

SEQUENCE SIMILARITY BETWEEN THYROID SELF-PROTEIN AND HEPATITIS C VIRUS POLYPROTEIN: possible triggering mechanism of autoimmune thyroiditis

Maristella de Araújo Carvalho **SOUSA**¹, Raymundo **PARANÁ**² and Luís Jesuíno de Oliveira **ANDRADE**¹

Received 20/12/2015
Accepted 1/4/2016

ABSTRACT - Background - Exposure to viral antigens that share amino acid sequence similar with self- antigens might trigger autoimmune diseases in genetically predisposed individuals, and the molecular mimicry theory suggests that epitope mimicry between the virus and human proteins can activate autoimmune disease. **Objective** - The purpose of this study is to explore the possible sequence similarity between the amino acid sequences of thyroid self-protein and hepatitis C virus proteins, using databanks of proteins and immunogenic peptides, to explain autoimmune thyroid disease. **Methods** - Were performed the comparisons between the amino acid sequence of the hepatitis C virus polyprotein and thyroid self-protein human, available in the database of National Center for Biotechnology Information on Basic Local Alignment Search Tool. **Results** - The sequence similarity was related each hepatitis C virus genotype to each thyroid antigen. The similarities between the thyroid and the viral peptides ranged from 21.0 % (31 identical residues out of 147 amino acid in the sequence) to 71.0% (5 identical residues out of 7 amino acid in the sequence). **Conclusion** - Bioinformatics data, suggest a possible pathogenic link between hepatitis C virus and autoimmune thyroid disease. Through of molecular mimicry is observed that sequences similarities between viral polyproteins and self-proteins thyroid could be a mechanism of induction of crossover immune response to self-antigens, with a breakdown of self-tolerance, resulting in autoimmune thyroid disease.

HEADINGS - Hepacivirus. Thyroiditis. Amino acid sequence homology. Computational biology.

INTRODUCTION

Viral infections are often cited as a major environmental factor involved in autoimmune thyroid diseases (ATD). It remains to determine whether they are responsible for ATD or whether they are just innocent bystanders.

Exposure to viral antigens that share amino acid (AA) sequence similar with self-antigens might trigger autoimmune diseases in genetically predisposed individuals and the molecular mimicry theory suggests that epitope mimicry between the virus and human proteins can activate autoimmune disease.

Studies suggesting the existence in humans of polarized T helper (Th) cell subsets, coded as Th1 and Th2, with defined cytokine secretion profiles. The Th1 cells seem to be involved in organ-specific autoimmunity, such as ATD⁽¹²⁾. The triggering of autoimmunity secondary to infection or immunization is often

related to antigenic mimicry because only five to six AA are necessary to induce an immune response⁽²⁸⁾. Has been showed that molecular mimicry between viral antigen may be involved in the pathogenesis of ATD, as autoimmune thyroiditis, in patients with hepatitis C virus (HCV) infection, as a result of cross-immune response to homolog viral antigen⁽²⁵⁾.

Autoimmune disorders are associated with chronic HCV infection and notably autoimmune and non-autoimmune thyroiditis is common in people with HCV infection, and the molecular mimicry mechanism could be trigger⁽⁴⁾. Therefore, autoimmune thyroiditis requires investigation in patients with HCV infection. Although the high incidence in patients with autoimmune thyroiditis HCV infection treated with interferon- α has been demonstrated in previous studies, the correlation between autoimmune thyroiditis in untreated patients remains under debate⁽³⁾.

Declared conflict of interest of all authors: none
Disclosure of funding: no funding received

¹ Programa de Pós-Graduação em Ciências da Saúde, Universidade Estadual de Santa Cruz, Ilhéus, BA, Brasil; ² Faculdade de Medicina, Universidade Federal da Bahia, Salvador, BA, Brasil.

Correspondence: Luis Jesuíno de Oliveira Andrade. Rua Nações Unidas, 511. Centro – CEP 45600-673 – Itabuna, BA, Brasil. E-mail: luis_jesuino@yahoo.com.br

Bioinformatics program currently available to construct the structure analyses of HCV component provides an essential framework for virus assembly and understanding of your molecular mechanisms. Similarly, molecular modeling has facilitated the understanding of molecular mimicry as a result of cross-immune response to similar bacterial or viral antigen with thyroid self-protein human.

Until recently, little structure information about HCV protein and thyroid self-protein human were available and molecular models has been built on partial structures, with assembly guided by biochemical data. The purpose of this study is to explore the possible sequence similarity between the AA sequences of thyroid self-protein, which are potential B- and T-cell epitopes of these antigens and proteins of HCV, using databanks of proteins and immunogenic peptides to explain ATD.

