

# Decoding the relationship between cow's milk proteins and development of type 1 diabetes mellitus

Luís Jesuino de Oliveira Andrade<sup>1</sup>  
<https://orcid.org/0000-0002-7714-0330>

Gabriela Correia Matos de Oliveira<sup>2</sup>  
<https://orcid.org/0000-0002-8042-0261>

Luísa Correia Matos de Oliveira<sup>3</sup>  
<https://orcid.org/0000-0001-6128-4885>

Alcina Maria Vinhaes Bittencourt<sup>4</sup>  
<https://orcid.org/0000-0003-0506-9210>

Yvana Baumgarth<sup>1</sup>  
<https://orcid.org/0009-0003-2419-9077>

Luís Matos de Oliveira<sup>5</sup>  
<https://orcid.org/0000-0003-4854-6910>

<sup>1</sup> Departamento de Saúde, Universidade Estadual de Santa Cruz, Ilhéus, BA, Brasil

<sup>2</sup> Programa Saúde da Família, Salvador, BA, Brasil

<sup>3</sup> Ecole Supérieure des Sciences et Technologies de l'Ingénierie de Nancy – Polytech Nancy, France

<sup>4</sup> Faculdade de Medicina, Universidade Federal da

Bahia, Salvador, BA, Brasil

<sup>5</sup> Escola Bahiana de Medicina e Saúde Pública, Salvador, BA, Brasil

## ABSTRACT

**Objective:** To analyze *in silico* the evidence of molecular mimicry between human beta-cell autoantigens and cow's milk proteins as a potential type 1 diabetes mellitus (T1DM) trigger.

**Materials and methods:** The *in silico* analysis was performed using bioinformatics tools to compare the amino acid sequences of cow's milk proteins (bovine serum albumin [BSA] and beta-lactoglobulin [BLG]) and human beta-cell autoantigens (glutamic acid decarboxylase-65 [GAD-65], insulin, and zinc transporter 8 [ZnT8]). The structural and functional characteristics of the proteins were analyzed to identify potential molecular mimicry mechanisms. **Results:** The results of the *in silico* analysis showed significant sequence similarity between BSA/BLG and GAD-65/human insulin/ZnT8, ranging from 19.64% to 27.27%. The cow's milk proteins evaluated shared structural features with the beta-cell antigens selected for comparison, indicating a potential for molecular mimicry between these proteins. **Conclusion:** The findings of this study provide further evidence for a potential role of cow's milk proteins in triggering T1DM. The *in silico* analysis suggests that molecular mimicry mechanisms between cow's milk proteins and human beta-cell antigens may contribute to the autoimmune response leading to T1DM.

## Keywords

Type 1 diabetes mellitus; cow's milk; autoantigens; molecular mimicry

## Correspondence to:

Luis Jesuino de Oliveira Andrade  
 Universidade Estadual de Santa Cruz,  
 Campus Soane Nazaré de Andrade  
 Rod. Jorge Amado, Km 16, Salobrinho  
 45662-900 – Ilhéus, BA, Brasil  
[luis\\_jesuino@yahoo.com.br](mailto:luis_jesuino@yahoo.com.br)

Received on June/14/2023  
 Accepted on Feb/8/2024

DOI: 10.20945/2359-4292-2023-0248

Preprint: June 22, 2023 (<https://doi.org/10.21203/rs.3.rs-3084369/v1>)

## INTRODUCTION

Molecular mimicry refers to the structural similarity of two or more molecules that may result in the human immune system mistaking a harmless substance for a pathogen. This is a common mechanism in the development of autoimmune diseases, where the immune system attacks healthy tissues and organs, leading to chronic inflammation and tissue damage (1).

Type 1 diabetes mellitus (T1DM) is an autoimmune disease caused by the destruction of pancreatic beta cells, resulting in insufficient production of insulin. The exact pathogenesis of T1DM remains unknown, but several factors have been implicated, including genetic predisposition, environmental triggers, and host

immune factors (2,3). Recent studies have suggested a molecular mimicry mechanism as a potential environmental trigger for T1DM development. Cow's milk proteins have been identified as potential trigger antigens in molecular mimicry involving glutamic acid decarboxylase-65 (GAD-65), human insulin, and zinc transporter 8 (ZnT8) in the development of T1DM (4). Cow's milk is a complex mixture of proteins, lipids, and carbohydrates that serves as a source of nutrition for newborn calves. However, some of the proteins in cow's milk have structural similarities to human proteins, which could trigger an autoimmune response in susceptible individuals. High levels of antibodies to cow's milk proteins – such as bovine serum albumin

(BSA) and beta-lactoglobulin (BLG) – have been associated with T1DM (5). Improved humoral immune response, characterized by increased IgG and IgA antibodies to protein fragments found in cow's milk, have been described in individuals with a recent diagnosis of T1DM (6).

The existing bioinformatics software for conducting structural analysis of protein components found in cow's milk is a crucial tool for assembling these proteins and comprehending their molecular mechanisms. Along the same lines, molecular modeling has aided in comprehending the concept of molecular mimicry, which occurs because of a cross-immune response to multiple antigens that resemble GAD-65, human insulin, and ZnT8 antigens. An essential prerequisite for effectively modeling similarity is the presence of a sufficiently similar protein sequence. Thus, the molecular mimicry between GAD-65/human insulin/ZnT8 and cow's milk proteins is a well-documented phenomenon that may contribute to the pathogenesis of T1DM by inducing cross-reactive immune responses against beta-cell antigens. These observations highlight the importance of dietary interventions and antigen-specific therapies in the prevention and treatment of autoimmune diseases.

