EFFECT OF NON-CODING RNA ON POST-TRANSCRIPTIONAL GENE SILENCING OF ALZHEIMER DISEASE

Krishnan Namboori P.K. , Arun Mohan, Radhagayathri. K.U., Varun Gopal K., Vasavi. C.S., Premkumar P., Deepagopakumar, K.I. Ramachandran

Computational Chemistry Group (CCG),

Computational Engineering and Networking, Amrita Vishwa Vidyapeetham, Ettimadai, Coimbatore-641105, Tamilnadu, INDIA

n_krishnan@cb.amrita.edu

ABSTRACT

A large amount of hidden biological information is contained in the human genome, which is not expressed or revealed in the form of proteins; the usual end product form of gene expression. Instead, most of such information is in the form of non-coding RNAs (ncRNAs). ncRNAs correspond to genes that are transcribed, but do not get translated into proteins. This part of the genome was, till recently, considered as 'junk'. The term 'junk' implied lack of any discernible function of these RNA. More than 98% of the human genomic size encompasses these non-coding RNAs. But, recent research has evidently brought out the indispensible contribution of non-coding RNA in controlling and regulating gene expression. ncRNA such as siRNAs and microRNAs have been reported to greatly help in causing post-transcriptional gene silencing (PTGS) in cells through RNA interference (RNAi) pathway. In this work, we have investigated the possibility of using siRNAs and microRNAs to aid in gene silencing of early onset Alzheimer's disease genes.

Alzheimer's disease specific mutations and their corresponding positions in mRNA have been identified for six genes; Presenilin-1, Presenilin-2, APP (amyloid beta precursor protein), APBB3, BACE-1 and PSENEN.

Small interfering RNAs (siRNAs) that can cause PTGS through RNA interference pathway have been designed. RNA analysis has been done to verify complementarity of antisense siRNA sequence with target mRNA sequence. Interaction studies have been done computationally between these antisense siRNA strands and seven Argonaute proteins. From the interaction studies, only one of the seven Argonaute proteins; 1Q8K, was found to have interaction with the siRNAs indicating the importance and uniqueness of this particular protein in RISC (RNA induced silencing complex).

The interaction studies have been carried out for the microRNAs also. Out of the 700 mature human microRNAs collected, 394 microRNAs have been identified to show partial complementarity with their target sequence on PSEN-1 mRNA. Of these 394, five microRNAs have shown partial complementarity to early onset Alzheimer's disease specific mutations in PSEN-1 mRNA. Interaction studies have been done between these microRNAs and Argonaute proteins. Thus, design, characterization and analysis of ncRNAs that contribute to post transcriptional gene silencing of Alzheimer's disease have been achieved.

INTRODUCTION

An estimation of 'United Nation population projection' reminds that the number of people older than 80 years will become approximately 370 million by the year 2050. Current estimation is that about half the strength of people older than 85 years are affected with dementia. This statistics warns that within 50 years, above 100 million people will suffer from dementia.

The most common cause of dementia is Alzheimer's disease (AD). Alzheimer's disease is caused by the deposition of amyloid plaques. Amyloid plaques are produced by the conversion of amyloid precursor protein (APP) into - amyloid (A). Mutations in genes encoding amyloid precursor protein (APP), presenilin-1 (position in chromosome 14), and presenilin-2 (position in chromosome 1), APBB3, BACE1 and PSENEN causes Alzheimer's disease. Thus mutated protein formed ultimately leads to Alzheimer's disease.