METHODS

Were performed the comparison between the AA sequence of the HCV poly-protein and thyroid self-protein human, available in the database of NCBI (National Center for Biotechnology Information) on Basic Local Alignment Search Tool (BLAST2p)⁽¹⁾. The BLAST was introduced as a sequence alignment heuristic that was an order of magnitude faster than earlier approaches for analyzing biological information. Very quickly, this software became a landmark enabling technique for bioinformatics. Thus, the BLAST refers to a program used to generate alignments between a nucleotide or protein sequence, referred to as a query and nucleotide sequences against other database of nucleotide, referred to as subject sequences.

Was used expect BLAST program for performing the comparisons. The expect value is a parameter that describes the number of hits one can "expect" to see by chance when searching a database of a particular size (expect value considered as statistical significance <0.05). It decreases exponentially as the score of the match increases. The low expect value, or the closer it is to zero, the more "significant" the match. However, identical short alignments have relatively high E values. This is because the calculation of the E value takes into account the length of the query sequence. These high expect value make sense because shorter sequences have a higher probability of occurring in the database purely by chance. The expect value can also be used as a convenient way to create a significance threshold for reporting results.

Sequence Similarity searches - Protein database search and analysis

One protein-protein sequence alignment method, BLAST2p program, was used to search for sequence similarity between HCV and thyroid self-proteins. This method is limited to searching for linear epitope similarities, which will miss three-dimensional conformational similarities and possible cross-reactivity between proteic and non-proteic epitopes, since most of molecular mimicry is likely to involve T-cell mediation, and T cells generally recognize linear peptides 8–20 AA in length.

HCV genotypes examined

Were evaluated the following HCV genotypes, with the respective NCBI sequence identification number (GI): hepatitis C virus subtype 1a (HCV1a) (GI:212675089), hepatitis C virus subtype 1b (HCV1b) (GI:190684416), hepatitis C virus subtype 1c (HCV1c) (GI:81992806), hepatitis C virus subtype 2a (HCV2a) (GI:557164188), hepatitis C virus subtype 2b (HCV2b) (GI:312299482), hepatitis C virus subtype 2c (HCV2c) (GI:395147275), hepatitis C virus subtype 2d (HCV2d) (GI:845452), hepatitis C virus subtype 2e (HCV2e) (GI:295684768), hepatitis C virus subtype 2f (HCV2f) (GI:295684656), hepatitis C virus subtype 3a (HCV3a) (GI:75566979), hepatitis C virus subtype 3b (HCV3b) (GI:290352647), hepatitis C virus subtype 3g (HCV3g) (GI:395147253), hepatitis C virus subtype 4a (HCV4a) (GI:110294850), hepatitis C virus subtype 4c (HCV4c) (GI:1369771), hepatitis C virus subtype 4d (HCV4d) (GI:576294948), hepatitis C virus subtype 4f (HCV4f) (GI:156469741), hepatitis C virus subtype 5a (HCV5a) (GI:262036445), hepatitis C virus subtype 6a (HCV6a) (GI:324330489), hepatitis C virus subtype 6h (HCV6h) (GI:295702049), hepatitis C virus subtype 6k (HCV6k) (GI:82617876), and hepatitis C virus subtype 6n (HCV6n) (GI:82617878).

The following proteins of the HCV have been examined: C (capsule, core protein), E1 (envelope glycoprotein 1), E2 (envelope glycoprotein 2, NS1), p7, NS2 (non-structural protein 2), NS3 (non-structural protein 3, protease/helicase), NS4a (non-structural protein 4a), NS4b (non-structural protein 4b), NS5a (non-structural protein 5a), NS5b (non-structural protein 5b, RNA polymerase), RNA_dep_RNAP (RNA-dependent RNA polymerase), HELICc (Helicase superfamily c-terminal domain), DEXDc (DEAD-like helicases superfamily), Peptidase_S29 (hepatitis C virus NS3 protease), and RdRP_3 (Viral RNA dependent RNA polymerase).

Tissue-specific antigens examined of thyroid gland

Were evaluated the following tissue-specific thyroid self-protein: thyroglobulin (Tg) (GI:55770862), and thyroid peroxidase (TPO) (GI:332078490), and thyrotropin receptor (TSHr) (GI:5705928).

The following proteins of the Tg polyprotein have been examined: TY super family (Thyroglobulin type I repeats), CO esterase (Carboxylesterase family), and GCC2_GCC3 (trinucleotide repeat polymorphism). The following proteins of the TPO polyprotein have been examined: EGF_CA (Calcium-binding EGF-like domain), and CCP (complement control protein). The following proteins of the TSHr polyprotein have been examined: 7tm_2 super family (7 transmembrane receptor - Secretin family), 7tm_1 (7 transmembrane receptor - rhodopsin family), and LRR_5 (Leucine rich repeats).

