Journal of Genetic and Environmental Resources Conservation, 2014, 2(1): 78-82.
www.jgerc.com

Structure Information for protein CD19 alignment of leukemia patient in Iraq

Najwa Sh. Ahmed[1]*, Saad M. Nada[2] and Zainib H.A. Al-Rikabi[3]

[1,2]Biotechnology Research Center, Al-Nahrain University, Baghdad and [3]Genetic Engineering and Biotechnology Institute, Baghdad University, Iraq.
*Corresponding author: najwaun@yahoo.com

Abstract

The aim of this work was study the conformation of protein *CD19* in leukemia patients in Iraq (Baghdad city). To achieve this goal, blood samples were collected from 50 leukemia (25 ALL and 25 CLL) and 50 samples of healthy, DNA was isolated and the *CD19 gene* (A 863 bp fragment) were amplified by using specific primers for exon2 of this gene, and nucleotide sequences of *CD19* gene by Macro gene company, USA, then analyzed using BLAST program which is available at the National Center Biotechnology Information (NCBI) online and Raptorx software. The DNA sequencing results of flank sense of *CD19* gene about 99% compatibility were found for that gene from 50 (ALL and CLL) cases patients with wild type of gene**.** The difference may be attributed to insertion 1900 C nucleotide in position +48 of exon 2 of *CD19* gene resulted in the replacement of a serine (TCC) residue into isoluseine (ATC), and deletion 1904 C in position + 49 of exon 2, this mutation which result change of codon from GCA to GCC but no changes translate to amino acid (Alanine to Alanine). The second mutation (Deletion nucleotide) amended influence the first mutation (Insertion nucleotide) and did not lead to a change all amino acid (Framshift). Raptorx software (http:reptox.uchicago.eDu/predict) for drawing structure protein (helixes, B-sheets and coils) of all sample, appear the result conformation of protein CD19 for healthy and patient, the lack of similarity of the healthy with patient conformation of protein compared with 3D structure by aligning it to structurally similar protein structures in PDB. In conclusion, our case study suggests that the insertion C and deletion C in exon 2 of the *CD19* gene lead to difference in the order of the amino acids and the emergence of site side of acids and the difference in 3D structure protein leads to variation in the function of CD 19.

Keywords: Raptor X, Framshift mutation, Acute lymphocyte leukemia, Chronic lymphocyte leukemia, Protein database.

Introduction

The human CD19 antigen is a 95kd transmembrane glycoprotein belonging to the immunoglobulin superfamily (Carter and Barrington, 20042; Thierry-Mieg, Thierry-Mieg, 2006). The gene contains 15 exons and codes for the CD19 molecule with 556 amino acids. Structurally, the gene contains an unusually short 5'-untranslated region. The proximal CD19 promoter lacks a TATA box, and its major start sites are found within 50 bp of the initiation codon (Kehrl *et al.*, 1994). CD19 is classified as a type I transmembrane protein, with a single transmembrane domain, a cytoplasmic C-terminus, and extracellular N-terminus. During B-cell differentiation in the bone marrow, the surface molecule CD19 appears early and remains on the B cell until it differentiates into a plasma cell. Four proteins on the surface of mature B cells CD19, CD21, CD81, and CD225 form the CD19 complex, which signals in conjunction with the B-cell antigen receptor, thereby decreasing the threshold for receptor-dependent signaling (Carter and Fearon, 1992; Ishiura *et al.*, 2010). The types of leukemia also can be grouped based on the type of white blood cell that is affected, Leukemia that affects lymphoid cells is called lymphoid, lymphocytic, or lymphoblastic leukemia. Leukemia that affects myeloid cells is called myeloid, myelogenous, or myeloblastic leukemia. CD19 is a biomarker for normal and neoplastic B cells, as well as follicular dendritic cells. Though it is not known if CD19 contributes directly to B cell carcinogenesis, its expression is highly conserved on most B cell tumors (Vardiman *et al.*, 2009). It is expressed in most acute lymphoblastic leukemias (ALL), chronic lymphocytic leukemias (CLL) and B cell lymphomas. In fact, the majority of B cell malignancies express *CD19* at normal to high levels (80% of ALL, 88% of B cell lymphomas and 100% of B cell leukemias).