Expression of mutated genes can be prevented by gene silencing techniques. Gene silencing is done at the transcriptional and post transcriptional levels. Post transcriptional gene silencing results in the destruction or inactivation of the mRNA of the intended gene. This prevents the mRNA from producing an active gene product; a protein. The most common mechanism of post transcriptional gene silencing (PTGS) is the RNA interference (RNAi), wherein the mRNA inactivation is done by means of non-coding RNAs (ncRNAs). The term non-coding RNA (ncRNA) means RNA that does not encode a protein. Such RNAs can have specific biological functions. For example, tRNA which is non-coding has specific function in transport of amino acids during translation. Noncoding RNA has a role in heterochromatic silencing, silencing of transposable elements (TEs), unpaired DNA in meiosis, and developmentally excised DNA [3]. Most of the functions of ncRNAs are unknown. Non-coding RNAs are formed at a higher rate than messenger RNAs [4]. Some of the different ncRNAs are small interfering RNAs (siRNAs), microRNAs (miRNAs), piwi interacting RNAs (piRNAs), transfer RNAs (tRNAs), ribosomal RNAs (rRNAs) and so on.

Double stranded non-coding RNAs (dsRNA; >200 nucleotides length) can initiate RNA interference pathway, once injected into the cell. Initially, the dsRNA is cleaved into smaller segments by an RNase enzyme called Dicer. These small segments (20-25 nucleotides length) are called small interfering RNAs (siRNAs). Each double stranded siRNA has a sense strand and an antisense strand. The sense strand is identical to a portion of the target mRNA sequence and the antisense strand is complementary to this sequence. The antisense strand then binds to the Argonaute protein present in the RNA induced silencing complex (RISC) of the cell cytoplasm. The RISC is a multiprotein endonuclease complex that catalyzes the destruction of the target mRNA, in the RNA interference pathway.

The Argonaute protein has been identified to have a major role in RNA interference pathway. When double stranded siRNA is introduced into the cell, one strand of siRNA; the antisense strand binds with Argonaute protein. When binding of the antisense strand occurs the sense strand gets unwound from the siRNA duplex and gets dissolved. Even though sense

strand is a part of mRNA, mRNA does not get dissolved because its dimensionality is very high compared to sense strand. If an mRNA complementary to Argonaute bound RNA is detected, then it can bind to the mRNA complementarily through base pairing between target mRNA sequence and Argonaute bound antisense siRNA. This binding would thereby cleave the mRNA at its central position and eventually degrade it. Thus the gene corresponding to that code is silenced at the mRNA level and protein encoding by that mRNA cannot be further done [1]. siRNA can be created in test tube conditions or can be synthesized from the double stranded RNA viruses. Computationally, siRNA can be designed in such a way that its sense strand is identical to the target sequence present inside mRNA and antisense strand complementary to this target sequence.

MicroRNAs of length 21-22 nucleotides are found in eukaryotes that can act through RNA interference (RNAi). A complex of miRNA and enzymes can break down the target mRNA. Here, the miRNA is complementary to the target mRNA and blocks the mRNA from being translated or accelerates its degradation [6].

Small interfering RNA (siRNA) and microRNA (miRNA) are anchored into specific binding pockets of Argonaute protein and guide it to target mRNA molecules to prevent decoding of protein information during translation. Argonaute proteins are evolutionarily conserved. They can be phylogenetically subdivided as Ago subfamily and the Piwi subfamily. Ago proteins are ubiquitously expressed. They can bind to siRNAs or miRNAs to guide post-transcriptional gene silencing. Piwi proteins are mostly expressed in the germ line and these proteins can associate with piRNAs to facilitate silencing of mobile genetic elements [7]. Thus, Argonaute proteins can interact or bind to small non-coding RNAs with its functional domains and control protein synthesis by affecting messenger RNA stability and they can also participate in the production of a new class of small RNAs called Piwi-interacting RNAs [8].

Several domains of Argonaute proteins have already been identified in RNA interference. The PAZ domain constitutes 110 amino-acids and is named so because it is

contained in Piwi, Argonaute and Zwille/Pinhead proteins [10]. The Tudor domain exhibits a conserved negatively charged surface which can interact with the C-terminal Arg and Gly-rich tails of proteins [11]. Thermo nuclease domain can catalyze the hydrolysis of both DNA and RNA at 5' position of the phosphodiester bond and give 3'-mononucleotides and dinucleotides. Two arginines and one glutamate are implicated in the catalytic mechanism. [12].

