Journal of Molecular Biology and Molecular Imaging

Open Access

Austin Publishing Group

Research Article

Systemic Lupus Erythematosus: Localization of Dnase, Trypsin-Like, and Metalloprotease Active Centers in the Protein Sequence of NGTA3-Pro-Dnase Monoclonal Light Chain of Human Antibodies

Timofeeva AM, Buneva VN and Nevinsky GA*
Institute of Chemical Biology and Fundamental Medicine, The Siberian Division of Russian Academy of Sciences, Lavrentiev Ave. 8, 630090, Russia

***Corresponding author:** Georgy A. Nevinsky, Institute of Chemical Biology and Fundamental Medicine, The Siberian Division of Russian Academy of Sciences, Lavrentiev Ave. 8, 630090, Russia

Received: April 25, 2020; **Accepted:** July 24, 2020; **Published:** July 31, 2020

Abstract

Canonical enzymes usually possess only one active site and catalyze only one chemical reaction. The completely unpredictable result was obtained four autoantibodies with catalytic activities. Monoclonal light chains NGTA3-pro-DNase (MLC), unlike classical enzymes, possess three enzymatic activities. NGTA3-pro-DNase hydrolyzes myelin basic protein (MBP) due to its serine-like and metalloprotease activities as well as split DNA by DNase activity. Possible homology of NGTA3-pro-DNase protein sequence with those of DNase I, as well as several known human serine-like, Zn^{2+}- and Ca^{2+}-dependent proteases were analyzed. The fragments of the protein sequence of MLC responsible for specific binding of myelin basic protein, chelation of metal ions, amino acid residues of active centers involved directly in the catalysis of two proteolytic reactions, as well as DNA hydrolysis were found. Computer simulation, a possible three-dimensional structure of NGTA3-pro-DNase, has shown its good structural similarity with the crystal structure of a catalytic antibody having a serine protease active site, the difference 1.79 Å. Possible potential complexes with oligopeptide and oligonucleotide were constructed by analysis of MLC structural similarity with the crystal structure of HIV-1 neutralizing antibodies in complex with its v3 loop oligopeptide antigen and complex of antibody 64M-2 Fab with d(TpT). The data obtained greatly expands the classical ideas about the possible biological functions of antibodies.

Keywords: Systemic lupus erythematosus; Monoclonal light chains; Proteolytic activity; DNase Activity; Abzyme catalytic centers

Abbreviations

Abs: Antibodies; AA: Amino Acid; ABZs: Abzymes; AID: Autoimmune Disease; AV: Average Value; CAPN: Calpain Protease; MLC: Monoclonal Light Chain; MBP: Myelin Basic Protein; SLE: Systemic Lupus Erythematosus; EDTA: Ethylenediaminetetraacetate; PMSF: Phenylmethylsulfonyl Fluoride.

Introduction

Antibodies (Abs) against chemically stable analogs of transition states of chemical reactions and natural auto-Abs demonstrating enzymatic activities (abzymes, or ABZs), are well described (for review see Refs [1-6]). IgGs and/or IgAs and IgMs abzymes that hydrolyze ATP, DNA, RNA, oligosaccharides, peptides, and proteins were revealed the blood sera of patients with various autoimmune diseases (AIDs): systemic lupus erythematosus (SLE), Hashimoto's thyroiditis, polyarthritis, multiple sclerosis (MS), lymphoproliferative disease, polyneuritis, malignant tumors, as well as with three viral diseases - viral hepatitis, tick-borne encephalitis and human immunodeficiency ([1-6] and references therein). It is very important that antibodies from the blood of healthy donors usually did not demonstrate perceptible catalytic activities [1-6]. Enzymatic activities of ABZs are easy to test at the onset of the development of various AIDs when other immunological, biochemical, and medical indicators of these pathologies are usually not differed statistically significantly from those for healthy donors. The increase ABZs enzymatic activities in time can be used both for the early diagnosis of various AIDs and for assessing the depth of autoimmune processes [1-6].

SLE is a systemic autoimmune pathology at which occurs disorganization of connective tissues, comprising damage of skin and visceral capillaries [7]. Polyclonal auto-Abs (IgGs and/or IgAs, IgMs) hydrolyzing DNA, RNA [8–11], MBP [12–15], and oligosaccharides [16–19] were revealed in the blood of patients with SLE.

Monoclonal light chains (MLCs) possessing only DNase and MBP-hydrolyzing activities were obtained using the kappa light chain cDNA library of peripheral blood Abs of patients with SLE [20–28]. All individual MLCs with DNase activity demonstrated very many pH optima, individual dependencies on the concentration of different metal ions [24,25]. The affinities of all MLCs to DNA were about 2–3 orders of magnitude higher than that for human DNase I. Similarly to canonical metalloproteases; several MLCs were inhibited only by EDTA demonstrating only metal-dependent activity [26]. Enzymatic activity of other MLCs having typical serine-like or thiol-like proteases activity was suppressed respectively only by PMSF or

Citation: Timofeeva AM, Buneva VN and Nevinsky GA. Systemic Lupus Erythematosus: Localization of Dnase, Trypsin-Like, and Metalloprotease Active Centers in the Protein Sequence of NGTA3-Pro-Dnase Monoclonal Light Chain of Human Antibodies. J Mol Biol & Mol Imaging. 2020; 6(1): 1031.

by iodoacetamide

In contrast to canonical enzymes, absolutely unusual was NGTA3-pro-DNase has three sites combined in one active center: serine-like, metal-dependent, and DNase activities [29]. Taking this into account, it was particularly interesting to what extent the combined in one variable part of the light chain three different active centers of unusual NGTA3-pro-DNase may be similar or different in comparison with those for classical enzymes.

In this paper, we have analyzed for the first time the protein sequence of NGTA3-pro-DNase and its homology with those for various peptidases and DNases in order to identify MLC sequences responsible for serine-like, metalloprotease, and DNase activities.

Materials and methods

Obtaining and analyzing monoclonal light chains

We used the phage cDNA library of recombinant kappa-type light chains of patients with SLE described earlier [21]. The cDNA library was a gift from S. Paul and S. Plaquet (University of Texas Houston Medical School, USA). All operations to obtain phage particles and monoclonal antibodies are described previously [21,26-29].

DNA and protein sequences of NGTA2-Me-pro-Tr

Determination of nucleotide sequence corresponding to NGTA3-pro-DNase was performed similarly to [21]; M13 reverse (5'- GGAAACAGCTATGACCATG-3') and FdSeq1 (5'-GAATTTTCTGTATGAGG-3') primers were used for sequence analysis. Nucleotide sequence VL-fragments of DNA were found using automatic sequenator CEQTM 2000XL DNA (Beckman) and special kit "F"CEQ DTCS. The amino acid sequence corresponding to the DNA sequence was found using base data IgBLAST and V-base. The nucleotide sequence of recombinant MLC has been submitted to the GenBank (gb-admin@ncbi.nlm.nih.gov); Data Bank with accession number nuc_4.sqn nuc_1 KP713344 for NGTA3-pro-DNase.