In this study, we analyzed *in silico* the evidence of molecular mimicry between GAD-65/human insulin/ZnT8 and cow's milk proteins as a potential T1DM trigger.

## MATERIALS AND METHODS

To analyze the potential link between GAD-65/human insulin/ZnT8 and cow's milk proteins as a trigger for T1DM, we used the Protein Data Bank (PDB) database to obtain the structural information of the proteins and the VectorBuilder website (<https://en.vectorbuilder.com/>) to evaluate the similarity between these proteins.

The first step was to obtain the crystal structure data of GAD-65, human insulin, ZnT8, and cow's milk proteins (BSA and BLG) from the PDB database. Subsequently, the UniProt tool was used to visualize the sequence of these proteins.

We performed an immunoinformatic prediction of GAD-65, human insulin, ZnT8, and cow's milk proteins epitopes using the NetMHCpan v4.1 software (<https://services.healthtech.dtu.dk/service.php?NetMHCpan-4.1>) (7), based on artificial neural

networks, following the recommended methodology and filtering high-affinity epitopes based on an  $IC_{50} \leq 50$  nM and percentile rank  $\leq 0.20$  for the respective alleles analyzed. Subsequently, we performed a comparison of the epitopes to verify the hypothesis of cross-recognition. The epitope similarity analyses were performed online using VectorBuilder and complemented with Pairwise (<https://www.rcsb.org/alignment>), an extension of the PDB program, for conformational similarity analyses. NetMHCpan 4.1 is a server that uses artificial neural networks to filter high-affinity epitopes (8).

For the extraction of epitopes through NetMHCpan 4.1, we used the FASTA format obtained from the UniProt protein database website, selecting 11mer peptides and the human HLA-A\*01:01 allele.

Pairwise structural alignment is a method of comparing two proteins by aligning their amino acid sequences and identifying similarities and differences in their structures. In the analysis of the beta-cell autoantigens (GAD-65, human insulin, and ZnT8) and cow's milk proteins (BSA and BLG), we performed a pairwise alignment between the amino acid sequences of each protein to determine their degree of structural similarity. Once the sequences were obtained, pairwise structural alignment compared each amino acid in the two sequences to find matches and gaps in the sequence and then calculated the optimal alignment based on the highest number of matching amino acids.

The degree of similarity between the proteins was quantified based on the number and location of similar features in the structures. This method of analysis using pairwise structural alignment allows for a detailed comparison between two proteins and can provide insight into the similarities and differences in their structure and function.

For comparison of candidate epitopes, we used those that showed the strongest binding, programmed in NetMHCpan 4.1 between 0.5 and 2.0. The similarity between the sequences of GAD-65 epitopes, human insulin, and ZnT8 epitopes was evaluated using the UniProt BLAST tool.

To establish structural mimicry, we used UniProt's Align tool, which is based on a simple approach of maximizing the TM score and the quantitative evaluation of the root-mean-square deviation (RMSD). The server is available for free at <https://www.uniprot.org/align>.

**Ethical aspects**

According to Resolution CNS 510/2016, our study did not require approval from an ethics committee, since it aimed at deepening the theoretical understanding of situations that arise spontaneously and contingently in medical practice.

**RESULTS**

For the prediction of epitopes for GAD-65, we selected four candidate epitopes, all of which exhibited a threshold indicating stronger binding peptides (weighted binding [WB] 0.693-1.313) (Table 1).

For the prediction of epitopes for human insulin, two candidate epitopes were selected, both of which

exhibited a threshold indicative of stronger binding peptides (WB 1.153-1.892) (Table 2).

For the prediction of epitopes for ZnT8, four candidate epitopes were selected, two of which exhibited a threshold of stronger binding peptides (WB 0.346-1.984) (Table 3).

For the prediction of epitopes for BSA, four candidate epitopes were selected, all of which exhibited a threshold indicative of stronger binding peptides (WB 0.587-1.987) (Table 4).

For the prediction of epitopes for BLG, four candidate epitopes were selected, all of which exhibited a threshold indicative of stronger binding peptides (WB 0.587-1.987) (Table 5).

**Table 1.** Prediction of epitopes for GAD-65

Pos	MHC	Peptide	Core	Of	Gp	Gl	Ip	Il	Icore	Identity	Score_EL	%Rank_EL	BindLevel
36	HLA-A*01:01	FTGGIGNKLCALLY	FTGGIGNLY	0	7	5	0	0	FTGGIGNKLCALLY	Sequence	0.1947730	0.772	<= WB
124	HLA-A*01:01	KSFDRSTKVIDFHY	KSFDRSTHY	0	7	5	0	0	KSFDRSTKVIDFHY	Sequence	0.0979520	1.302	<= WB
125	HLA-A*01:01	SFDRSTKVIDFHYP	SFDRSTFHY	0	6	4	0	0	SFDRSTKVIDFHY	Sequence	0.0009260	19.556	
126	HLA-A*01:01	FDRSTKVIDFHYPN	FTKVIDFHY	0	1	3	0	0	FDRSTKVIDFHY	Sequence	0.0000250	82.500	
127	HLA-A*01:01	DRSTKVIDFHYPNE	DIDFHYPNE	0	1	5	0	0	DRSTKVIDFHYPNE	Sequence	0.0000090	93.571	
128	HLA-A*01:01	RSTKVIDFHYPNEL	RSTKVPNEL	0	5	5	0	0	RSTKVIDFHYPNEL	Sequence	0.0001060	56.600	
129	HLA-A*01:01	STKVIDFHYPNELL	STHYPNELL	0	2	5	0	0	STKVIDFHYPNELL	Sequence	0.0004860	28.167	
130	HLA-A*01:01	TKVIDFHYPNELLQ	THYPNELLQ	0	1	5	0	0	TKVIDFHYPNELLQ	Sequence	0.0000330	78.182	
131	HLA-A*01:01	KVIDFHYPNELLQ	KVIDELLOE	0	4	5	0	0	KVIDFHYPNELLQ	Sequence	0.0000340	77.727	
132	HLA-A*01:01	VIDFHYPNELLQ	VIDELLQ	0	3	5	0	0	VIDFHYPNELLQ	Sequence	0.2223700	0.693	<= WB
205	HLA-A*01:01	FTYEIAPVFLLEY	FTYEIALEY	0	6	5	0	0	FTYEIAPVFLLEY	Sequence	0.0967680	1.313	<= WB