MATERIALS AND METHODS

Mutation records on genes that lead to Alzheimer's disease have been collected from 'Human Genome Mutation Database' (HGMD) [13]. mRNA sequences of these AD causing genes have been collected from NCBI database [14].

Mutations have been identified in genes that code for presenilin-1, presenilin-2, amyloid precursor protein (APP), BACE-1 and PSENEN proteins. These mutations in genes are copied into open reading frame (ORF) of mRNA during transcription. Thus, the corresponding mutations in ORF of mRNA have been identified. The target regions on mRNA into which siRNAs can bind have been identified using siRNA target finder tool [15]. Sense strands of siRNA of length 20-25 nucleotides have been designed that are identical to the mutated regions of mRNA. Antisense strands of siRNA have been designed by considering complementarity with the sense strands. Thus the designed complementary antisense siRNA strand is also complementary to the mutated region of target mRNA. Hyperchem tool has been used to build all antisense siRNAs insilco [17].

Guanine and cytosine content (GC content) of all antisense siRNAs have been identified. If GC content of one strand of siRNA is greater than 60% it cannot become a functional siRNA because bond between guanine and cytosine is a triple bond. These bonds are much stronger compared to A-U or A-T double bonds and each strand of siRNA can have a tendency to form loops by itself . This may prevent the strands from binding to their target mRNA. Therefore GC percentage has been checked for each siRNA strand.

RNA analysis of the antisense siRNA strands has been done using RNAfold tool to identify self-complementarity among nucleotides, secondary structure formation tendency with respect to base pair distance within the sequence and thermodynamic energy differences among different antisense siRNA sequences. Complementarity of bases between antisense siRNA strand and the target mRNA sequence and thermodynamic energy of antisense siRNA strand bound to target mRNA sequence were found using RNAcofold tool [16].

Seven Argonaute proteins; 1Q8K, 1S12, 1S13, 2E6N, 2HQE, 2HQX and 3BDL, in human were collected from Protein Data Bank (PDB) database [18]. These proteins have been subjected to primary and secondary structure analyses. Secondary structure, class and species details of proteins were collected from SCOP database [20]. Percentage of alpha helix and beta turns were identified for each proteins using SOPMA tool [21]. Prosite tool was used to find domains in proteins [22]. A domain is a region in the protein, where an active site or motif can be found out. ProtParam tool was used to find molecular weight, theoretical PI, hydrophilicity and stability of proteins [23]. If GRAVY (Grand Average Hydropathicity) is positive, then protein is hydrophobic and if negative, protein is hydrophilic. Antisense siRNAs were docked with the seven Argonaute proteins present in human using CDOCKER tool of Accelrys Discovery studio [24].

Mature human microRNAs were collected from miRBase database to analyze whether they can bind to Alzheimer's disease specific mutated regions on target mRNA [25]. Out of 700 mature human microRNAs, 394 microRNAs have been identified using microINSPECTOR tool which showed partial complementarity with their target sequence on PSEN-1 mRNA [26]. Of these 394 microRNAs, five microRNAs have shown partial complementarity to early onset Alzheimer's disease specific mutations in PSEN-1 mRNA. Interaction studies have been done between these microRNAs and Argonaute proteins.

RESULTS AND DISCUSSIONS

Antisense strands of siRNA of length 20 nucleotides have been designed complementary to nine early onset Alzheimer's disease specific mutated regions in mRNA encoded by PSEN-1 gene. Two siRNAs show high GC content in its antisense strand (above 60%). siRNA with GC content greater than 60% are non-functional. So, it cannot be consider in post transcriptional gene silencing. 46 functional siRNAs are designed against 72 Alzheimer's disease specific mutated regions in mRNA encoded by PSEN-1 gene. 53 functional siRNAs have been designed for 81 mutated regions in mRNA encoded by PSEN-1 gene. Details of siRNAs designed against mutated regions on mRNA of PSEN-2, APP, APBB3, BACE-1 and PSENEN genes are given in Table 7.1. The sequences that have high base-pair complementarity among themselves are not feasible to be considered in post transcriptional gene silencing because if these strands form rigid secondary structures by themselves, then their binding sites to the target mRNA may decrease. Fig.7.1 represents thermodynamic free energy of antisense strands designed complimentary to mutated region of PSEN-1 mRNA. X-axis represents thermodynamic free energy of each antisense strand complimentary to the target sequence in the mRNA and Y-axis represents codon position numbers where mutation occurred in the mRNA. Antisense strands with relatively strong internal structures are not feasible in post transcriptional gene silencing because they may make antisense strands incapable of binding to their target mRNA sequences. Thermodynamic free energy of antisense siRNA strand bound to target mRNA sequence has been identified as thermodynamically supportive.