A possible homology between NGTA3-pro-DNase and classical enzymes was verified by the ClustalW2 server (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The percent of homology between sequences was calculated using lalign (http://www.ch.embnet.org/software/LALIGN_form.html). The analysis of MLC gene type was performed using BLASTN 2.2.27+ server (http://www.ncbi.nlm.nih.gov/igblast/igblast.cgi). DNA sequence of NGTA3-pro-DNase was closest to those for previously described IgG light chains: IGKJ1*01 (100% of identity), IGKJ4*01 (95.7%), IGKJ4*02 (91.2%); IGKV1-5*03 (79.8% of identity), IGKV1-5*02 (78.4%), and IGKV1-5*01 (78.4%).

The spatial structure of MLC

The spatial structure of NGTA3-pro-DNase was found using the crystal structure of abzyme with a serine protease activity (pdb id 1eap chain a). The good structural similarity was found; the difference between the structures (RMSD) is 1.79.

The potential complex of NGTA3-pro-DNase with oligopeptide was built using structural similarity of MLC with the crystal structure of an HIV-1 neutralizing antibody 50.1 in complex with its v3 loop oligopeptide antigen. A potential complex of MLC with dinucleotide was found using structural similarity of NGTA3-pro-DNase with antibody 64M-2 Fab complexed with d (TpT).

Results

Homology of sequences of different protein

The nucleotide and amino acid sequences of NGTA3-pro-DNase (further designated as MLC) were determined (Figure 1).

Interestingly, only the same enzymes from various mammals usually demonstrate a high homology (70-95%). NGTA3-pro-DNase possesses three different activities: serine-like, metalloprotease, and DNase. [30]. The homology of different proteins with the same enzymatic activities turned out to be significantly lower than the same enzymes from various mammals [30]. The homology between four serine proteases (elastase, trypsin, chymotrypsin, and trypsinogen) is low and varies from 34.4% to 41% identity of amino acids (AAs) [30].

The levels of homologies (the identity of AAs) between NGTA3-pro-DNase and classical proteases are varying from 24.5 to 25.4%, while similarities (non-identical amino acids with highly and moderately conserved properties) is significantly higher: from 42.5 to 48% [30].

Protein sequences of thirteen human metalloproteases Zn^{2+}-(MMPs) [31-34] and Ca^{2+}-dependent proteases [35,36] are described earlier. NGTA3-pro-DNase is neither Zn^{2+}- nor Ca^{2+}-dependent enzyme since it is activated by ions of various metals [29]. Its maximal activity was observed in the presence of Ca^{2+} ions (100%), while other metal ions activate MLC not much [29]. The identity of MMP1 AAs with the other twelve MMPs (MMP25, MMP24, MMP14, MMP7-MMP12, MMP3, and MMP2) varies from 30.3 to 61.5% [30]. Since the low level of homology between Zn^{2+}-dependent proteases, it was difficult to expect a high level of homology between MLC and these classical metalloproteases. The level of homology between the Zn^{2+}-dependent proteases and MLC ranged from 24.6 to 27.2%, the average value (AV) is 25.7±0.8%, which is comparable in the homology between some metalloproteases (30.6-31.3%). The similarity between all Zn2+- proteases, and MLC was significantly higher: 40.0-48.6%, AV is 44.5±0.3% [30].

The number of AAs of nine described Ca2+-dependent proteases (Calpains or CAPNs with numbers 1-3, 5-10) varies from 640 to 821

A DNA sequence of NGTA3-pro-DNase

1- ATGGCCGACATCCAGCCCAATCTCCATCCACCCTGTCTGCATCAGTGGGAG
ACAGAGTCACCATCAGTTGCCGGGCCAGTCCGGTATTACATGGTTGGCCTGTTAT
CACGACATTGCAGGGAAAGCCCCTTAACTCCTGATCTAGGCGTCTACTTTACAAA
CTGGGGTCCCATCTATGTGCAGCGGCCATGGCTCTGGGACTCAGTTCAATCTCA
CCATCAACACCCTGCGGCCTGGTTCTGCAACTTACTACTGTCGACAGTAATACAG
TTATCCCTGAACGTTCGGCCAAGGGACCAAGGTGGAAATCAAACGAACTGTGGC
TGCACCATCTGTCTTCATCTTCCCGCCATCGCAGTTGAAATCTGGAACTGCCTCT
GTTGTGTGCCTGCTGAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTG
GACGCCCTCCAATCGGGCTCCCAGGAGAGTGTGCAGAAGCAGGACAGTCGGGA
CAGCGCCTACAGCCTCATCAGCACCCTGACGCTAACGAAAGCCGCCTACCAGAA
ACCACCGAGTCTACCCCTGCGAAGTCACCCATCAGGACCACCGCCTTGGGAGGU
AGAGAGAACAAGCTCGCACTCCCCAAGCATGATTTCTGCCGCACCGACGGTGGA
GAGCTTATCGGTAGTAGCGGTCACGAGCCAGCGGCTGCCTTACTCGGTCTCACT
CGCGCCGAAATCGGATAGAAAGTCGAGAAACTCGGCCGCTTCGCTCTCATTCAT
CGAGAACCTGTCGTTTTAATTCTCGCTGCACCCCGAACTGGAGGTCTTAACGAAA
CTAGUGCAGCTAGTAAGCCAGCTGCAGGGTTCGGTCAACATACTCGCGCTCTTG
TTGTAGGAGCTAATAAGCTCCTAGCTAGAACTCCCAGTCGTGCCGCACTCGGATT
CGGAAACGTTGTAGTTAGTCGTCTAGCACAAGTA - 954

B Protein sequence of NGTA3-pro-DNase

1 - MADIQPNLHPPCLHQWETESPSVAGPVRYYMVGLLSRHCRESPLTPDLGVYF
TNWGPIYVQRPWLWDSVQSHHQHPAAWFCNLLLSTVIQLSLNVRPRDQGGNQTN
CGCTICLHLPAIAVEIWNCLCCVPAELLSQRGQSTVEGGRPPIGLPGECAEAGQSG
QRLQPHQHPDANESRLPETTESTPAKSPIRTTAFGGRENKLALPKHEFRTDGGELIG
SSGHEPAAAFLGLTRAEIGSKVEKLGRFALIHREPVVFILAAPRTGGLNETSAASKP
AAGFGQHTERALVVGANKLLARTPSRAALGFGNVVVSRLAQV - 318

Figure 1: DNA (A) and protein (B) sequences of NGTA3-pro-DNase.

