Parvovirus b19 and auto antibodies reactive with ssDNA in lupus disease: bioinformatics analysis and hypothesis

**Abstract**

In this article, the possible link between Parvovirus B19 infection and Lupus Disease is hypothesized and the validation of possibility checked through structural features of single stranded (ss) viral DNA. The binding sites on ssDNA are thymidine pen tamer as it is confirmed by X-ray crystallography data. Five of them are necessary for recognition and three for binding, and the structure of ss Parvovirus B19 found in literature indicates that. We have also found confirmation for this structure using an interactive programming environment MATLAB®, commonly applied in many technical fields for data analysis indicating the base number where we have at least five consecutive T’s. These segments are potential sites for initial anti ssDNA antibody binding and ssDNA hydrolysis. In a further step, we have located the CpG islands, through the online sequence analysis tool CpGPPlot/CpGReport with the specific algorithm parameters such as region length (>50 base pairs), GC% (>30%), and observed/expected CpG ratio (>60%). It has been shown that upon viral infection these unmethylated viral CpGs within DNA or synthetic oligodeoxynucleotides (ODNs) can stimulate immune cells via TLR9, leading to production of antibodies against the virus. Since the virus is retained in the body fluids after infection, its re-appearance could initiate the production of anti-ssDNA auto antibodies with not only neutralizing but also hydrolytic activity in order to completely eliminate the viral antigen. Thus, anti-ssDNA auto antibodies in lupus with hydrolytic activity may be of predictive, diagnostic and prognostic significance in this systemic autoimmune disease.

**Keywords:** parvovirus b19, lupus, anti-ssDNA auto antibodies, DNA hydrolysis, toll like receptors (tlr) bioinformatics

**Introduction**

Systemic lupus erythematosus (SLE) is an autoimmune disease which is difficult to diagnose due to need to employ multiple criteria that can vary from one patient to another. One of the hallmarks of lupus disease is the secretion of anti-DNA autoantibodies, the titer of which is increased during a flare of symptoms.1,2 Reactive B-cells are considered to play a central role in the pathogenesis of SLE and multiple autoimmune diseases, although so far nobody knows what the lupus B-cell really is.3,4 Despite emerging new data on cooperative functions between inflammasomes, DNA-sensing proteins, Toll like receptors (TLR), and signal transduction to transcription factors (NF-kB) after viral or bacterial infections as potential triggering factors for anti-DNA antibody immunoglobulin gene-rearrangement and antibody secretion, the precise mechanisms that promote alterations in B-cell tolerance, hyperactivity and production of anti-DNA antibodies remain incompletely defined.5,6 Nevertheless, anti-DNA autoantibodies are considered to be structural, functional and pathogenic entities not only in SLE but also in the broad spectrum of diseases with reactive B-lymphocytes.7,8 Both earlier and quite recent research studies have confirmed that capture of antigens is not the only function of the antibodies and that they can be multifunctional, e.g., predictive, prognostic and protective.9,10 Catalytic antibodies, known as enzymes have been predicted and since then, the entire spectrum has been discovered, especially in autoimmune diseases.11,12 Despite the expanded research within last years, though hydrolytic activity of anti-DNA antibodies has been proven by several groups, their prognostic and pathogenic values are not yet fully elucidated.13-15 One of the crucial reasons for that was the lack of sensitive and precise, real-time methods for measurement of their activity, which seems to be at least, conceptually solved recently.16,17,18 Accordingly, this makes the cryptic nature of anti-DNA autoantibodies, extremely intriguing. Therefore, there are so many unanswered questions within this field, such as: Catalytically active antibodies are found in the sera and milk of pregnant and lactating women and considered to provide a maternal strategy against microbial attack of the fetus and newborns.19,20 In addition, to anti-DNA antibodies, antibodies towards RNA, NMP, NDP and NTP as well as antibodies with proteolytic activity have been found in human sera of healthy and diseased people suggesting that the organism is trying to fight a microbial agent. Microbial DNA (bacterial and viral) is a known immunogen as some authors suggest that a high frequency of methylated CpG motifs in microbial DNA playing stimulatory role in anti-DNA antibody production and lupus flare acting through toll-like receptor 9 (TLR-9) on/in B cells and plasmacytoid dendritic cells.21-23 The basis for such theorizing are the results of studies showing anti-DNA antibody binding to peptide-mimicking antigens synthetically designed, some of which mimic viral or bacterial proteins.16