Source: NetMHCpan 4.1 (7).

**Table 2.** Prediction of epitopes for human insulin

Pos	MHC	Peptide	Core	Of	Gp	Gl	Ip	Il	Icore	Identity	Score_EL	%Rank_EL	BindLevel
1	HLA-A*01:01	FVNQHLGSHL	FVNQHSHL	0	5	2	0	0	FVNQHLGSHL	Sequence	0.0011170	17.514	
2	HLA-A*01:01	VNQHLCGSHLV	VNQHSHLV	0	4	2	0	0	VNQHLCGSHLV	Sequence	0.0003960	31.333	
3	HLA-A*01:01	NQHLCGSHLVE	NQHSHLVE	0	3	2	0	0	NQHLCGSHLVE	Sequence	0.0000400	75.000	
4	HLA-A*01:01	QHLGSHLVEA	QHLSHLVEA	0	3	2	0	0	QHLGSHLVEA	Sequence	0.0001070	56.400	
5	HLA-A*01:01	HLGSHLVEAL	HLGSHLVI	0	8	2	0	0	HLGSHLVEAL	Sequence	0.0003370	34.000	
6	HLA-A*01:01	LCGSHLVEALY	LSHLVEALY	0	1	2	0	0	LCGSHLVEALY	Sequence	0.0539070	1.892	<= WB
7	HLA-A*01:01	CGSHLVEALYL	CSHLVEALY	0	1	1	0	0	CGSHLVEALY	Sequence	0.0003190	35.000	
8	HLA-A*01:01	GSHLVEALYLV	GSHLVEALY	0	0	0	0	0	GSHLVEALY	Sequence	0.0004570	29.138	
9	HLA-A*01:01	SHLVEALYLV	SVEALYLV	0	1	2	0	0	SHLVEALYLV	Sequence	0.0000440	73.462	
10	HLA-A*01:01	HLVEALYLVCG	HLVEALYLG	0	8	2	0	0	HLVEALYLVCG	Sequence	0.0000200	85.714	
11	HLA-A*01:01	LVEALYLVCGE	LVEALYLVCGE	0	7	2	0	0	LVEALYLVCGE	Sequence	0.0000150	89.286	
12	HLA-A*01:01	VEALYLVCGER	VEALYLV	0	7	2	0	0	VEALYLVCGER	Sequence	0.0001340	51.774	
13	HLA-A*01:01	EALYLVCGERG	EALYLVGERG	0	5	2	0	0	EALYLVCGERG	Sequence	0.0000130	90.714	
14	HLA-A*01:01	ALYLVCGERGF	ALYLVGERGF	0	4	2	0	0	ALYLVCGERGF	Sequence	0.0002710	37.933	
15	HLA-A*01:01	LYLVCGERGF	LVCGERGF	0	1	2	0	0	LYLVCGERGF	Sequence	0.0005280	26.946	
16	HLA-A*01:01	YLVCGERGFY	YLVGERGFY	0	2	2	0	0	YLVCGERGFY	Sequence	0.1179300	1.153	<= WB
17	HLA-A*01:01	LVCGERGFY	LVCGERGFY	0	3	1	0	0	LVCGERGFY	Sequence	0.0020240	12.530	
18	HLA-A*01:01	VCGERGFYTP	VCGERGFY	0	0	0	0	0	VCGERGFY	Sequence	0.0000790	62.381	
19	HLA-A*01:01	CGERGFYTPK	CGERGFYK	0	8	2	0	0	CGERGFYTPK	Sequence	0.0000360	76.818	
20	HLA-A*01:01	GERGFYTPK	GERGFYTPK	0	5	2	0	0	GERGFYTPK	Sequence	0.0000350	77.273	

Source: NetMHCpan 4.1 (7).



