Creation of Point Mutations in NALP3 gene by site directed mutagenesis

Creation of Point Mutations in NALP3 gene by site directed mutagenesis

Research · November 2015

DOI: 10.13140/RG.2.1.3545.2882

1 author:

Umer Naveed Chaudhry
The University of Edinburgh
63 PUBLICATIONS **163** CITATIONS

Creation of Point Mutations in *NALP3* gene by site directed mutagenesis

Genetic engineering 7.5 hp

(2008.10.09 - 2008.10.24)

Report 2

Authors:
Shoaib Ashraf
a08ashsh@student.his.se
M.Sc Molecular biology

Umer Naveed Chaudhary
a08chaum@student.his.se
M.Sc Molecular biology

Farhan Basit
a08farba@student.his.se
M.Sc Molecular biology

Lab supervisors:
Magnus Fagerlind
magnus.fagerlind@his.se

Jonathan Holmén
jonathan.holmen@his.se

Rathan Joy Komban
rathan.joy.komban@his.se

School of Life Sciences
Skövde University
BOX 408
SE-541 28 Skövde
Sweden

Abstract

Mutagenesis is the replacement of nucleotides at the DNA level and one of the novel technologies used to produce it is by site directed mutagenesis. *NALP3* is a gene which encodes the protein cryopyrin. The protein is produced as an immune response, however if there is some mutation in the *NALP3* gene the outcomes are different autoinflammatory diseases. Till date more than forty different mutations have been identified for this gene but among them the most important and frequently occurring diseases are: chronic infantile neurological cutaneous and articular syndrome (CINCA)/NOMID neonatal onset multisystem inflammatory disease, Muckle-Wells Syndrome (MWS), and familial cold urticaria (FCU). All of these diseases have some common inflammatory signs as urticaria, fever and arthralgia. The *NALP3* gene contains 3100bp nucleotide sequence, 1550 amino acids and is located on chromosome position 1. In the current study we have created a point mutation in the *NALP3* gene by replacing arginine with leucine at position 260. This mutation has been observed causing the above discussed diseases. To get the desired mutation, site directed mutagenesis was performed by using mutated primers. The mutated strand was then further sequenced and analyzed. Sequencing results were very interesting and there were multiple mutations instead of a single point mutation. These strange results may be due to a number of reasons including faulty primers, improper plasmid purification, and other pipetting errors. However when the primers were aligned even though there were different dissimilarities between the mutant and the wild DNA, yet designed mutation was observed indicating the success of our project.

Introduction

Mutagenesis is the development and production of mutations by changing the DNA base sequence at a specific site in a genome. Mis-sense mutations correspond to those mutations in which the nucleotide altering the codon results in the production of a different amino acid (Reece, 2004) and site directed mutagenesis is the changing of the codon at DNA level (Glick and Pasternak, 2003). Different mutations in the NACHT-, Leucine rich repeat (LRR) and PYD-containing proteins (cryopyrin) could induce autoinflammatory syndrome-1 (*CIAS-1*)/cryopyrin cause autoinflammatory diseases as Muckle Wells syndrome (MWS), familial cold autoinflammatory syndrome (FCAS), chronic infantile neurological cutaneous and articular (CINCA) syndrome also called as neonatal, onset multisystem inflammatory (NOMID) disease. Currently there are more than forty mutations which are recognized in the *CIAS-1* gene (Ivona *et al*., 2007). The mutation R260L is a missense mutation due to which the CINCA syndrome occurs at exon-3 of *CIAS-1* gene, which is responsible for the production of the cryopyrin protein (Jeanette *et al*., 2008). The cryopyrin associated periodic system (CAPS) is a complex of autosomal dominant inherited diseases and their genes are expressed on the leukocytes moving in the peripheral circulation (Anna and Jos, 2007). Different mutations occurring in the *CIAS-1* gene (that encodes (NALP3)/PYPAF1 (pyrin-containing Apaf-1–like protein) 1, 2/Cryopyrin which is the member of the recently discovered NALP/PYPAF, subfamily of the CATERPILLER (caspase-recruitment domain (CARD), transcription enhancer, R (purine)-binding, pyrin, lots of LRR) protein family 3,4) results in different problems (Hoffman *et al*., 2001; Harton *et al*., 2002). The LRRs are found to be associated in recognition of the proteins derived from the pathogens like the heat shock proteins (hsp70) and thus provoke immune responses (Girardin *et al*., 2002). The *NLRP3* gene which encodes the pyrin like protein has a LRR motif, a nucleotide binding site domain (NBS), and a pyrin domain. The cryopyrin protein is a part of inflammasome which corresponds to the apoptosis-associated-speck-like protein called PYCARD/ASC, which has a domain for caspase recruitment. ASC which is linked to cryopyrin activates the nuclear factor $_{k}$B (NF-$_{k}$B) (Martinon *et al*., 2006).