**Methods, results and discussion**

One of the candidates for the appearance of anti-ssDNA antibody could be the human ssDNA virus, Parvovirus B19, which has all possible characteristics14,15 of a causative/triggering agent in (at least
a small fraction) of lupus disease spectrum (Figure 1A) (Figure 1B). Parvovirus B19 can be found in most humans; however, symptoms only appear in a few patients. Additionally, when symptoms are displayed, they closely mimic lupus symptoms. The striking difference is that during infective disease, patients’ blood contains interferon gamma and during developed lupus disease, interferon alfa secreted for evolution and growth of plasmacytoid dendritic cells, indicating skewing in cytokine production after transition into autoimmune disease. The spectrum of lupus symptoms and the heterogeneous nature of the disease suggest it is probably caused and triggered by multiple environmental factors (similar to cancer) and therefore, requires individual diagnosis, therapy and prevention. Perhaps, prevention of that fraction of lupus caused/triggered by parvoviruses may be possible by early vaccination against the virus, which failed so far. There are emerging computational data on Rational Vaccine Design for some DNA viruses.

Figure 1 Parvovirus B19 DNA Sequences. A) Human parvovirus B19 5,594 bp linear genomic complete sequence, isolate C39 from plasma (FN598217.1; GI: 270118453). B) Linear, non-segmented, ssDNA, ~5kb. Most of the strands packaged seem to be (+) sense, but AAVs package equal amounts of (+) and (-) strands, and all seem to package at least a proportion of (+) sense strands. The ends of the genome have palindromic sequences of ~115nt which form “hairpins”. These structures are essential for the initiation of genome replication.

We have initial computational data on analysis of Parvovirus B19 ssDNA sequence indicating that it’s DNA could have been one potential reason for production, and the target for binding of anti-ssDNA antibodies produced in lupus disease Figure 2. For that purpose, we used an interactive programming environment MATLAB®, commonly applied in many technical fields for data analysis. A code was written for analysis of the nucleotide sequence of Parvovirus B19 complete DNA sequence obtained from gene bank (bases 1 through 5591). Nucleotide sequence(s) were represented by their designated single-letter code (A, C, G, or T). This sequence was then loaded into the MATLAB® for data analysis. If at least 5 consecutive T’s (Target nucleotides) are detected, each of these consecutive T’s were assigned a single numerical value of “1”, while remaining nucleotides (A, C, G or less than 5 consecutive T’s) were assigned a “0” value. All the 1’s were plotted using MATLAB® plotting feature, thus indicating the base number where we have at least five consecutive T’s.

Figure 2 Distribution of at least five consecutive thymidine within Paroviral B19 sequence (bases 1-5594). The X-axis are the bases from 1-5594. If 5 consecutive T’s are present, then the base number from which we have 5 consecutive T’s is indicated by the blue line. We can see the exact base number where these T’s begin and label the axis as desired.

The X-axis are the bases from 1-5594. If 5 consecutive T’s are present, then the base number from which we have 5 consecutive T’s is indicated by the blue line. We can see the exact base number where these T’s begin. Specific segments with at least 5 consecutive T’s in Parvovirus B19 genome are represented. These segments are potential sites for initial anti DNA antibody binding and hydrolysis. According to literature data, 5t’s are necessary for recognition; at least three of them are necessary for binding (Figure 3 & Figure 4).

Figure 3 Locations of potential thymidine “hot spots” for anti-ssDNA antibody binding. Parvovirus B19 genome segment regions with 5 consecutive T’s.