**Table 3.** Prediction of epitopes for ZnT8

Pos	MHC	Peptide	Core	Of	Gp	Gl	Ip	Il	Icore	Identity	Score_EL	%Rank_EL	BindLevel
1	HLA-A*02:01	MEFLERTYLVN	MFLERTYLV	0	1	1	0	0	MEFLERTYLV	Sequence	0.0000200	64.375	
2	HLA-A*02:01	EFLERTYLVND	EFLERTYLV	0	0	0	0	0	EFLERTYLV	Sequence	0.0000030	87.500	
3	HLA-A*02:01	FLERTYLVNDK	FLERTYLVK	0	8	2	0	0	FLERTYLVNDK	Sequence	0.0000640	47.500	
105	HLA-A*02:01	AHLLIDLTSFL	ALIDLTSFL	0	1	2	0	0	AHLLIDLTSFL	Sequence	0.4557030	0.451	<= SB
106	HLA-A*02:01	HLLIDLTSFLL	HLLIDLTSL	0	8	2	0	0	HLLIDLTSFLL	Sequence	0.0928300	1.984	<= WB
165	HLA-A*02:01	RLLYPDYQIQA	RLLDYQIQA	0	3	2	0	0	RLLYPDYQIQA	Sequence	0.5398340	0.346	<= SB
166	HLA-A*02:01	LLYPDYQIQAT	LLYPDIQAT	0	5	2	0	0	LLYPDYQIQAT	Sequence	0.1093580	1.760	<= WB
144	HLA-A*02:01	ALLSILCIWV	ALLSILIWV	0	6	1	0	0	ALLSILCIWV	Sequence	0.1026360	1.848	<= WB
228	HLA-A*02:01	SISVLISAL	SISVLISAL	0	0	0	0	0	SISVLISAL	Sequence	0.0970800	1.927	<= WB
229	HLA-A*02:01	ISVLI\$ALI	ISVLI\$ALI	0	0	0	0	0	ISVLI\$ALI	Sequence	0.0018660	13.946	

Source: NetMHCpan 4.1 (7).

**Table 4.** Prediction of epitopes for BSA

Pos	MHC	Peptide	Core	Of	Gp	Gl	Ip	Il	Icore	Identity	Score_EL	%Rank_EL	BindLevel
1	HLA-A*02:01	MKWVTFISLLL	MVTFISLLL	0	1	2	0	0	MKWVTFISLLL	Sequence	0.0004440	24.417	
2	HLA-A*02:01	KWVTFISLLL	KTFISLLL	0	1	2	0	0	KWVTFISLLL	Sequence	0.0014550	15.440	
3	HLA-A*02:01	WVTFISLLLLF	WVISLLLLF	0	2	2	0	0	WVTFISLLLLF	Sequence	0.0000300	58.636	
4	HLA-A*02:01	VTFISLLLLFS	VTFISLLLL	0	0	0	0	0	VTFISLLLL	Sequence	0.0000250	61.250	
5	HLA-A*02:01	TFISLLLLFSS	TFISLLLLSS	0	7	2	0	0	TFISLLLLFSS	Sequence	0.0000220	63.125	
6	HLA-A*02:01	FISLLLLFSSA	FLLLLFSSA	0	1	2	0	0	FISLLLLFSSA	Sequence	0.0022180	13.003	
7	HLA-A*02:01	ISLLLLFSSAY	ILLFSSAY	0	1	2	0	0	ISLLLLFSSAY	Sequence	0.0000460	52.333	
8	HLA-A*02:01	SLLLLFSSAYS	SLLLLSSAYS	0	4	2	0	0	SLLLLFSSAYS	Sequence	0.0008130	19.359	
9	HLA-A*02:01	LLLLLFSSAYS	LLFSSAYS	0	1	2	0	0	LLLLLFSSAYS	Sequence	0.0002880	28.667	
10	HLA-A*02:01	LLLLFSSAYSRG	LLFSSAYSG	0	8	2	0	0	LLLLFSSAYSRG	Sequence	0.0000500	51.000	
11	HLA-A*02:01	LLFSSAYSRGV	LLFSSAYSV	0	8	2	0	0	LLFSSAYSRGV	Sequence	0.1057850	1.801	<= WB
12	HLA-A*02:01	LFSSAYSRGVF	LFSSAYSRF	0	8	2	0	0	LFSSAYSRGVF	Sequence	0.0000100	73.750	
91	HLA-A*02:01	HTLFGDELCKV	TLFGDELKV	1	7	1	0	0	TLFGDELCKV	Sequence	0.0917910	1.997	<= WB
92	HLA-A*02:01	TLFGDELCKVA	TLFGDELKV	0	7	1	0	0	TLFGDELCKV	Sequence	0.0997340	1.891	<= WB
420	HLA-A*02:01	KLGEYGFQNAL	KLYGFQNAL	0	2	2	0	0	KLGEYGFQNAL	Sequence	0.3687980	0.587	<= WB
421	HLA-A*02:01	LGEYGFQNALI	LEYGFQNAL	0	1	1	0	0	LGEYGFQNAL	Sequence	0.0000170	66.667	

Source: NetMHCpan 4.1 (7).