Other B cell malignancies, in contrast, show diminished CD19 levels (Cooper *et al.*, 2004; Yang *et al.*, 2005). Although CD19 expression is observed in normal plasma cells, malignant plasma (myeloma) cells isolated from multiple myeloma patients have been shown to lack CD19 expression, while isolates from premyeloma patients show a mix of both CD19- and CD19 + plasma cells (Ginaldi *et al.*, 1998). Homozygous frame shift mutations of the cd19 gene have been documented to result in truncation of the three key cytoplasmic tyrosine residues. Patients with this type of *CD19* gene mutations showed normal numbers of precursor and total B cells, but decreased numbers of CD27 + memory B cells and CD5+ B cells. Their B cells exhibit normal levels of CD81 and CD225, but decreased CD21 and very low to undetectable *CD19* (Poe *et al.*, 2012; van Zelm *et al.*, 2006). The aim of this work was study the conformation of protein *CD19* in leukemia patients of Iraq and relationship with function of CD19 protein.

Materials and Methods

Samples and DNA extraction: Whole blood samples were obtained from 50 Iraqi patients affected by Leukemia (25 ALL and 25 CLL, age ranged from 4 to 25 years) and also obtained from 50 healthy used as a control group, who were admitted to the National Center of Haematological Diseases (Al-Mustanisyria University). The disease was clinically diagnosed by the consultant medical staff at the centre. In total, 4 ml whole blood was collected into an EDTA- tube, DNA was extracted from the samples by DNA extraction kit (Wizard® Genomic DNA Purification Kit, Promega, Madison, WI, USA) according to the manufacturer's protocol.

Detection of Gene *CD19* by Using PCR: Detection of *CD19* gene was conducted by using primers for amplification of for *exon 2 CD19* gene. A fragment 865 bp of *CD19* was amplified using a forward primer (CD19F: 5'-GGAAGGTATGTCCAAAGGGCA-3') and a reverse primer (CD19R:5'-GCCACCTAGGTCCGAAACAT-3') (Primers set supplied by alpha DNA Company, Canada). The PCR amplification was performed in a total volume of 25µl containing 1.5µl DNA, 12.5 µl Go Taq green master mix 2X (Promega corporation, USA), 1µl of each primer (10 pmol) then the volume was completed with 25µl of nucleases free water. The thermal cycling conditions were done as follows: Denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 35s, 62°C for 1 min and 72 °C for 1min with final incubation at 72 °C for 7 min using a thermal Cycler (Gene Amp, PCR system 9700; Applied Biosystem). The PCR products were separated by 1.5% agarose gel electrophoresis and visualized by exposure to ultraviolet light (302nm) after ethidium bromide staining.

Sequencing and Sequence Alignment: Sequencing of exon 2 of *CD19* gene was performed by Macro gen company, USA. Homology search was conducted using Basic Local Alignment Search Tool (BLAST) program which is available at the National Center Biotechnology Information (NCBI) online at (http:// www.ncbi.nlm.nih.gov) and BioEdit program. The results were compared with data obtained from Gene Bank published ExPASY program which is available at the NCBI online.