Figure 4 Locations of potential thymidine “hot spots” for anti-ssDNA antibody binding. Parvovirus B19 genome segment regions with 6 or more consecutive T’s.
To locate the CpG islands, the online sequence analysis tool CpG Plot/CpG Report was used with the specific algorithm parameters such as region length (>50 base pairs), GC% (>30%), and observed/expected CpG ratio (>60%) Figure 5. The CpGP lot function identifies and plots CpG islands in nucleotide sequence(s), and CpG Report identifies and report CpG-rich regions in nucleotide sequence(s). The observed number of CpG patterns is the count of the number of times a ‘C’ is found followed immediately by a ‘G’. The expected number of CpG patterns is the number of CpG dinucleotide you would expect to see in that window based on the frequency of C’s and G’s. Each CpG region longer that 50 base pars is represented in the bottom plot. Upon viral infection these unmethylated viral CpGs within DNA or synthetic oligodeoxynucleotides (ODNs) have been shown to stimulate immune cells via TLR9 leading to production of antibodies against the virus.

Although viral anti IgGs, and anti IgMs are distinctive in acute stage of the disease, the virus is known to withdraw into the body’s fluids and do not appear for the long time, or never. It is possible that with the brake of tolerance and appearance of the lupus flares, which include the reoccurrence of the virus in the bloodstream, somewhat different anti-ssDNA autoantibodies are designed by host immune system in order not only to neutralize viral ssDNA, but to toxically and hydrolytically eliminate it from the body. Thus, the structure of parvo viral B19 DNA becomes attractive for the immune system of lupus patients indicating and worsening the flares. Therefore, anti-ssDNA autoantibodies in lupus with hydrolytic activity could have been of predictive, diagnostic and prognostic significance in this systemic autoimmune disease. Further studies are necessary to confirm this hypothesis and its significance.

It is even of higher significance when it is known that Parvovirus B19 is not the only human non-circular ssDNA virus and that it has been confirmed to be integrated into human genome similar to HERVS (human endogenous retroviruses), by using the rolling-hairpin mechanism of replication.\textsuperscript{30‒33} Of general significance is the evidence that Parvovirus gives the rise to theory of horizontal gene transfer evolution (through evolutionary tree) of bidnaviridae that infect the
silkworm, using the same replication mechanism. The fact that there is still not a good vaccine for Parvovirus B19 indicates that these genomic events between viral and host DNA might be a contributing variable to its failure. However, there is the lack of literature data on polydT and CpG island analyses within the context conceptualized in our hypothesis in both circular and non-circular ssDNA viral species. Genomic features of gemycircularviruses HV-GcV1 and HV-GcV2 and of circular single-stranded DNA (ssDNA) virus, HV-CV1, including hairpin structure and predicted open reading frames (ORF) are given recently. Both newly discovered viruses contain at least one poly thymidine pentamer. The discovery of viruses belonging to ssDNA strains is still going on, making the classifications the matter of future changes. These also applies to the mechanisms of viral ssDNA integration which are not fully understood neither totally controlled. The conditions governing how the cell chooses which mode of integration to employ are unclear Table 1.