Whenever there is some problem in this gene the results may be as following: chronic infantile neurological cutaneous and articular syndrome (CINCA)/NOMID, Muckle-Wells Syndrome (MWS), and familial cold urticaria (FCU). As a result due to some problem in this gene (*NALP3*) it starts producing inflammasomes which then activates the caspase-1 and in turn elevates the secretion of Interleukin-1 β and ultimately the process of inflammation is triggered (Yu *et al*., 2006; Ebun *et al.*, 2002). Pyrin domain (PYD) enhances interactions between the proteins containing them which ultimately form a complex for mediating signals. The LLRs, which is member of the CATERPILLAR family and subfamily NOD (nucleotide binding oligomerization domain), responds to intracellular stimulation and can initiate an immune response. The three most important auto-inflammatory diseases caused by mutations in the *CIAS-1* gene, FCU, MWS, and CINCA/NOMID have some same inflammatory responses as

fever, urticaria and arthralgia. Those patients having CINCA syndrome have the most severe conditions starting from the neonatal life as polymorphonuclear meningitis (PMN) which results in arthropathies and neuropathies which leads to deafness, and visual problems. In MWS deafness of the neurosensory organs, amyloidosis, and arthritis have been observed. FCU is the mildest among all the diseases and it is aggravated in when the individuals are exposed to cold (Benedicte *et al*., 2004).

Three mutations were found in exon-3 like R260L, R260W, and R260P. The codon 260 (CGA) is a hotspot of these mutations. Originally, this codon encodes for arginine but after mutation it encodes for leucine, tryptophan and proline, respectively. The substitute of arginine with these amino acids alters the structure and function of protein which disturbs its normal inflammatory function and cause disease. The purpose of this study was to create PCR based site directed mutagenesis (R260L) by replacing arginine with leucine in the *NALP3* gene by switching the nucleotide codon CGA to CTA which was further analyzed by PCR and sequencing.

Materials and Methods

Mutagenic primer design

To design the template specific mutagenic primer, a web based primer deigning software; QuikChange® Primer Design Program (Stratagene, 2008) was used. The mutagenic oligonucleotide primer was designed according to the desired mutation (Appendix I). The specifications and working of the primer were as follows:

The desired mutation was present in both of the mutagenic primers and it annealed to the opposite plasmid strand on the same sequence. The length of the primer was 35 nucleotide bases long, with a melting temperature (Tm) of 78.10°C. 10-15 bases of correct nucleotide sequences were included on both sides of the desired mutation so that the mutation was in the middle of the primer. The primer contained at least 40% GC content and terminated in one or more C or G bases. The forward primer used was 5'-gttctatatccactgtctagaggtgagccttgtga-3'; where as the reverse primer used was 5'-tcacaaggctcacctctagacagtggatatagaac-3'.