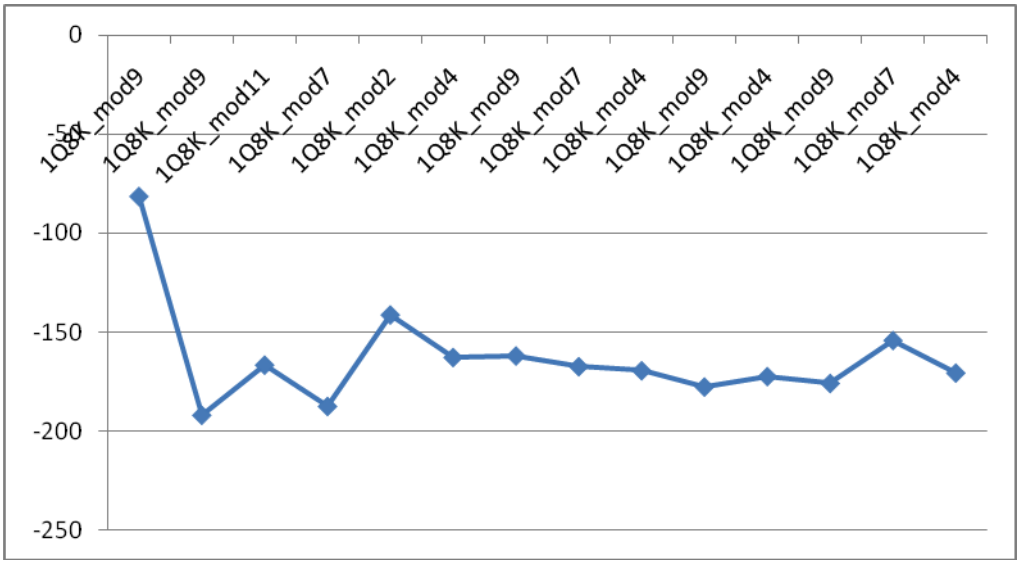
Table-1 Details of siRNAs designed against mutated regions on mRNA of PSEN-2, APP, APBB3, BACE-1 and PSENEN genes

Gene names encoding mutated mRNA	Number of functional siRNAs designed	Average GC percentage	Complementarity of anti-sense strand of siRNA towards target mRNA	Codon position numbers on ORF of target mRNA in which their complimentary antisense strands shown interaction with Argonaute protein
Presenilin -1	53	40%	full	35,116,123,143, 214, 206, 245, 386, 377, 378, 431, 435, 436, 213, 267, 268, 269, 278, 213, 219, 365, 280, 139, 120, 209,115, 409, 282, 284, 394, 384, 318, 139, 222, 113
Presenilin -2	3	52.35%	full	122
APP	10	38.1%	full	692,693,665,677,678,713, 714,715,716,717
APBB3	1	46.67%	full	318
BACE1	2	45.72%	full	118,452
PSENEN	1	42.9%	full	90

Interaction studies have been done computationally between these antisense strands and Argonaute proteins. Interaction of these antisense strands has been reported with only one out of the seven Argonaute proteins i.e. 1Q8K. It can be suggested that, antisense siRNA strands that do not show interaction with the antisense strands are not feasible in post transcriptional gene silencing because only those antisense siRNA strands that form a complex

with Argonaute can bind to the target mRNA. This binding cleaves the target mRNA or prevents it from being used as a translation template. The result of computation of interaction energy between antisense siRNA strands, complementary to mutated regions on PSEN-1 mRNA, and conformers of Argonaute protein 1Q8K has been included in Fig.7.1.

