Određivanje epitopa autoantigena tetraspanina 7 za vezanje antitijela u dijabetesu tipa I

Kraus, Gloria

University of Zagreb

Faculty of Science

Department of Biology

Gloria Kraus

Mapping of autoantibody epitopes of the novel tetraspanin 7 autoantigen relevant in type I diabetes

Određivanje epitopa autoantigena tetraspanina 7 za vezanje antitijela u dijabetesu tipa I

Graduation Thesis

Zagreb, 2016.

This graduation thesis was conducted at the Bonifacio laboratory group of the Center for Regenerative Therapies Dresden, Cluster of Excellence, TU Dresden, under supervision of Dr. Anne Eugster and cosupervision of Dr. Petra Korać, Asst. Prof. The thesis was submitted to the evaluation of the Department of Biology at the Faculty of Science, University of Zagreb in order to obtain the title of Master of Molecular Biology (mag.biol.mol.).

I would like to thank Dr. Ezio Bonifacio, Prof., for giving me the opportunity to write my thesis in his group at the Center for Regenerative Therapies Dresden, Dr. Anne Eugster for being my supervisor and for her advices which guided me through this study, Msc. Denise Müller for being an immense help and for her patient explanations, and the entire Bonifacio group for accepting me, helping me, and advising me.

I would like to thank dr.sc. Petra Korać, Asst.Prof., for cosupervising, advising, and guiding for the science projects to come.

Finally, I would like to thank all of my friends and family just for being them.

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek Diplomski rad

ODREĐIVANJE EPITOPA AUTOANTIGENA TETRASPANINA 7 ZA VEZANJE ANTITIJELA U DIJABETESU TIPA I

Gloria Kraus

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Dijabetes mellitus tipa 1 (T1DM) je autoimuna bolest koja je specifična za gušteraču, a karakterizira ju nedostatak proizvodnje inzulina. Najčešći biomarkeri koji se koriste u dijagnozi i predviđanju nastanka bolesti su autoantitijela iz seruma za četiri najznačajnija antigena beta stanica: inzulin, GAD65, IA2 i ZnT8. Međutim, primijećeno je da još jedan protein koprecipitira s četiri spomenuta autoantigena. Istraživanja su pokazala da se radilo o glikoziliranom membranskom proteinu molekularne težine 38 kDa koji je nazvan Glima 38. Glima 38 je kasnije identificiran kao tetraspanin 7 (TSPAN7). TSPAN7 pripada proteinskoj obitelji tetraspanina te se sastoji od sedam domena: tri citoplazmatske (C1 do C3), četiri transmembranske (TM1 do TM4) i dvije ekstracelularne (E1 i E2), a u gušterači se nalazi specifično u stanicama Langerhansovih otoka. Hipoteza ovog istraživanja glasi da je imunoprecipitacijskim testom luciferaze moguće kvantificirati vezanje autoantitijela na proteinske fragmente TSPAN7. Cilj ovog istraživanje je bio utvrditi na koje domene se vežu autoantitijela kako bi se detaljnije mogli opisati autoimuni procesi u osobama koje boluju od dijabetesa tipa I. Glavna metoda korištena u ovom istraživanju je bila imunoprecipitacijski test luciferazom (LIPS), gdje su se proteinski fragmenti TSPAN7 obilježeni nanoluciferazom testirali na 41 uzorku seruma pacijenata koji su sadržavali autoantitijela na cjeloviti TSPAN7 i 20 uzoraka seruma koji nisu sadržavali autoantitijela na cjeloviti TSPAN7. Rezultati sugeriraju da se najvjerojatnije u domenama TSPAN7 E2 i C3 nalaze epitopi za vezanje autoantitijela te da domena TM4 također ima ulogu ili tako što i sama sadržava epitope ili tako što smješta domene E2 i C3 u pravilnu konformaciju.

(45 stranica, 13 slika, 5 tablica, 29 literaturnih navoda, jezik izvornika: engleski)

Rad je pohranjen u Središnjoj biološkoj knjižnici

Ključne riječi: dijabetes mellitus tipa 1, T1DM, tetraspanin 7, TSPAN7, autoantitijela, imunoprecipitacijski test luciferazom, LIPS, epitopi

Voditelj: dr. sc. Anne Eugster

Suvoditelj: doc. dr. sc. Petra Korać

Ocjenitelji: doc. dr. sc. Petra Korać, doc. dr. sc. Tomislav Ivanković, izv. prof. dr. sc. Željka Vidaković-Cifrek

Rad prihvaćen: 1. prosinca 2016.

BASIC DOCUMENTATION CARD

University of Zagreb
Faculty of Science
Department of Biology

Graduation Thesis

MAPPING OF AUTOANTIBODY EPITOPES OF THE NOVEL TETRASPANIN 7 AUTOANTIGEN RELEVANT IN TYPE I DIABETES

Gloria Kraus

Rooseveltov trg 6, 10000 Zagreb, Croatia

Type 1 diabetes mellitus (T1DM) is an organ-specific autoimmune disease, which is characterized by a deficiency in endogenous pancreatic insulin production. The typical biomarkers that predict type I diabetes are serum autoantibodies against beta-cell antigens: insulin, GAD65, IA2, and ZnT8. Along with the four known autoantigens, autoantibodies were also detected to another potential autoantigen, a 38 kDa glycosylated membrane protein known as Glima 38. Glima 38 was later identified as tetraspanin 7 (TSPAN7). TSPAN7 is a member of the tetraspanin protein family, it has seven domains: three cytoplasmic (C1 to C3), four transmembrane (TM1 to TM4), and two extracellular (E1 and E2), and in pancreas it is found specifically in the islets of Langerhans. Hypothesis of this study is that by using luciferase immunoprecipitation systems (LIPS) assay, binding of autoantibodies to TSPAN7 protein fragments can be observed and quantified. The aim of this study was to map autoepitopes of the TSPAN7 in order to provide a more detailed description of autoimmune responses in T1DM individuals. The main method was the luciferase immunoprecipitation systems (LIPS) assay, where nanoluciferase-tagged truncated TSPAN7 forms were tested against 41 serum samples that had autoantibodies against full-length TSPAN7 and 20 samples that did not have autoantibodies against full-length TSPAN7. The results suggest that E2 and C3 TSPAN7 domains are the most likely candidates that carry autoepitopes, and that TM4 could play a role in the binding as well, either by carrying autoepitopes, or by positioning E2 and C3 in the proper conformation.

(45 pages, 13 figures, 5 tables, 29 references, original in: English)

Thesis deposited in the Central Biological Library

Key words: type I diabetes mellitus, T1DM, tetraspanin 7, TSPAN7, autoantibodies, luciferase immunoprecipitation systems assay, LIPS, autoepitopes

Supervisor: Dr. Anne Eugster

Cosupervisor: Dr. Petra Korać, Asst. Prof.

Reviewers: Dr. Petra Korać, Asst. Prof., Dr. Tomislav Ivanković, Asst. Prof., Dr. Željka Vidaković-Cifrek, Assoc. Prof.

Thesis accepted: December 1st, 2016

CONTENT

1. INTRODUCTION

1.1. Type I diabetes mellitus

Type 1 diabetes mellitus (T1DM) is an organ-specific autoimmune disease, which is characterized by a deficiency in endogenous insulin production. It arises as the result of autoreactive immune cell-mediated destruction of insulin-producing beta-cells in the pancreas (Weenink and Christie, 2006). This destruction can progress over many years and ultimately results in metabolic abnormalities starting with impaired glucose tolerance and then progressing to symptomatic hyperglycemia (Taplin and Barker, 2008). The clinical symptoms start to manifest only when most beta-cells have lost function and presumably have been destroyed (Skyler, 2013). However, T1DM usually has a preclinical phase characterized by the presence of autoantibodies to antigens of the pancreatic beta-cells (Ziegler et al., 2013). The primary genetic risk for developing T1DM is the inheritance of certain major histocompatibility complex (MHC) alleles (HLA-DR3/DR4, HLA-DQ8/DQ2) (Weenink and Christie, 2006). However, there are other numerous genes that also influence the disease onset, and usually a small number of genes have large effects and a large number of genes have small effects (Michels et al., 2015). T1DM comprises the majority of cases of diabetes in childhood, and it is one of the most prevalent childhood chronic diseases (Taplin and Barker, 2008). The incidence of type I diabetes has increased worldwide, especially in children and in developed countries (Achenbach et al., 2005).

1.2. Prediction and prevention of type I diabetes

1.2.1. Genetic factor – human leukocyte antigen (HLA)

The human leukocyte antigen (HLA) gene region is located on the short arm of the chromosome 6 (6p21) and it is the most important of the multiple gene loci that affects susceptibility to type I diabetes. A number of alleles of class II, DR and DQ HLA genes, are particularly important for the susceptibility (Lipponen et al., 2010) (Figure 1). HLA genes are co-dominant and follow a Mendelian form of transmission in families and, therefore, both alleles are expressed at a given HLA locus. The HLA antigens comprise two molecular classes of cell surface glycoproteins differing in structure, function and tissue distribution. There is a strong connection between HLA haplotypes and T1DM, considering the fact that more than 90% of T1DM patients carry either HLA DR3-DQ2 or DR4-DQ8 haplotypes, a group of genes inherited together from a single parent. Children carrying high-risk HLA

genotypes have a higher risk for early disease onset and more frequent development of islet autoantibodies in infancy. The risk of developing islet autoantibodies by age 2 years is 20% in individuals with the high-risk HLA DR3-DQ2/DR4-DQ8 or HLA DR4-DQ8/DR4-DQ genotypes compared with 2.7% in offspring without these genotypes. Overall, 50% of islet autoantibody-positive individuals have at least one of these genotypes (Achenbach et al., 2005) (Figure 2). In cases of adult-onset type I diabetes genes have smaller impact, as shown by patients with disease onset after 20 years of age have lower frequencies of the high-risk HLA DR and DQ haplotypes (Bonifacio, 2015). Interestingly, the HLA allele DQB1*0602 confers dominant protection of developing T1DM (Michels et al., 2015).

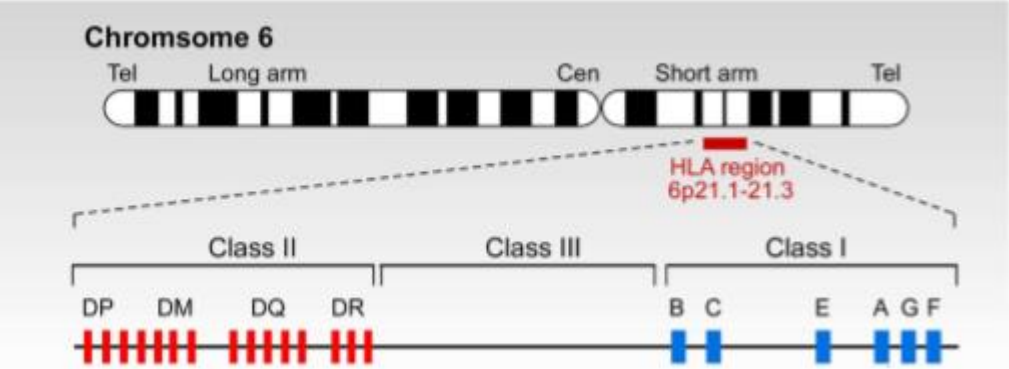


Figure 1. The human leukocyte antigen (HLA) gene region located on the short arm of the chromosome 6 (6p21). It bears two classes of HLA genes (Class I and Class II), where the Class III is located in the same region, but it does not belong to the HLA genes. Class II genes are most important for susceptibility to T1DM (Zhang et al., 2014).

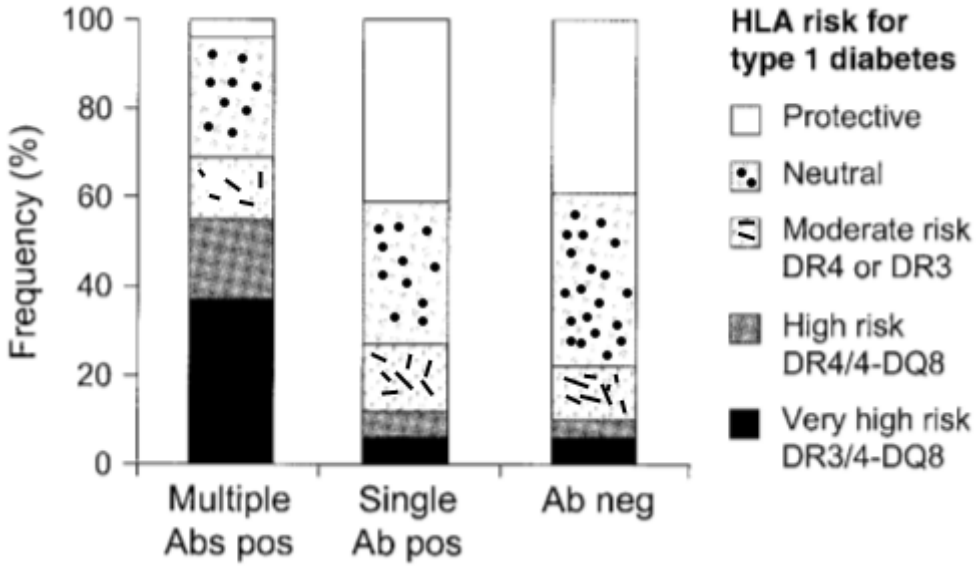


Figure 2. Frequency of the human leukocyte antigen (HLA) genotypes. Children who develop multiple islet autoantibodies (Multiple Abs pos) have HLA genotypes that are found in T1DM. Nearly 60% of Multiple Abs pos children have either very high risk DR3/4-DQ8 or high risk DR4/4-DQ8 HLA genotypes, whereas children who developed single autoantibodies (Single Ab pos) have HLA genotypes similar to children who are autoantibody negative (Ab neg). Approximately 30% of Single AB pos and Ab neg children have very high risk or high risk HLA genotypes (Achenbach et al., 2005).

1.2.2. Autoantibodies to beta-cell antigens

The typical biomarkers that predict type I diabetes are serum autoantibodies against beta-cell antigens (Bonifacio, 2015). There are four major humoral autoantigens that have been identified in type 1 diabetes by defining the specificity of autoantibodies in the disease: pre-pro-insulin, GAD65 (glutamate decarboxylase), IA2 (receptor-type tyrosine-protein phosphatase-like N), and ZnT8 (zinc transporter-8) (McLaughlin et al., 2016). Detection of multiple islet autoantibodies in children who are genetically at risk marks a preclinical stage of type I diabetes, and therefore, it can be said that development of multiple islet autoantibodies in children predicts type I diabetes (Ziegler et al., 2013). In genetically at-risk individuals, the islet autoantibody seroconversion is greatest between 9 months and 2 years of age and occurs earliest for pre-pro-insulin autoantibodies (Ziegler et Bonifacio, 2012). Ziegler et al. (2013) reported that progression to type I diabetes 10-year after islet seroconversion in

children with multiple islet autoantibodies was 69.7%, whereas in children with a single islet autoantibody it was 14.5%. Overall, over 80% of children with multiple islet autoantibodies progress to symptomatic, insulin-requiring diabetes within 15 years (Giannopolou et al., 2015). Diabetes risk by 15 years of age was 0.4% in children with no autoantibodies, 12.7% in children with a single islet autoantibody, 61.6% in children with two islet autoantibodies, and 79.1% in children with three islet autoantibodies (Ziegler et al., 2013) (Figure 3). The same study reports that variation in progression time is associated with the age of seroconversion (progression is faster for children who had islet autoantibody conversion younger than 3 years), genetic markers (progression is faster for children with the HLA genotype DR3/DR4-DQ8), sex (progression seems to be faster for girls), and the type of islet autoantibody. Further, autoantibody titer is another contributor. Achenbach et al. (2004) reported that genetically at-risk individuals have significantly higher diabetes risk if they have high-titer islet cell antibodies. However, the situation is somewhat different in adults. In cases of adult-onset T1DM, number of islet autoantibodies is less than in cases when T1DM develops in childhood, and many patients only present GADAs (Bonifacio, 2015). Autoantibodies tend to develop sequentially rather than simultaneously (Taplin and Baker, 2008). Autoantibodies to insulin (IAA) tend to appear first, followed by autoantibodies to GAD65 (GADA) and IA2. GADAs and IAAs, when both present, are the most frequent antibodies in childhood, while GADA itself is the hallmark of adult-onset type I diabetes (Bonifacio, 2015). Since IA2 autoantibodies are associated with a higher risk than GADAs or IAAs in single autoantibody patients (Ziegler et al., 2013), detection of IA2 autoantibodies alone during childhood is very specific for the development of T1DM (Bonifacio, 2015). Detection of autoantibodies is important because T1DM risk stratification models that include autoantibody characteristics may improve T1DM stratification and risk assessment (Achenbach et al., 2004).