Mutant strand synthesis

Forward and reverse primers each of 35 nucleotide length were used for site directed mutagenesis by replacing the amino acid arginine with leucine on the position number 260 of the *NALP3* gene by switching the nucleotide codon CGA to CTA. The components of the reaction mix are mixed as described in the

QuickChange®II Site-Directed Mutagenesis Kit (Stratagene, 2008). The thermocycler was run according to the protocol by Stratagene, with 16 cycles and 55 ºC annealing temperature. The amount of template used was 25ng and 125ng of both forward and reverse primers were utilized.

Dpn I Digestion of mutant plasmids

The mutant non-methylated DNA was obtained by digestion of the methylated wild type DNA with endonuclease *Dpn-1* by using the QuickChange®II Site-Directed Mutagenesis Kit (Stratagene, 2008).

Transformation of XL1-Blue Supercompetent Cells

The *DpnI*-treated plasmid DNA was then transformed into *Escherichia coli* (XL1-Blue) supercompetent cells, plated on two ampicillin containing LB agar plates at 37^{o}C for 16 hours. Two colonies, one from each plate, were transferred to a falcon tube having 5 ml LB broth containing ampicillin which was then incubated overnight at 37^{o}C. All the procedures were according to the QuickChange®II Site-Directed Mutagenesis Kit (Stratagene, 2008).

Mutant plasmid DNA purification

For harvesting and purification of mutant plasmid DNA from the overnight culture of bacteria the protocol given in the Qiagen Spin Miniprep Kit was adopted. This kit is based on the alkali preparation of plasmid DNA. Ten milliliter of culture was collected and from it plasmid was purified. Purified plasmid was stored at -20^{o}C.

DNA concentration of purified Mutant Plasmid DNA

The DNA concentration was measured using the Nanodrop, ND-1000 (Thermo technologies, USA)

Sequencing PCR

After dilution of the plasmid, sequencing PCR was performed by using a left sequencing primer 5`-ACACACGACTGCGTCTCATC-3` which started and ended at 503rd nucleotide and 524th nucleotide position, respectively. As a template 270ng of the mutated purified plasmid was used, the annealing temperature was 50˚C, and the PCR was subjected to 40 cycles as described in BigDye terminator v 1.1cycle sequencing kit (Applied Biosystems, 2008).

Purification of the Amplicon

The amplicon was purified using DyeEx 2.0 Spin Kit (Qiagen, 2008) to remove the unincorporated dye terminator from sequencing reaction. DyeEx 2.0 Spin Kit contains prehydrated gel filtration resin and can be readily used.

Sequencing

The purified mutated plasmid DNA product was then heat dried for 11-14 minutes at 90^{o}C which was further resuspended in 15µl of TSR (template suppression reagent) according to the BigDye terminator v1.1cycle sequencing kit and then sequenced by using ABI 310 Genetic Analyzer (GMI, USA).

Bioinformatics analysis

The nucleotide sequence obtained from ABI 310 Genetic Analyser (GMI, USA) was then aligned with the wild type *NALP3* gene using the software Lalign (Expasy, 2008).