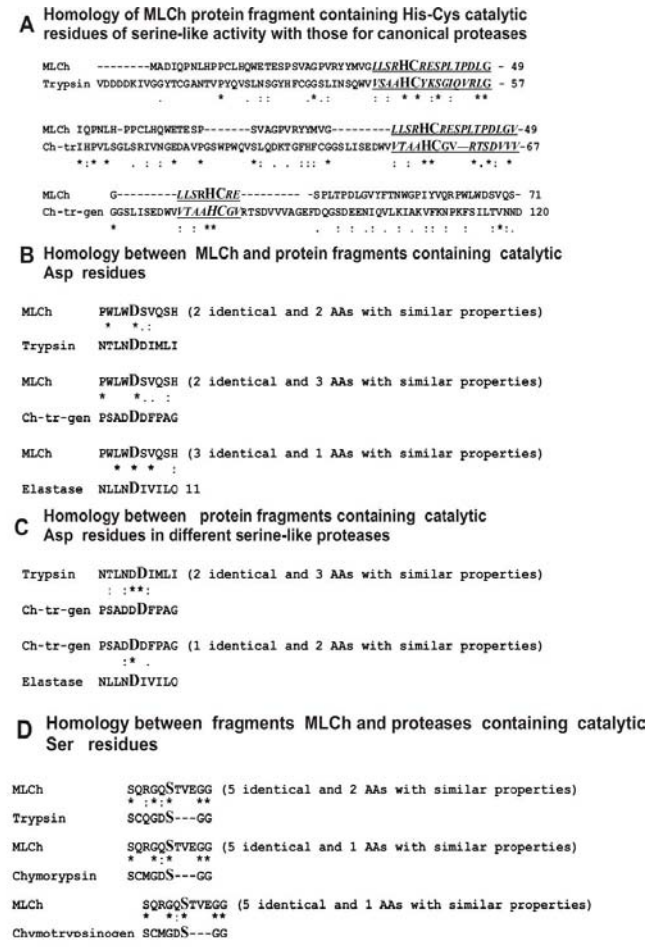

Figure 2: Homology of fragments of MLC protein sequences containing catalytic His-Cys (A), Asp (B), and Ser (D) with corresponding clusters of trypsin, chymotrypsin, elastase, and chymotrypsinogen containing the same catalytic residues. Panel C demonstrates homology between sequences of several canonical serine-like proteases containing catalytic Asp. Homology of Panel A sequences was analyzed using ClustalW2, while Panels B-D by LALIGN server. The amino acids identical between two pairs of sequences are marked with an asterisk (*), while non-identical amino acids with highly (colon; A) and moderately (point; A) or together highly and moderately conserved properties with a colon (:); Panels B-D.

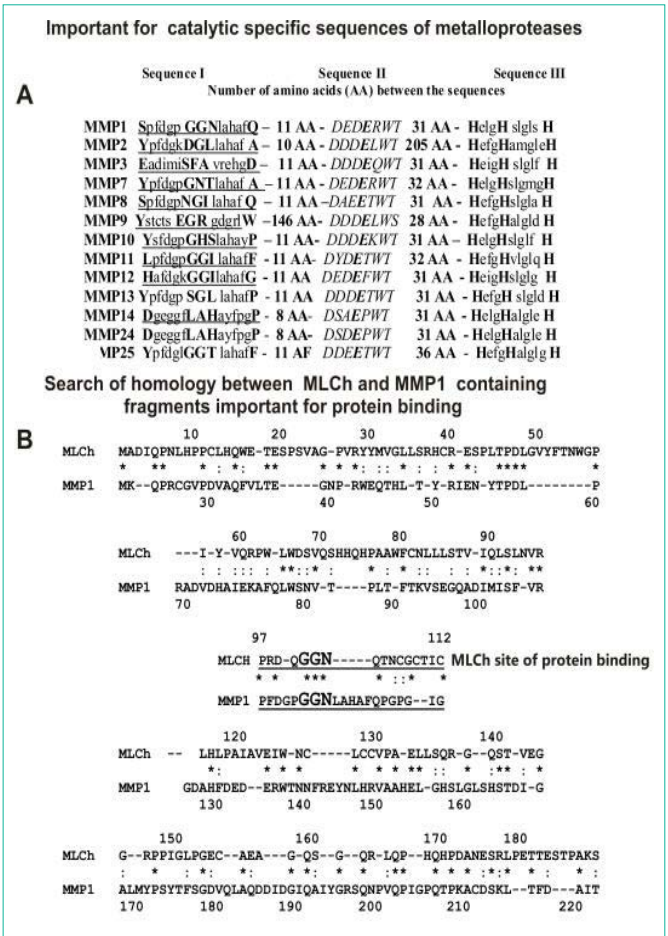

Figure 3: Three specific fragments of protein sequences of thirteen human metalloproteases, which are important for binding of protein-substrates (sequence I), catalysis of the reaction of protein substrate hydrolysis (sequence II) and chelating of metal ions (sequence III) (A). Homology (LALIGN server) between the universal protein protein-binding sequences of MMP1 with hypothetical MBP-binding site of NGTA3-pro-DNase (B). The amino acids identical between two pairs of sequences are marked with an asterisk (*), while non-identical amino acids with highly and moderately conserved properties with a colon (:).

[36]. The maximum identity and similarity are observed for several calpains (%): CAPN1-CAPN2 (63.0 and 87.1), CAPN2-CAPN8 (62.0 and 85.8), and CAPN1-CAPN8: (59.0 and 83.2). The minimal identity and similarity were found for other CAPNs (%): CAPN2-CAPN7 (25.2 and 52.7), CAPN3-CAPN7 (24.9 and 51.1), CAPN7-CAPN10 (26.1 and 52.1), CAPN7-CAPN8 (26.3 and 52.9), CAPN1-CAPN7 (26.7 and 53.0) and CAPN7-CAPN9 (26.7 and 51.6). Overall, the level of AAs identity for nine CAPNs varies from 24.9 to 63%, and the similarity from 51.1 to 87.1% (the AV of identity is 37.9 ± 11, and the similarity is 63.2 ± 10.8%).

The identity of AAs between all CAPNs and MLC varied from 24.5 to 27.5% (AV 25.8±0.9%), and the similarity from 44.3 and to 48.9% (AV 47.1±1.7%). Thus, the homology of all the CAPNs and NGTA3-pro-DNase is comparable to that for several Ca^{2+}-proteases themselves (24.9-26.7%), as well as with the sequence of the MLC with some of Mg^{2+}- and Ca^{2+}-dependent proteases. Thus, the overall level of homology between various protease enzymes with the same functions can be relatively low. However, the homology between the specific fragments of these enzymes forming their catalytic centers is usually significantly higher, since they are more important for the formation of their active centers and the catalysis.

Since NGTA3-pro-DNase having DNase, serine-like, and metal-dependent activities is a typical light chain, its sequence must combine specific elements of protein sequences corresponding to the light chain V-genes and any canonical DNases, serine-like, and metalloproteases. The active centers of various abzymes are usually formed by variable regions on the light chains of auto-Abs [2-6]. Approximately half of the complete NGTA3-pro-DNase sequences (~150-156 AAs) shapes its variable region. All known serine-like proteases are different in their lengths and are longer than MLC: elastase (267 AAs), chymotrypsin (245 AAs), and trypsin (229 AAs) [37]. All zinc-dependent proteases are also more longer proteins containing from 403 to 707 AAs [31-34]. The lengths of CAPNs vary from 640 to 821 AAs [36]. In the hydrolysis of proteins by calcium-