Exposed protein databases

HCV genotypes

Regarding HCV1a, the ID contained 3023 AA protein sequences⁽¹⁹⁾. In the case of HCV1b, the ID had 3010 AA protein sequences⁽¹⁹⁾. The HCV1c study allowed us to obtain

an ID that comprised 3011AA protein sequences⁽¹⁸⁾. Concerning HCV2a, we were able to assemble an ID of 3033 AA protein sequences⁽²¹⁾. The HCV2b, the ID contained 191 AA protein sequences⁽²⁹⁾. In the case of HCV2c, the ID had 3033 AA protein sequences⁽²⁷⁾. The HCV2d, the ID contained 149 AA protein sequences⁽³⁴⁾. The HCV2e study allowed us to obtain ID of 119 AA protein sequences⁽³³⁾. In the case of HCV2f, the ID had 93 AA protein sequences⁽³³⁾. Concerning HCV3a, we were able to assemble an ID of 3021 AA protein sequences⁽¹⁸⁾. The HCV3b study allowed us to obtain an ID that comprised 111 AA protein sequences⁽³¹⁾. The HCV3g study allowed us to obtain an ID that comprised 3022 AA protein sequences⁽²⁷⁾. In the case of HCV4a, the ID had 3009 AA protein sequences⁽¹³⁾. The HCV4c, the ID contained 319 AA protein sequences⁽³⁰⁾. In the case of HCV4d, the ID had 3006 AA protein sequences⁽³⁶⁾. The HCV4f study allowed us to obtain an ID that comprised 3010 AA protein sequences⁽¹⁶⁾. The HCV5a, the ID contained 186 AA protein sequences⁽¹⁵⁾. The HCV6a study allowed us to obtain an ID of 3019 AA protein sequences⁽³⁶⁾. In the case of HCV6h, the ID had 101 AA protein sequences⁽⁷⁾. The HCV6k study allowed us to obtain an ID that comprised 3015 AA protein sequences⁽²²⁾. The HCV6n, the ID contained 3015 AA protein sequences⁽²²⁾.

Tissue-specific antigens of thyroid gland

Regarding Tg, the ID contained 2768 AA protein sequences⁽⁹⁾. In the case of TPO, the ID had 933 AA protein sequences⁽⁸⁾. In the case of TSHr, the ID had 253 AA protein sequences⁽¹⁴⁾ (Figure 1).

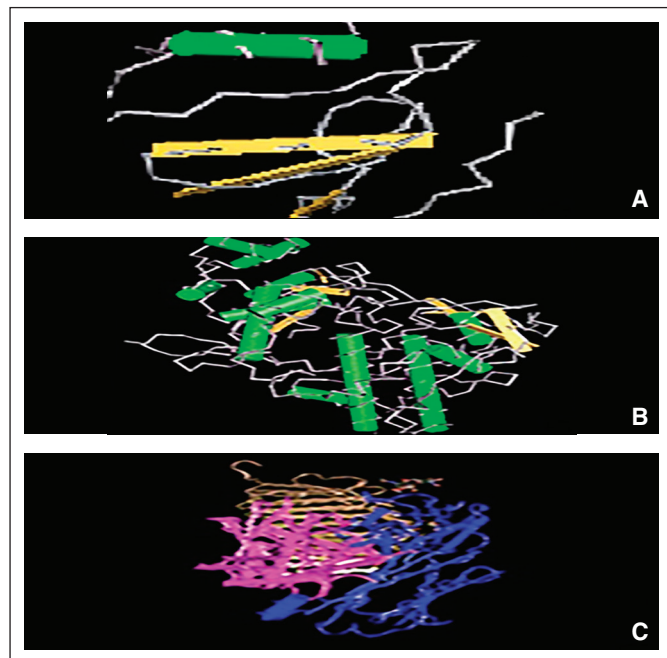


FIGURE 1. Sequence analysis and homology modeling of Tg, TPO and TSHr Structure of the membrane domain homology model. A) Tg; B) TPO; C) TSHr Source <http://www.ncbi.nlm.nih.gov/protein>

RESULTS

The sequence similarity was related each HCV genotype to each thyroid antigen. The similarities between the thyroid and the viral peptides ranged from 21.0% (31 identical residues out of 147AA in the sequence) to 71.0% (5 identical residues out of 7 AA in the sequence).

The AA sequences identities between the Tg, TPO, TSHr and the HCV1a poly-protein showed 60%, 33% and 30% sequences similarities respectively. The sequences similarities are shown in Figure 2. AA appears in standard single letter code. + indicates conserved or semi-conserved substitutions.

The AA sequences identities between the Tg, TPO, TSHr and the HCV1b poly-protein showed 35%, 33% and 58% sequence similarities respectively. The sequences similarities are shown in Figure 3. AA appears in standard single letter code. + indicates conserved or semi-conserved substitutions.