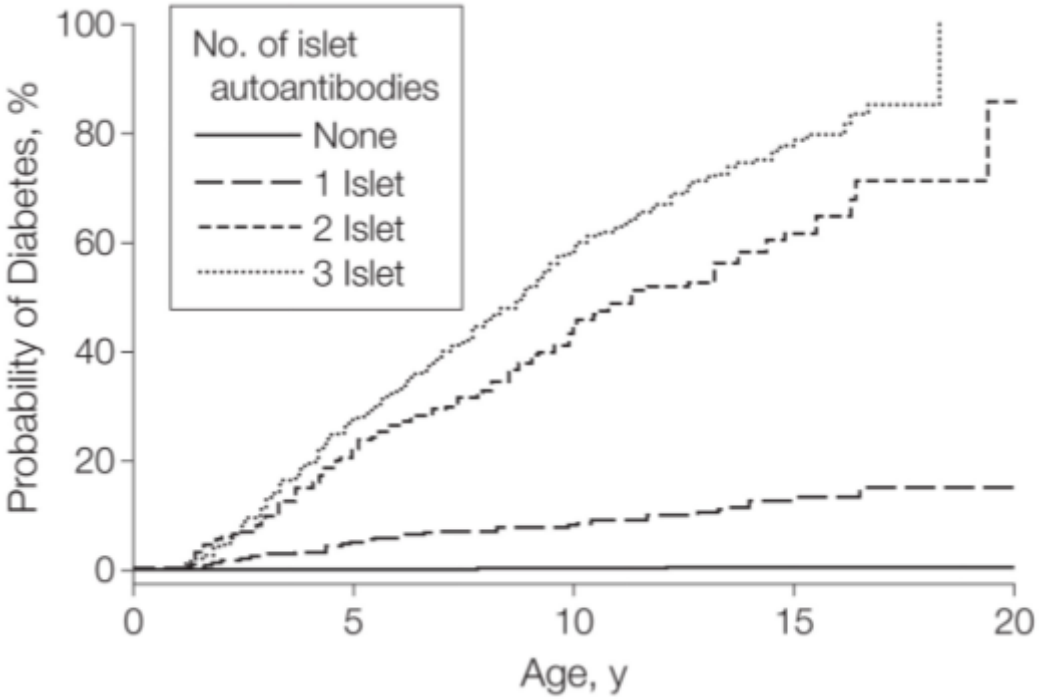


Figure 3. Development of T1DM in children based on the number of islet autoantibodies. Diabetes risk by 15 years of age was 0.4% in children with no autoantibodies, 12.7% in children with a single islet autoantibody, 61.6% in children with two islet autoantibodies, and 79.1% in children with three islet autoantibodies (Ziegler et al., 2013).

1.2.2.1. Discovery of another autoantigen

Along with the four known autoantigens, during several years and by several groups autoantibodies were also detected to a fifth potential autoantigen, a 38 kDa glycosylated membrane protein known as Glima 38 (Aanstoot et al., 1996; Roll et al., 2000; Winnock et al., 2001). It was only recently that the true identity of Glima 38 was revealed. Mass spectrometry analysis showed that only tetraspanin 7 matched the known properties of Glima 38 (McLaughlin et al., 2016).

1.2.3. Time points for intervention

There are three possible time points for intervention in T1DM: prior to any evidence of autoimmunity (primary), after the development of islet autoantibodies (secondary), and shortly after onset of the disease (tertiary). The intervention involves dietary intervention that aims to interrupt putative environmental triggers of T1DM, antigen-specific therapies that aim

to control the autoimmune response, and other immunotherapies. The goal of T1DM detection in primary and secondary time points, before the disease onset, is to arrest the immune process by modulating the autoimmune response, and possibly prevent or delay clinical disease (Skyler, 2013). Secondary time point is characterized by usage of various biomarkers, such as genetic factors and antibodies to beta-cell antigens. The usage of biomarkers should provide benefits that span from learning about the disease process to preventing complications such as diabetic ketoacidosis at the diagnosis of diabetes, as well as possible prevention of diabetes entirely. However, it should be noted that biomarkers change with age, and predicting T1DM in adults is different to its prediction in children (Bonifacio, 2015).

1.3. Tetraspanin protein family

Tetraspanins are small integral membrane proteins, of approximately 30 kDa (Skaar et al., 2015), and they protrude 4-5 nm above the membrane (Hemler, 2005). They share four characteristic transmembrane domains (TM1 to TM4), and mainly have three cytoplasmic and two extracellular domains (one short – E1, and one long – E2) (Hemler, 2005) (Figure 4). E2 is divided into constant and variable regions, where constant region contains structurally conserved A, B and E alpha-helices, which provide a putative dimerization interface. There is a region of hypervariability between B and E helices, which demonstrates conformational fluctuations. Within the variable region reside critical protein-protein interaction sites. TM1, TM3 and TM4 typically contain polar residues (asparagine, glutamine, and glutamic acid) (Hemler, 2003), which are highly conserved between the tetraspanin family (Maecker et al., 1997), and are believed to be involved in stabilizing the overall tetraspanin structure. Transmembrane domain interactions also stabilize the conformation of the E2 domain (Hemler, 2003). E1 domain fits into a groove of E2 and transmembrane domains form a coil-coiled structure, which is stabilized by hydrogen bonds involving the polar residues (Charrin et al., 2014). Unlike the transmembrane domains, extracellular domains show greater sequence divergence. There are, however, three cysteines in E2, which are located at defined distances from the TM regions in the majority of tetraspanin family members. Two of the mentioned cysteines are located in a conserved CCG motif about 50 amino acids after TM3. The third cysteine is often preceded by a glycine and is fixed as 11th amino acids upstream of TM4 (Maecker et al., 1997). The two latter cysteine residues contribute to two crucial disulphide bonds within the E2 (Hemler, 2005). Many tetraspanins are posttranslationally modified by N-linked glycosylation or palmitoylation (Skaar et al., 2015). They are involved in regulating cell morphology, motility, tissue invasion, fusion and signaling, in a number of

tissues as for example in the brain and the immune system, but they also play an important role in tumors (Hemler, 2005). Many members of the tetraspanin family associate specifically and directly with other proteins, and also with other tetraspanins, thereby generating a cascade of interactions (Charrin et al., 2014). The association of tetraspanin with the integrin family allows tetraspanins to regulate integrin-dependent cell migration and adhesion (Berditchevski, 2001). The tetraspanins could be described as "molecular facilitators" that bring together large molecular complexes and allow them, through stabilization, to function more efficiently (Maecker et al., 1997). Individual tetraspanin proteins are often expressed at 30 000 – 100 000 or more copies per cell (Hemler, 2003). There are 33 members of the tetraspanin family expressed in mammals, 37 in *Drosophila melanogaster* and 20 in *Caenorhabditis elegans*, as well as tetraspanin-like proteins expressed in plants (Charrin et al., 2014).

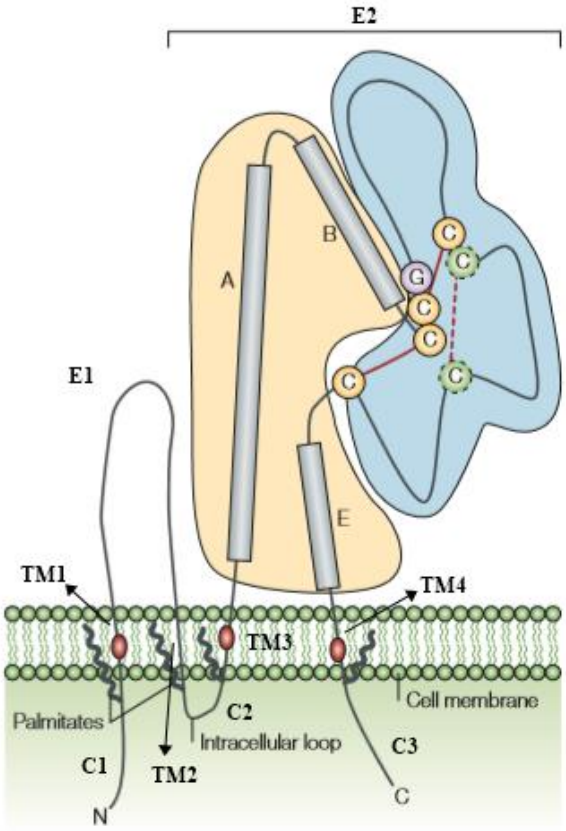


Figure 4. Structural features of tetraspanins. Tetraspanins contain four transmembrane domains, two extracellular loops (E1 and E2) with a very short intracellular loop (C2) flanked by relatively short N-terminal and C-terminal cytoplasmic tails (of approx. 8–21 amino acids) (C1 and C3, respectively). The E2 is subdivided into a constant region (yellow, containing α-helices A, B and E), and a variable region (blue). All tetraspanins contain a CCG motif after the B helix, and two other conserved cysteines (yellow), which are arranged to form two intramolecular disulphide bonds (red lines). Nearly all tetraspanins also contain membrane proximal cysteines that undergo palmitoylation (Hemler, 2005).

1.4. Tetraspanin 7

Tetraspanin 7 (TSPAN7, A15, MXS1, CD231, MRX58, CCG-B7, TM4SF2, Glima 38) (Skaar et al., 2015) is one member of the tetraspanin family. The human TSPAN7 gene is located on the short arm of X-chromosome at position Xp11.4 (Kent et al., 2002). When it was called Glima 38 protein, one study, suggested it to be membrane bound based on its insolubility and fractionation with the detergent phase (Aanstoot et al., 1996). TSPAN7 is posttranslationally modified by N-glycosylation on its 22 kDa protein core, and it is a N-

linked glycoprotein of the biantennary complex-type (Aanstoot et al., 1996; Roll et al., 2000). N-linked carbohydrates contribute to more than about 40% of the relative molecular weight of TSPAN7. TSPAN7 is likely to contain a minimum of five carbohydrate chains and it is, therefore, heavily glycated (Roll et al., 2000). Roll et al. (2000) reported that within 1 hour TSPAN7 is fully glycated *in vivo*. The same study reported the presence of terminal sialic acids in the carbohydrates of TSPAN7, as well as that it does not contain any O-linked carbohydrates. Carbohydrate modifications make TSPAN7 resistant to proteolysis (Roll et al., 2000). Autoantibodies to TSPAN7 recognize both, the glycated and the non-glycated forms, suggesting that the 22 kDa protein core, not the carbohydrates, is the part carrying autoepitopes. Linearized form of TSPAN7 is not recognized by autoantibodies, which suggests that TSPAN7 autoepitopes might be conformational rather than linear (Roll et al., 2000).

1.4.1. Tetraspanin 7 and T1DM

In the pancreas, TSPAN7 is found specifically in the islets of Langerhans (McLaughlin et al., 2016). Autoantibodies to TSPAN7 have been detected in 19-38% of Type I diabetic patients, with significantly higher prevalence (up to 50%) in children (McLaughlin et al., 2016). They are almost exclusively found in new-onset children patients with other beta-cell autoantibodies, and they seem to appear late in the preclinical phase of type I diabetes (Walther et al., 2016). They mark the preclinical of type I diabetes, but they do not provide much additional information for the prediction or classification of type I diabetes because they do not add to the predictive power of other autoantigens (Winnock et al., 2001). However, since TSPAN7 autoantibodies were not analyzed in adult onset diabetes, it is possible that autoantibodies to TSPAN7 may be useful in diagnosis of adult autoimmune diabetes (Walther et al., 2016). The same study reported that some autoantibodies might bind to epitopes within the external domains of the protein, and such antibodies could directly mediate beta cell death or affect beta cell function. Further characterization of TSPAN7 could provide a more detailed description of autoimmune responses that develop in individuals with T1DM, which would be necessary for guiding the selection of autoantigen-specific immunotherapy (McLaughlin et al., 2016). Hypothesis of this study is that by using luciferase immunoprecipitation systems (LIPS) assay, binding of autoantibodies to TSPAN7 protein fragments can be observed and quantified. The aim of this study is to map TSPAN7 autoantibody epitopes (autoepitopes).

2. MATERIALS AND METHODS

2.1. Materials

For the research, 61 serum samples from patients either diagnosed with T1DM, or that are first degree relatives of T1DM patients were used. From those, 41 samples were positive for antibodies against full length TSPAN7, and 20 were negative and used as negative control. The full length TSPAN7 positive samples were selected from the DiMelli cohort for new onset T1DM patients (Thümer et al., 2010) and negative from the TeenDiab cohorts of first degree relatives of T1DM patients (Ziegler et al., 2012A; Ziegler et al., 2012B). The ethical committees of Bavaria or the Ludwig-Maximillians University approved the studies, which were carried out in accordance with the Declaration of Helsinki. Informed, written consent was obtained from patients or parents of participants. For TSPAN7 protein fragment production and isolation, human embryonic kidney (HEK) 293T cell line was used, which originated from a fetus, and the cells have a very complex karyotype.

2.2. Methods

2.2.1. Molecular cloning – vector design for the production of truncated tetraspanin 7

2.2.1.1. Polymerase chain reaction and purification

In order to produce truncated forms of TSPAN7, the first step was to perform PCR by using the source of full-length TSPAN7 as a template to obtain TSPAN7 DNA that had different TSPAN7 domains. By using specific 5' and 3' primers with restriction sites for molecular cloning (Table 1), the following DNA fragments of TSPAN7 were obtained: C1-TM1-E1-TM2 (225 bp), corresponding to amino acids from 1 to 75, C1-TM1-E1 (168 bp), corresponding to amino acids 1 to 56, TM1-E1-TM2-C2 (207 bp), corresponding to amino acids 17 to 86, E1-TM1-C2 (135 bp), corresponding to amino acids 41 to 86, TM2-C2-TM3-E2 (471 bp), corresponding to amino acids 57 to 213, C2-TM3-E2 (414 bp), corresponding to amino acids 76 to 213, TM3-E2-TM4-C3 (495 bp), corresponding to amino acids 87 to 249, E2-TM4-C3 (437 bp), corresponding to amino acids 113 to 249, and TM4-C3 (111 bp), corresponding to amino acids 214 to 249 (Figure 5). As a template for the PCR reactions, pCMV6-AC-TSPAN7-IRES-GFP-Puro (c = 87.7 ng/µl) containing the full length sequence of TSPAN7 was used. For the PCR approx. 100 ng of the template, (or 1 µl of DEPC-treated H_2O (Thermo Scientific) for the negative control), 10 µl of 5x PrimeSTAR Buffer (Mg2+ plus) (Takara Bio), 4 µl of 10 mM dNTPs (Takara Bio), 1 µl of 10 µM 5' – primer, 1 µl of 10

µM 3' – primer, 0.5 µl of PrimeSTAR HS DNA Polymerase (Takara Bio), and 32.5 µl of DEPC-treated H_2O (Thermo Scientific) were mixed resulting in a final reaction volume of 50 µl. The PCR cycling protocol was the following: 30 cycles of denaturation at 98 °C for 10 s, annealing at 60 °C for 5 s, and extension at 72 °C for 1 min/kb, depending on the fragment size (Table 1), followed by 10 min final extension at 72 °C and storage at 4°C. 3 µl of the PCR reaction supplied with 6x DNA loading dye (Thermo Scientific) was analyzed on a 2% agarose gel along with 5 µl of 100 bp Gene Ruler (Thermo Scientific) as size control. After gel validation, PCR fragments were purified by using the QIAquick® PCR Purification Kit (Qiagen). After PCR purification, the presence and quantity of the fragments were analyzed on 2% agarose gels.

Table 1. The list of PCR primers with BamHI and NotI restriction sites for cloning truncated forms of TSPAN7 into the pCMV6-AC-Nluc-IRES-GFP-Puro vector.