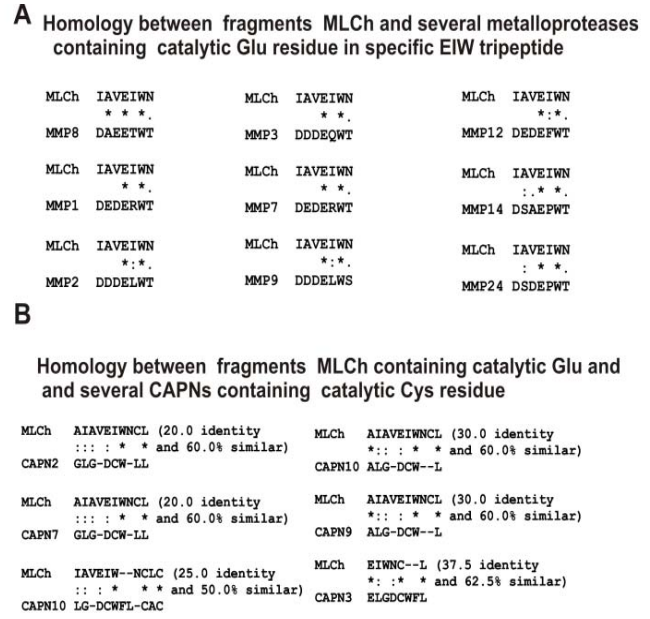

Figure 4: Homology between MLC sequence containing catalytic Glu with several MMPs containing catalytic Glu (A) and several CAPNs containing Cys catalytic residue (B). The amino acids identical between two pairs of sequences are marked with an asterisk (*), while non-identical amino acids with highly and moderately conserved properties with a colon (:).

dependent proteases are involved three AAs: Asn (N), His (H), and Cys (C). In nine different calpains, Cys residue can be in positions from 97 (CAPN9) to 359 (CAPN6) [36]. Depending on the CAPN, His may be far from Cys at a distance from 133 to 205 amino acids and Asn - 179-229 AAs. The spatial rapprochement of these AAs during the formation of catalytic centers of calpains occurs during the formation of tertiary structures of these proteases. In the case of NGTA3-pro-DNase, all its three enzymatic activities (DNase, metalloprotease, and serine-like) should be formed by ~150-156 AAs of MLC variable part. This implies possible significant differences in the organization of MLC active centers compared with those of canonical enzymes.

Analysis of a possible location of the serine-like active site

Ser residue of serine-like proteases catalyzes the hydrolysis of the peptide bonds due to existing of a specific catalytic triad, including serine (S), histidine (H), and aspartate (D), [37,38]. The D and H residues are necessary to make Ser of the proteolytic site more nucleophilic. Notwithstanding, a relatively low homology between complete sequences of various serine-like proteases (see above), all of them contain several more homologous and similar protein fragments, which are important for the catalysis. The analysis of homology between NGTA3-pro-DNase and serine-like proteases was carried out using the variable part of the MLC sequence (150-156 AAs).

The search of homology between MLC and different sequences of three serine-like proteases (three different sequence alignments) containing the same AAs residues was carried out. Only one fragment of good homology between NGTA3-pro-DNase and specific sequences of proteases containing catalytic His38 residue was found

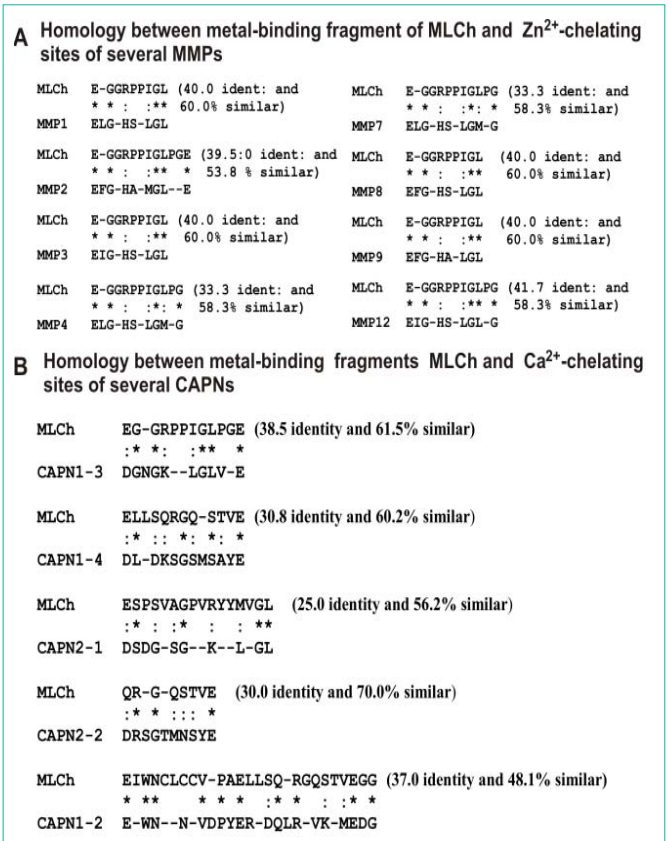

Figure 5: Homology between the sequence of the metal-chelating site of MLC and Mg2+-binding sites of several MMPs (A) and different Ca2+-binding clusters of CAPN1 and CAPN2 (the second number after the SAPN number means the number of the calcium ion binding site) (B) The amino acids identical between two pairs of sequences are marked with an asterisk (*), while non-identical amino acids with highly and moderately conserved properties with a colon (:).

(Figure 2A). The identity and similarity of His-sequence of MLC and four proteases (16 AAs) were respectively (%): trypsin (35.8 and 58.8), chymotrypsin (38.5 and 42.6), chymotrypsinogen (38.5 and 46.2), and elastase (50.0 and 66.7). The identity and similarity of these 16-mer fragments of these classical proteases (47.1-57.1 and 64.3-85.7%, respectively) is, to some extent, comparable with these parameters characterizing these proteases homology with the sequence of MLC (35.8-50 and 42.6-66.7%).

The Asp of the D-H-S triad accepts the proton of serine through the histidine residue. Only one relative short sequence of MLC containing an important for catalysis Asp67 residue was found (for example, Figure 2B). Interestingly, the homology of this MLC fragment and corresponding sequences of four canonical serine-like enzymes is comparable with that for these proteases themselves (for example, Figure 2C).

The short sequence of MLC containing Ser140 was only one demonstrating maximal identical and similar AAs to those of Ser-containing sequences, which are impotent for hydrolysis of proteins by four serine-like proteases (Figure 2D). This sequence of MLC is disposed on a distance of 72 AAs from Asp67 residue, which is average between such distances for chymotrypsinogen (66 AAs) and elastase (84 AAs).

The identity of MLC Ser-sequence with those of canonical proteinases is 54.5%, while similarity varied from 54.6 to 63.6%. The identity (70-84%) and similarity (90-100%) of Ser containing sequences of classical proteases noticeably higher.

Analysis of a possible location of MBP binding sites

In all MMPs, the AAs involved in the hydrolysis of the proteins are formed by short clusters containing 7-15 AAs, and there is data on protein-binding sites of MMPs [31-35], but not for of CAPNs. The analysis of a possible similarity of protein-binding sites of various MMPs and NGTA3-pro-DNase has been carried.

All human MMPs contain three protein sequences, which are important for protein substrates binding (type I sequence), catalysis of the hydrolysis (type II sequence), and the chelating of metal ions (type III sequence) [31-34]. Figure 3 shows these three sequences (I-II), including the central tripeptides that are most important for protein recognition, as well as the terminal AAs of these sequences (shown in bold). 15-mer sequences of the first type (I) in different MMPs have a very different level of AA's identity, which varies from 18.2 (MMP3-MMP13) to 92.3% (MMP1-MMP11). It is interesting that the identity of AAs of MMP3 with six other MMPs varies from 18.2 to 27.3%, and the similarity is only from 40 to 66.7%. Thus, protein recognition sites of different MMPs can be very different in their sequences. However, such homology is likely to be sufficient, since these MMPs usually recognize a wide variety of proteins mainly due to the interaction of the carbonyl C=O groups of metalloproteases peptide bonds with AAs of hydrolyzing proteins [31-34]. Apparently, for the complexation of protein-substrates with metalloproteases, the spatial structure of these globular molecules and the relative location in enzyme catalytic sites of various AAs responsible for the recognition is more important.