Statistical analysis:1- The statistical analysis is a very important final step in the research to analyses and evaluate the obtained results. Medical statistics of this study was conducted via computer based statistical program which was: $\mathbf{X^2}$ for Windows computer package. The statistical analysis tests which used in this were as follows: P value <0.001 is considered a significant correlation. 2- MEGA5*:* The number of base substitutions per site from between sequences is shown. Standard error estimate (s) are shown above the diagonal. Analyses were conducted using the Kimura 2-parameter model (Kimura, 1980). The analysis involved 4 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 466 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura *et al.*, 2011). 3- Raptorx software: (http: reptox.uchicago.eDu/ predict) for drawing structure protein (helixes, B-Sheets and coils) of amino acid. For a specific target RaptorX first aligns it to each of the templates using the single-template alignment algorithm. Then RaptorX predicts the alignment quality and ranks all the templates by predicted quality descendingly. If the target is not suitable for multiple-template threading, RaptorX builds a 3D model for the target from the pairwise alignment with the highest predicted quality. Otherwise, RaptorX runs multiple-template threading for the target and builds a corresponding 3D model.

Results and Discussion

CD19 gene was successfully amplified using specific PCR primers for exon 2. Figure (1) showed PCR amplification of exon 2 of the *CD19* where a specific product at 862 bp was observed. Sequencing of this gene was performed to detect variant insertion C at position +48 and deletion C at position +49 which related to development of disease. The *CD19* gene from 50 leukemia patients (ALL and CLL) shows 99% compatibility with the wild type sequences of *CD19* gene from Gene Bank as shown in figure (2), there are one insertion C at position + 1900, that cause a serine (TCC) to Isoluseine (ATC) substitution in codon 48, there is a highly significant between Leukemia and incidence of

insertion of + 48 C position in exon 2 of *CD19* gene (X^2 = 100, *P*>0.001), Figure (3) shows the translation of *CD19* gene of all groups (healthy and patient) to a protein sequence, deletion C at position 1904 of exon 2, this mutation which result change of codon from (GCA) to (GCC) but no changes translate to amino acid (Alanine to Alanine) in codon 49. The second mutation (Deletion nucleotide) amended influence the first mutation (Insertion nucleotide) and did not lead to a change all amino acid (Framshift).

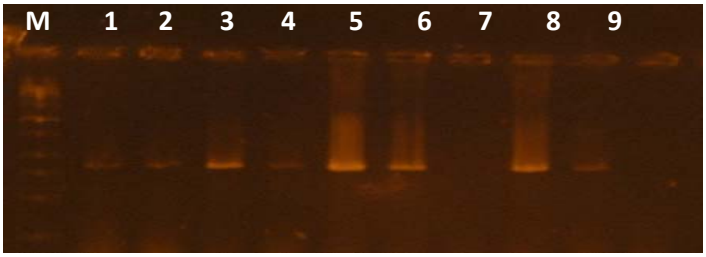

Figure (1): Agarose gel electrophoresis for amplified *CD19* gene of lymphocyte belonging to healthy, ALL, and CLL patients was done. Bands were fractionated by electrophoresis on a 1.5 % agarose gel (2 h., 5V/cm, 1X Tris-acetic buffer) and visualized under U.V. light after staining with ethidium bromide staining. Lane M: 1000bp ladder, Lane: 1, 2, 3 (Healthy), Lane: 4, 5, 6 (ALL), Lane: 7, 8, 9 (CLL).