**Table 5.** Prediction of epitopes for BLG

Pos	MHC	Peptide	Core	Of	Gp	Gl	Ip	Il	Icore	Identity	Score_EL	%Rank_EL	BindLevel
1	HLA-A*02:01	MKCLLLALALT	MLLLALALT	0	1	2	0	0	MKCLLLALALT	Sequence	0.0006540	21.079	
2	HLA-A*02:01	KCLLLALALT	KLLALALT	0	1	2	0	0	KCLLLALALT	Sequence	0.0024390	12.527	
3	HLA-A*02:01	CLLLALALTCG	CLLLALALG	0	8	2	0	0	CLLLALALTCG	Sequence	0.0000050	82.500	
46	HLA-A*02:01	SLLDAQSAPL	SLLDAQSAL	0	8	1	0	0	SLLDAQSAPL	Sequence	0.3697980	0.585	<= WB
47	HLA-A*02:01	LLDAQSAPLR	LLDAQSAPL	0	0	0	0	0	LLDAQSAPL	Sequence	0.0027550	11.884	
9	HLA-A*02:01	ALTCGAQAL	ALTCGAQAL	0	0	0	0	0	ALTCGAQAL	Sequence	0.0955580	1.947	<= WB
10	HLA-A*02:01	LTCGAQALI	LTCGAQALI	0	0	0	0	0	LTCGAQALI	Sequence	0.0039000	10.255	
148	HLA-A*02:01	ALEKFDKAL	ALEKFDKAL	0	0	0	0	0	ALEKFDKAL	Sequence	0.1231090	1.610	<= WB
149	HLA-A*02:01	LEKFDKALK	LEKFDKALK	0	0	0	0	0	LEKFDKALK	Sequence	0.0000190	65.000	
150	HLA-A*02:01	EKFDKALKA	EKFDKALKA	0	0	0	0	0	EKFDKALKA	Sequence	0.0004510	24.271	
151	HLA-A*02:01	KFDKALKAL	KFDKALKAL	0	0	0	0	0	KFDKALKAL	Sequence	0.0164200	5.271	
152	HLA-A*02:01	FDKALKALP	FDKALKALP	0	0	0	0	0	FDKALKALP	Sequence	0.0000150	68.333	
153	HLA-A*02:01	DKALKALPM	DKALKALPM	0	0	0	0	0	DKALKALPM	Sequence	0.0000250	61.250	
154	HLA-A*02:01	KALKALPMH	KALKALPMH	0	0	0	0	0	KALKALPMH	Sequence	0.0003200	27.548	
155	HLA-A*02:01	ALKALPMHI	ALKALPMHI	0	0	0	0	0	ALKALPMHI	Sequence	0.3445090	0.643	<= WB
156	HLA-A*02:01	LKALPMHIR	LKALPMHIR	0	0	0	0	0	LKALPMHIR	Sequence	0.0000590	48.600	
157	HLA-A*02:01	KALPMHIRL	KALPMHIRL	0	0	0	0	0	KALPMHIRL	Sequence	0.1059820	1.799	<= WB
158	HLA-A*02:01	ALPMHIRLS	ALPMHIRLS	0	0	0	0	0	ALPMHIRLS	Sequence	0.0096390	6.856	

Source: NetMHCpan 4.1 (7).

### Analyses of epitope similarities between the human beta-cell antigens and cow's milk proteins

The percentage of similarity between the strongest binding GAD-65 epitope and the BSA epitopes was 46.50%, while the percentage of similarity between the strongest binding GAD-65 epitope and the BLG epitopes was 21.83% (Figures 1 and 2).

The percentage of similarity between the strongest binding human insulin epitope and the BSA epitopes was 14.49%, while the percentage of similarity between the strongest binding human insulin epitope and the BLG epitopes was 14.49% (Figures 3 and 4).



The percentage of similarity between the strongest binding ZnT8 epitope and the BSA epitopes was 54.87%, while the percentage of similarity between the strongest binding human insulin epitope and the BLG epitopes was 27.66% (Figures 5 and 6).

The conformational similarity analysis between cow's milk proteins and GAD-65/human insulin/ZnT8 proteins involved comparing their structures to identify homology and structural similarities. By evaluating the degree of structural overlap and comparing key interaction residues, we were able to gain insights into potential functional and immunological implications. This allowed us to explore potential cross-reactivity and identify structural similarities that may be involved in autoimmune responses or other biological interactions related to the onset of T1DM, as demonstrated below. Our results from the conformational similarity analysis between cow's milk

proteins and GAD-65/human insulin/ZnT8 proteins showed a structural comparison.

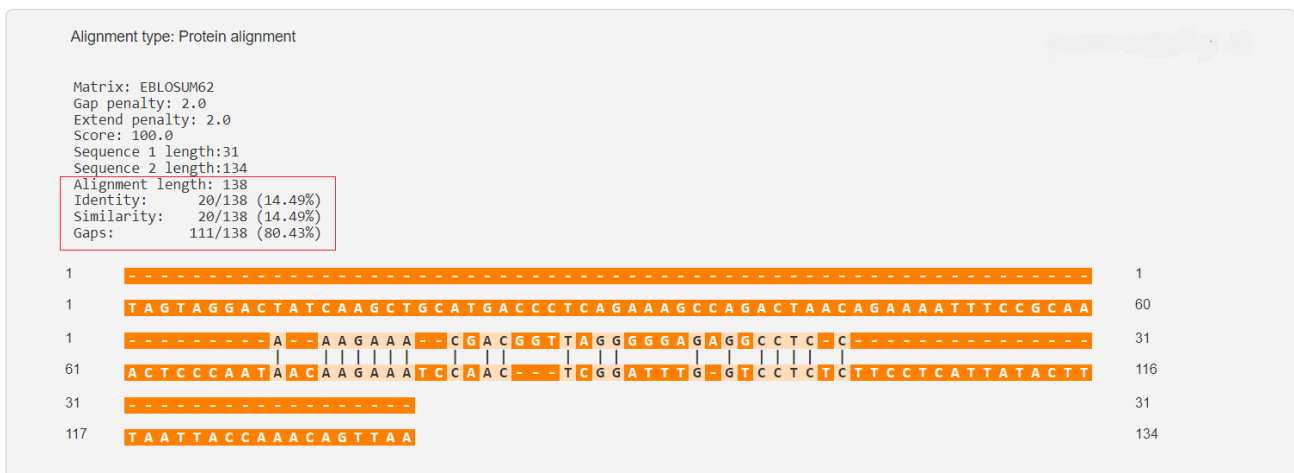
### Analysis of conformational similarities between the human beta-cell antigens and cow's milk proteins

The pairwise structure alignment analysis showed a sequence similarity between GAD-65 and BSA, with a 46.50% identity (Figure 7).