Score	Expect	Method	Identities	Positives	Gaps
20.4 bits(41)	5.5	Compositional matrix adjust.	6/10(60%)	8/10(80%)	0/10(0%)
Query 657	RCPTDCEKQR	666			
	RCPTDC ++				
Sbjct 580	RCPTDCFRKH	589			
Score	Expect	Method	Identities	Positives	Gaps
22.3 bits(46)	0.097	Compositional-based stats	14/47(30%)	21/47(44%)	2/47(4%)
Query 2134	APPCPKLLREEVSRVGLHEYPVGSQLPCEPEPDVAULTSMLTDPHS	2180			
	+PPC+ +E FRV + LP + + T + T PSH				
Sbjct 26	SPPCE - -CHQEEEDFRVICKDIQRIPSLPSTQTLKLIETHLRTIPSH	70			
Score	Expect	Method	Identities	Positives	Gaps
19.6 bits(39)	3.2	Compositional matrix adjust.	7/21(33%)	13/21(61%)	0/21(0%)
Query 2505	VEEACSLTPPHSAKSKFGYGA	2525			
	++ + TP ++K+ FG GA				
Sbjct 237	IDHDI AFTPQSTSKAAFGGGA	257			

Query: HCV1a. Sbjct: Tg, TPO, TSHr. +: conservative substitution.

FIGURE 2. The AA sequences homology between HCV1a and Tg; TPO; TSHr.

Score	Expect	Method	Identities	Positives	Gaps
20.4 bits(41)	5.7	Compositional matrix adjust.	9/26(35%)	12/26(46%)	0/26(0%)
Query 2121	FFSEVDGVRLLHR Y A PACRPLLREEVT	2146			
	F+++ CR LLREE T				
Sbjct 2156	FYADTQSCTHSL QGQNCRLLLREPEAT	2181			
Score	Expect	Method	Identities	Positives	Gaps
20.0 bits(40)	2.8	Compositional matrix adjust.	11/33(33%)	16/33(48%)	0/33(0%)
Query 792	GVVPLLL LLLALPPRAY AMDREMAASCG GAVFV	824			
	G+ PL+ LLA P + D+ M +FV				
Sbjct 534	GLDPLIRGLLARPAKLQVQDQLMNEELTERLFV	566			
Score	Expect	Method	Identities	Positives	Gaps
17.3 bits(33)	4.0	Composition-based stats.	7/12(58%)	8/12(66%)	0/12(0%)
Query 1510	GMFDSSVLCECY	1521			
	GM SS CEC+				
Sbjct 21	GMGCSSPPCECH	32			

Query: HCV1b. Sbjct: Tg, TPO, TSHr. +: conservative substitution.

FIGURE 3. The AA sequences homologies between HCV1b and Tg; TPO; TSHr.

The AA sequences identities between the Tg, TPO, TSHr and the HCV subtype 2a polyprotein showed 36%, 32% and 55% sequence similarities respectively. The AA sequences similarities between the Tg, TPO, TSHr and the HCV2b polyprotein showed 41%, 41% and 35% sequence similarities respectively. The AA sequences similarities between the Tg, TPO, TSHr and the HCV2c polyprotein showed 67%, 32% and 55% sequences similarities respectively. The AA sequences between the Tg, TPO, TSHr and the HCV2d polyprotein showed 71%, 30% respectively, and no significant sequences similarities with TSHr. The AA sequence similarities between the Tg, TPO, TSHr and the HCV2e polyprotein showed 57%, 56% and 43% sequences similarities respectively. The AA sequences similarities between the Tg, TPO, TSHr and the HCV2f polyprotein showed 67%, 36% and 26% sequences similarities respectively. The percentage AA sequences identities detected between the Tg, TPO, TSHr and the HCV subtype 2 polyprotein are shown in Table 1.

The AA sequences similarities between the Tg, TPO, TSHr and the HCV subtype 3a polyprotein showed 34%, 39% respectively, and no significant sequences similarities with TSHr. The AA sequences similarities between the Tg, TPO, TSHr and the HCV3b polyprotein showed 25%, 47% and 31% sequences similarities respectively. The AA sequences similarities between the Tg, TPO, TSHr and the HCV3g polyprotein showed 40%, 40% respectively, and no significant sequences similarities with TSHr. The percentage AA sequences similarities detected between the Tg, TPO, TSHr and the HCV subtype 3 polyprotein are shown in Table 2.

TABLE 1. Percentage AA sequences structural similarity between HCV subtype 2 and thyroid-specific antigens

Thyroid-specific antigens	Hepatitis C virus subtype 2					
	HCV2a	HCV2b	HCV2c	HCV2d	HCV2e	HCV2f
Tg	36%	41%	67%	71%	57%	67%
TPO	32%	41%	32%	30%	56%	36%
TSHr	55%	35%	55%	NS	43%	26%

NS: no significant.