Fig. 7.1 **Interaction energy between antisense siRNA strands (complementary to mutated regions on PSEN-1 mRNA) and 1Q8K protein conformers**



Similarly, the results of computation of interaction energy between conformers of argonaute protein 1Q8K and antisense siRNA strands, complementary to mutated regions on PSEN-2 mRNA, APP mRNA and BACE-1 mRNA have been included in Fig.7.2, Fig.7.3 and Fig.7.4. All these interactions are exothermic and thermodynamically supportive.

Fig.7.2 **Interaction energy between antisense siRNA strands, (complementary to mutated regions on PSEN-2 mRNA) and 1Q8K protein conformers**

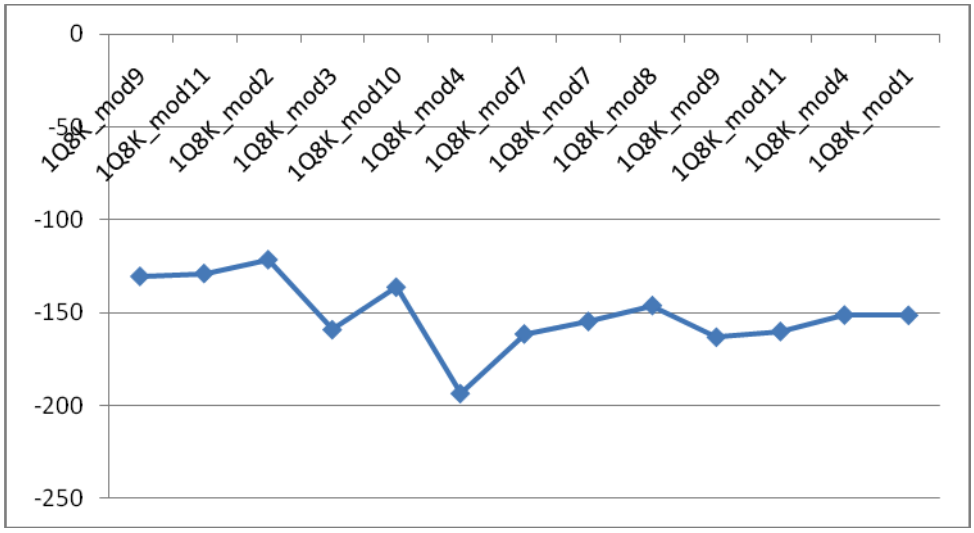


Fig.7.3 **Interaction energy between antisense siRNA strands complementary to mutated regions on APP mRNA and Argonaute proteins.**