A possible homology of MLC (156 AAs) and thirteen MMPs was analyzed. Figure 3B demonstrates the data of MLC sequence comparison with that for MMP1. In the protein sequence of the MLC was found only one 15-mer 102-VRPRDQGGNQTNCGC-116 fragment, which to some extent, homologous to sequences I responsible for the binding of protein substrates by all classical metalloproteases. Similarly to MMP1 (SPFDGPGGNLAHAFQ) this MLC sequence contains central GGN-tripeptide.

Possible localization of catalytic acids of Me2+-dependent active center

Glutamic acid (Glu) of MMP8 (and other metalloproteases) binds and activates the H2O molecule, which attacks and cleavages the peptide bond of protein substrates [31-34]. Type II protein sequences containing the catalytic Glu in all Zn2+-proteases are more homologous to each other than sequences of type I. The identity of AAs of these sequences in the case of 13 MMPs varied from 50 to 100% (AV 71.3±9.7%), while similarity from 70 to 100% (AV 87.2±8.5%).

The peculiarity of the sequence II in all metalloproteases, it is Trp (W) located across a single link from the rest of Glu (Glu-x-Trp). The distance between sequences I and II containing Glu residue for MMP14 and MMP24 is eight AAs. The search for the homology of the sequences of type II of all MMPs and protein sequence of MLC (156 AAs) revealed only one fragment of the light chain (IAVEIWN) containing Glu-x-Trp with the distance between sequence I and II

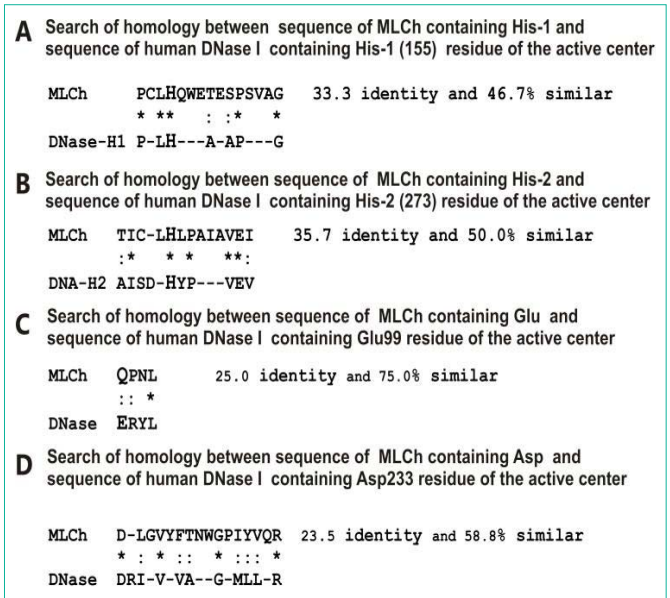

Figure 6: Search in the sequence MLC for histidine residues H1 (A) and H2 (B), as well as Glu (C) and Asp (D) corresponding to these catalytic residues in the sequences of human DNase I. The amino acids identical between two pairs of sequences are marked with an asterisk (*), while non-identical amino acids with highly and moderately conserved properties with a colon (:).

of 8 AAs, which is comparable with that for several MMPs. This fragment of MLC demonstrates homology with short sequences of all MMPs and maximal similarity with that for MMP8 (Figure 4A).

It could be assumed that the sequences containing catalytically active residues of the active centers of CAPNs and MMPs may have a noticeable level of homology. First, the analysis of the homology 12-mer fragments of nine CAPNs containing Cys was carried out. The identity of this CAPN1 sequence with those for nine other proteases ranged from 55.6-57.1 (CAPN5 and CAPN6) to 100% (CAPN2 and CAPN7); AV = 81.9±17.5%.

The analysis of possible homology of the 12-mer sequences surrounding the catalytic Glu in thirteen MMPs and catalytic Cys residues in nine CAPNs was also carried out. The identity of AAs varied from 25 to 50% (AV = 31.8±5.8%), and similarity from 33.3 to 53.8% (AV = 45.5±5.2%). This indicated that for metalloproteases of different types, the sequences surrounding catalytic AAs might have a perceptible level of homology.

The identity between CAPNs sequences containing Cys residue and a catalytic cluster of MLC containing Glu varied from 20 to 35% (AV = 27.0±6.3%), while similarity from 50 to 66.7% (AV = 59.3±4.7%) (for example, Figure 4B).

A feature of all Ca2+-proteases is that after the catalytic Cys, there is a Trp residue (for example, the sequence CAPN1 - GALGDCWLLAAI). All MMPs contain Trp residue through one AA from the Glu catalytic residue (for example, MMP1 - DEDERWT). Similar to all MMPs, the catalytic center of the MLC contains W residue (PAIAVEIWNCLC), one AA later the Glu residue. Interestingly, in spite that the MMP (Glu), CAPNs (Cys), and MLC (Glu) sites responsible for catalysis contain different AAs directly involved in catalysis, their specific sequences surrounding catalytic

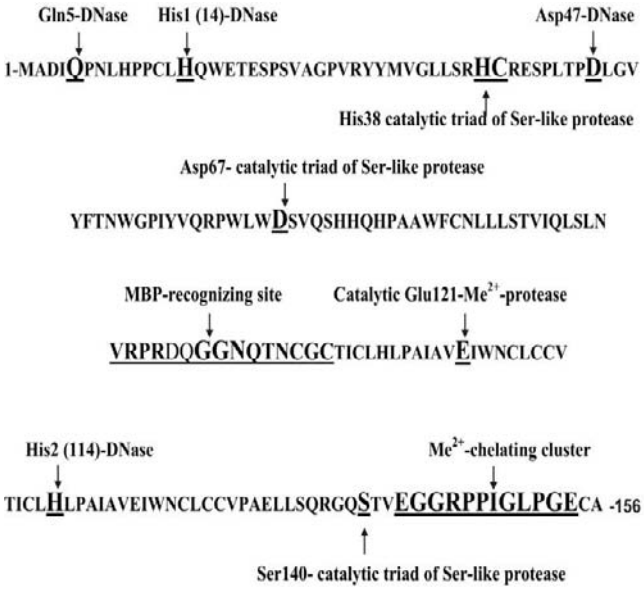

Figure 7: Hypothetical models of a possible location of AAs of a specific triad (D-H-S), which may be responsible for the serine-like activity of MLC; metal-chelating cluster (sequence III), the specific site of MBP recognition (sequence I), and Glu containing active site (sequence II) similar to those for Zn2+-dependent metalloproteases as well as catalytic residues responsible for DNase activity of MLC.

AAs have a very comparable identity and especially similarity.