Frame	Strand	Gaps	Identities	Expect	Score
	Plus/Plus	2/790(0%)	788/790(99%)	0.0()	1447 bits(783)

```
Query  181   CCCATGCCCACACctctctccctctctctcCACAGAGGGAGATAACGCTGTGCTGCAGTG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1717  CCCATGCCCACACCTCTCTCCCTCTCTCTCCACAGAGGGAGATAACGCTGTGCTGCAGTG  1776

Query  241   CCTCAAGGGGACCTCAGATGGCCCCACTCAGCAGCTGACCTGGTCTCGGGAGTCCCCGCT  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1777  CCTCAAGGGGACCTCAGATGGCCCCACTCAGCAGCTGACCTGGTCTCGGGAGTCCCCGCT  1836

Query  301   TAAACCCTTCTTAAAACTCAGCCTGGGGCTGCCAGGCCTGGGAATCCACATGAGGCCCCT  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1837  TAAACCCTTCTTAAAACTCAGCCTGGGGCTGCCAGGCCTGGGAATCCACATGAGGCCCCT  1896

Query  361   GGCCATC-TGGCTTTTCATCTTCAACGTCTCTCAACAGATGGGGGGCTTCTACCTGTGCC  419
             ||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1897  GGC-ATCCTGGCTTTTCATCTTCAACGTCTCTCAACAGATGGGGGGCTTCTACCTGTGCC  1955

Query  420   AGCCGGGGCCCCCCTCTGAGAAGGCCTGGCAGCCTGGCTGGACAGTCAATGTGGAGGGCA  479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1956  AGCCGGGGCCCCCCTCTGAGAAGGCCTGGCAGCCTGGCTGGACAGTCAATGTGGAGGGCA  2015

Query  480   GCGGTGAGGGCCGGGCTGGGGCAGGGGCAGGAGGAGAGAAGGGAGGCCACCATGGACAGA  539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2016  GCGGTGAGGGCCGGGCTGGGGCAGGGGCAGGAGGAGAGAAGGGAGGCCACCATGGACAGA  2075

Query  540   AGAGGTCCGCGGCCACAATGGAGCTGGAGAGAGGGGCTGGAGGGATTGAGGGCGAAACTC  599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2076  AGAGGTCCGCGGCCACAATGGAGCTGGAGAGAGGGGCTGGAGGGATTGAGGGCGAAACTC  2135

Query  600   GGAGCTAGGTGGGCAGACTCCTGGGGCTTCGTGGCTTCAGTATGAGCTGCTTCCTGTCCc  659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2136  GGAGCTAGGTGGGCAGACTCCTGGGGCTTCGTGGCTTCAGTATGAGCTGCTTCCTGTCCC  2195
```

Fig. (2): Sequencing of sense flanking the *C19* gene for leukemia (ALL and CLL) as compared with standard *CD19* obtained from Gene Bank, (A: 25 ALL; B: 25 CLL patients).

The results were compared with data obtained from Gene Bank published ExPASY program which is available at the NCBI online at www.ncbi.nlm.nih.gov. Translation of exon 2 *CD19* gen of all groups (healthy and patient) to a protein sequence, as shown in figure (3), than used Raptorx software (http:reptox.uchicago.eDu/predict) for drawing structure protein (helixes, B-Sheets and coils) of all samples, appear the result conformation of protein CD19 for healthy and patient, the lack of

similarity of the healthy with patient conformation of protein as in Figure (4), Shows the difference in the order of the amino acids and the emergence of site side of acids as in shape and the difference in body protein leads to variation in the function of CD 19.

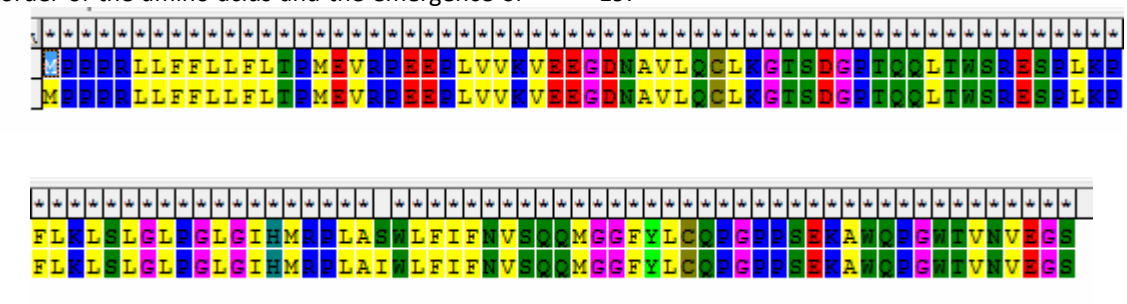

Figure (3): Translate of nucleotides to amino acid and different between healthy and patient group.