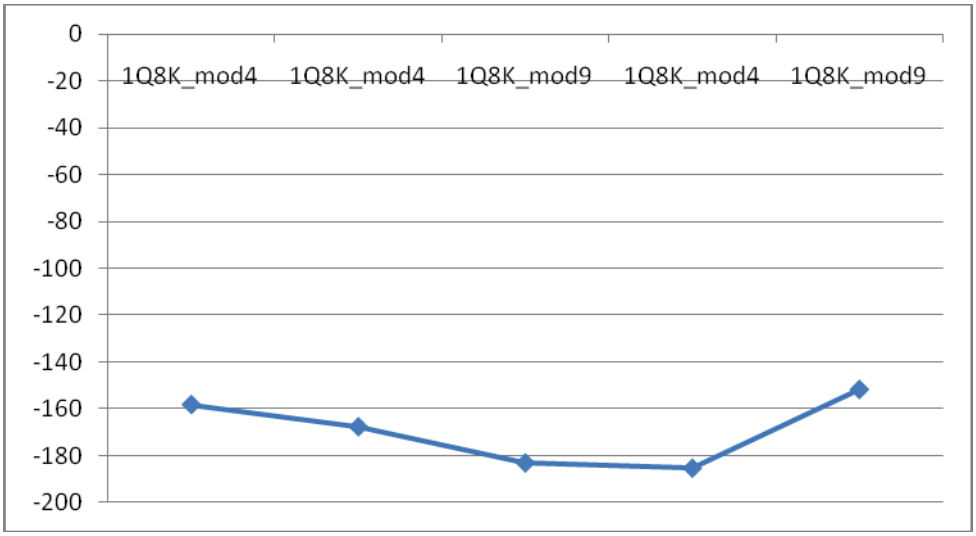
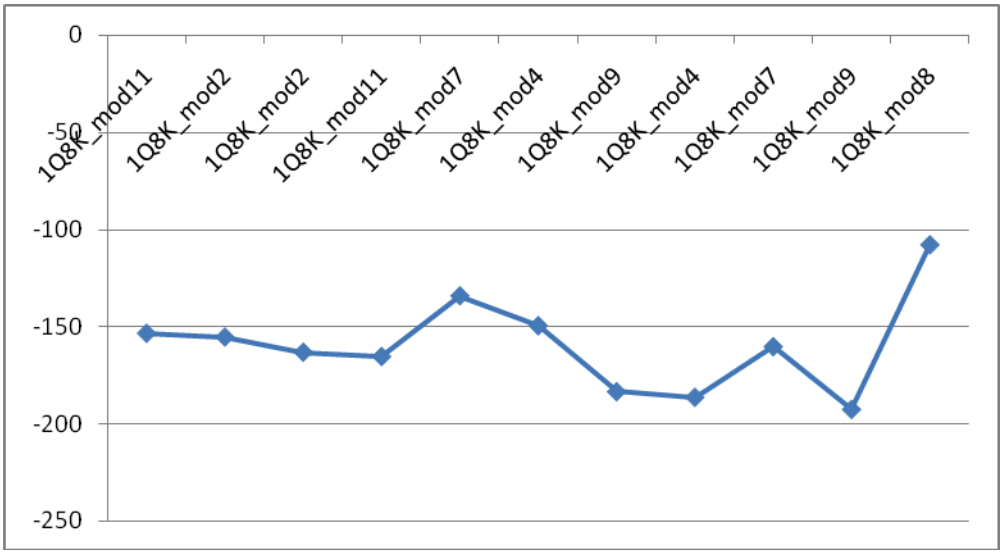


Fig. 7.4 Interaction energy between antisense siRNA strands complementary to mutated regions on BACE-1 mRNA and Argonaute proteins.



One antisense siRNA strand has been designed complimentary to mutated region of APBB3 mRNA and one designed for PSENEN mRNA. These two antisense siRNA have

shown interaction with 1Q8K protein. Table-2 shows interaction energy of antisense siRNA strands complimentary to mutated region of APBB3 mRNA and PSENEN mRNA.

Table-2 interaction energy of antisense siRNA strands complimentary to mutated region of APBB3 mRNA and PSENEN mRNA

Gene names encoding mutated mRNA	Protein conformer id	Codon position number in ORF of mRNA that is complementary to the designed antisense siRNA strand	Interaction energy(kcal/mol)
APBB3	1Q8K_mod4	318	-167.82
PSENEN	1Q8K_mod9	90	-187.89

Mature human microRNAs that are partially complementary to mutated regions of early onset Alzheimer's disease specific mutations in PSEN-1 mRNA are shown in Table-3. Complementarity is shown in table as dot-bracket notation. Dot symbol represents one nucleotide in the sequence. Each opening and closing bracket represents one base pair. MicroRNA and target sequence is separated by ampersand symbol (&).