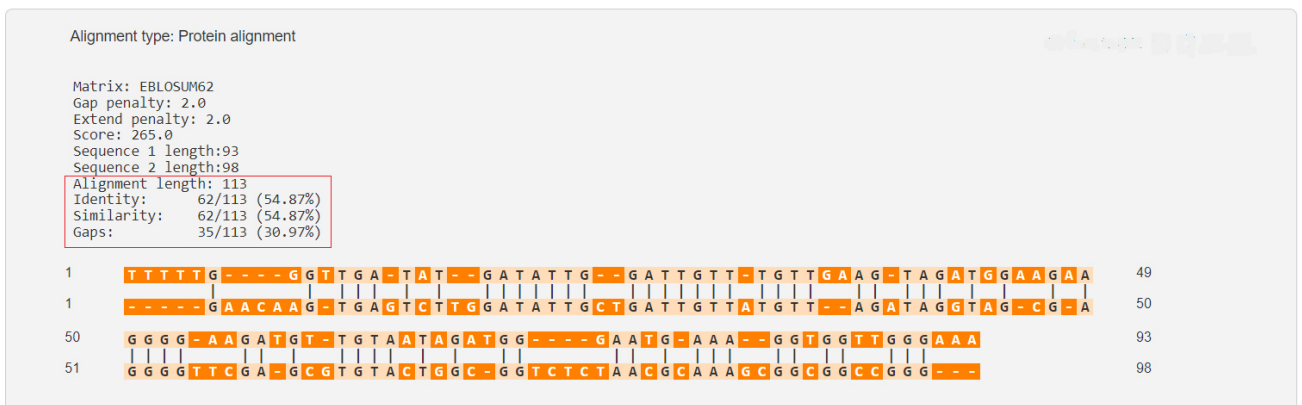
The pairwise structure alignment analysis showed a sequence similarity between GAD-65 and BLG, with a 40.00% identity (Figure 8).

The pairwise structure alignment analysis showed a significant sequence similarity between the human insulin and BLG, which shared 57.00% of sequence identity (Figure 9).

The pairwise structure alignment analysis showed structural similarity between human insulin and BSA, which shared 33.00% of sequence identity (Figure 10).

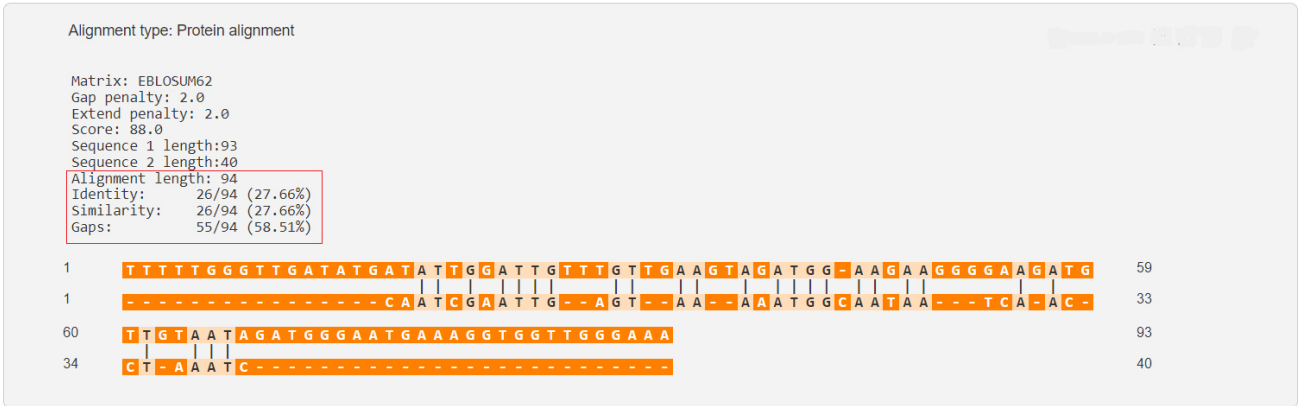


**Figure 4.** Human insulin (P01308 · INS\_HUMAN) versus BSA (P02769 · ALBU\_BOVIN).  
 Source: <https://en.vectorbuilder.com/tool/sequence-alignment.html>.

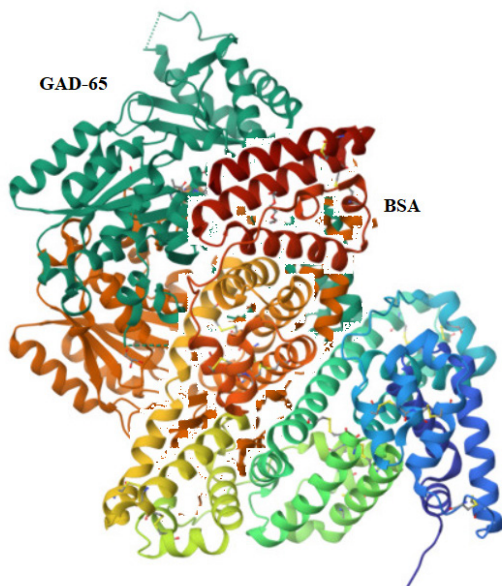


**Figure 5.** ZnT8 (Q8IWU4 · ZNT8\_HUMAN) versus BLG (P02754 · LACB\_BOVIN).  
 Source: <https://en.vectorbuilder.com/tool/sequence-alignment.html>.