Results and discussions

The mutated primers replaced the nucleotide G with T on the position number 260 in the cloned vector containing the desired gene of *NALP3*, as a result the amino acid arginine was replaced with leucine. Following successive digestion the mutant (non-methylated) DNA was revealed and uniform growth of the ampicillin resistant plasmid showed the plasmid was transformed into the *E. coli* competent cells (XL 1 blue). Excellent turbidity was observed after culturing of the two colonies in the LB broth which confirmed dense bacterial population. Nanodrop illustrated a concentration of 365ng/µl of the plasmid and the ratio of 260/280 was found to be 1.83 which confirmed the purity of plasmid DNA. The sequencing results showed numerous unidentified nucleotides (N,s) in the mutated *NALP3* gene (Appendix II), after removing the N,s the following sequence was revealed as shown in Figure 1.

```
TAAGGACACCGCACCGCAGTATAGGAGCAGGAGCTTCTGGCCTCGGCAAGACCAAGACGTGTGAGA
GCCCCGTGAGTCCCATTAAGATGGAGTTGCTGTTTGACCCCGATGATGAGCATTCTGAGCCTGTGC
ACACCGTGGTGTTCCAGGGGGCGGCAGGGATTGGGAAAACAATCCTGGCCAGGAAGATGATGTTGG
ACTGGGCGTCGGGGGACACTCTACCAAGACAGGTTTTGACTATCTGTTTCTATATTCCACTGGTCT
AAGAGGGTGAGCCCTTTGGTGACACAAGAAGGAAGCCCTGGGGGGACCCTGGATTCAATTGAAGCC
TTGCTTGGCCCCCGACCCCAAAACCCCAACCCATTCCCACAAAGATTCGTTGGAGAAAAACCCCTT
CCAAAATTCCTTCTTTCTTCATGGGACCGGGCTTTCCAATGAAGCTGTAAGGTTGCCCTTTTGACA
AGCACATTAGGGACCCGCTTTTGCACTTGACTTGGGAAAAAGGCCAAGCGGG
```

Figure 1. This shows the sequence results after replacing the unidentified nucleotides (N,s) manually by using the web tool; Chromas lite, Softpedia, 2008 .

Specific base calls were made and the N,s were then substituted manually by their respective nucleotides from the mutant DNA by using the chromatogram. Many homonucleotide runs such as repetitive sequences were observed which act as hotspots for mutations (Gordenin *et al*., 1998). As a whole the chromatogram furnished a justified sequencing trace with crisp clean bands which are separated without any ambiguity resulting in a proper base call with little baseline noise, however in the first hundred nucleotides low peaks were identified, and may be observed as a result of the DNA polymerase slippage effect of *Pfu* due to secondary structures (Salganik *et al*., 1990). An abnormal peak height of polyA was observed at the 2nd nucleotide which due to over condensation was not able to separate as a result of some interfering nucleotide, may be due to any sort of contamination due to improper purification of the sample. The polyT initially is not being amplified and is shown at the basal line may be due to contamination. Many broken peaks are also observed indicating a non-specific primer which may be binding to more than one DNA templates. The broken peak distortions kept on increasing which ultimately swamped the whole chromatogram after the 140th nucleotide. Merger of many different peaks were also seen due to fact of presence of more than one template DNA due to contamination. Finally the mutant DNA sequence was aligned against the wild type *NALP3* gene by using Lalign (Appendix III). The web based bioinformatics tool revealed multiple mutations along with a number of dissimilarities in both wild type and mutant strands, yet the desired mutation CGA was replaced with CTA at the nucleotide number 780 (Appendix III). The multiple point mutations found may be due to gene conversion between different homologous sequences (Igor *et al*., 2003). Gel electrophoresis should be performed to reveal the size of the plasmid DNA, and if the desired plasmid DNA size is obtained only then sequencing should be performed, however these types of abrasions can be avoided by counter checking each step which might be one of our major mistakes. The current project would be further utilized to check the expression and structure of proteins produced from this mutant *NALP3* gene.

Reference

Anna S. and Jos W.M. (2007) Pathogenesis of familial periodic fever syndromes or hereditary auto inflammatory syndromes. *Am J Physiol Regul Integr Comp Physiol, 292,* (p. 86–98).

Applied Biosystems. (2008) BigDye® Terminator v1.1 Cycle Sequencing Kit.

Benedicte N., Isabelle C., Anne-M., Jerome F., Christine B. and Loredana L. (2004) Molecular basis of the spectral expression of *CIAS1* mutations associated with phagocytic cell-mediated auto inflammatory disorders CINCA/NOMID, MWS and FCU. *Blood, 103,* (p. 7).