Table 1 Group II - ssDNA viruses

<table>
<thead>
<tr>
<th>Family</th>
<th>Genus</th>
<th>Species</th>
<th>CpG</th>
<th>Polyd T</th>
<th>Disease Association in Humans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anelloviridae</td>
<td>Alphatorquevirus</td>
<td>Torque teno virus (TTV)</td>
<td>Yes</td>
<td>Yes</td>
<td>May be associated with hepatitis, pulmonary diseases, hematologic disorders, myopathy, multiple sclerosis and lupus.</td>
</tr>
<tr>
<td></td>
<td>Betatorquevirus</td>
<td>Torque teno mini virus (TTmV)</td>
<td>Yes</td>
<td>Yes</td>
<td>Not known</td>
</tr>
<tr>
<td></td>
<td>Gammatorquevirus</td>
<td>Torque teno mini virus (TTmDV)</td>
<td>Yes</td>
<td>Yes</td>
<td>Associated with encephalitis, diarrhea and sewage. Also found in cerebrospinal fluid and brains of patients with multiple sclerosis.</td>
</tr>
<tr>
<td>Genomoviridae</td>
<td>Gemygorvirus</td>
<td>Sewage derived gemygorvirus 1</td>
<td>Yes</td>
<td>Yes</td>
<td>Associated with encephalitis, diarrhea and sewage. Also found in cerebrospinal fluid and brains of patients with multiple sclerosis.</td>
</tr>
<tr>
<td></td>
<td>Gemykibivirus</td>
<td>Human associated gemykibiviruses 1-5</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gemyvongvirus</td>
<td>Human associated gemyvongvirus 1 (DB1)</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclassified</td>
<td>Divergent GeVs and CRESS-DNA virus (CV)</td>
<td>Yes</td>
<td>Yes</td>
<td>Associated with pleuropneumonia and pericarditis</td>
</tr>
<tr>
<td>Parvoviridae</td>
<td>Bocaparvovirus</td>
<td>Human bocavirus (HBoV)</td>
<td>Yes</td>
<td>Yes</td>
<td>Acute respiratory illness &amp; gastroenteritis</td>
</tr>
<tr>
<td></td>
<td>Dependoparvovirus</td>
<td>Adeno-associated virus 1&amp;2</td>
<td>Yes</td>
<td>Yes</td>
<td>Not known; potential to prevent and reverse autoimmune conditions</td>
</tr>
<tr>
<td></td>
<td>Erythroparvovirus</td>
<td>Parvovirus B19</td>
<td>Yes</td>
<td>Yes</td>
<td>Fifth disease and skin lesions. Implicated as the precipitating agent of several autoimmune disorders including rheumatoid arthritis, systemic lupus, antiphospholipid syndrome, systemic sclerosis and vasculitides.</td>
</tr>
<tr>
<td></td>
<td>Protoparvovirus</td>
<td>Human bufavirus (BuV)</td>
<td>Yes</td>
<td>No</td>
<td>Gastroenteritis</td>
</tr>
<tr>
<td></td>
<td>Tusavirus 1</td>
<td></td>
<td></td>
<td></td>
<td>Gastroenteritis</td>
</tr>
<tr>
<td></td>
<td>Tetraparvovirus</td>
<td>Human parvovirus 4 (PARV4)</td>
<td>Yes</td>
<td>Yes</td>
<td>Associated with influenza-like syndrome, encephalitis, acceleration of HIV disease, and foetal hydrops</td>
</tr>
</tbody>
</table>

Describes some common ssDNA viral infections and possibility for exhibition of CpG islands and polydT sequences with at least one of polydT (pentamers) in viral genome, which opens the possibility for formation and binding of catalytic anti-ssDNA antibodies and their involvement in autoimmune mechanisms.

Conclusion and future work

Based upon our bioinformatics data analysis it seems reasonable to hypothesize that ss viral DNA is a possible immunogen with a high frequency of polydT and unmethylated CpG motifs in viral DNA playing stimulatory role in hydrolytic anti-DNA antibody production, and lupus flare initiation, acting through toll-like receptor 9 (TLR-9) on/in B cells and plasmacytoid dendritic cells. Aotsuka et al., Teodorescu et al., have suggested in their clinical studies that appearance of anti-ssDNA might be regarded the predictive sign of the incoming flare of the disease. However, it is not yet accepted by clinicians and as it could be a helpful marker, we shall check this possibility in future works, in order to try to answer the intriguing remaining questions: What is the exact role of anti-ssDNA antibodies in lupus and other autoimmune diseases? What part of their molecule possesses catalytic activity: Fab fragment or either heavy or light chain. Does the Fc
fragment inhibit or slow down enzymatic activity as it is proposed by others, or is it causing that effect by its non-specific binding to the cells?\(^1\)\(^2\) Some healthy individuals as well as SLE patients produce anti-DNA antibodies, which can be isolated from the serum.\(^3\)\(^4\) However, the normal individuals do not appear to be affected and their antibodies are not hydrolytically active.\(^5\) Hydrolytic activity has only been seen in anti-DNA antibodies produced in the disease state.\(^6\)\(^7\) Why are they then present in normal individuals? Is it individual and so, why?

**Acknowledgments**

None.

**Conflicts of interest**

Author declares that there is no conflicts of interest.

**References**