TSPAN7 fragment	Primer name (restriction site)	Primer sequence	DNA fragment size (bp)
C1-TM1-E1-TM2 (225)	5' - TSPAN7_C1 (BamHI)	5' - ggcgtGGATCCATGGCATCGAGGAGAATGGA	225
	3' - TSPAN7_TM2 (NotI)	5' - gatcaacGCGGCCGCTTATTATTAAAACAGGCCAAAGACAACAA	
C1-TM1-E1 (168)	5' - TSPAN7_C1 (BamHI)	5' - ggcgtGGATCCATGGCATCGAGGAGAATGGA	168
	3' - TSPAN_E1 (NotI)	5' - gatcaacGCGGCCGCTTATTATTTGTGGAGTTCTCGGCAATAA	
TM1-E1-TM2-C2 (207)	5' - TSPAN7_TM1 (BamHI)	5' - ggcgtGGATCCACCCTCCTCATCATCTACTC	207
	3' - TSPAN7_C2 (NotI)	5' - gatcaacGCGGCCGCTTATTATTATGGGCTACCACGACATGTAGCAAAG	
E1-TM2-C2 (135)	5' - TSPAN7_E1 (BamHI)	5' - ggcgtGGATCCAAACTTACTCTGGGCACCTA	135
	3' - TSPAN7_C2 (NotI)	5' - gatcaacGCGGCCGCTTATTATTATGGGCTACCACGACATGTAGCAAAG	
TM2-C2-TM3-E2 (471)	5' - TSPAN7_TM2 (BamHI)	5' - ggcgtGGATCCAATGCTCCCTATGTGCTCAT	471
	3' - TSPAN7_E2 (NotI)	5' - gatcaacGCGGCCGCTTATTATTACATGTTAGTCTCCATGAAAC	
C2-TM3-E2 (414)	5' - TSPAN7_C2 (BamHI)	5' - ggcgtGAATTCGGATGCTTTGCTACATGTCGTGGTA	414
	3' - TSPAN7_E2 (NotI)	5' - gatcaacGCGGCCGCTTATTATTACATGTTAGTCTCCATGAAAC	
TM3-E2-TM4-C3 (495)	5' - TSPAN7_TM3 (BamHI)	5' - ggcgtGGATCCTGGATGCTGAAACTGTATGC	495
	3' - TSPAN7_C3 (NotI)	5' - gatcaacGCGGCCGCTTATTATTACACCATCTCATACTGAT	
E2-TM4-C3 (437)	5' - TSPAN_E2 (BamHI)	5' - ggcgtGGATCCCGTCATGAGATCAAGGACAC	437
	3' - TSPAN7_C3 (NotI)	5' - gatcaacGCGGCCGCTTATTATTACACCATCTCATACTGAT	
TM4-C3 (111)	5' - TSPAN_TM4 (BamHI)	5' - ggcgtGGATCCGGAATCATCGCTGGAGTGGC	111
	3' - TSPAN7_C3 (NotI)	5' - gatcaacGCGGCCGCTTATTATTACACCATCTCATACTGAT	

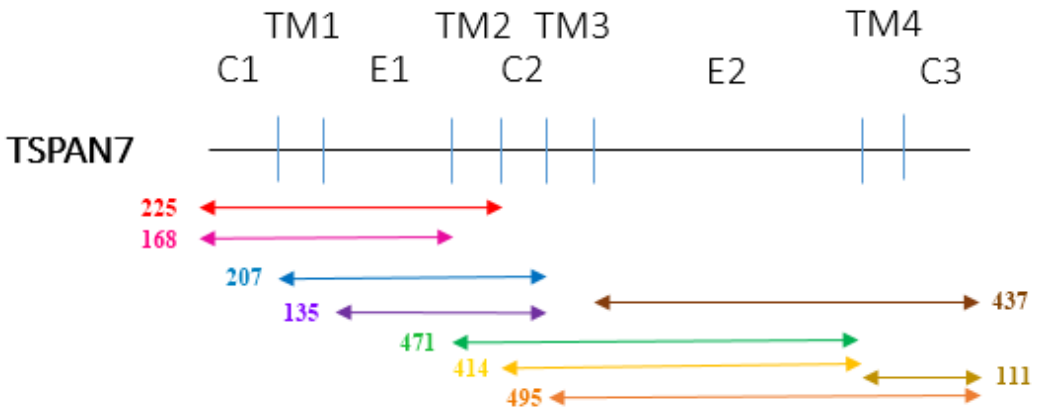


Figure 5. Truncated forms of TSPAN7. C1-TM1-E1-TM2 (225, red), AA 1 – 75; C1-TM1-E1 (168, pink), AA 1 – 56; TM1-E1-TM2-C2 (207, blue), AA 17 – 86; E1-TM1-C2 (135, purple), AA 41 – 86; TM2-C2-TM3-E2 (471, green), AA 57 – 213; C2-TM3-E2 (414, yellow), AA 76 – 213; TM3-E2-TM4-C3 (495, orange), AA 87 – 249; E2-TM4-C3 (437, brown), AA 113 – 249; TM4-C3 (111, gold), AA 214 – 249.

2.2.1.2. Restriction digestion and purification

All TSPAN7 DNA fragments obtained by PCR and the pCMV6-AC-Nluc-IRES-GFP-Puro vector (Figure 6) were digested with restriction enzymes BamHI and NotI in order to be suitable for ligation. Since the BamHI and NotI enzymes do not exert complete activity in the same buffer, the DNA was first digested with the BamHI enzyme. For each reaction, approximately 1 µg TSPAN7 fragment and approximately 1 µg vector DNA were used, 2 µl of BamHI enzyme (Thermo Scientific), 4 µl of 10x BamHI buffer (Thermo Scientific) and distilled H_2O (dH_2O) up to 40 µl. The reaction was incubated for 1.5 h at 37 °C. Afterwards, the DNA was precipitated using 60 µl of dH_2O and 70 µl of isopropanol, centrifuged 6 min at 13000 rpm, washed by adding 100 µl of 70% ethanol, centrifugated 5 min at 13000 rpm, and finally dried in an Eppendorf Vacufuge® Plus vacuum concentrator (Eppendorf) for 20 min at 45 °C. The DNA was then resuspended in 2 µl of NotI enzyme (Thermo Scientific), 4 µl of 10x buffer O (Thermo Scientific) and 34 µl of dH_2O. The reaction was again incubated for 1.5 h at 37 °C. After incubation the TSPAN7 fragment DNA was precipitated and dried as described before and resuspended in 15 µl of dH_2O, while the vector DNA was further dephosphorylated by adding 1 µl of Fast Alkaline Phosphatase (FastAP) enzyme (Thermo Scientific) into the reaction mix. The reaction was incubated first for 15 min at 37 °C and the enzyme was heat inactivated for 5 min at 75 °C. The vector DNA (pCMV6-AC-Nluc-IRES-GFP-Puro) was then purified from an 0.7% agarose gel by using the MinElute® Gel

Extraction Kit (Qiagen). Both the precipitated TSPAN7 fragment and the purified vector DNA were analyzed on the 2% agarose gel.

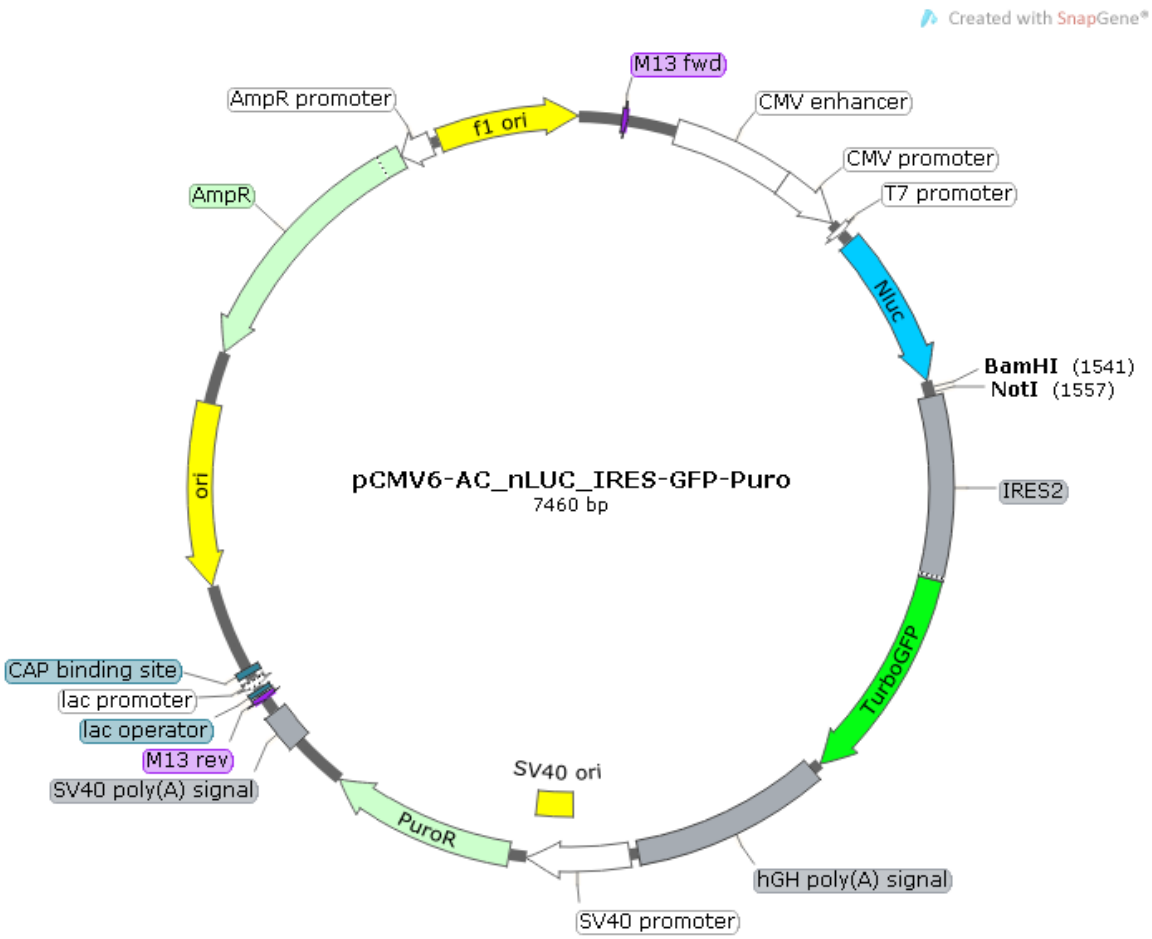


Figure 6. pCMV6-AC-Nluc-IRES-GFP-Puro vector used as destination vector for cloning TSPAN7 fragments. pCMV6 is the promotor, AC is the ampicillin resistance cassette, Nluc is the nanoluciferase tag, IRES is an internal ribosome entry site used for independent GFP synthesis, GFP is green fluorescent protein and Puro is puromycin resistance cassette.

2.2.1.3. Ligation and electroporation

To ligate the digested PCR and vector fragments, I used 100 ng of the vector DNA, 6 ng of the TSPAN7 fragment DNA (fragments 168, 135, and 111), or 9 ng of the TSPAN7 fragment DNA (fragments 225 and 207), or 18 ng of the TSPAN7 fragment DNA (fragments 414, 437, 471, and 495), 1 µl of 10x T4 DNA ligase buffer (Thermo Scientific), 1 µl of T4 DNA ligase (Thermo Scientific), and dH_2O up to 10 µl. The reaction was then incubated for 2 h at room temperature (RT). Afterwards 5 µl of each sample were dialyzed against water through EMD Millipore MF-Millipore™ nitrocellulose membrane filters, (0.025 µm pore size) (Fisher

Scientific) for 15 min at RT. During that time, 900 µl of SOC medium was added to a 1.5 ml microtube for each sample, and pre-warmed the microtubes at 37 °C. For each sample, 30 µl of *Escherichia coli* GB2005 electro competent cells were placed on ice to thaw. After thawing, down the cells were shortly span down. When the dialysis was complete, DNA was added to the *E.coli* cells. The complete bacterial sample was resuspended, transferred to pre-cooled cuvette, and placed into the Eporator (Eppendorf). After electroporation at 1150V, 100 µl of the pre-warmed SOC medium was added to the cuvette and resuspended. The reaction mix from the cuvette was put into a 1.5 ml microtube with the rest of the remained pre-warmed SOC medium. The samples were placed in Thermomixer comfort (Eppendorf) to incubate for 1 h at 37 °C at 900 rpm. After incubation, the samples were centrifuged 3 min at 6000 rpm. Afterwards, 800 µl of supernatant was removed and the rest of the pellet was resuspended in the remaining supernatant by pipetting. The complete cell suspension was plated on agar plates containing ampicillin (Amp^+) and incubated overnight at 37 °C in the CO_2-Incubator CB-210 (Binder, Tuttlingen, Germany).

2.2.1.4. Minipreparation of plasmid DNA isolation

The following day colonies were picked (10 colonies per one fragment sample if possible), and each colony was placed into a 2 ml microtube containing 1 ml of LB Amp^+ medium. The microtubes were incubated at 37 °C and 900 rpm for approximately 8 h. After incubation, 100 µl of the bacterial culture of each colony was transferred into 1.5 ml microtubes and stored them at 4 °C. The remaining 900 µl were centrifuged for 2 min at 6000 rpm at RT. The supernatant was drained and the pellet was resuspended in 200 µl of the P1 resuspension buffer containing RNAse A (Qiagen) by vortexing for approximately 5 minutes. 200 µl of the P2 lysis buffer (Qiagen) was added, and the samples were mixed by inverting the microtubes and left at RT for 5 min. 200 µl of the P3 neutralization buffer (Qiagen) was added the suspension was mixed by inverting microtubes. The samples were centrifuged for 10 min at 13000 rpm at RT. After centrifugation, the supernatant was added to 1.5 ml microtubes prelaid with 500 µl isopropanol. The samples were centrifuged for 5 min at 13000 rpm at RT. The supernatant was drained, and the pellet was washed with 300 µl of 70% ethanol. The samples were again centrifuged for 5 min at 13000 rpm at RT. The supernatant was again drained, and the microtubes were then placed in the Eppendorf Vacufuge® Plus vacuum concentrator (Eppendorf) for 20 min at 45 °C. After the pellets were dry, they were resuspended in 30 µl of dH_2O. The isolated plasmid DNA containing the appropriate TSPAN7 fragments was then digested with restriction enzymes listed in Table 2., and checked

on 2% agarose gels. An aliquot of the samples containing the correct construct were sent for sequencing to the Max Planck Institute of Molecular Cell Biology and Genetics and the rest stored at 4 °C. If the correct plasmid DNA sequence was confirmed by sequencing, the plasmid DNA was purified for the use in the cell culture. 60 µl of each culture containing appropriate plasmid DNA was put it into a 250 ml flask containing 100 ml of LB Amp^+ medium. The suspension was incubated overnight at 37 °C and at 200 rpm. The subsequent steps of the purification followed the EndoFree® Plasmid Maxi Kit protocol (Qiagen). The concentrations of each sample were measured with the NanoPhotometer Implen P-Class (Implen, Munich, Germany) and are given in the Table 3.

Table 2. Clones analyses after the plasmid DNA isolation from *E.coli* GB2005. The list of restriction enzymes for each plasmid DNA, used buffers in final concentrations, and size (bp) of the expected fragments on the agarose gel after digestion.

TSPAN7 fragment	Restriction enzymes	Buffer	Fragment size after digestion (bp)
C1-TM1-E1-TM2 (225)	BamHI	2x Tango (Thermo Scientific)	7445
	NotI		239
C1-TM1-E1 (168)	BamHI	2x Tango (Thermo Scientific)	7445
	NotI		183
TM1-E1-TM2-C2 (207)	BamHI	2x Tango (Thermo Scientific)	6966
	NotI		223
E1-TM1-C2 (135)	BamHI	2x Tango (Thermo Scientific)	7445
	NotI		150
TM2-C2-TM3-E2 (471)	SacI	1x SacI buffer (Thermo Scientific)	6903
			760
C2-TM3-E2 (414)	SacI	1x SacI buffer (Thermo Scientific)	6892
			703
TM3-E2-TM4-C3 (495)	PvuII	1x Buffer G (Thermo Scientific)	3728
			2514
			1714
E2-TM4-C3 (437)	PvuII	1x Buffer G (Thermo Scientific)	3723
			2513
			1633
TM4-C3 (111)	BamHI	2x Tango (Thermo Scientific)	7445
	NotI		126

Table 3. Concentrations (ng/µl) of plasmid DNA for each TSPAN7 fragment after EndoFree® Plasmid Kit (Qiagen) isolation.