Ebun A., Fabio M., Philip N., Hawkins., John B., Ross., Daniel C., Swan., David R., Booth. and Helen J. L (2002) Association of Mutations in the *NALP3/CIAS1/PYPAF1*Gene With a Broad Phenotype Including Recurrent Fever, Cold Sensitivity, Sensorineural Deafness, and AA myloidosis. *Arthritis & rheumatism, 46,* (p. 2445–2452).

Expasy. (2008) Available at: http://www.ch.embnet.org/ software/ LALIGN_form. html. [Collected 2008.10.25]

Martinon F., Pétrilli V., Mayor A., Tardivel A. and Tschopp J. (2006) Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature 440*, (p. 237-241).

Girardin S.E., Sansonetti P.J. and Philpott D.J. (2002) Intracellular vs extracellular recognition of pathogens Common concepts in mammals and flies. *Trends Microbiol. 10,* (p. 193-199).

Glick B.R. and Pasternak J.J. (2003). Molecular Biotechnology

Gordenin D.A. and Resnick M.A. (1998) Yeast ARMs (DNA at-risk motifs) can reveal sources of genome instability. *Mutation Research, 400*, (p. 45–58)

Harton J.A., Linhoff M.W., Zhang J. and Ting J.P. (2002). Cutting edge: CATERPILLER: a large family of mammalian genes containing CARD, pyrin, nucleotidebinding, and leucine-rich repeat domains. *J Immunol, 169,* (p. 4088-4093).

Hoffman H.M., Mueller J.L., Broide D.H., Wanderer A.A. and Kolodner R.D. (2001) Mutation of a new gene encoding a putative pyrin-like protein causes familial cold autoinflammatory syndrome and Muckle- Wells syndrome. *Nat Genet, 29,* (p. 301-305).

Igor B.R. and Youri I. (2003) Theoretical analysis of mutation hotspots and their DNA sequence context specificity. *Mutation Research, 544*, (p. 65–85).

Ivona A., Christopher D.P., Elaine F.R., James L.M., Julie Le. and Richard DK. (2007) Novel *CIAS1* Mutations in North American Patients and a New Cryopyrin Model. *Arthritis & Rheumatism, 56,* (p. 1273–1285).

Jeanette M.W., Tanja P.O. and Koichi S.K. (2008) NLR proteins: integral members of innate immunity and mediators of inflammatory diseases. *Journal of Leukocyte Biology, 83,* (p. 13-30.)

Yu J.W., J Wu., Zhang Z., Datta P., Ibrahimi I., Taniguchi S., Sagara J., Fernandes A. and Alnemri E. (2005) Cryopyrin and pyrin activate caspase-1, but not NF-κB, via ASC oligomerization. *Cell Death and Differentiation, 1038,* (p. 236–249).

QIAGEN. (2002) DyeEx® 2.0 Spin Kit. Available at: http://www1.qiagen.com/literature/handbooks/literature.aspx?id=1000066. [collected 2008.10.02]
QIAGEN. (2006) QIaprep® miniprep handbook. Available at: http:// www1.qiagen.com / literature/ handbooks/literature.aspx?id=1000248. [collected 2008.10.02]

Reece R.J. (2004) Analysis of Genes and Genomes. ISBN: 0-470-84379-9 New York,

Salganik R.I., Dianov G. and Medvedev A. (1990) Cluster of point mutations redetermined by a quasipalindromic nucleotide sequence in plasmid pBR322 DNA. *FEBS Lett, 261, (p.* 28–30).

Softpedia. (2008) Available at:http://www.softpedia.com/get/Science-CAD/Chromas-Lite.shtml. [collected 2008.10.02].

Strategene. (2008) Available at:http://www.stratagene.com/tradeshows/feature.aspx?fpId=118. [collected 2008.09.12].

Stratagene. (2008) QuikChange® II site-directed mutagenesis kit. Available at: http://www.stratagene.com/manuals/200523.pdf. [collected 2008.10.02].