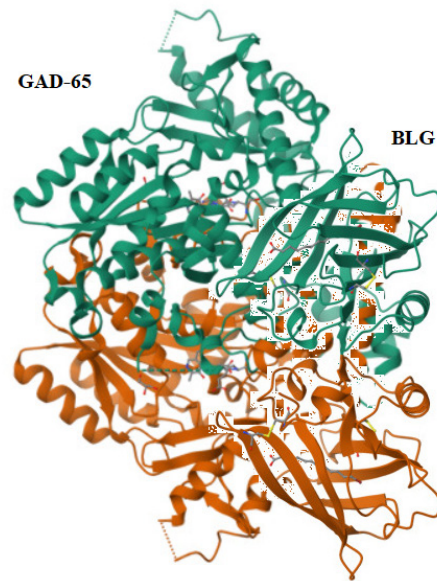




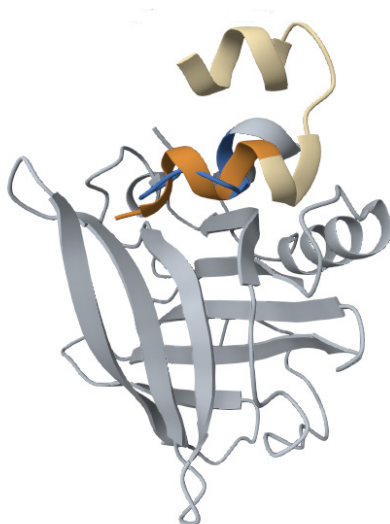
**Figure 6.** ZnT8 (Q8IWU4 · ZNT8\_HUMAN) versus bovine serum albumin (P02769 · ALBU\_BOVIN)  
 Source: <https://en.vectorbuilder.com/tool/sequence-alignment.html>.



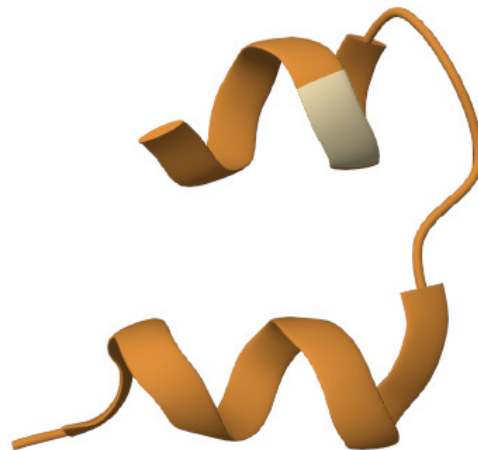
**Figure 7.** Structure alignment analysis between GAD-65 and BSA.  
 Source: <https://www.rcsb.org/alignment>.



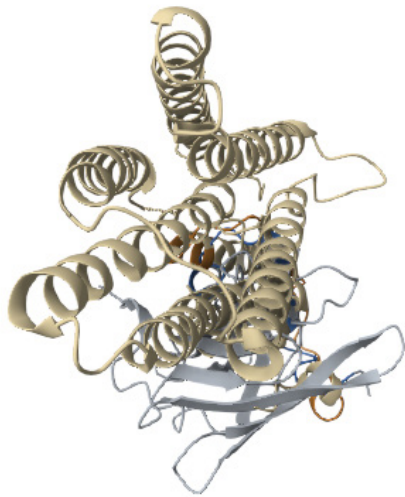
**Figure 8.** Structure alignment analysis between GAD-65 and BLG.  
 Source: <https://www.rcsb.org/alignment>.



**Figure 9.** Structure alignment analysis between human insulin and BLG.  
 Source: <https://www.rcsb.org/alignment>.



**Figure 10.** Structure alignment analysis between human insulin and BSA.  
 Source: <https://www.rcsb.org/alignment>.



**Figure 11.** Structure alignment analysis between ZnT8 and BLG.  
Source: <https://www.rcsb.org/alignment>.



**Figure 12.** Structure alignment analysis between ZnT8 and BSA.  
Source: <https://www.rcsb.org/alignment>.

Human insulin and BSA shared structural similarities in their tertiary structures, as both proteins have alpha helices and beta sheets. Additionally, both proteins have a similar fold in their binding domains. Overall, they have a RMSD value of 2.00 Å, indicating good similarity.

The pairwise structure alignment analysis showed a significant sequence similarity between the ZnT8 and BLG, which shared 41.00% sequence identity (Figure 11). The structural similarity between human insulin and BSA showed 18.29% sequence identity (Figure 12).

Despite potential sequence differences, conformational similarity and sequence identity provided valuable information. Although amino acid sequences may differ between pairs of proteins, and a simple analysis of the amino acid sequences may result in false-positive results, we did not exclude the analysis of similarity between cow's milk proteins and GAD-65/human insulin/ZnT8 proteins.

## DISCUSSION

The molecular mimicry mechanism between cow's milk proteins and human insulin and GAD-65 provides a potential environmental trigger for T1DM development in genetically predisposed individuals. The present study investigated the possibility of the cow's milk proteins BSA and BLG sharing homology with GAD-65, human insulin, and ZnT8 and causing T1DM.

In 1987, Damian proposed that the resemblance of an antigen protein to that of the host was a mechanism

to evade recognition and elimination by the host (9). This phenomenon, known as molecular mimicry, may contribute to autoimmune diseases by triggering an adverse autoimmune response through cross-reactivity between protein epitopes and antigens present in the body (10).