TSPAN7 fragment	Concentration (ng/µl)
C1-TM1-E1-TM2 (225)	1326
C1-TM1-E1 (168)	167.0
TM1-E1-TM2-C2 (207)	990.0
E1-TM1-C2 (135)	1093
TM2-C2-TM3-E2 (471)	659.0
C2-TM3-E2 (414)	1037
TM3-E2-TM4-C3 (495)	46.60
E2-TM4-C3 (437)	811.0
TM4-C3 (111)	331.0

2.2.2. Transfection of HEK 293T cells

2.2.2.1. Day 1 – seeding cells

Prior to each usage of the HEK cell medium (DMEM 1g/L glucose without L-glutamine (Lonza, Basel, Switzerland), 1:100 Penicillin-Streptomycin (Gibco), 2mM L-glutamine (Lonza, Basel, Switzerland), 10% heat inactivated and filtered Fetal Bovine Serum ultra-low IgG (Gibco)), it needed to be pre-warmed to 37 °C in a waterbath. To prepare HEK 293T cells for transfection, cells had to be detached from the surface by adding 1 ml of trypsin 0.25% EDTA (Gibco) to the plate. Afterwards, 4 ml of HEK cell medium was added and the clumps were resuspended by pipetting. The suspension was transferred to a 15 ml Falcon tube. 5 ml of HEK cell medium was added again, and the suspension was transferred to the same Falcon tube. The cells were centrifuged for 10 min at 300*g* at RT to remove the trypsin. The supernatant was aspired, the pellet was resuspended in 1 ml of HEK cell medium, and filled up to 10 ml with the HEK cell medium. The cell suspension was mixed by inverting the Falcon tube and 10 µl of the cell suspension was taken out for counting. The 10 µl of the cell suspension was mixed with 10 µl of 0.4% trypan blue dye suspension (Gibco), and 10 µl of that suspension was placed into a counting chamber ZK06 (A. Hartenstein, Würzburg, Germany) and cells were counted under the light microscope CKX41 (Olympus). 600 000 cells per well in a 6-well plate (Thermo Scientific) were seeded in duplicate in 3 ml of the HEK cell medium. The cells were incubated overnight at 37 °C in CO_2-Incubator CB-210 (Binder, Tuttlingen, Germany).

2.2.2.2. Day 2 – transfection

The cells seeded the day before needed to be 40-60% confluent prior to transfection. The medium was removed from the wells, and 1.5 ml of fresh HEK cell medium was added to each well. The transfection reaction was prepared by mixing 2 µg of the appropriate plasmid DNA containing the epitopes and DMEM 1g/L glucose without L-glutamine, resulting in a final volume of 100 µl in one 1.5 ml microtube for the positive control and for each well. As a negative control 100 µl of DMEM 1g/L glucose without L-glutamine was used. To each microtube I added 20 µl of transfection reagent polyethylenimine (PEI) (Polysciences, Warrington, PA, USA). The tubes were vortexed and shortly span down. The samples were incubated for 10 min at RT. After incubation, 600 µl of the HEK cell medium was added to each sample to stop the transfection reaction, and everything was added to the appropriate well with cells. The plates were placed back in the CO2-Incubator CB-210 (Binder, Tuttlingen, Germany) at 37 °C.

2.2.2.3. Day 3 – transfection validation

The cells were placed under a Nikon Eclipsi Ti fluorescent microscope of 470 nm wave-length. Since the translation of the green fluorescent protein (GFP) is controlled by the IRES element in the vector, it is independent of the TSPAN7 protein fragment synthesis, but serves as a marker of a successful transfection, because both proteins are translated from the same mRNA. Thus, if the cells emitted green signal, it confirmed that the plasmid DNA containing the epitopes entered the cells.

2.2.2.4. Day 4 – protein isolation

The medium was removed from each well and 400 µl of trypsin was added to detach the cells. To each well, 1 ml of the HEK cell medium was added, the clumps were resuspended by pipetting, and cell suspension was transferred into a separate 15 ml Falcon tube. Each well was flushed twice with 1 ml of the HEK cell medium and transferred into the corresponding Falcon tubes. The samples were then centrifuged for 10 min at 300*g* at RT. The supernatant was removed and the cell pellets were resuspended in 1 ml of 1x PBS and transferred into a separate 1.5 ml microtube. The samples were centrifuged again for 10 min at 400*g* at 4 °C. The supernatant was removed and the cell pellets were lysed in 100 µl of protein lysis buffer (0.02M Tris pH 8, 0.14M NaCl, 1mM EDTA, 1% Triton X100, 1:100 Phosphatase inhibitor, 1:100 Protease inhibitor, dH_2O). The samples were placed 30 min on ice. During that

incubation period, the lysate was mixed twice more. Afterwards, the samples were centrifuged for 10 min at 13000 rpm at 4 °C. The supernatant was transferred into a separate pre-cooled 1.5 ml microtube and the luciferase light emission activity (protein activity) was measured. The Nluc-tagged TSPAN7 protein fragments were diluted in 1x TBST buffer (Tris buffered saline pH 7.2 plus 0.1% Tween 20). Each Nluc-tagged TSPAN7 protein fragment sample was first diluted 100 times, and the 1:100 dilution was diluted again 10, 100 and 1000 times. The final protein dilutions were 10^{-3}, 10^{-4} and 10^{-5}. The Luciferase protein activity was measured in white 96-well microplates (Perkin Elmer, Waltham, MA, USA), where I used 1x TBST as a negative control. The dilutions of the samples were measured in triplicates. Into each well 25 µl of 1x Passive lysis buffer (Promega), 5 µl of the protein sample, and 25 µl of 1x Nluc Glo luciferase substrate (Promega) were added. The Luciferase protein activity was measured with Glomax 96 Microplate reader (Promega) by using the standard CellTiterGlo protocol (Promega). After the Luciferase protein activity measurement, the undiluted protein samples were stored at -80 °C.

2.2.3. Luciferase immunoprecipitation systems (LIPS) assay

In the luciferase immunoprecipitation systems (LIPS) assay an antigen is fused to a luciferase protein. The recombinant antigen is then incubated with antibodies, where antibodies bind to the recombinant antigen. The antibody-antigen complex is then transferred to a filter plate containing antibody capturing reagents. The antibody capturing beads bind both free immunoglobulins and antibodies bound to the recombinant antigen, while free unbound antigens are removed by several washing steps. The relative amount of antibody bound to the recombinant antigen can be determined by measuring the light emitted after addition of the substrate for the luciferase (Burbelo et al., 2015) (Figure 7).

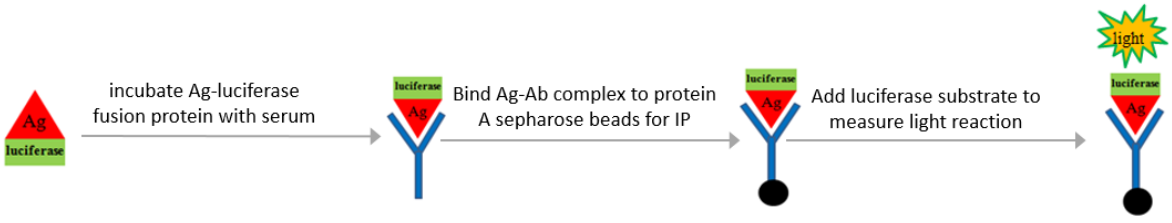


Figure 7. LIPS workflow (Figure from Dr. Michelle Ashton, AG Bonifacio, CRTD, TU Dresden, Cluster of Excellence).

The patient serum was diluted 1:5 in 1x TBST, and the 10 µl of serum was added in duplicates to transparent round-bottom Microplate 96/U (Eppendorf). The activity of the Nluc-tagged TSPAN7 protein fragments was measured prior to the LIPS assay as described in the section 2.2.2.4. To each well 25 µl of the TSPAN7 protein fragment was added, containing 5 million counts per s of Nluc-tagged antigen diluted in 1x TBST. The plates were shortly spun down, mixed on a plate shaker for 30 s at 300 rpm, and incubated in the dark at room temperature for 2 h. Protein-A Sepharose (GE Healthcare, Freiburg, Germany), in equivalent to 1.5 mg per one well sample, was pre-swollen in 1x TBST buffer containing 0.1% BSA low IgG (Life Technologies, Darmstadt, Germany). During the plate incubation, the sepharose beads were centrifuged for 5 min at 500g at room temperature with acceleration 7 and deceleration 5 mode on the centrifuge. The supernatant was removed, the 15 ml Falcon tube with the sepharose beads was filled up with 1xTBST buffer containing 0.1% low IgG BSA, and centrifuged again as described above. This procedure was repeated once more. After the second wash, the Falcon tube was filled up to the required volume. To each well of a UniFilter-96 GF/C microfiltration plate (Perkin Elmer, Waltham, MA, USA), 50 µl of the sepharose beads suspension was added. To the same microfiltration plate 25 µl of antigen-antibody complex was added to each well. The plate was then incubated for 1 h on a shaker at 300 rpm and at 4 °C. After incubation, the plate was washed 10 times with 200 µl of 1x TBST per well. The bottom lid was removed, the bottom was dried, and sealed again with white adhesive bottom seal (Perkin Elmer, Waltham, MA, USA). On the microfiltration plate 25 µl of 1x Passive Lysis Buffer and 25 µl of 1x Nluc Glo luciferase substrate were added to each well. The Nluc-tagged protein activity was measured as described in the chapter 2.2.2.4.

2.2.4. Inhibition luciferase immunoprecipitation systems (LIPS) assay

In the inhibition LIPS assay there were two versions: the first version (v1) contained untagged full-length TSPAN7 along with the Nluc-tagged TSPAN7 protein fragments, and the second (v2) contained protein lysis buffer (used in the protein isolation) along with the Nluc-tagged TSPAN7 protein fragment. The two versions of the assay were conducted for each tested Nluc-tagged TSPAN7 protein fragment. The assay was designed to confirm specific antigen-antibody binding, and rule out antibody binding to the Nluc fragment or the fusion peptide. For a specific binding, there should be a drop in the Nluc protein activity in the first version of the assay compared to the second version because untagged and tagged protein in the first version competed for the antibodies present in the patient serum, and antibodies should bind

untagged protein as well. The protocol was the same as for the LIPS assay described in the section 2.2.3.

3. RESULTS

3.1. Production of tetraspanin 7 truncated forms

As the result of molecular cloning, the plasmid DNA containing the TSPAN7 fragments was produced and digested with restriction enzymes as listed in the Table 2. to see whether the cloning was successful. For all fragments but 495 and 437 there was only one band shown on the gel, instead of two, because the larger band about 7kb was at the very top of the 2% agarose gel. For all samples there were bands correlated with the expected fragments after the digestion (Figure 8). In summary, restriction digests showed that all the TSPAN7 DNA fragments were successfully cloned into the pCMV6-AC-Nluc-IRES-GFP-Puro vector (Figure 6). In total, nine DNA plasmids were produced, containing different TSPAN7 fragments that were used to transfect HEK cells to produce the fusion proteins. The list of the DNA plasmids and their abbreviated names are given in the Table 4.

Table 4. List of plasmid DNA containing the TSPAN7 fragments obtained after molecular cloning and their abbreviated names.

Plasmid DNA with TSPAN7 fragments	Abbreviated name
pCMV6-AC-Nluc-C1-TM1-E1-TM2-IRES-GFP-Puro	225
pCMV6-AC-Nluc-C1-TM1-E1-IRES-GFP-Puro	168
pCMV6-AC-Nluc-TM1-E1-TM2-C2-IRES-GFP-Puro	207
pCMV6-AC-Nluc-E1-TM2-C2-IRES-GFP-Puro	135
pCMV6-AC-Nluc-TM2-C2-TM3-E2-IRES-GFP-Puro	471
pCMV6-AC-Nluc-C2-TM3-E2-IRES-GFP-Puro	414
pCMV6-AC-Nluc-TM3-E2-TM4-C3-IRES-GFP-Puro	495
pCMV6-AC-Nluc-E2-TM4-C3-IRES-GFP-Puro	437
pCMV6-AC-Nluc-TM4-C3-IRES-GFP-Puro	111

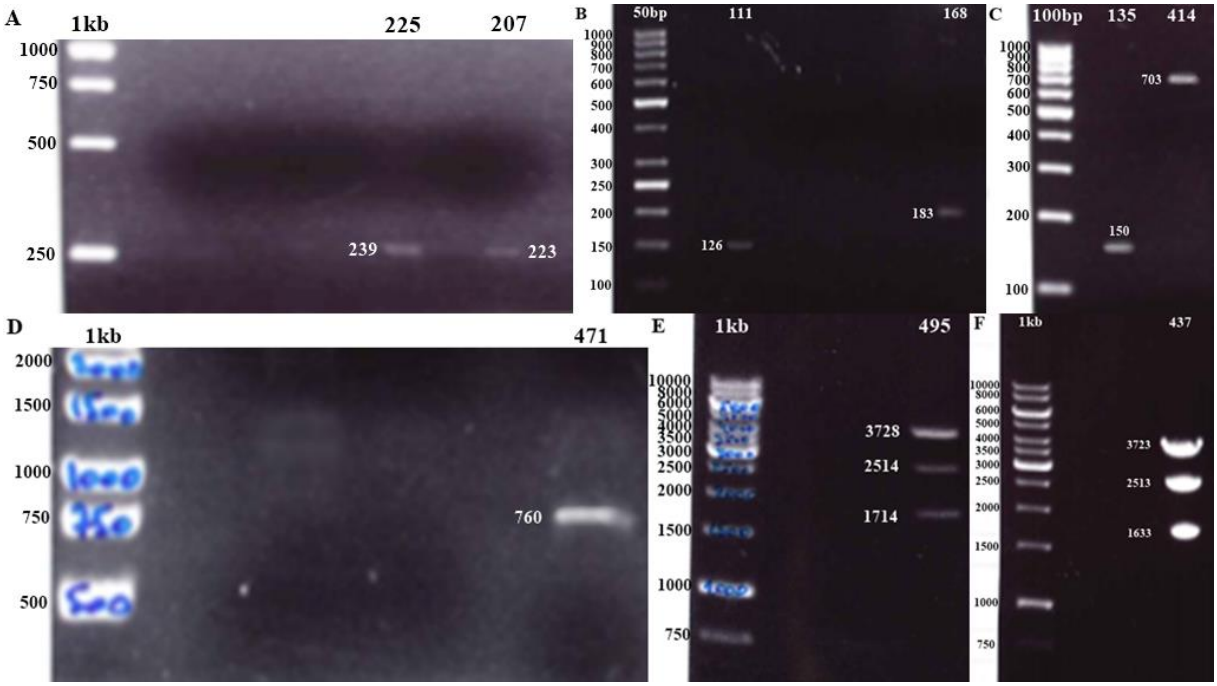


Figure 8. Restriction digestion of vector DNA on 2% agarose gels **A**. For 225 and 207 fragments, bands of approximately 250 bp correspond to the expected 239 and 223 bp fragments, respectively; **B**. For 111 and 168 fragments, bands of approximately 150 and 200 bp, respectively, correspond to the expected 126 and 183 bp fragments, respectively; **C**. For 135 and 414 fragments, bands between 100 and 200 bp and of approximately 700 bp, respectively, correspond to the expected 150 and 703 bp fragments, respectively; **D**. For 471 fragment, a band of approximately 750 bp, correspond to the expected 760 bp fragment; **E**. For 495 fragment, bands of approximately 1500, 2500, and 3500 bp correspond to the expected 1714, 2514, and 3728 bp fragments; **F**. For 437 fragment, bands of approximately 1500, 2500, and 3500 bp correspond to the expected 1633, 2513, and 3723 bp fragments.