Appendices

Appendix I

Wild type primer: 5'-gttctatatccactgtcgagaggtgagccttgtga-3'

Mutagenic primer design

Prime	Primer Sequence (5' to 3')
9t	5'-gttctatatccactgtctagaggtgagccttgtga-3'
9t_antisense	5'-tcacaaggctcacctctagacagtggatatagaac-3'

Oligonucleotide information:

Primer Name	Length (nt.)	Tm	Duplex Energy at 68°C	Energy Cost of Mismatches
g779t	35	78.10°C	-41.85 kcal/mole	7.4%
g779t_antisense	35	78.10°C	-46.65 kcal/mole	6.0%

Primer-template duplexes:

Primer Name	Primer-Template Duplex
g779t	5'-gttctatatccactgtctagaggtgagccttgtga-3' |||||||||||||||| |||||||||||||||| agacaagatataggtgacagctctccactcggaacactgtg
g779t_antisense	tctgttctatatccactgtcgagaggtgagccttgtgacac |||||||||||||||| |||||||||||||||| 3'-caagatataggtgacagatctccactcggaacact-5'

Appendix II

DNA sequencing result

TAAGNACACCGCACCGCAGNANAGGAGCAGGAGCTTCTGGCCTCGGCAAGACCAAGACGTGTGAGA
GCCCCGTGAGTCCCATTAAGATGGAGTTGCTGTTTGACCCCGATGATGAGCATTCTGAGCCTGTGC
ACACCGTGGTGTTCCAGGGGGCGGCAGGGATTGGGAAAACAATCCTGGCCAGGAAGATGATGTTGG
ACTGGGCGTCGGGGGACACTCTACCAAGACAGGTTTTGACTATCTGTTTCTATATTCCACTGGTCT
AAGAGGGTGAGCCCTTTGGTGACACANGAAGGANGCCNTGGGGGGACCCTGGATTCNATNGAAGCC
TTGCTTGGCCCCCGACCCCAAAACCCCANCCCNNTCCCACAAAGATTCGNTGGAGAAAAACCCCTT
CCAAAATTCCTNCTTTCTNCATGGGACCGGGCTTTCCAATGAAGCTGNNAGGNTGCCCTTTTGACN
AGCNCATTAGGGACCNGCTTTTGCACTNNACTTGGGAAAANGGCCAAGCGGG

DNA sequencing result by removing N,s

TAAGGACACCGCACCGCAGTATAGGAGCAGGAGCTTCTGGCCTCGGCAAGACCAAGACGTGTGAGA
GCCCCGTGAGTCCCATTAAGATGGAGTTGCTGTTTGACCCCGATGATGAGCATTCTGAGCCTGTGC
ACACCGTGGTGTTCCAGGGGGCGGCAGGGATTGGGAAAACAATCCTGGCCAGGAAGATGATGTTGG
ACTGGGCGTCGGGGGACACTCTACCAAGACAGGTTTTGACTATCTGTTTCTATATTCCACTGGTCT
AAGAGGGTGAGCCCTTTGGTGACACAAGAAGGAAGCCCTGGGGGGACCCTGGATTCAATTGAAGCC
TTGCTTGGCCCCCGACCCCAAAACCCCAACCCATTCCCACAAAGATTCGTTGGAGAAAAACCCCTT
CCAAAATTCCTTCTTTCTTCATGGGACCGGGCTTTCCAATGAAGCTGTAAGGTTGCCCTTTTGACA
AGCACATTAGGGACCCGCTTTTGCACTTGACTTGGGAAAAAGGCCAAGCGGG