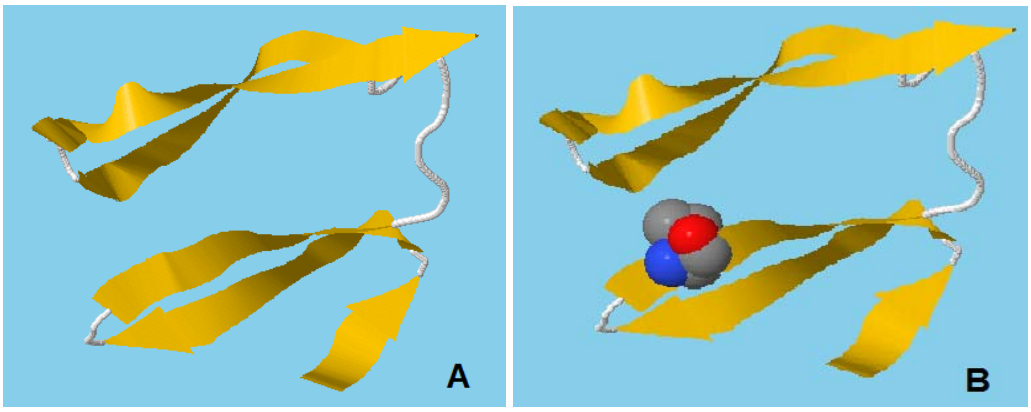

Figure (4): conformation of protein CD19 for leukemia (B) and healthy (A) of Baghdad / Iraq population by using Raptoxt software.

CD19 is a member of the Ig superfamily expressed on the surface of B lymphocytes, and may play a pivotal role in B-cell differentiation and activation (Carter and Fearon, 1992). On mature B cells, the protein complex consisting of CD19, CD21, CD81 and CD225, signals in conjunction with the B-cell antigen receptor upon antigen encounter, and is necessary to decrease the threshold for receptor dependent signaling (Van Noesel *et al.*, 1992). Patients with mutations in the CD19 (MIM 107265) gene have reduced immunoglobulin (Ig) serum levels, reduced memory B cell numbers and impaired antibody responses to vaccinations (Van Zelm *et al.*, 2006; Kanegane *et al.*, 2007). Hypogammaglobulinemic patients from two unrelated families were shown to have defective CD19, resulting from mutations in the CD19 gene, describe a new case of CD19 deficiency with two novel mutations in the CD19 gene (van Zelm *et al.*, 2006). Van Zelm *et al.,* (2006) and Kanegane, *et al.,* (2007) described a CD19 deficient patient, who was compound heterozygous for two novel mutations, these mutations lead to loss of CD19 membrane expression, and result in an antibody deficiency syndrome, characterized by recurrent (mainly) bacterial infections. The mutation in the splice acceptor site of exon 6 results in skipping of exon 6, placing exon 7 out-of frame after exon 5, which leads to a premature stop codon at the intracellular amino acid 320. In line with this, the patient's B cells completely lack membrane CD19 expression. Furthermore, the lack of all intracellular domains responsible for signal transduction makes any expressed protein non-functional (van Zelm *et al.*, 2010). van Zelm, *et al.,*(2011) showed to 156G.C mutation resulted in the replacement of a tryptophan residue into a cysteine (W52C), the affected tryptophan is located in domain 1 of CD19 and corresponds to nucleotide mutation 123G.C and amino acid change W41.C in exon 2 of CD19 of antibody deficiency.

Conclusions

Our study Shows the difference in the order of the amino acids and the emergence of site side of acids and the difference in body protein leads to variation in the function of CD 19 and association

of mutation in *CD19* with leukemia in Baghdad, Iraqi.

References

Carter, R.H. and Barrington, R.A. 2004. Signaling by the CD19/CD21 complex on B cells. Curr. Dir. Autoimmun, 7: 4–32.

Carter, R.H. and Fearon, D.T. 1992. CD19: lowering the threshold for antigen receptor stimulation of B lymphocytes. Science, 256: 105-7.