3.2. Fusion proteins have sufficient luciferase activity for the LIPS assay

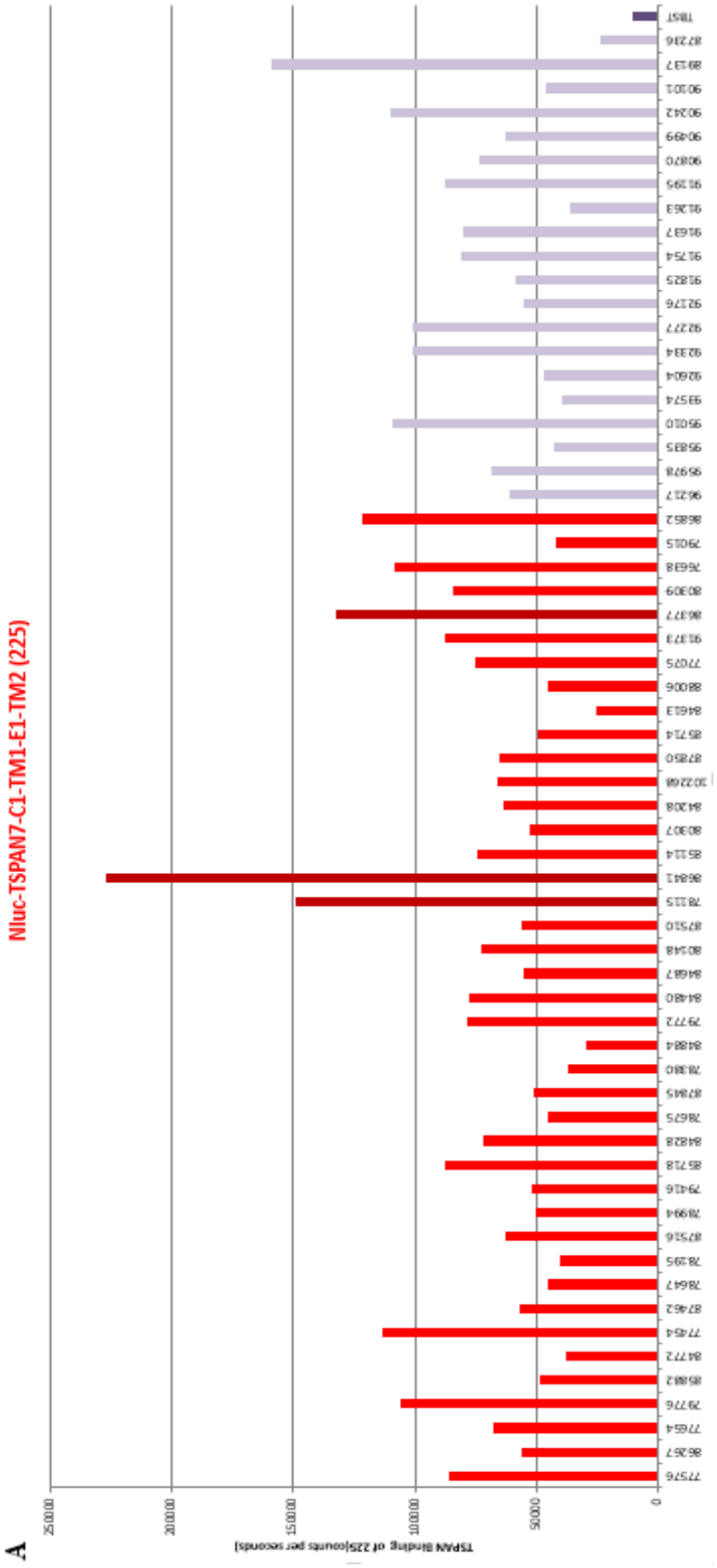
Immediately prior to the LIPS assay, the luciferase light emission activity (protein activity) of Nluc-tagged TSPAN7 protein fragments was measured. The results showed that the fusion proteins 207, 135 and 437 had the highest activity, while the fusion protein 471 the lowest (Table 5). However, the protein activity in all samples was sufficient for the LIPS assay.

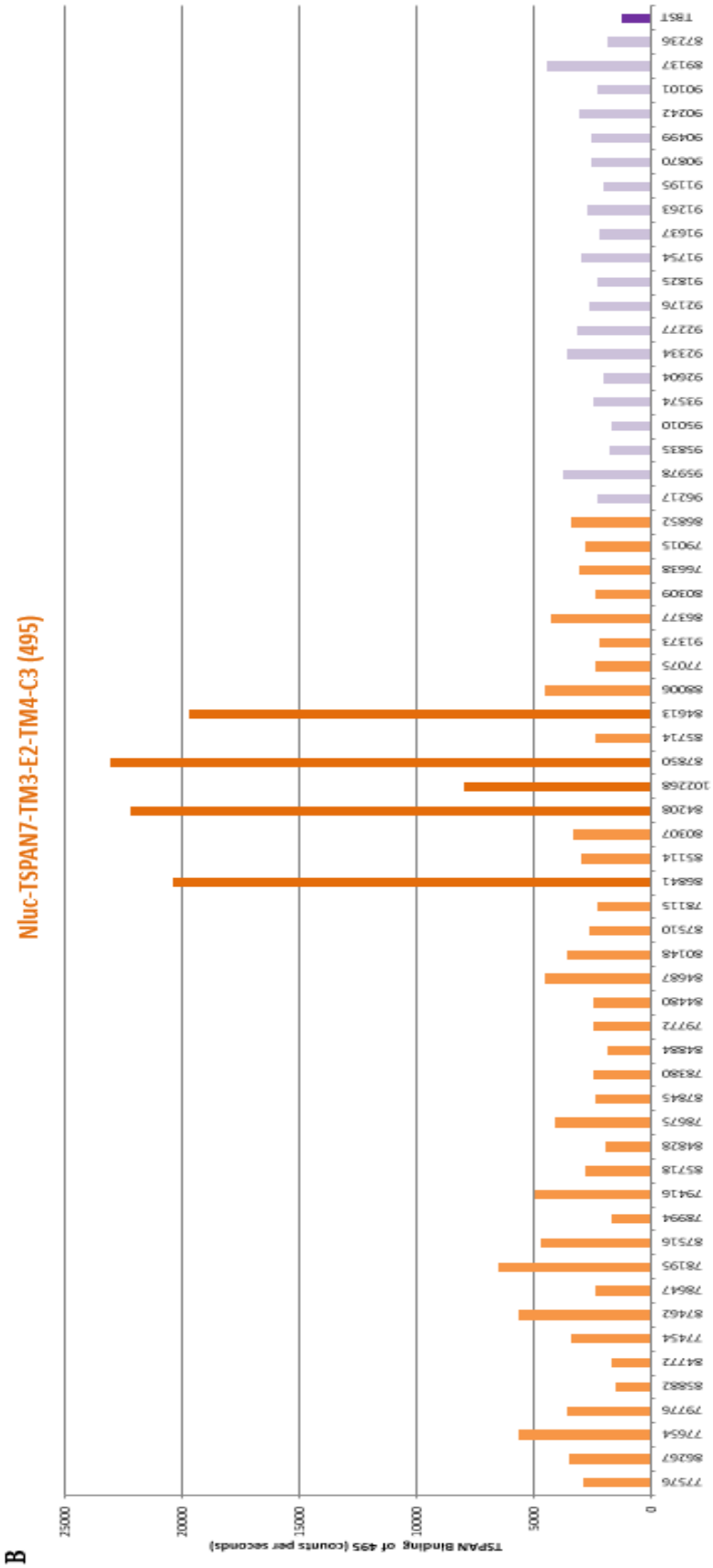
Table 5. Luciferase light emission activity (protein activity) of Nluc-tagged TSPAN7 protein fragments given in counts per s (cps)/µl.

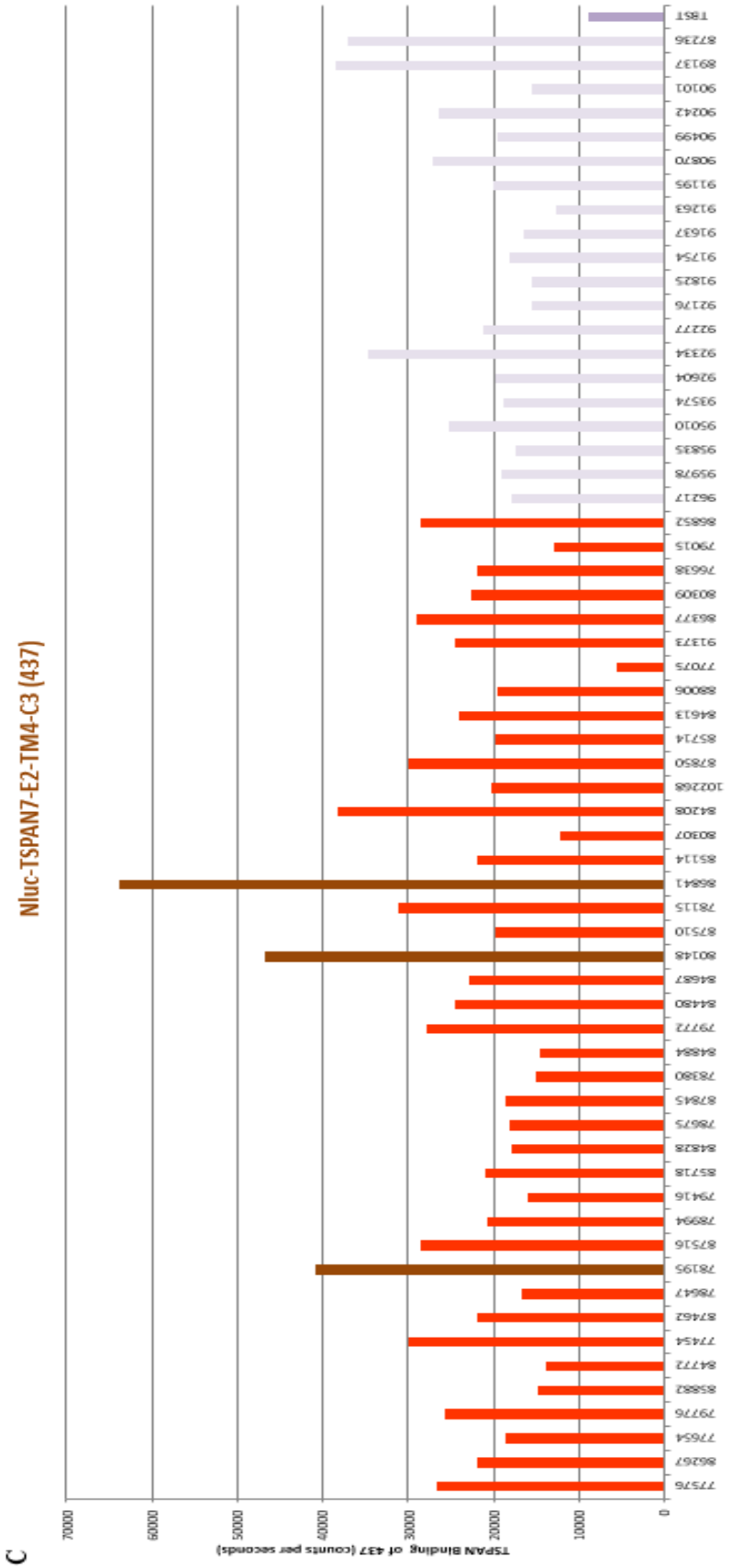
TSPAN7 fragment	Protein activity (counts per s/µl)
C1-TM1-E1-TM2 (225)	2191755
C1-TM1-E1 (168)	16488029
TM1-E1-TM2-C2 (207)	58633824
E1-TM2-C2 (135)	70715686
TM2-C2-TM3-E2 (471)	581924
C2-TM3-E2 (414)	9212820
TM3-E2-TM4-C3 (495)	13844731
E2-TM4-C3 (437)	55787286
TM4-C3 (111)	2761858

3.3. Screening patients serum samples for autoantibodies against TSPAN7

41 patient serum samples that had antibodies to the full-length TSPAN7 served as the source of antibodies for TSPAN7 epitope testing. As negative controls, 20 patient serum samples that did not have antibodies to the full-length TSPAN7 and 1xTBST were used. Six out of the nine TSPAN7 protein fragments, were bound by antibodies in patient sera, as shown by higher counts detected in the LIPS assay using some of the positive serum samples as compared to the negative controls. Fragments C1-TM1-E1-TM2 (225), TM3-E2-TM4-C3 (495), E2-TM4-C3 (437), and TM4-C3 (111) gave stronger positive signals through antibody binding (Figure 9), while fragments C1-TM1-E1 (168), and TM1-E1-TM2-C2 (207) gave positive signals through antibody binding that were barely higher than the background signals (Figure 10), as determined by the appropriate cut-offs. Three fragments, E1-TM2-C2 (135), TM2-C2-TM3-E2 (471), and C2-TM3-E2 (414), gave no positive signals through antibody binding compared to the counts in the negative controls (Figure 11). The six fragments and the corresponding sera, which gave positive signals, were further tested in the inhibition LIPS assay.


A
Nluc-TSPAN7-C1-TM1-E1-TM2 (225)
TSPAN Binding of 225(counts per seconds)
250000
200000
150000
100000
50000
0


B
Nluc-TSPAN7-TM3-E2-TM4-C3 (495)
TSPAN Binding of 495 (counts per seconds)
25000
20000
15000
10000
5000
0
1851
87236
89137
90101
90242
90499
90870
91195
91263
91637
91754
91825
92176
92277
92334
92604
93574
95010
95835
95978
96217
86852
79015
76638
80309
86377
91373
77075
89006
84613
85714
87850
102268
84208
80307
85114
86841
78115
87510
80148
84687
84480
79772
84884
78380
87845
78675
84828
85718
79416
78994
87516
78195
78647
87462
77454
84772
85882
79776
77654
86267
77576


C
Nluc-TSPAN7-E2-TM4-C3 (437)
TSPAN Binding of 437 (counts per seconds)
70000
60000
50000
40000
30000
20000
10000
0
TBST
87236
89137
90101
90242
90499
90870
91195
91263
91637
91754
91825
92176
92277
92334
92604
93574
95010
95835
95978
96217
86852
79015
76638
80309
86377
91373
77075
88006
84613
85714
87850
102268
84208
80307
85114
86841
78115
87510
80148
84687
84480
79772
84884
78380
87845
78675
84828
85718
79416
78994
87516
78195
78647
87462
77454
84772
85882
79776
77654
86267
77576

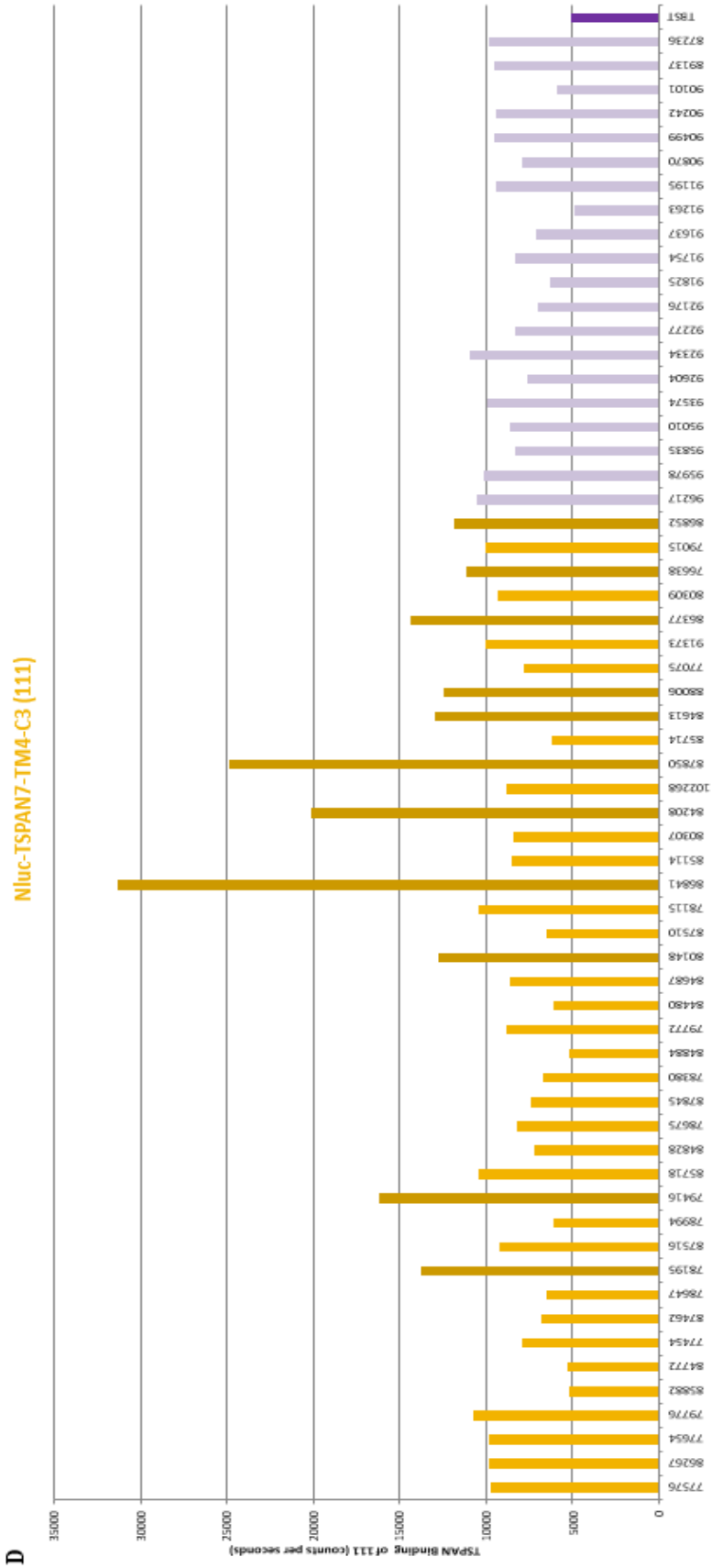
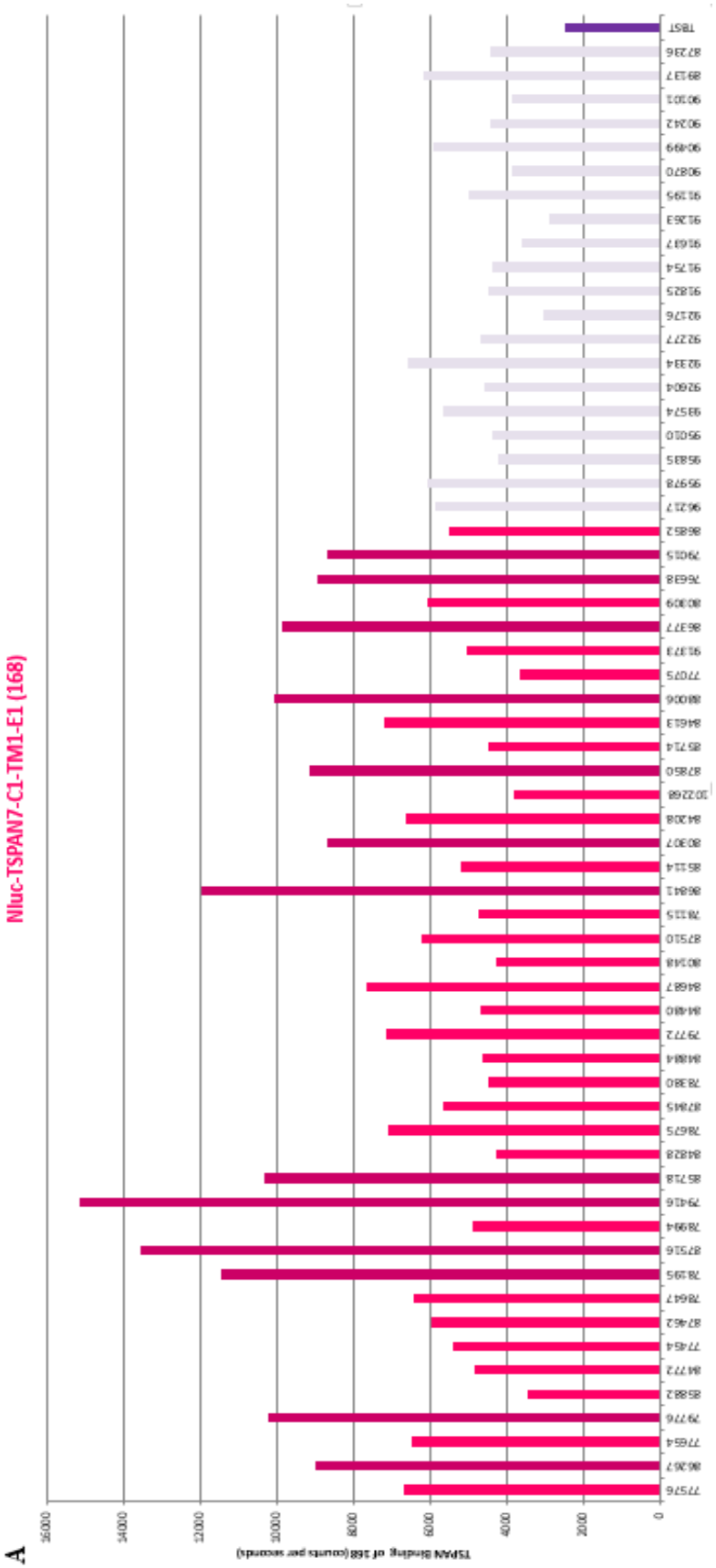

