was sub-cloned at upstream of the inverted fragment, which was considered as a sense strand.

![Figure 2. Schematic presentation of wheat γ-gliadins-specific hpRNA-inducing cassette cloned into pTG19-T cloning vector with ~4350bp-long.](image)

In the cloning of the direct fragment, the correct colonies were screened out by standard blue/white screening system. The DNA plasmid with 360bp direct fragment was confirmed by NheI restriction digestion (Figure 3(A)). The inverted fragment was sub-cloned during a ligation-restriction step to the plasmid DNA with direct insert by double digestion of PstI/XbaI restriction sites. In Figure 3(B), the plasmid DNA, containing the inverted repeats (~4350bp in length, Lane2), was distinguished from the plasmid harboring only direct fragment (~3250bp in length, Lane3). The final size of DNA plasmids harboring the siRNA-inducing cassette (~1500bp in length) was validated by HindIII/NheI double restriction digestions (Figure 3(C)).

4. Discussion

RNAi technologies make it possible to effectively down-regulate the target mRNA of γ-gliadin multigenes from wheat, which causes celiac disease symptoms in genetically predisposed individuals. Thus, we have designed and developed an efficient and a specific siRNA-inducing cassette as a critical step for targeting wheat γ-gliadins. We have followed the strategy based on 200bp in sense and antisense orientation with a ~160bp sequence of none potent siRNA-containing region of γ-Gliadin direct insert as a spacer in between. To achieve the high performance siRNA-inducing cassette, the previously reported endosperm-specific γ-gliadin strong promoter was used to drive siRNA-inducing cassette. The nucleotide alignment results validated the sequence data of γ-gliadin promoter and direct inserts with high homology identities of 99% and 99.97%, respectively. Here, six potential and consecutively arranged siRNA sites were predicted in the inverted repeat region using computational approaches. The consequences of the present study demonstrated that these potential siRNA sites have high efficacy and performance values for triggering of RNAi.

Acknowledgements

This work was financially supported by Department of Plant Breeding and Biotechnology, University of Tabriz, Tabriz, Iran.
Natl Acad Sci