Cooper, L.J.; Al-Kadhimi, Z.; DiGiusto, D.; Kalos, M.; Colcher, D.; Raubitschek, A.; Forman, S.J. and Jensen, M.C. 2004. Development and application of CD19-specific T cells for adoptive immunotherapy of B cell malignancies. Blood Cells Mol. Dis., 2004, 33(1): 83–89.

Ginaldi, L.; De Martinis, M.; Matutes, E.; Farahat, N.; Morilla, R. and Catovsky, D. 1998. Levels of expression of CD19 and CD20 in chronic B cell leukaemias. J. Clin. Pathol., 51(5): 364–369.

Ishiura, N.; Nakashima, H.; Watanabe, R.; Kuwano, Y.; Adachi, T.; Takahashi, Y.; Tsubata, T.; Okochi, H.; Tamaki, K. and Tedder, T.F. 2010. Differential phosphorylation of functional tyrosines in CD19 modulates B-lymphocyte activation. Eur J Immunol, 40(4): 1192–1204.

Kehrl, J.H.; Riva, A.; Wilson, G.L. and Thevenin, C. 1994. Molecular mechanisms regulating CD19, CD20 and CD22 gene expression. Immunol. Today, 15(9): 432–436.

Poe, J.C.; Minard-Colin, V.; Kountikov, E.I.; Haas, K.M. and Tedder, T.F. 2012. A c-Myc and Surface cd19 signaling amplification loop promotes B cell lymphoma development and progression in mice. J. Immunol., 189(5): 2318–2325.

Kanegane, H.; Agematsu, K.; Futatani, T.; Sira, M.; Suga, K.; Sekiguchi, T.; Van Zelm M. and T Miyawaki T. 2007. Novel mutations in a Japanese patient with CD19 deficiency. Genes Immunity, 8, 663–670.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Molec. Evol., 16: 111-120.

Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M. and Kumar, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molec. Biol. Evol., 28: 2731-2739.

Thierry-Mieg, D. and Thierry-Mieg, J. 2006. AceView: a comprehensive cDNA-supported gene and transcripts annotation. Genome Biol., 7(Suppl 1:S12):11–14.

van Noesel, C.J.; Lankester, A.C. and van Lier, R.A. 1993. Dual antigen recognition by B cells. Immunol. Today, 14: 8–11.

van Zelm, M.C.; Reisli, I.; van der Burg, M.; Castano, D.; van Noesel, C.J.M. and van Tol, M.J.D. 2006. An antibody-deficiency syndrome due to mutations in the CD19 gene. North England J. Med., 354: 1901–1912.

van Zelm, M.C.; Smet, J.; Adams, B.; Mascart, F.; Schandene, L.; Janssen, F.; Ferster, A.; Kuo, C.C.; Levy, S. and van Dongen, J.J. 2010. CD81 gene defect in humans disrupts CD19 complex formation and leads to antibody deficiency. J. Clin. Invest., 120, 1265–1274.

Vardiman, J.W.; Thiel, J. and Arber, D.A. 2009. The 2008 revision of the world health organization (WHO) classification of myeloid neoplasms and acute leukemia, ration ate and important changes. Blood, 114: 937-951.

Yang, W.; Agrawal, N.; Patel, J.; Edinger, A.; Osei, E.; Thut, D.; Powers, J. and Meyerson, H. 2005. Diminished expression of CD19 in B-cell lymphomas. Cytometry B. Clin. Cytom., 63(1): 28–35.

Zelm, M.; Smet, J.; Burg, M.; Ferster, A.; Le, P.; Schandene, L.; Dongen, J. and Mascart, F. 2011. Antibody deficiency due to a missense mutation in CD19 demonstrates the importance of the conserved tryptophan 41 in immunoglobulin superfamily domain formation. Human Molec. Genet., 20(9): 1854–1863.