D
Nluc-TSPAN7-TM4-C3 (111)
TSPAN Binding of 111 (counts per seconds)
35000
30000
25000
20000
15000
10000
5000
0
TBST
87236
89137
90101
90242
90499
90870
91195
91263
91637
91754
91825
92176
92277
92334
92604
93574
95010
95835
95978
96217
86852
79015
76638
80309
86377
91373
77075
88006
84613
85714
87850
102268
84208
80307
85114
86841
78115
87510
80148
84687
84480
79772
84884
78380
87845
78675
84828
85718
79416
78994
87516
78195
78647
87462
77454
84772
85882
79776
77654
86267
77576

Figure 9. Nluc-tagged TSPAN7 fragments that gave stronger positive signals through antibody binding compared to negative controls in the LIPS assay. The sera samples which gave positive signals are shown in darker shade in each graph and are used in inhibition LIPS assay. **A.** C1-TM1-E1-TM2 (225) fragment gave positive signals in three serum samples with approx. cutoff at 150000 cps; **B.** TM3-E2-TM4-C3 (495) fragment gave positive signals in four serum samples with approx. cut-off at 5000 cps; **C.** E2-TM4-C3 (437) fragment gave positive signals in three serum samples with approx. cutoff at 40000 cps; **D.** TM4-C3 (111) fragment gave positive signals in 11 serum samples with approx. cutoff at 12000 cps.


A
Nluc-TSPAN7-C1-TM1-E1 (168)
TSPAN Binding of 3B8 (counts per seconds)
16000
14000
12000
10000
8000
6000
4000
2000
0
TEST
87236
89137
90101
90242
90499
90870
91195
91363
91687
91754
91825
92176
92277
92334
92604
93574
95010
95835
95978
96217
86852
79015
76638
80309
86377
91373
77075
88006
84613
85714
87850
102268
84308
80307
85114
86841
78115
87510
80148
84687
84480
79772
84884
78380
87845
78675
84828
85758
79416
78994
87516
78395
78647
87462
77454
84772
85882
79776
77654
86367
77576

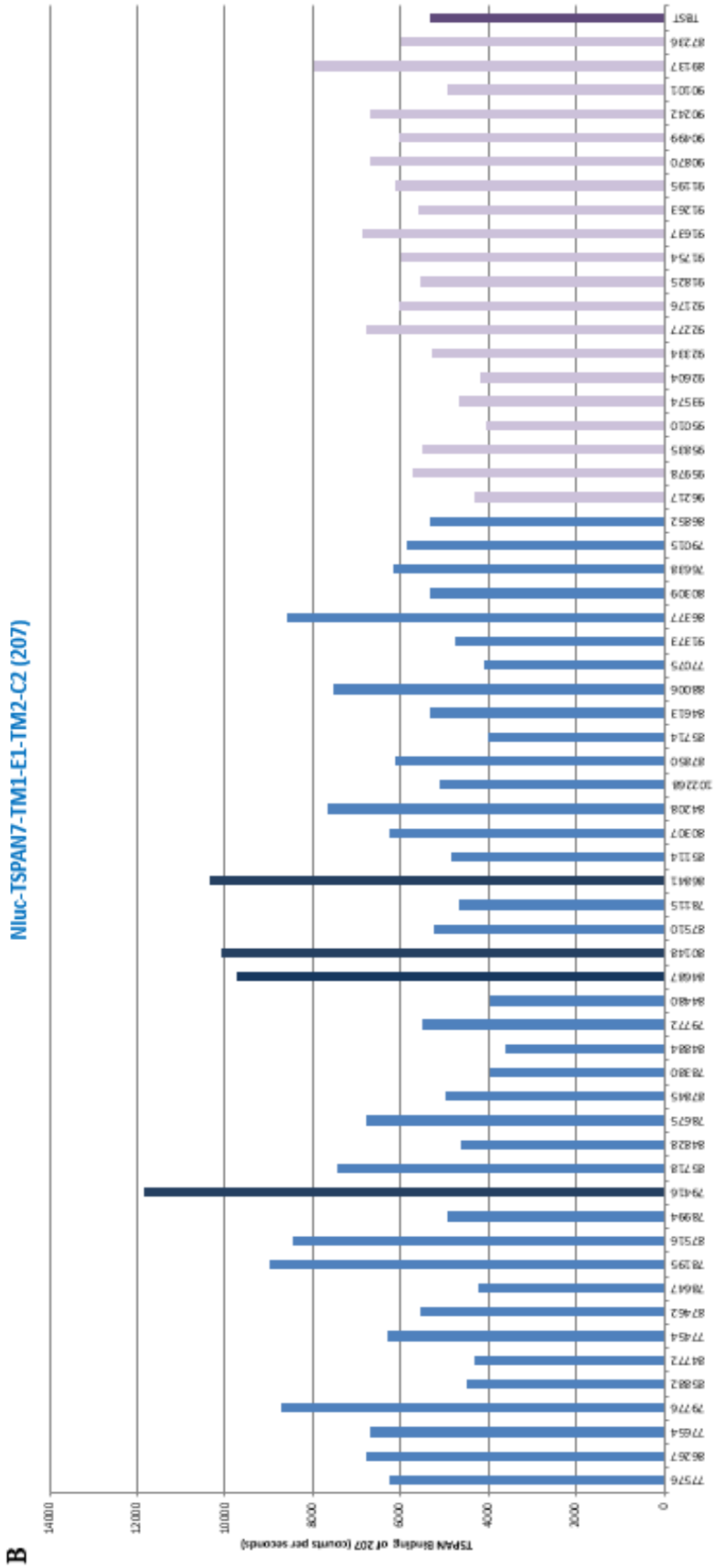
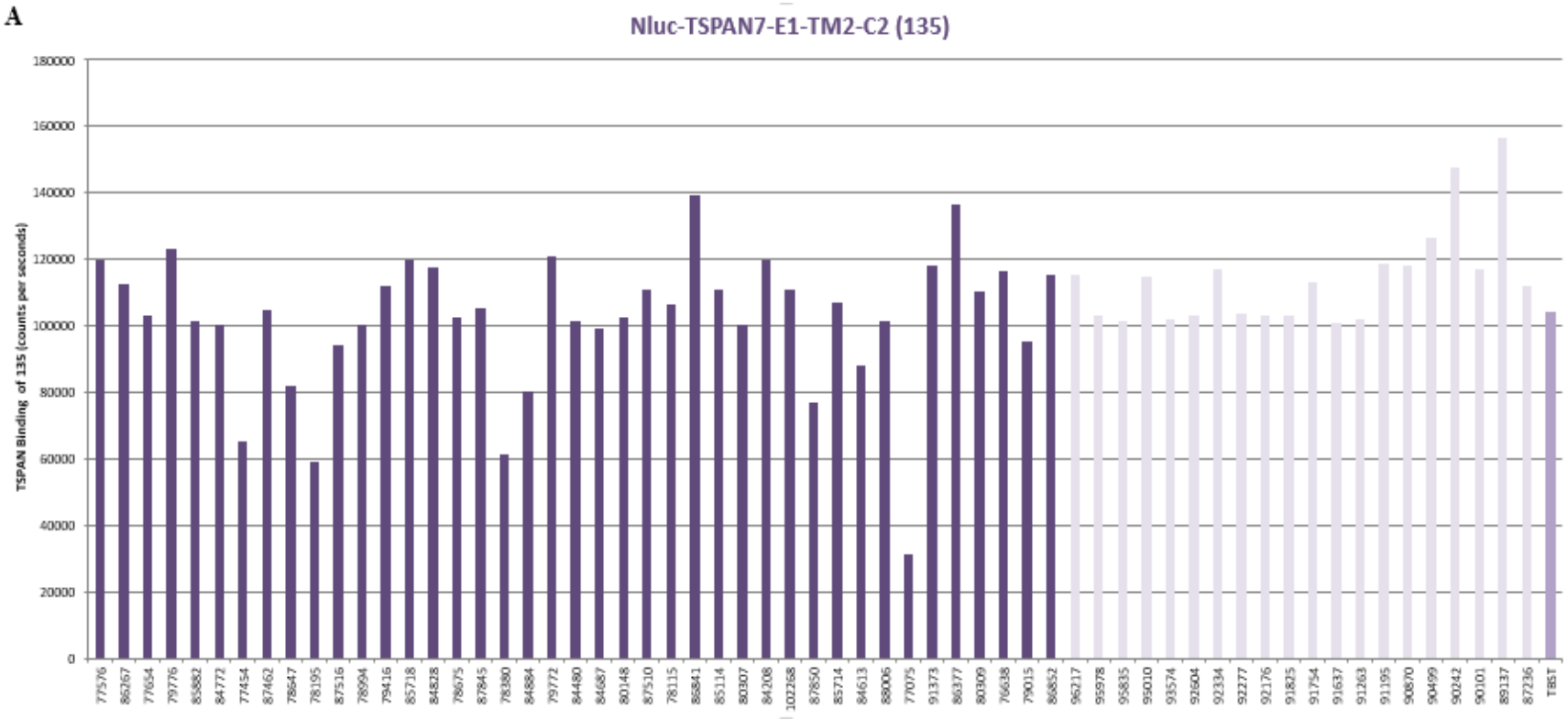
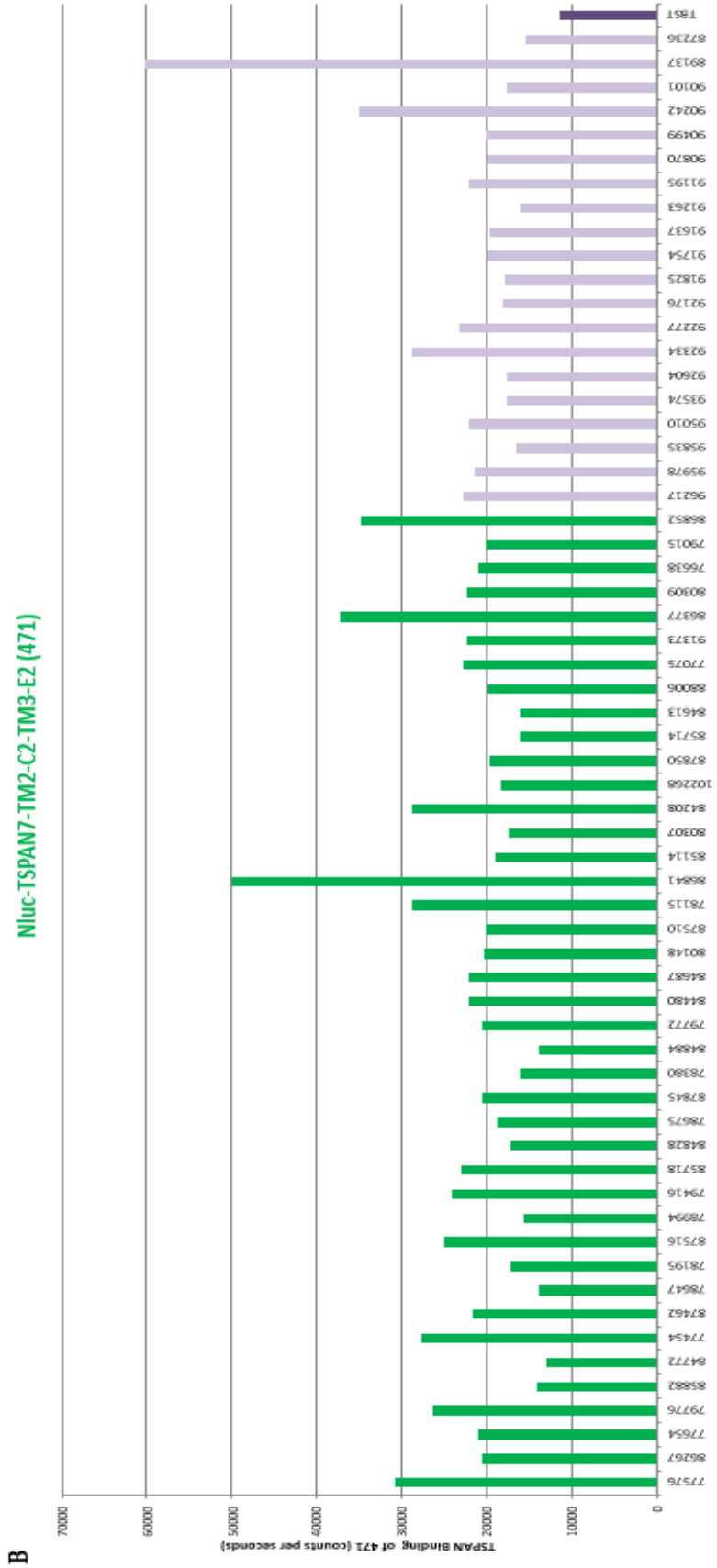

B
Nluc-TSPAN7-TM1-E1-TM2-C2 (207)
TSPAN Binding of 207 (counts per seconds)
0
2000
4000
6000
8000
10000
12000
14000

Figure 10. Nluc-tagged TSPAN7 fragments that gave weaker positive signals through antibody binding compared to negative controls in the LIPS assay. The sera samples which gave positive signals are shown in darker shade in each graph and are used in inhibition LIPS assay. **A.** C1-TM1-E1 (168) fragment gave positive signals in 13 serum samples with approx. cutoff at 8000 cps; **B.** TM1-E1-TM2-C2 (207) fragment gave positive signals in four serum samples with approx. cutoff at 10000 cps.