Table-3 Details of mature human microRNAs that are partially complementary to mutated regions of early onset Alzheimer's disease specific mutations in PSEN-1 mRNA

Name of mature human microRNA	codon change on target PSEN-1 gene due to mutation	Complementarity of microRNA with target sequence that include mutation in PSEN-1 mRNA	Thermodynamic free energy of heterodimer formed. (Kcal/mol)	Argonaute protein name	Interaction energy of microRNA with Argonaute protein (Kcal/mol)

hsa-mir-214	ATT-AAT	Complementary microRNA: ACAGCAGGCACAGACAGGCAGU Target mRNA sequence with mutation position:- AAUGCUGCCAUCAUGAUCAGUGUCAAUGUUGU ((((((((((.((((((((((&...))))))....)).)).)))))..))))))	-21.24	1Q8K_mod4	-168.348
hsa-miR-371	TTT-ATT	Complementary microRNA: AAGUGCCGCCAUCUUUUGAGUGU Target mRNA sequence with mutation position:- AGUCAGCAUUAUACCCGGAAGGAUGGGCAGCU ....((.(((.(((((((.((((&.....)))).)))))))...))).)).	-17.80	No interaction	Nil
hsa-mir-450b-3p	GGT-GAT	Complementary microRNA:- UUGGGAUCAUUUUGCAUCCAUA Target mRNA sequence with mutation position:- UGGAAUUUUGAUGUGGUGGGAAUGAUUUCCAU .((((((((((((.(((((((.&.....))).)))))))))).)))))).	-17.91	No interaction	Nil
hsa-mir-936	CAC-TAC	Complementary microRNA: ACAGUAGAGGGAGGAAUCGCAG Target mRNA sequence with mutation position:- UUGAUGUGGUGGGAAUGAUUUCCAUUUACUGG .((((((((((((..((((((.&....))))))..... ..))))).)))))))).	-18.25	1Q8K_mod4	-144.331
hsa-mir-202	CTT-CAT	Complementary microRNA:- AGAGGUAUAGGGCAUGGGAA Target mRNA sequence with mutation position:- GUCAUCCAUGCCUGGCAUAUUAUAUCAUCUCU ((((((((((((((((((..&....))))))))).... ...)))))).))).	-20.37	No interaction	Nil

CONCLUSION

Antisense siRNA strands that show comparatively stronger interaction with argonaute protein may be able to form RNA-protein complex. This complex can recognize target mRNA sequence by considering base pair complementarity and can bind through hydrogen bonds. Antisense siRNA strands designed with comparatively low GC content, and their interaction with protein argonaute indicates feasibility of designed siRNA strands to be used in RNA interference pathway to silence AD specific mutations. Antisense siRNA strands have shown more complementarity with their target mRNA sequence than mature human microRNAs. These strands are more thermodynamically stable than human microRNAs when bound with their target mRNA sequence. This implies that siRNA strands that are designed are more efficient to be used in post transcriptional gene silencing than human microRNAs.

REFERENCE

[1] Hammond S.M., Caudy A.A., Hannon G.J. *"Post-transcriptional gene silencing by double-stranded RNA"*. Nat. Rev. Genet. 2:110-119(2001).

[2] Yoo-Hun Suh and Frederic Checler , "Amyloid Precursor Protein, Presenilins, and -Synuclein: Molecular Pathogenesis and Pharmacological Applications in Alzheimer's Disease", pharmacological reviews, vol.54 no.3 pages 469-525, (2002).

[3] M. Zaratiegui, D. Irvine, R. Martienssen, "Noncoding RNAs and Gene Silencing", *Cell*, Volume 128, Issue 4, Pages 763-776(2007)

[4] Mattick JS, Makunin IV. "Non-coding RNA", Hum Mol Genet., 15 Spec No 1, R17-29 (2006).

[5] Gisela Storz," An Expanding Universe of Noncoding RNAs", Science, vol. 296, no. 5571, pp. 1260-1263, (2002).

[6] Wu L, Belasco JG). "Let me count the ways: mechanisms of gene regulation by miRNAs and siRNAs". Mol. Cell vol. 29 (1): pages 1–7, (2008).

[7] Höck J, Meister G.," The Argonaute protein family", Genome Biol., 9(2):210, (2008).

[8] Gyorgy Hutvagner & Martin J. Simard," Argonaute proteins: key players in RNA silencing", Nature Reviews Molecular Cell Biology 9, doi: 10.1038/nrm2321, 22–32, (2008).