Analysis of the possible location of site chelating metal ions

X-ray diffraction analysis of Zn2+-peptidases revealed the general structure of protein sequences of zinc-binding clusters: HxxxHxxxxxH (Figure 3A) [31-34]. However, MLC is not Zn2+- or Ca2+dependent enzyme; it is activated by different metal ions [29]. Therefore, it was difficult to expect that the MLC AA cluster recognizing metal ions may be very similar to those for classical Zn2+- or Ca2+-dependent metalloproteases. Considering this, in the beginning, the search for a possible sequence of type III was based on an analysis of the possible homology of any MLC sequences and clusters of thirteen MMPs chelating metal ions, as well as the ability of such clusters of the MLC to bind metal ions. Such chelating clusters of MLC could, in principle, contain from one to three of several hydrophilic AAs (Ser, Asp, Tyr, Thr, Glu, Gln, Asn, Gln, His, and Lys), which are capable of forming donor-acceptor bonds with various metal ions.

The identity of AAs of metal ion binding sites of Zn^{2+}-binding metalloproteases varies from 44.4 to 55.6%, and similarity from 55.6 to 66.7%. The sequences of Zn^{2+}-binding sites of several MMPs were homologous to different N-, while other to C-terminal fragments of MLC. However, the sequences of Zn^{2+}-binding sites of 10 out of 13 MMPs were homologous to different short fragments of the same 9-12-mer sequence of MLC: EGGRPPIGLPGE. Several examples of homology are given in Figure 5A. The identity of different length fragments of MLC and ten MMPs varied from 33.3 to 50.0% (AV = 39±3.1%) and similarity from 53.8 to 62.5% (AV = 59.1±2.2%).

The Ca^{2+}-chelating sequences of several CAPNs are known, but the number of them in one enzyme can vary from two to seven. However, some CAPNs more often contain from one to three of such sequences [36]. Interestingly, depending on the CAPNs Ca^{2+}-binding sequences may be very different in length (8-32 AAs) and AAs composition. Many of these sequences are mostly not homologous, demonstrating short homologous fragments containing only 3-4 AAs. The level of AAs identity of several Ca^{2+}-chelating sites more often is relatively low, ≤20–25.5%. Only several sequences demonstrated higher homology. For example, in the case of CAPN1 and CAPN2, the identity varied from 33 to 75%, while the similarity from 58.3 to 83.3%. This data indicated that the Ca^{2+}-binding sequences of these enzymes could be of very different length and AA composition.

Then, analysis of possible homology between the Ca2+-binding sites of CAPN1 (and CAPN2) and Zn^{2+}-chelating clusters of MMPs was carried out. Overall, the identity of the AA residues of these sequences of CAPN1 and CAPN2 with Zn^{2+} clusters of thirteen MMPs varied from 25 to 50%, and similarity from 33.3 to 53.8% (AVs = 31.8±5.8 and 45.5±5.2%, respectively). Thus, the homology between the Ca^{2+}-sites of CAPN1 and CAPN2 and Zn^{2+}-clusters of thirteen MMPs is quite comparable despite the fact that they chelate ions of different metals.

Taking into account the relatively low level of homology between the sequences of Ca^{2+}-binding sites of various CAPNs, it was difficult to expect their good homology with a universal Me^{2+}-binding cluster of MLC. Several Ca^{2+}-binding sequences of CAPN1 and CAPN2 demonstrated homology with different fragments of a long MLC sequence: 132-ELLSQRGQSTVEGGRPPIGLPGE-154 (Figure 5B). However, the same AAs of the MLC sequence, which were found in all analyzed pairs of MLC and Ca^{2+}-binding sites of two CAPNs, was the fragment EGGRPPIGLPGE. This sequence of MLC was the same, which was revealed from the comparison of MLC with Zn^{2+}-binding sites of 13 MMPs. No other MLC Me^{2+}-binding site homologous simultaneously to Zn^{2+}-clusters of MMPs and Ca^{2+}-chelating sites of CAPNs were found. Overall, the level of homology of the Me^{2+}-chelating site of MLC with clusters of CAPNs and MMPs recognizing metals ions is very comparable (Figure 5).

The sequence found by us in MLC contains 12 AAs (EGGRPPIGLPGE). Metal-binding sequences of various proteases more often contain from 8 to 12 amino acids and three AAs interacting with ions. In Zn^{2+}-peptidases, three His residues of 11-mer HxxxHxxxxxH sequence chelate metal ions. [31-34]. In MLC sequence 1-EGGRPPIGLPGE-12, Glu (E1), Arg (R3), and Glu (E12) AAs can interact with metal ions. However, it is known that it is not necessary that only Arg (R3) can participate in the chelating of metal ions. For example, the chelating site of calcium ions may include carbonyl and/or amino groups of peptide bonds of Gln, Ile, Val, and Arg residues [39]. It cannot be excluded that any of AAs of PPIGLPG fragment of 12-mer sequence of MLC is able to form bonds with metal ions due to the carbonyl (C = O) or amino group of the peptide bond (O = C-NH-).

Analysis of amino acid residues responsible for DNase activity

The homology between the protein sequences of human DNase I with this enzyme of various animals is relatively high (mouse, 77.8%; bovine, 71.6%; and rat 45.5% of the identity). Homology of MLC with human DNase I (26.7% of identity and 48.3% of similarity) was

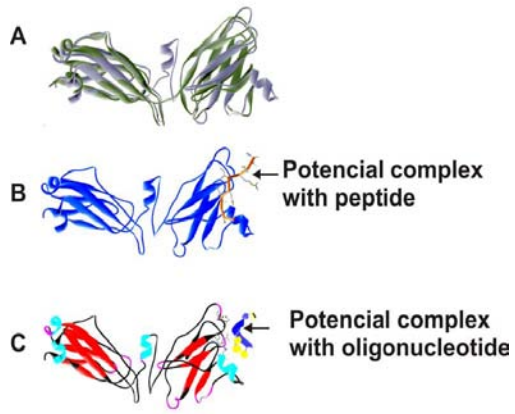

Figure 8: The comparison of spatial structures of MLC and catalytic antibody with a serine protease active site (pdb id 1eap chain a) was performed using the crystal structure of serine-like proteolytic Abs; the difference between the structures is 1.79 Å (A). A potential complex of MLC with peptide was obtained using its structural similarity with the crystal structure of an HIV-1 neutralizing antibody 50.1 in complex with its v3 loop peptide antigen (B). A potential complex of MLC with dinucleotide was found using its structural similarity with antibody 64M-2 Fab complexed with d(TpT) (C).

estimated. From the crystal structures of human DNase I complexes with oligonucleotide four important for catalysis AAs were found: Glu99, His155, Asp 233, and His273. It has been suggested that His155 (H1) and His273 (H2) forming hydrogen bonds with Glu99 and Asp233, respectively, provide common acid-base catalysis of the hydrolysis reaction [40, 41]. These AA residues are conserved for all DNases.

First, we tried to find possible sequences of MLC homologous to His155 (H1) and His273 (H2) of DNase I. Figure 6 demonstrates homology of two sites of MLC with the sequence of human DNase I containing these AAs. The MLC sequence containing H1 demonstrated 33.3% of identical and 46.7% of similar AAs with DNase I cluster containing His155 (H1) (Figure 6A). The identity and similarity of the second found site of MLC containing H2 (His114) with the DNase I cluster were 35.7 and 50%, respectively.

We tried to find a site in the protein sequence of MLC containing AA residue, which can have the role of Glu99 in the case of DNase I. Only a short fragment of MLC containing only four AA residues demonstrated homology with the DNase sequence containing Glu99 (Figure 6C). In addition, in this fragment containing Glu of DNase I was changed for Gln in MLC. But, Gln of MLC, similar to Glu of DNase I can form hydrogen bonds.