A


Nluc-TSPAN7-E1-TM2-C2 (135)
TSPAN Binding of 135 (counts per seconds)
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
77576
86267
77654
79776
85882
84772
77454
87462
78647
78195
87516
78994
79416
85718
84828
78675
87845
78380
84884
79772
84480
84687
80148
87510
78115
85841
85114
80307
84208
102268
87850
85714
84613
88006
77075
91373
86377
80909
76638
79015
86852
95217
95978
95835
95010
93574
92604
92334
92277
92176
91825
91754
91637
91263
91195
90870
90499
90242
90101
89137
87236
TBST


B
Nluc-TSPAN7-TM2-C2-TM3-E2 (471)
TSPAN Binding of 471 (counts per seconds)
70000
60000
50000
40000
30000
20000
10000
0
TBS1
87236
89137
90101
90242
90499
90870
91195
91263
91637
91754
91825
92176
92277
92334
92604
93574
95010
95835
95978
96217
86852
79015
76638
80309
86377
91373
77075
88006
84613
85714
87850
102268
84208
80307
85114
86841
78115
87510
80148
84687
84480
79772
84884
78380
87845
78675
84828
85718
79416
78994
87516
78195
78647
87462
77454
84772
85882
79776
77654
86267
77576

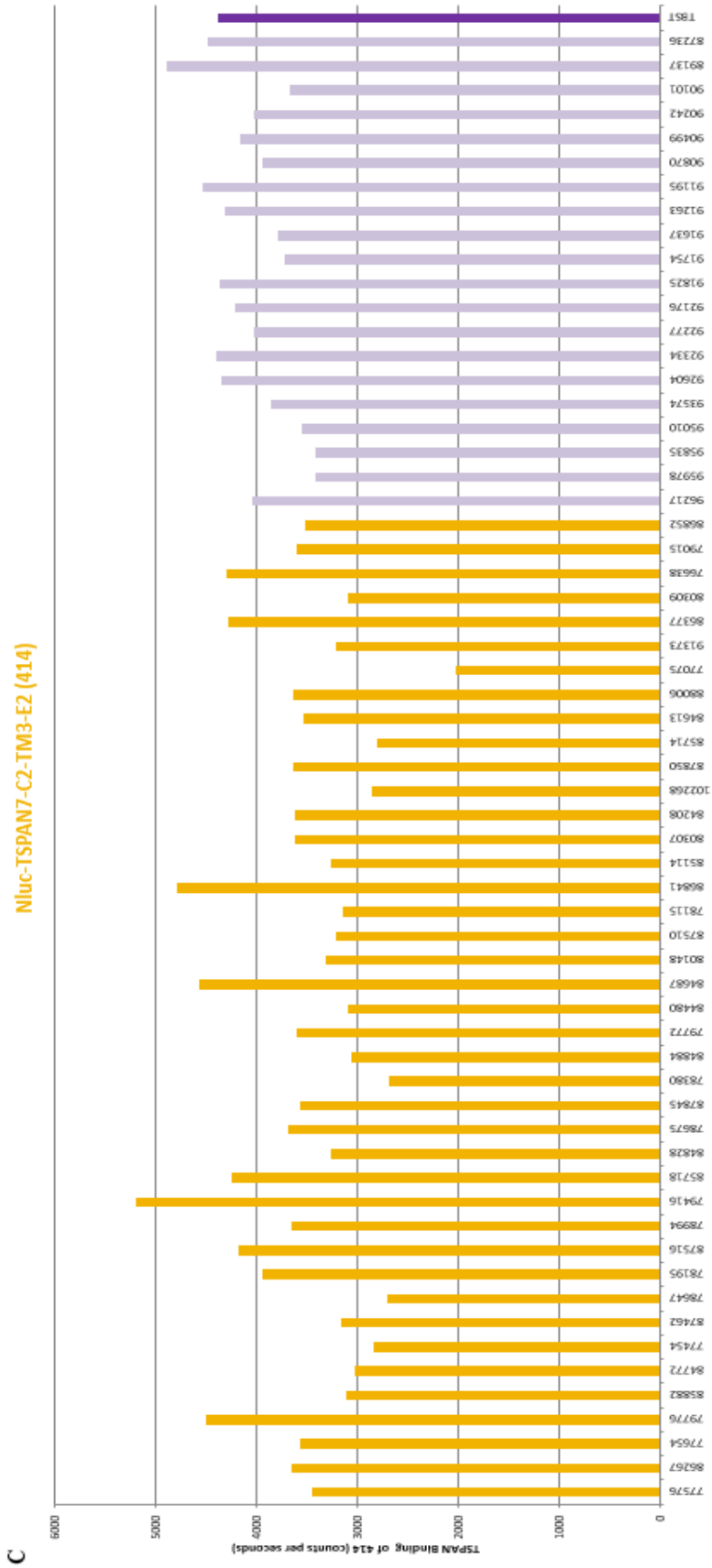
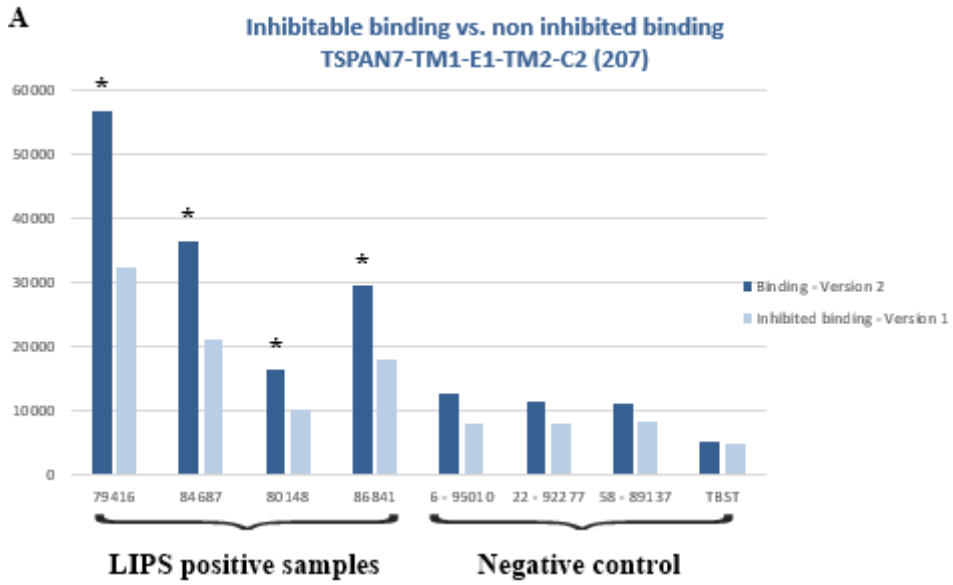

C
Nluc-TSPAN7-C2-TM3-E2 (414)
TSPAN Binding of 414 (counts per seconds)
0
1000
2000
3000
4000
5000
6000
T85T
87236
89137
90101
90242
90499
90870
91195
91263
91637
91754
91825
92176
92277
92334
92604
93574
95010
95835
95978
96217
86852
79015
76638
80309
86377
91373
77075
88006
84613
85714
87850
102268
84208
80307
85114
86841
78115
87510
80148
84687
84480
79772
84884
78380
87845
78575
84828
85718
79416
78994
87516
78195
78647
87462
77454
84772
85882
79776
77654
86267
77576

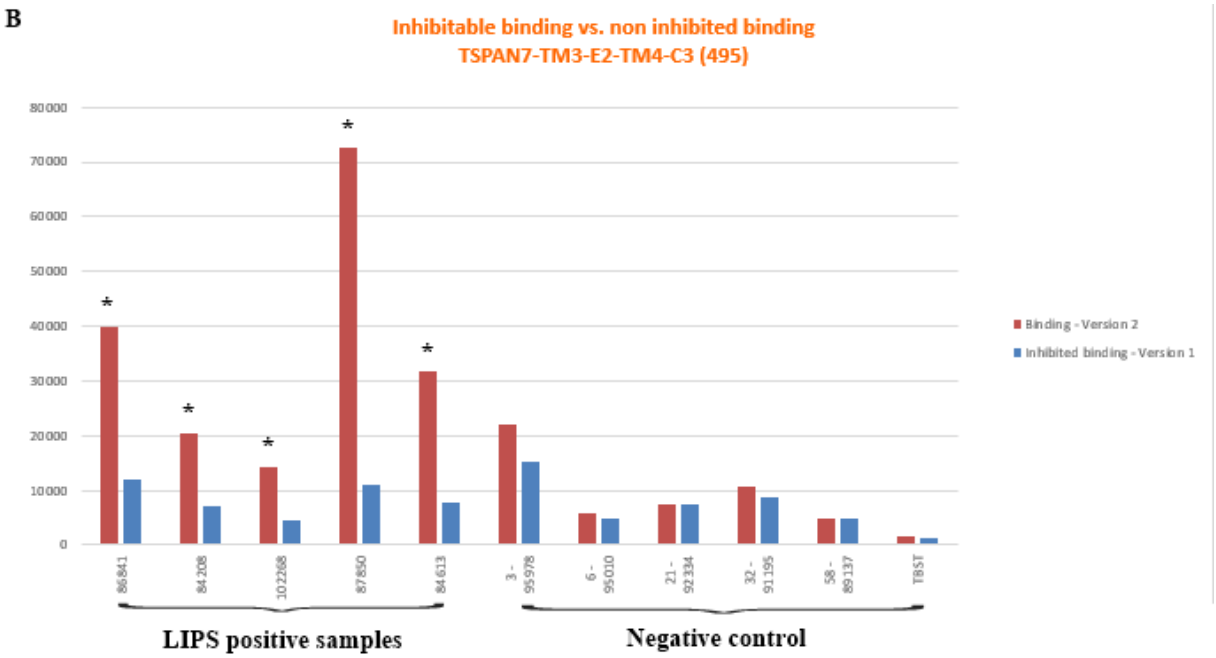
Figure 11. Nluc-tagged TSPAN7 protein fragments which gave no positive signals through antibody binding in serum samples in the LIPS assay. **A.** E1-TM2-C2 (135); **B.** TM2-C2-TM3-E2 (471); **C.** C2-TM3-E2 (414). They were not tested in inhibition LIPS assay.

3.4. Inhibition luciferase immunoprecipitation systems (LIPS) assay

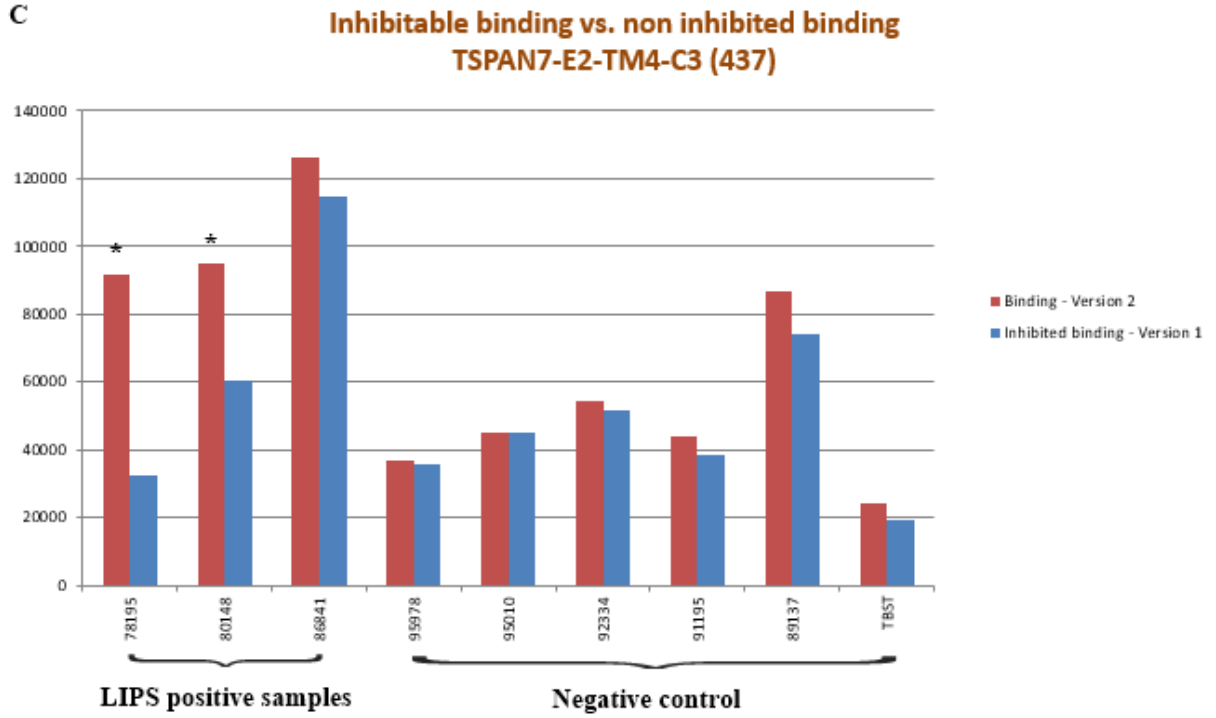
The TSPAN7 fragment TM3-E2-TM4-C3 (495) for all tested serum samples gave observably higher signals through antibody binding in v2 than in v1 of the assay. The fragments E2-TM4-C3 (437) and TM4-C3 (111) gave higher signals through antibody binding in v2 compared to v1 for some of the tested serum samples, while the fragment TM1-E1-TM2-C2 (207) gave higher signals through antibody binding for all tested serum samples in v2, but the difference in counts between v2 and v1 was not as big (Figure 12). For the fragments C1-TM1-E1-TM2 (225) and C1-TM1-E1 (168), there was no observable difference in signals through antibody binding between v2 and v1 (Figure 13).



B


Inhibitable binding vs. non inhibited binding
TSPAN7-TM3-E2-TM4-C3 (495)
80000
70000
60000
50000
40000
30000
20000
10000
0
*
*
*
*
*
86841
84208
102268
87850
84613
3 - 95978
6 - 95010
21 - 92334
32 - 91195
58 - 89137
TBST
LIPS positive samples
Negative control
Binding - Version 2
Inhibited binding - Version 1

C


Inhibitable binding vs. non inhibited binding
TSPAN7-E2-TM4-C3 (437)
140000
120000
100000
80000
60000
40000
20000
0
*
*
78195
80148
86841
95978
95010
92334
91195
89137
TBST
LIPS positive samples
Negative control
Binding - Version 2
Inhibited binding - Version 1