Appendix III

```
Comparison of:
(A) ./wwwtmp/.25917.1.seq Wild type 3105 bp                                  - 3105
nt
(B) ./wwwtmp/.25917.2.seq Mutant type 514 bp                                 - 514
nt
 using matrix file: DNA (5/-4), gap-open/ext: -14/-4 E(limit)   0.05

  82.8% identity in 518 nt overlap (517-976:2-514); score: 1257 E(10000): 3.9e-96

        520       530       540       550       560       570
Wild   AAGGAGCACCGGAGCCAGCAGGAGAGGGAGCAGGAGCTTCTGGCCATCGGCAAGACCAAG
       ::::: ::::: : :  :::: : ::: ::::::::::::::::: ::::::::::::::
```

```
Mutant AAGGA-CACCGCACC--GCAGTATAGG-AGCAGGAGCTTCTGGCC-TCGGCAAGACCAAG
              10          20         30        40         50

        580       590       600       610       620       630
Wild   ACGTGTGAGAGCCCCGTGAGTCCCATTAAGATGGAGTTGCTGTTTGACCCCGATGATGAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Mutant ACGTGTGAGAGCCCCGTGAGTCCCATTAAGATGGAGTTGCTGTTTGACCCCGATGATGAG
         60        70        80        90       100       110

        640       650       660       670       680       690
Wild   CATTCTGAGCCTGTGCACACCGTGGTGTTCCAGGGGGCGGCAGGGATTGGGAAAACAATC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Mutant CATTCTGAGCCTGTGCACACCGTGGTGTTCCAGGGGGCGGCAGGGATTGGGAAAACAATC
        120       130       140       150       160       170

        700       710       720       730        740       750
Wild   CTGGCCAGGAAGATGATGTTGGACTGGGCGTCGGGG-ACACTCTACCAAGACAGGTTT-G
       :::::::::::::::::::::::::::::::::: :::::::::::::::::::: :
Mutant CTGGCCAGGAAGATGATGTTGGACTGGGCGTCGGGGGACACTCTACCAAGACAGGTTTTG
        180       190       200       210       220       230

          760        770         780         790          800
Wild   ACTATCTGTT-CTATAT-CCACTG-TCGA-GAGG-TGAGCCTT---GTGACACAG--AGG
       :::::::::: :::::: :::::: :: : :::: :::::: :   ::::::::   :::
Mutant ACTATCTGTTTCTATATTCCACTGGTCTAAGAGGGTGAGCCCTTTGGTGACACAAGAAGG
        240       250       260       270       280       290

            810         820                 830       840
Wild   AG--CCTGGGGGA--CCTGA--TCA--TGA------GCTGCTGCCCCGACCC--AAACCC
       :   ::::::::   ::::   :::  :::      :::    ::::::::::  ::::::
Mutant AAGCCCTGGGGGGACCCTGGATTCAATTGAAGCCTTGCTTGGCCCCCGACCCCAAAACCC
        300       310       320       330       340       350

        850         860           870          880         890
Wild   --ACCCAT--CCACAA-GAT-CGT--GAGAAAACCC---TCCAGAATCC--TCTTCCT-C
         ::::::  :::::: ::: :::  ::::::: ::   :::: ::: :  :::: :: :
Mutant CAACCCATTCCCACAAAGATTCGTTGGAGAAAAACCCCTTCCAAAATTCCTTCTTTCTTC
        360       370       380       390       400       410

           900           910        920          930         940
Wild   ATGG-AC--GGCTT--CGATGA-GCTGCAAGGT-GCC--TTTGACGAGCACATA--GGAC
       :::: ::  :::::  : :::: :::: ::::: :::  :::::: :::::::   ::::
Mutant ATGGGACCGGGCTTTCCAATGAAGCTGTAAGGTTGCCCTTTTGACAAGCACATTAGGGAC
        420       430       440       450       460       470

             950          960       970
Wild   C-GCTCT-GCACT-GACT--GGCAGAAGGCCGAGCGGG
       : ::: : ::::: ::::  :: : :::::: ::::::
Mutant CCGCTTTTGCACTTGACTTGGGAAAAAGGCCAAGCGGG
        480       490       500       510
```