From searching for the homology of DNase sequence containing catalytically important Asp233, a longer MLC sequence (16 AAs) surrounding Asp residue was found (Figure 6D) The identity of AAs of this sequence of MLC with a similar sequence of DNase was 23.5%, while similarity – 58.8%. Overall, all four AAs of MLC can form hydrogen bonds, while two His residues can provide general acid-base catalysis. MLC may be considered as a new DNase having some structural similarities with canonical DNases I.

Computer modeling of MLC three-dimensional structure

The computer simulation was used for the analysis of the possible three-dimensional structure of MLC. A good structural similarity of MLC was found with the crystal structure of a catalytic antibody having a Ser-protease active site (PDB Id 1eap chain A). The difference (RMSD) between these structures is 1.79 Å (Figure 8A).

We have constructed a potential MLC complex with oligopeptide using its structural similarity with the crystal structure of the HIV-1 neutralizing Abs in the complex with its v3 loop oligopeptide antigen. The predicted binding site of oligopeptide was disposed of from 14 to 91 AAs of MLC.

A potential complex of MLC with dinucleotide was constructed using its structural similarity with antibody 64M-2 Fab complexed with d(TpT). The predicted binding site corresponds to 12- 85 AAs of MLC.

Thus, using computer simulation, it was confirmed that MLC could form complexes with peptides and oligonucleotides similar to the other previously described light chains of Abs, recognizing proteins and DNA.

Discussion

The data of [26-29] indicate the extreme diversity of anti-MBP and anti-DNA recombinant kappa light chains in their affinity for MBP and DNA, respectively. In contrast, canonical enzymes, anti-MBP, and anti-DNA ABZs demonstrate other very different physicochemical properties. In this article, the analysis of the possible location of the active centers in the NGTA3-pro-DNase with two different proteases and DNase activities was carried out for the first time.

NGTA3-pro-DNase is a very unusual abzyme having two alternative proteolytic (serine-like and metalloprotease) and DNase active sites. The length of four serine-like proteases, 13 various Zn^{2+}-dependent metalloproteases, and described CAPNs is remarkably or significantly greater than that of MLC. In addition, specific AAs, which are important for catalysis of protein and DNA hydrolysis by canonical enzymes, sequences responsible for substrate recognition, and AA clusters chelating metal ions, are localized in different parts throughout the length of these enzymes. Therefore, it was unexpected that all these many structural elements corresponding to three different combined active centers can be located in the relatively short variable part of MLC (150-156 AAs).

In serine-like proteases, there is a specific catalytic triad - Asp-His-Ser [37,38]. The sequence of MLC loop containing catalytic His is homologous to those of three serine-like proteases (trypsin, chymotrypsin, and elastase) was revealed (Figure 2A). Only one sequence of MLC containing Asp (D) demonstrated homology with Asp-containing sequences of canonical proteases (Figure 2B). Only the MLC sequence containing Ser140 has high homology with specific Ser-sequences of classic serine-like proteases (Figure 2C). The data of a possible location of AAs, which may be responsible for the creation of a serine-like triad (Asp-His-Ser) of MLC, are summarized using a hypothetical model (Figure 7).

Data on the binding sites of proteins by CAPNs in the literature are absent. In the protein sequence of MLC, there is only one 15-mer fragment (95-VRPRDQGGNQTNCGC-109), which homologous to sequences I responsible for the binding of proteins in the case of all classical Zn2+-dependent metalloproteases (Figure 4).

Only one MLC sequence containing Glu-x-Trp fragment (KSNELWLHH) was found to be homologous to the sequences of the active centers of thirteen MMPs. The Zn^{2+}-binding sites of MMPs and Ca^{2+}-chelating clusters are homologous to EGGRPPIGLPGE sequence of MLC (Figure 5).

The data concerning the search for AAs responsible for the serine-like activity and three types of sites important for the binding of proteins, the recognition of metal ions, and directly catalysis are summarized in Figure 7.

Using computer modeling, we found a good structural similarity of MLC with the crystal structure of catalytic antibody having a serine protease active site (Figure 8A). In addition, potential zones of MLC interacting with the oligopeptide (Figure 8B) and dinucleotide (Figure 8C) were revealed. The predicted residues of the binding sites interacting with these relatively short ligands are located in the N-terminal sequence from 12 to 91 AAs of MLC sequence. It should be noted that in the computer simulation, oligopeptide and dinucleotide were used as ligands of MLC. It should be assumed that such short ligands should not form contacts with all amino acid residues, which may be involved in the recognition by MLC of extended proteins and nucleic acids. In addition, such an analysis can probably reveal mainly only AAs involving in ligand recognition, but not in the catalysis of the reactions. It is interesting that both oligopeptide and dinucleotide are close in the complex with MLC with AA residues from 12 to 91 from N-terminus. This may indicate in favor that this zone of MLC is, to some extent, maybe universal for the recognition of proteins and DNA. It is known that in solution, many enzymes and proteins have potent conformational lability and greatly change their structure after their binding with ligands. Therefore it may be assumed that MLC can possess similar conformational lability. In order to realize their three enzymatic activities, MLC at the recognition stage can use the same sequences for binding of DNA and proteins, and then depending on the nature of the ligand, there maybe its various specific conformational changes providing more strong interaction and effective catalysis.

Acknowledgments

This research was maintained by the Russian Science Foundation 19-15-00145. We are very grateful to Paul S. and Plaquet S. (University of Texas Houston Medical School, USA) for providing us with the cDNA library of SLE patients.