TABLE 2. Percentage AA sequences structural similarity between HCV subtype 3 and thyroid-specific antigens

Thyroid-specific antigens	Hepatitis C virus subtype 3		
	HCV3a	HCV3b	HCV3g
Tg	34%	25%	40%
TPO	39%	47%	40%
TSHr	-	31%	-

The AA sequences similarities between the Tg, TPO, TSHr and the HCV subtype 4a polyprotein showed 28%, 26% and 63% sequences similarities respectively. The AA sequences similarities between the Tg and HCV4c polyprotein showed no significant, whereas the AA sequences similarities between the TPO, TSHr and HCV4c polyprotein showed 21% and 50% sequences similarities respectively. The AA sequences similarities between the Tg, TPO, TSHr and the HCV4d polyprotein showed 45%, 27% and 27% sequences similarities respectively. The AA sequences similarities between the Tg, TPO, TSHr and the HCV4f polyprotein showed 38%, 28% and 50% sequences similarities respectively. The percentage AA sequences similarities detected between the Tg, TPO, TSHr and the HCV subtype 4 polyprotein are shown in Table 3.

TABLE 3. Percentage AA sequences structural similarity between HCV subtype 4 and thyroid-specific antigens

Thyroid-specific antigens	Hepatitis C virus subtype 4			
	HCV4a	HCV4c	HCV4d	HCV4f
Tg	28%	NS	45%	38%
TPO	26%	21%	27%	28%
TSHr	63%	50%	27%	50%

NS: no significant.

The AA sequences similarities between the Tg, TPO, and the HCV5a polyprotein showed 44% and 24% respectively, whereas the AA sequences similarities between the TSHr and HCV5a polyprotein showed no significant of sequences similarities (Figure 4). AA appears in standard single letter code. + indicates conserved or semi-conserved substitutions.

Score	Expect	Method	Identities	Positives	Gaps
20.4 bits(41)	0.30	Composition-based stats.	7/16(44%)	9/16(56%)	0/16(0%)
Query 118	GRPAV I PDREVLVYQOF	133			
	G P + P R E + Q F				
Sbjct 971	GLPPLFPFPREAFAEQF	986			
Score	Expect	Method	Identities	Positives	Gaps
14.6 bits(26)	1.5	Compositional matrix adjust.	16/66(24%)	23/66(34%)	1/66(1%)
Query 37	KCLNRLKPTLTGPTPLLYRLGAVQNEITLTHPI TKYIMACMSADLEVI TSTWVLVGGVVA	96			
	+ L + P PLL LG + + +TK + LE IT + V				
Sbjct 112	RNLTYIDPDALKELPLLKFLGIFNTGLKMPDLTKVYSTDIFFILE -ITDNPYMTSIPVN	170			
Query 97	ALAAYC	102			
	A C				
Sbjct 171	AFQGLC	176			

Query: HCV5a. Sbjct: Tg, TPO, TSHr. +: conservative substitution.

FIGURE 4. The AA sequences identities between HCV5a and Tg; TPO; TSHr.

The AA sequences identities between the Tg, TPO, TSHr and the HCV subtype 6a polyprotein showed 53%, 50% and 35% sequences similarities respectively. The AA sequences similarities between the Tg, TPO, and the HCV6h polyprotein showed 53%, 28% respectively, whereas the AA sequences similarities between TSHr and HCV subtype 6h polyprotein showed no significant sequences similarities. The AA sequences similarities between the Tg, TPO, TSHr and the HCV6k polyprotein showed 27%, 25% and 47% sequences similarities respectively. The AA sequences similarities between the Tg, TPO, TSHr and the HCV6n polyprotein showed 26%, 31% and 63% sequences similarities respectively (Table 4).

TABLE 4. Percentage AA sequences structural similarity between HCV subtype 6 and thyroid-specific antigens

Thyroid-specific antigens	Hepatitis C virus subtype 6			
	HCV6a	HCV6h	HCV6k	HCV6n
Tg	53%	53%	27%	26%
TPO	50%	28%	25%	31%
TSHr	35%	NS	47%	63%

NS: no significant.

DISCUSSION

Several mechanisms have been proposed to explain induction of autoimmunity by infection and three mechanisms appear be important in individuals susceptible to developing autoimmune disease: molecular mimicry, polyclonal T cell activation, and MHC class II antigen induction⁽²⁴⁾.

This study suggests the possible role of molecular mimicry between the AA sequences of thyroid self-protein and proteins of HCV in development of ATD. The hypothesis that foreign antigens of proteins HCV sharing sequence similarity with the thyroid self-protein might drive autoimmunity, via a molecular mimicry mechanism, has been already suggested⁽²³⁾.