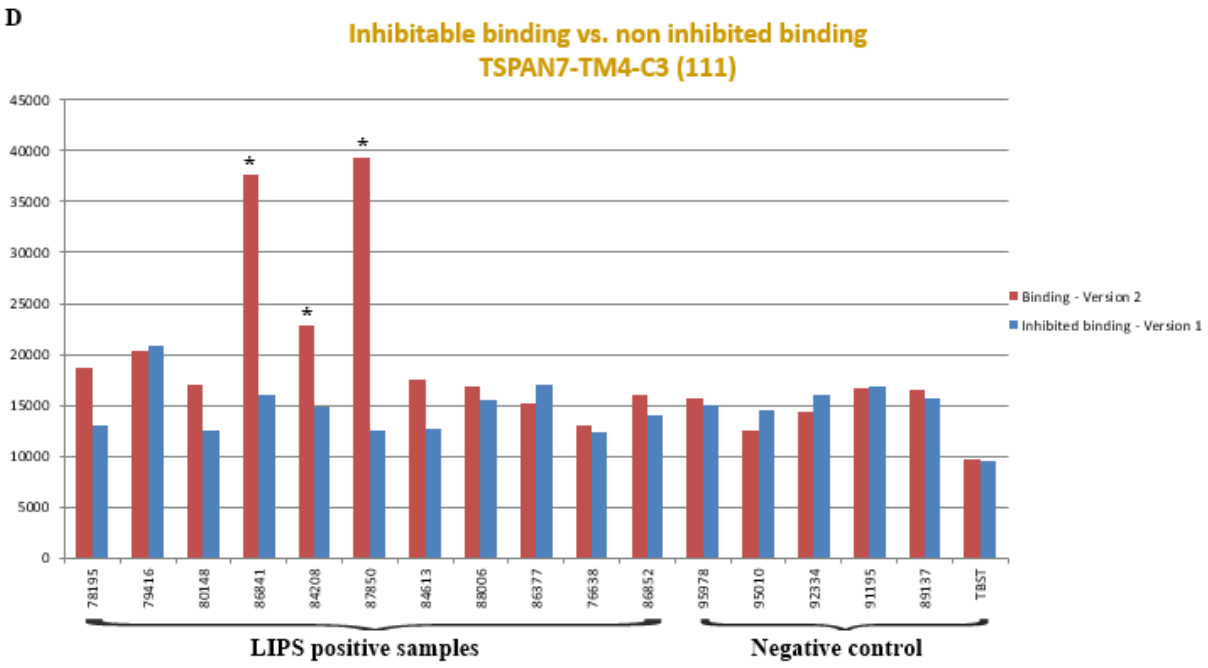
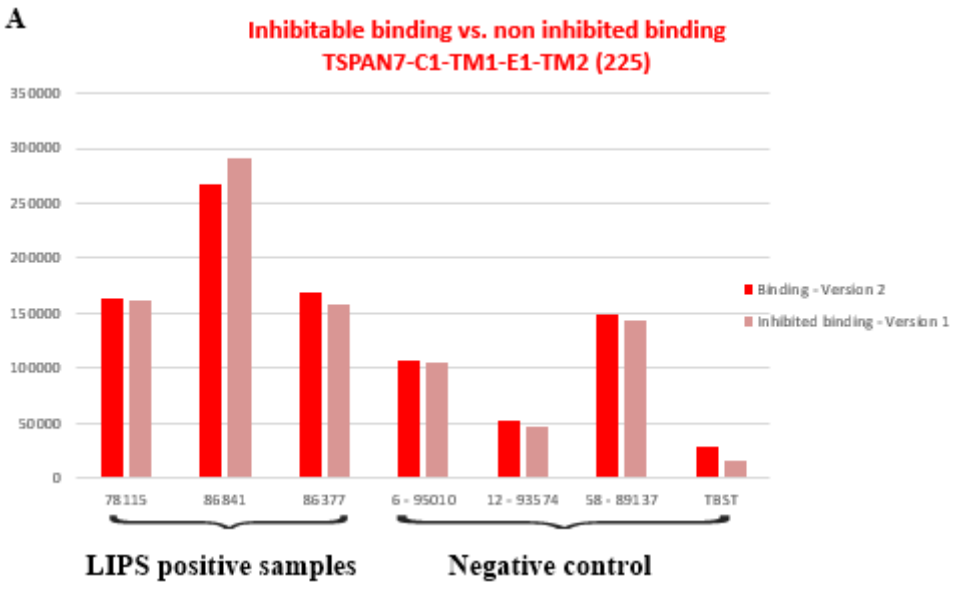


Figure 12. TSPAN 7 protein fragments that showed positive inhibition to the binding of Nluc-TSPAN fragments to some or all tested serum samples, which confirmed specific antibody-TSPAN7 protein fragment binding. **A.** TM1-E1-TM2-C2 (207); **B.** TM3-E2-TM4-C3 (495); **C.** E2-TM4-C3 (437); **D.** TM4-C3 (111).



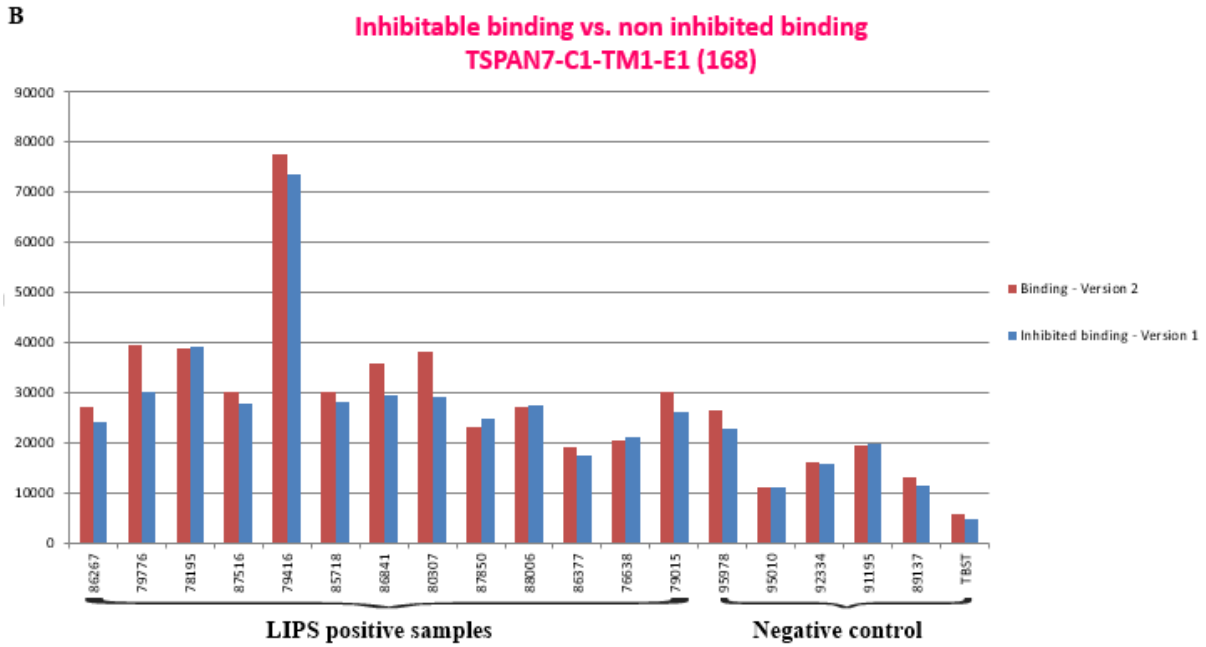


Figure 13. TSPAN 7 protein fragments that showed no inhibition to the binding of Nluc-TSPAN fragments to serum samples, which suggests that antibodies bound Nluc fragment or Nluc fusion peptide instead of TSPAN7 protein fragment. **A.** C1-TM1-E1-TM2 (225); **B.** C1-TM1-E1 (168).

4. DISCUSSION

In the LIPS assay, antibodies from the serum samples bound six TSPAN7 fragments. The indicator of the positive binding were the counts produced through binding of antibodies in the patient serum samples to TSPAN7 fragments, which were observably higher than the counts of the negative controls. Out of the six TSPAN7 fragments that produced positive signals when tested against patient serum samples, only three showed inhibited binding in the inhibition LIPS assay. In the inhibition LIPS assay there were two versions of the assay. In the first version, Nluc-tagged and untagged TSPAN7 fragments competed for antibodies from the sera samples. If the antibody binding was specific for a protein fragment rather than for Nluc tag or to a part overlapping Nluc and TSPAN7 (Nluc fusion peptide), the second version of the assay, which contained only Nluc-tagged TSPAN7 fragments, should give rise to observably higher counts than the first version. From the results it can be seen that fragments TM3-E2-TM4-C3 (495), E2-TM4-C3 (437), and TM4-C3 (111) gave strong positive signals through antibody binding in the LIPS assay, which indicated that antibodies from patient serum samples bound those Nluc-tagged fragments. With the inhibition LIPS assay, it was confirmed that the antibodies indeed bound TSPAN7 fragments rather than Nluc tag or the fusion peptide due to the significant drop in counts in v1 compared to the v2. The results of the LIPS and inhibition LIPS assays suggest that E2 and C3 TSPAN7 domains are the most likely candidates that carry autoepitopes, and that TM4 domain could play a role in the binding as well. Because most serum samples that had antibodies against full-length TSPAN7 did not have increased binding of antibodies to TSPAN7 truncated forms, it can be suggested that the major epitopes are lost when only fragments of the TSPAN7 are expressed. This finding supports the theory about conformational, rather than linear, TSPAN7 autoepitopes (Roll et al., 2000). It also suggests that other domains, such as transmembrane domains, are necessary for proper TSPAN7 conformation and overall protein stabilization, as proposed by Hemler (2003). In order to confirm that indeed E2 and C3 TSPAN7 domains carry autoepitopes, these domains should be used to replace the same domains in another member of the tetraspanin family, with similar structure, but which does not bind antibodies in the sera of T1DM patients. In that case, the tetraspanin protein carrying E2 or C3 TSPAN7 domains should be able to achieve proper conformation. If E2 and C3 domains of another tetraspanin protein would be replaced with E2 and C3 TSPAN7 domains, respectively, it could be confirmed that the positive binding of antibodies is specific to E2 and/or to C3 TSPAN7 domains. In order to further narrow down the autoepitopes, site directed mutagenesis of the

TM3-E2-TM4-C3 (495) fragment would be a logical next step. By specifically mutating amino acids in E2 and C3 regions, positions of amino acids that carry autoepitopes could be uncovered.

5. CONCLUSION

E2 and C3 TSPAN7 domains are the likely candidates that carry autoepitopes in type I diabetes. TM4 could also play a role either by carrying autoepitopes itself, or by positioning E2 and C3 in proper conformation. When only fragments of the TSPAN7 protein are expressed, instead of the full-length protein, major epitopes are lost.

6. LITERATURE

Aanstoot H. J., Kang S. M., Kim J., Lindsay L., Roll U., Knip M., Atkinson M., Mose-Larsen P., Fey S., Ludvigsson J., Landin M., Bruining J., Maclaren N., Akerblom H. K ., Baekkeskov S. (1996): Identification and characterization of glima 38, a glycosylated islet cell membrane antigen, which together with GAD65 and IA2 marks the early phases of autoimmune response in type 1 diabetes. J. Clin. Invest. **97**: 2772–2783.

Achenbach P., Bonifacio E., Koczwara K., Ziegler A. G. (2005). Natural history of type I diabetes. Diabetes. **54(2)**: 25-31.

Achenbach P., Warncke K., Reiter J., Naserke H. E., Williams A. J. K., Bingley P. J., Bonifacio E., Ziegler A. G. (2004): Stratification of type I diabetes risk on the basis of islet autoantibody characteristics. Diabetes. **53**: 384-392.

Berditchevski F. (2001): Complexes of tetraspanins with integrins: more than meets the eye. J. Cell Sci. **114**: 4143-4151.

Bonifacio E. (2015): Predicting type I diabetes using biomarkers. Diabetes Care. **38**: 989-996.

Burbelo P. D., Lebovitz E. E., Notkins A. L. (2015): Luciferase immunoprecipitation systems for measuring antibodies in autoimmune and infectious diseases. Transl. Res. **165(2)**: 325-335.

Charrin S., Jouannet S., Boucheix C., Rubinstein E. (2014): Tetraspanins at a glance. J. Cell Sci. **127**: 3641-3648.

Giannopolou E. Z., Winkler C., Chmiel R., Matzke C., Scholz M., Beyerlein A., Achenbach P., Bonifacio E., Ziegler A. G. (2015): Islet autoantibody phenotypes and incidence in children at increased risk for type I diabetes. Diabetologia. **58**: 2317-2323.

Hemler M. E. (2003): Tetraspanin proteins mediate cellular penetration, invasion and fusion events and define a novel type of membrane microdomain. Annu. Rev. Cell Dev. Biol. **19**: 397-422.

Hemler M. E. (2005): Tetraspanin functions and associated microdomains. Nature Rev. Mol. Cell Biol. **6**: 801-811.

Kent W. J., Sugnet C. W., Furey T. S., Roskin K. M., Pringle T. H., Zahler A. M., Haussler D. (2002): The human genome browser at UCSC. Genome Res. **12(6)**: 996-1006. http://genome.ucsc.edu/

Lipponen K., Gombos Z., Kiviniemi M., Siljander H., Lempainen J., Hermann R., Veijola R., Simell O., Knip M., Ilonen J. (2010): Effect of HLA class I and class II alleles on

progression from autoantibody positivity to overt type I diabetes in children with risk-associated class II genotypes. Diabetes. **59**: 3253-3256.

Maecker H. T., Todd S. C., Levy S. (1997): The tetraspanin superfamily: molecular facilitators. FASEB J. **11**: 428-442.

McLaughlin K. A., Richardson C. C., Ravishankar A., Brigatti C., Liberati D., Lampasona V., Piemonti L., Morgan D., Feltbower R. G., Christie M. R. (2016): Identification of Tetraspanin-7 as a Target of Autoantibodies in Type 1 Diabetes. Diabetes. **65(6)**: 1690-1698.

Michels A., Zhang L., Khadra A., Kushner J. A., Redondo M. J., Pietropaolo M. (2015): Prediction and prevention of type I diabetes: update on success of prediction and struggles at prevention. Pediatr. Diabetes. **16**: 465-484.

Roll U., Turck C. W., Gitelman S. E., Rosenthal S. M., Nolte M. S., Masharani U., Ziegler A. G., Baekkeskov S. (2000): Peptide mapping and characterisation of glycation patterns of the glima 38 antigen recognised by autoantibodies in type I diabetic patients. Diabetologia. **43**: 598-608.

Skaar K., Korza H. J., Tarry M., Sekyrova P., Högbom M. (2015): Expression and subcellular distribution of GFP-tagged human tetraspanin proteins in *Saccharomyces cerevisiae*. PLoS One. **10(7)**: e0134041.

Skyler J. S. (2013): Primary and secondary prevention of type I diabetes. Diabet. Med. **30(2)**: 161-169.

Taplin C., Barker J. M. (2008): Autoantibodies in type I diabetes. Autoimmunity. **41(1)**: 11-18.

The UniProt Consortium. (2015): UniProt: a hub for protein information. Nucleic Acids Res. **43**: D204-D212.

Thümer L., Adler K., Bonifacio E., Hofmann F., Keller M., Milz C., Munte A., Ziegler A. G. (2010) German new onset diabetes in the young incident cohort study: DiMelli study design and first-year results. Rev Diabet Stud **7**: 202-208.

Walther D., Eugster A., Jergens S., Gavrisan A., Weinzierl C., Telieps T., Winkler C., Ziegler A. G., Bonifacio E. (2016): Tetraspanin 7 autoantibodies in type I diabetes. Diabetologia. **59**: 1973-1976.

Weenink S. M., Christie M. R. (2006): Autoantibodies in diabetes. In: Pollard K. M. (ed.) Autoantibodies and autoimmunity: molecular mechanisms in health and disease. Weinheim, WILEY-VCH Verlag GmbH & Co. KGaA, pg. 321-349.

Winnock F., Christie M. R., Batstra M. R., Aanstoot H-J., Weets I., Decochez K., Jopart P., Nicolaij D., Gorus F. K., The Belgian Diabetes Registry. (2001): Autoantibodies to a 38-kDa glycosylated islet cell membrane-associated antigen in (pre)type 1 diabetes. Diabetes Care. **24**: 1181-1186.

Zhang G. L., Keskin D. B., Lin H. N., Lin H. H., DeLuca D. S., Leppanen S., Milford E. L., Reinherz E. L., Brusic V. (2014): Human leukocyte antigen typing using a knowledge base coupled with a high-throughput oligonucleotide probe array analysis. Front. Immunol. **5**: 597.

Ziegler A. G., Bonifacio E. (2012): Age-related islet autoantibody incidence in offspring of patients with type I diabetes. Diabetologia. **55**: 1937-1943.

Ziegler A. G., Bonifacio E., Babydiab-Babydiet Study Group. (2012A): Age-related islet autoantibody incidence in offspring of patients with type 1 diabetes. Diabetologia **55(7)**: 1937-1993.

Ziegler A. G., Meier-Stiegen F., Winkler C., Bonifacio E.; the TEENDIAB Study Group. (2012B): Prospective evaluation of risk factors for the development of islet autoimmunity and type 1 diabetes during puberty – TEENDIAB: study design. Pediatr Diabetes. **13**: 419-424.

Ziegler A. G., Rewers M., Simell O., Simell T., Lempainen J., Steck A., Winkler C., Ilonen J., Veijola R., Knip M., Bonifacio E., Eisenbarth G. S. (2013): Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. JAMA. **309(23)**: 2473-2479.

CURRICULUM VITAE

I was born on May 21st, 1993, in Zagreb. In 2011, I enrolled Undergraduate school of Molecular biology at the Faculty of Science, University of Zagreb. During that time, I participated in the Night of Biology, for which I also received a Special Rector's Award in 2012. In order to improve my laboratory skills, I volunteered in Genos Ltd. in the fields of protein glycosylation and DNA analysis. In 2014, I enrolled Graduate school of Molecular biology. I completed laboratory practice under supervision of Dr. Kristina Majsec in the field of plant biology and molecular genetics. I was also an assistant in laboratory practice for the Molecular genetics course led by Dr. Ivana Ivančić Baće, Asst. Prof. In 2015, I enrolled in Computational biology module on the final year of my Graduate studies. In March, I went on Erasmus + student internship in Center for Regenerative Therapies Dresden, Cluster of Excellence, TU Dresden, where I also wrote my thesis. In August, 2016, I attended Research Summer School of Statistical Omics (RSSSO) in Split to study bioinformatics and statistics and machine learning.