References

1. Keinan E. Catalytic Antibodies. Germany, Wiley-VCH. 2005.
2. Nevinsky GA, Buneva VN. Natural catalytic antidodies-abzymes. Keinan E., editor. In: Catalytic Antibodies; Wiley VCH, Germany. 2005; 503–567.
3. Nevinsky GA. Natural catalytic antibodies in norm and in autoimmune diseases. Brenner KJ editor. In: Autoimmune diseases: symptoms, diagnosis and treatment. Nova Science Publishers Inc., New York, USA. 2010; 1-107.
4. Nevinsky GA. Natural catalytic antibodies in norm and in HIV-infected patients. Kasenga FH editor. In: Understanding HIV/AIDS Management and Care—Pandemic Approaches the 21st Century. InTech, Rijeka, Croatia. 2011; 151–192.
5. Nevinsky GA. Autoimmune processes in multiple sclerosis: production of harmful catalytic antibodies associated with significant changes in the hematopoietic stem cell differentiation and proliferation. Gonzalez-Quevedo A editor. In: Multiple Sclerosis. InTech, Rijeka, Croatia. 2016; 100–147.
6. Nevinsky GA. Catalytic antibodies in norm and systemic lupus erythematosus. Khan WA editor. In: Lupus. InTechOpen, Rijeka, Croatia. 2017; 41-101.
7. Hhachn BCh. Internal Diseases. Braunvald EE, Isselbakher KD, Petersdorf RG, Wilson DD, Martin DB, and Fauchi AS editors. Moscow; Medicine. 1996, 1–407.
8. Shuster AM, Gololobov GV, Kvashuk OA, Bogomolova AE, Smirnov IV, Gabibov AG. DNA hydrolyzing autoantibodies. Science. 1992; 256: 665-667.
9. Andrievskaya OA, Buneva VN, Baranovskii AG, Gal'vita AV, Benzo ES, Naumov VA, et al. Catalytic diversity of polyclonal RNA-hydrolyzing IgG antibodies from the sera of patients with systemic lupus erythematosus. Immunol Lett. 2002; 81: 191-198.
10. Andrievskaya OA, Buneva VN, Naumov VA, Nevinsky A. Catalytic heterogeneity of polyclonal RNA-hydrolyzing IgM from sera of patients with lupus erythematosus. Med Sci Monit. 2000; 6: 460–470.
11. Buneva VN, Andrievskaia OA, Romannikova IV, Gololobov GV, Iadav RP, Iamkovoi VI, et al. Interaction of catalytically active antibodies with oligoribonucleotides. Mol Biol (Mosk). 1994; 28: 738-743.
12. Bezuglova AM, Konenkova LP, Doronin BM, Buneva VN, Nevinsky GA. Affinity and catalytic heterogeneity and metal-dependence of polyclonal myelin basic protein-hydrolyzing IgGs from sera of patients with systemic lupus erythematosus. J Mol Recognit. 2011; 24: 960-974.
13. Bezuglova AM, Konenkova LP, Buneva VN, Nevinsky GA. IgGs Containing light chains of the lambda- and kappa- type and of all subclasses (IgG1-IgG4) from the sera of patients with systemic lupus erythematosus hydrolyze myelin basic protein. Int Immunol. 2012; 24: 759-770.
14. Bezuglova AM, Konenkova LP, Doronin BM, Buneva VN, Nevinsky GA. Affinity and catalytic heterogeneity and metal-dependence of polyclonal myelin basic protein-hydrolyzing IgGs from sera of patients with systemic lupus erythematosus. J Mol Recognit 2011; 24: 960-74.
15. Legostaeva GA, Polosukhina DI, Bezuglova AM, Doronin BM, Buneva VN, Nevinsky GA. Affinity and catalytic heterogeneity of polyclonal myelin basic protein-hydrolyzing IgGs from sera of patients with multiple sclerosis. J Cell Mol Med. 2010; 14: 699-709.
16. Savel'ev AN, Eneyskaya EV, Shabalin KA, Filatov MV, Shabalin KA. Autoantibodies with amylolytic activity. Prot Pept Lett. 1999; 6: 179-184.
17. Ivanen DR, Kulminskaya AA, Ershova NA, Eneyskaya EV, Shabalin KA, Savel'ev AN, et al. Human autoantibodies with amylolytic activity. Biologia. 2002; 11: 253-260.
18. Savel'ev AN, Kulminskaya AA, Ivanen DR, Nevinsky GA, Neustroev KN. Human antibodies with amylolytic activity. Trends Glycosci Glycotechnol. 2004; 16: 17-31.
19. Neustroev KN, Ivanen DR, Kulminskaya AA, Brumer IH, Saveliev AN, Nevinsky GA. Amylolytic activity and catalytic properties of IgM and IgG antibodies from patients with systemic lupus erythematosus. Hum Antibodies. 2003; 12: 31-34.
20. Tyutyulkova S, Gao QS, Thompson A, Rennard S, Paul S. Efficient vasoactive intestinal polypeptide hydrolyzing autoantibody light chains selected by phage display Biochim Biophys Acta. 1996; 1316: 217-223.
21. Paul S, Tramontano A, Gololobov G, Zhou YX, Taguchi H, Karle S, et al. Phosphonate ester probes for proteolytic antibodies. J Biol Chem. 2001; 276: 28314-28320.
22. Clackson T, Hoogenboom HR, Griffiths AD, Winter G. Making antibody fragments using phage display libraries. Nature. 1991; 352: 624–628.
23. McCafferty J, Fitzgerald KJ, Earnshaw J, Chiswell DJ, Link J, Smith R, et al. Selection and rapid purification of murine antibody fragments that bind a transition-state analog by phage display. Appl Biochem Biotechnol. 1994; 47: 157–173.
24. Botvinowskaya AV, Kostrikina IA, Buneva VN, Nevinsky GA. Systemic lupus erythematosus: molecular cloning of several recombinant DNase monoclonal kappa light chains with different catalytic properties. J Mol Recognit. 2013; 26: 450-460.

25. Kostrikina IA, Buneva VN, Nevinsky GA. Systemic lupus erythematosus: Molecular cloning of fourteen recombinant DNase monoclonal kappa light chains with different catalytic properties. Biochim Biophys Acta. 2014: 1840: 1725-1737.

26. Timofeeva AM, Buneva VN, Nevinsky GA. Systemic lupus erythematosus: molecular cloning and analysis of 22 individual recombinant monoclonal kappa light chains specifically hydrolyzing human myelin basic protein. J Mol Recognit. 2015; 28: 614-627.

27. Timofeeva AM, Buneva VN, Nevinsky GA. Systemic lupus erythematosus: molecular cloning and analysis of recombinant monoclonal kappa light chain NGTA1-Me-pro with two metalloprotease active centers. Mol Biosyst. 2016; 12: 3556-3566.

28. Timofeeva AM, Ivanisenko NV, Buneva VN, Nevinsky GA. Systemic lupus erythematosus: molecular cloning and analysis of recombinant monoclonal kappa light chain NGTA2-Me-pro-ChTr possessing two different activities-trypsin-like and metalloprotease. Int Immunol. 2015: 27: 633-645.

29. Bezuglova AV, Buneva VN, Nevinsky GA. Systemic lupus erythematosus monoclonal light chains of immunoglobulins against myelin basic protein possess proteolytic and DNase activities. Russ J Immunol. 2011; 5: 215-227.

30. Timofeeva AM, Nevinsky GA. Systemic lupus erythematosus: possible localization of trypsin-like and metalloprotease active centers in the protein sequence of the monoclonal light chain (NGTA2-Me-pro-Tr. Biotechnol. Appl. Biochem. 2019.

31. Solov'eva NI. Matrix metalloproteinases and their biological functions. Bioorg Khim (Mosk). 1998; 24: 245-255.

32. Sequences available at site:

33. MMP sequences.

34. Verma RP, Hansch C. Matrix metalloproteinases (MMPs): chemical-biological functions and (Q)SARs. Bioorg Med Chem. 2007; 15: 2223-2268.

35. Whittaker M, Floyd CD, Brown P, Gearing AJ. Design and therapeutic application of matrix metalloproteinase inhibitors. Chem Rev. 1999; 99: 2735-2776.

36. UniProt. Calpain sequences.

37. Polgár L. The catalytic triad of serine peptidases. Cell Mol Life Sci. 2005; 62: 2161-2172.

38. Hou H, Wang S, Zhu X, Li Q, Fan Y, Cheng D, Li B. A novel calcium-binding peptide from Antarctic krill protein hydrolysates and identification of binding sites of calcium-peptide complex. Food Chem. 2018; 243: 389-395.

39. Jones SJ, Worrall AF, Connolly BA. Site-directed mutagenesis of the catalytic residues of bovine pancreatic deoxyribonuclease I. J Mol Biol. 1996; 264: 1154-1163.

40. Gueroult M, Picot D, Abi-Ghanrm J, Hartman B, Baaden M. How cations can assist DNase I in DNA binding and hydrolysis PLoS Comput. Biol. 2010; 6: 1-11.