Thyroid disorders are common in patients with chronic HCV infection. Anti-thyroid antibodies are present in 5% to 17%, and thyroid disease occurs in 2% to 13% of patients with HCV infection⁽²⁾. Thyroid dysfunction and autoimmune thyroiditis were observed during antiviral therapy with interferon-alpha (IFN- α) and ribavirin therapy especially in patients with previous positivity for antithyroid antibodies. IFN- α therapy for HCV infection may induce thyroid changes or dysfunction in 2.5% to 20% of treated patients. Up to 40% become thyroid antibody positive and these side-effects can interfere with effective management of HCV⁽³²⁾.

The fact of a microorganisms develop antigen determinants that resemble antigens of the host as a mean of avoiding recognition and elimination by the host was first proposed

in 1964 by Damian⁽¹¹⁾. The molecular mimicry could explain the development of autoimmune diseases, through of cross reactions between epitopes of invading microorganisms and antigens present in the body that promote an adverse autoimmune response⁽¹⁰⁾.

In this work, we analyzed the sequence similarity between the AA sequences of the several subtypes HCV polyprotein and three tissue-specific antigens of human thyroid. We found that HCV and thyroid auto-antigens share AA sequence similarity, wherein some similar regions contain epitopes of both HCV and thyroid auto-antigens very high. A similar study showed that the similarities between the thyroid and the viral peptides were similar to ours, and the frequency of the sequence similarity increased up to 100%, when conservative substitutions were included in the analysis⁽²³⁾. In our study high similarities were found between Tg and HCV polyprotein subtype 2d, while the lowest percentage of sequence similarity between TPO and HCV polyprotein occurred with subtype 4c.

Given the sequence similarity between HCV sequence and thyroglobulin sequence and the HCV is classified in six major genotypes and over 70 subtypes, has been observed at possible association between genotype viral and susceptibility of patients in developing of ATD⁽¹⁷⁾. The presence of anti-Tg antibodies in the clinical course of the autoimmune thyroiditis in HCV infection treated with IFN- α may be related to the high number of similarities between the Tg antigen and the HCV polyprotein⁽⁶⁾. In our study significant similarities were found between Tg and HCV polyprotein subtype 2d, while the lowest percentage of sequence similarity between Tg and HCV polyprotein occurred with subtype 3b.

The TPO, that plays a key role in the biosynthesis of thyroid hormones, contains 933 AA protein sequences, and most of the extracellular region of TPO shows a high degree of sequence similarity with HCV polyprotein⁽²⁰⁾. AA sequences similarities between HCV polyprotein and TPO have been reported⁽²³⁾, however the similar viral peptides have not been examined for immune-pathogenicity. A study examining 20 amino acid-length peptides showed 41.7 to 58.3% sequence similarities between TPO and HCV5a, increased to 75.0% when including the conservative residues and identical residues⁽⁵⁾. Our results showed larger sequence similarity between TPO and HCV polyprotein subtype 2e, while the lowest percentage of sequence similarity between TPO and HCV polyprotein occurred with subtype 4c, and HCV polyprotein subtype 5a.

The TSHr regulates thyroid growth and differentiated functions, is also a member of G protein-coupled receptors and well known to undergo similar desensitization⁽²⁶⁾. Prior study already demonstrated molecular mimicry between the TSHr and HCV1a⁽³⁵⁾. Comparing the protein sequences of TSHr and HCV polyprotein, our study found similarities equality between TSHr and HCV polyprotein subtype 4a and 6n, while the lowest percentage of similarities between TSHr and HCV polyprotein occurred with subtype 5a. In our study there were no significant similarities between TSHr and HCV polyprotein subtype 2d, 3a, 3g and 6h.

Despite the limitations posed by the unavailability of complete proteome data for HCV proteins, because the BLAST is a method is limited to searching for linear epitope similarities, losing three-dimensional conformational similarities and possible cross-reactivity between proteic and non-proteic epitopes, the similarities between the AA sequences of thyroid self-protein, which are potential B- and T-cell epitopes of these antigens and proteins of HCV were successfully identified. Thus, these observed similarities could be functionally important in molecular mimicry, receptor binding and cell signaling events involved in autoimmunity, and has important implications for the understanding of the occurrence of autoimmune diseases.

In conclusion, it is important to draw attention that bioinformatics data suggest a possible pathogenic link between

HCV and ATD. Therefore, through of molecular mimicry is observed that sequences similarities between viral poly-proteins and self-proteins thyroid could be a mechanism of induction a cross over immune response to self-antigens resulting in autoimmune disease thyroid. Finally, our result adds not only new data, but also incorporates data to few already published studies.

Authors' contributions

Conception and design, writing the article, critical revision of the article, final approval of the article: Andrade LJO, Sousa MAC, Paraná R. Analysis and interpretation, data collection, statistical analysis: Andrade LJO, Sousa MAC. Overall responsibility: Andrade LJO.