[9] Xiaohua hu, Drexel University, Philadelphia, Pennsylvania, Yi pan, Georgia State University, Atlanta, Georgia. "Knowledge discovery in bioinformatics", Wiley Interscience, pages 113-144(2007)

[10] Cerutti L., Mian N., Bateman A. *"Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain"*. Trends Biochem. Sci. 25:481-482(2000).

[11] Selenko P., Sprangers R., Stier G., Buehler D., Fischer U., Sattler M. *"SMN tudor domain structure and its interaction with the Sm proteins"* Nat. Struct. Biol. 8:27-31(2001).

[12] Hynes T.R., Fox R.O. *"The crystal structure of staphylococcal nuclease refined at 1.7 A resolution"* .Proteins 10:92-105(1991).

[13] DN cooper, EV ball, and M krawczak, "The human gene mutation database", Nucl. Acids Research (1998).

[14] David L. Wheeler, Tanya Barrett, Dennis A. Benson, Stephen H. Bryant, Kathi Canese, Vyacheslav Chetvernin, Deanna M. Church, Michael DiCuccio, Ron Edgar, Scott Federhen, Lewis Y. Geer, Yuri Kapustin, Oleg Khovayko, David Landsman, David J. Lipman, Thomas L. Madden, Donna R. Maglott, James Ostell, Vadim Miller, Kim D. Pruitt, Gregory D. Schuler, Edwin Sequeira, Steven T. Sherry, Karl Sirotkin, Alexandre Souvorov, Grigory Starchenko,

Roman L. Tatusov, Tatiana A. Tatusova, Lukas Wagner and Eugene Yaschenko, “Database resources of the National Center for Biotechnology Information” Nucleic Acids Research (2006).

[15] Suvi Asikainen, Liisa Heikkinen, Garry Wong, “SiRNA Target Finder”, Bioscience Technology, (2002).

[16] I. L. Hofacker, W. Fontana, P. F. Stadler, L. S. Bonhoeffer, M. Tacker, P. Schuster “Fast folding and comparison of RNA secondary structures” , Monatshefte für Chemie / Chemical Monthly, Vol. 125, No. 2. (1994).

[17] “HyperChem Release 7 for Windows”, Hypercube Inc. Publication (2002).

[18] Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N.Bhat, Helge Weissig, Ilya N. Shindyalov and Philip E. Bourne, “The Protein Data Bank”, Nucleic Acids Research, Vol. 28, No. 1 235-242, (2000).

[19] Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS, ”UniProt: the Universal Protein knowledgebase” Nucleic Acids Res. 1;32(Database issue):D115-9, (2004).

[20] Murzin AG, Brenner SE, Hubbard T, Chothia C, “SCOP: a structural classification of proteins database for the investigation of sequences and structures”, J Mol Biol, vol. 247(4):536-40.(1995).

[21] Geourjon C, Deléage G, “SOPMA: a self-optimized method for protein secondary structure prediction”, Protein Eng. vol. 7(2):157-64,(1994).

[22] Hulo N, Bairoch A, Bulliard V et al., “The PROSITE database-Nucleic Acids ” Res. vol. 34(Database issue), D227-30, (2006).

[23] Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel R.D, Bairoch A, “Expasy: The Proteomics Server For In-Depth Protein Knowledge And Analysis”, Nucleic Acids Res. vol. 31:3784-3788,(2003).

[24] Accelrys Software Inc., Discovery Studio Modeling Environment, Release 2.1 , San Diego: Accelrys Software Inc., (2007)

[25] Sam Griffiths-Jones, Russell J. Grocock, Stijn van Dongen, Alex Bateman and Anton J. Enright “miRBase: microRNA sequences, targets and gene nomenclature”, Nucleic Acids Research, vol.34 (Database issue): D140 - D144, (2005).

[26] Ventsislav Rusinov, Vesselin Baev, Ivan Nikiforov Minkov and Martin Tabler, “MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence”, Nucleic Acids Research, vol. 33 (Web Server Issue): W696-W700, (2005).