Sousa MAC, Paraná R, Andrade LJO. Similaridade de seqüências entre auto-proteínas da tireoide e poliproteínas do vírus da hepatite C: possível mecanismo de desencadeamento da tireoidite auto-imune. *Arq Gastroenterol.* 2016,53(3):185-91.

RESUMO - Contexto - A exposição a antígenos virais que compartilham seqüência de aminoácidos semelhantes a auto-antígenos pode provocar doenças auto-imunes em indivíduos predispostos geneticamente, e a teoria do mimetismo molecular sugere que o mimetismo entre epítomos de vírus e proteínas humanas pode ativar doenças auto-imunes. **Objetivo** - O objetivo deste estudo foi explorar a possível semelhança entre as seqüências de aminoácidos de auto-proteínas da tireóide e proteínas do vírus da hepatite C, utilizando bancos de dados de proteínas e peptídeos imunogênicos, para explicar a doença auto-imune da tireóide. **Métodos** - Foram realizadas comparações entre as seqüências de aminoácidos de poliproteínas do vírus da hepatite C e auto-proteínas da tireóide humana, disponível na base de dados do National Center for Biotechnology Information no Basic Local Alignment Search Tool. **Resultados** - A semelhança de seqüências foi relacionada para cada genótipo de vírus da hepatite C e proteínas da tireóide. As semelhanças entre proteínas da tireóide e os peptídeos virais variaram de 21,0% (31 resíduos idênticos da seqüência de 147 aminoácidos) a 71,0% (cinco resíduos idênticos da seqüência de 7 aminoácidos). **Conclusão** - Dados de bioinformática sugerem uma possível ligação entre vírus da hepatite C e doença auto-imune da tireóide. Através de mimetismo molecular observa-se que as semelhanças entre as seqüências de poliproteínas virais e auto-proteínas da tireóide pode ser um mecanismo de indução de resposta imune resultando em doença auto-imune da tireóide.

DESCRITORES - Hepacivirus. Tireoidite. Homologia de seqüência de aminoácidos. Biologia computacional.

REFERENCES

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215:403-10.
- Andrade LJ, Atta AM, Atta ML, Mangabeira CN, Paraná R. Thyroid disorders in patients with chronic hepatitis C using interferon-alpha and ribavirin therapy. *Braz J Infect Dis.* 2011;15:377-81.
- Andrade LJ, Atta AM, D'Almeida Junior A, Paraná R. Thyroid dysfunction in hepatitis C individuals treated with interferon-alpha and ribavirin--a review. *Braz J Infect Dis.* 2008;12:144-8.
- Antonelli A, Ferri C, Pampana A, et al. Thyroid disorders in chronic hepatitis C. *Am J Med.* 2004;117:10-3.
- Benvenega S, Alesci S, Trimarchi F, Facchiano A. Homologies of the thyroid sodium-iodide symporter with bacterial and viral proteins. *J Endocrinol Invest.* 1999;22:535-40.
- Carella C, Mazziotti G, Amato G, Braverman LE, Roti E. Clinical review 169: Interferon-alpha-related thyroid disease: pathophysiological, epidemiological, and clinical aspects. *J Clin Endocrinol Metab.* 2004;89:3656-61.
- Chan DP, Lee SS, Lee KC. The effects of widespread methadone treatment on the molecular epidemiology of hepatitis C virus infection among injection drug users in Hong Kong. *J Med Virol.* 2011;83:1187-94.
- Cipollini M, Pastor S, Gemignani F, et al. TPO genetic variants and risk of differentiated thyroid carcinoma in two European populations. *Int J Cancer.* 2013;133:2843-51.
- Citterio CE, Rossetti LC, Souchon PF, et al. Novel mutational mechanism in the thyroglobulin gene: imperfect DNA inversion as a cause for hereditary hypothyroidism. *Mol Cell Endocrinol.* 2013;381:220-9.
- Cusick MF, Libbey JE, Fujinami RS. Molecular mimicry as a mechanism of autoimmune disease. *Clin Rev Allergy Immunol.* 2012;42:102-11.
- Damian, RT. Molecular mimicry: antigen sharing by parasite and host and its consequences. *Am Nat.* 1964; 98:129-49.
- De Carli M, D'Elios MM, Zancuoghi G, Romagnani S, Del Prete G. Human Th1 and Th2 cells: functional properties, regulation of development and role in autoimmunity. *Autoimmunity.* 1994; 18:301-8.
- Franco S, Tural C, Clotet B, Martínez MA. Complete nucleotide sequence of genotype 4 hepatitis C viruses isolated from patients co-infected with human immunodeficiency virus type 1. *Virus Res.* 2007;123:161-9.
- Graves PN, Tomer Y, Davies TF. Cloning and sequencing of a 1.3 KB variant of human thyrotropin receptor mRNA lacking the transmembrane domain. *Biochem Biophys Res Commun.* 1992;187:1135-43.
- Henquell C, Guglielmini J, Verbeeck J, et al. Evolutionary history of hepatitis C virus genotype 5a in France, a multicenter ANRS study. *Infect Genet Evol.* 2011;11:496-503.
- Hmaied F, Legrand-Abravanel F, Nicot F, et al. Full-length genome sequences of hepatitis C virus subtype 4f. *J Gen Virol.* 2007;88:2985-90.
- Huang MJ, Tsai SL, Huang BY, Sheen IS, Yeh CT, Liaw YF. Prevalence and significance of thyroid autoantibodies in patients with chronic hepatitis C virus infection: a prospective controlled study. *Clin Endocrinol (Oxf).* 1999;50:503-9.
- Kim CS, Seol SK, Song OK, Park JH, Jang SK. An RNA-binding protein, hnRNP A1, and a scaffold protein, septin 6, facilitate hepatitis C virus replication. *J Virol.* 2007;81:3852-65.
- Kuntzen T, Timm J, Berical A, et al. Naturally occurring dominant resistance mutations to hepatitis C virus protease and polymerase inhibitors in treatment-naïve patients. *Hepatology.* 2008;48:1769-78.
- Libert F, Ruel J, Ludgate M, et al. Thyroperoxidase, an auto-antigen with a mosaic structure made of nuclear and mitochondrial gene modules. *EMBO J.* 1987;6:4193-6.
- Lu J, Xiang Y, Tao W, et al. A novel strategy to develop a robust infectious hepatitis C virus cell culture system directly from a clinical isolate. *J Virol.* 2014;88:1484-91.
- Lu L, Nakano T, Li C, et al. Hepatitis C virus complete genome sequences identified from China representing subtypes 6k and 6n and a novel, as yet unassigned subtype within genotype 6. *J Gen Virol.* 2006;87:629-34.
- Martocchia A, Falaschi P. Amino acid sequence homologies between HCV polyprotein and thyroid antigens. *Intern Emerg Med.* 2007;2:65-7.
- Menconi F, Monti MC, Greenberg DA, et al. Molecular amino acid signatures in the MHC class II peptide-binding pocket predispose to autoimmune thyroiditis in humans and in mice. *Proc Natl Acad Sci U S A.* 2008 Sep 16;105:14034-9.
- Muratori L, Bogdanos DP, Muratori P, et al. Susceptibility to thyroid disorders in hepatitis C. *Clin Gastroenterol Hepatol.* 2005;3:595-603.
- Nagayama Y, Tanaka K, Hara T, et al. Involvement of G protein-coupled receptor kinase 5 in homologous desensitization of the thyrotropin receptor. *J Biol Chem.* 1996;271:10143-8.
- Newman RM, Kuntzen T, Weiner B, et al. Whole genome pyrosequencing of rare hepatitis C virus genotypes enhances subtype classification and identification of naturally occurring drug resistance variants. *J Infect Dis.* 2013;208:17-31.
- Oldstone MB. Molecular mimicry and immune-mediated diseases. *FASEB J.* 1998;12:1255-65.
- Palmer BA, Menton J, Levis J, Kenny-Walsh E, Crosbie O, Fanning LJ. The pan-genotype specificity of the hepatitis C virus anti-core monoclonal antibody C7-50 is contingent on the quasispecies profile of a population. *Arch Virol.* 2012;157:2235-9.
- Stuyver L, Wyseur A, van Arnhem W, et al. Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable samples. *Virus Res.* 1995;38:137-57.
- Tanimoto T, Nguyen HC, Ishizaki A, et al. Multiple routes of hepatitis C virus transmission among injection drug users in Hai Phong, Northern Vietnam. *J Med Virol.* 2010;82:1355-63.
- Tomer Y, Blackard JT, Akeno N. Interferon alpha treatment and thyroid dysfunction. *Endocrinol Metab Clin North Am* 2007;36:1051-66.
- Utama A, Tania NP, Dhenni R, et al. Genotype diversity of hepatitis C virus (HCV) in HCV-associated liver disease patients in Indonesia. *Liver Int.* 2010;30:1152-60.
- van Doorn LJ, Kleter GE, Stuyver L, et al. Sequence analysis of hepatitis C virus genotypes 1 to 5 reveals multiple novel subtypes in the Benelux countries. *J Gen Virol.* 1995;76:1871-6.
- Weiner AJ, Geysen HM, Christopherson C, et al. Evidence for immune selection of hepatitis C virus (HCV) putative envelope glycoprotein variants: potential role in chronic HCV infections. *Proc Natl Acad Sci U S A.* 1992;89:3468-72.
- Xu R, Tong W, Gu L, Li C, Fu Y, Lu L. A panel of 16 full-length HCV genomes was characterized in China belonging to genotypes 1-6 including subtype 2f and two novel genotype 6 variants. *Infect Genet Evol.* 2013;20:225-9.