The investigation of molecular mimicry using bioinformatics in autoimmune diseases plays a crucial role in the understanding and potential treatment of these complex conditions. By utilizing computational tools and techniques, researchers can analyze and identify similarities between self-antigens and microbial antigens, which may trigger autoimmune responses. Bioinformatics approaches offer a comprehensive and systematic examination of the immune system, enabling the identification of potential molecular mimics and their impact on autoimmunity. Recent studies have highlighted the significance of bioinformatics-based approaches in unraveling the molecular mechanisms underlying autoimmune diseases and discovering novel therapeutic targets (11,12). These findings emphasize the need for further research and collaboration in this field to improve diagnostic accuracy and develop personalized treatment strategies for individuals affected by autoimmune disorders.

One of the environmental factors that have been implicated in the development of T1DM is cow's milk. Cow's milk contains several proteins that may trigger an autoimmune response in susceptible individuals. Two of these proteins are BLG and BSA (13).

Several studies have investigated the role of molecular mimicry between cow's milk proteins and



GAD-65/human insulin/ZnT8 in the development of T1DM. Some studies have reported that cow's milk consumption during infancy is associated with an increased risk of developing T1DM later in life, while others have failed to find a significant association. One of the limitations of these studies is the difficulty in establishing an accurate measure of cow's milk consumption. Cow's milk is often used as an ingredient in many processed foods, hindering an accurate assessment of the amount of cow's milk consumed by an individual.

One of the earliest studies on molecular mimicry in T1DM was conducted by Elliott and cols. (14). The authors used monoclonal antibodies to examine the cross-reactivity between GAD-65 and several proteins, including BSA and BLG, the two most abundant proteins in cow's milk. The study found that some of the monoclonal antibodies created against BSA and BLG recognized GAD-65, indicating a structural similarity between the two proteins. However, the authors did not investigate the clinical implications of this observation.

Another study provided evidence for the role of molecular mimicry in the development of T1DM. The authors demonstrated that feeding cow's milk formula to infants who carried the HLA-DQB1\*0201 allele, a genetic risk factor for T1DM, increased the risk of developing islet autoantibodies. They also showed that the autoantibodies recognized an epitope in BSA that shared structural similarity with GAD-65. The authors concluded that cow's milk proteins could induce the production of autoantibodies against GAD-65 through molecular mimicry, which could contribute to the development of T1DM (15).

Notably, BSA has been shown to share structural similarities with GAD-65 (16). The molecular mimicry between cow's milk proteins and GAD-65 has been suggested as another potential trigger of T1DM. As an enzyme in pancreatic beta cells, GAD-65 catalyzes the conversion of glutamate to gamma-aminobutyric acid. Additionally, GAD-65 is a major autoantigen in T1DM, and autoantibodies to GAD-65 are detected in 60–70% of individuals with newly diagnosed T1DM, suggesting that the destruction of beta cells may be mediated by the immune response against GAD-65 (17).

Despite their different functions and structures, BSA and GAD-65 share a certain degree of amino acid similarity. Several studies have reported sequence

alignments and comparisons between BSA and GAD-65 using bioinformatics tools. A study using multiple sequence alignment to compare the amino acid sequences of BSA and GAD-65 found that the two proteins presented 29% sequence identity and 48% sequence similarity, with most of the similar residues located in the flexible loops and turns of the proteins (18). Our study, using pairwise structure alignment analysis, compared the amino acid sequences of BSA and GAD-65. The evaluation between these two proteins showed a lower percentage of sequence similarities than other studies in the literature, with most of the similar residues located in the flexible loops and bends of the proteins. In addition, we used homology modeling to generate a three-dimensional structure of BSA and compared it to the crystal structure of GAD-65. Our results showed that BSA shares an overall fold similar to that of GAD-65, which has not been described in previous studies.

We found no studies in the literature evaluating the similarity between GAD-65 and BLG. In our study, we compared the amino acid sequences of BLG and GAD-65, which showed 21.83% sequence identity and 40.00% sequence similarity between these two proteins.

Insulin and BLG, despite their distinct functions, share a resemblance in their amino acid sequences. Several studies have suggested that the amino acid sequence of BLG contains regions that resemble the amino acid sequence of human insulin, thereby leading to molecular mimicry. Molecular mimicry occurs when a foreign antigen (in this case, BLG) shares structural similarities with a self-antigen (in this case, human insulin), leading to cross-reactivity of T cells and autoimmunity (19).

Several *in silico* studies have evaluated the molecular mimicry between insulin and cow's milk proteins in triggering T1DM. These studies aimed to investigate the structural and immunological similarities between the two proteins and their potential role in the development of autoimmune reactions leading to diabetes. Previous studies have shown that antibody cross-reactivity between human insulin and BLG is possible, indicating that molecular mimicry may be a contributing factor (20). However, we found no *in silico* studies in the literature comparing the similarity between BLG and human insulin. Our study using pairwise structure alignment analysis compared the amino acid sequences of BLG and human insulin and found that these two

proteins share a 57.00% sequence identity and 14.49% sequence similarity and structurally show a required minimum distribution indicative of good similarity.

Reviewing the medical literature, we found no studies in which the similarity between human insulin and BSA was assessed. In our study, we compared the amino acid sequences of human insulin and BSA and found a sequence identity of 33.0% and a sequence similarity of 14.49% between these two proteins.

An analysis focused on discovering beta-cell-specific proteins associated with the regulatory pathway of secretion recently identified ZnT8 as a novel autoantigen (21). Found in the membrane of insulin-containing secretory granules, ZnT8 is responsible for transporting zinc ions from the cytosol into the vesicles (22). Additionally, ZnT8 has emerged as a potential target for humoral autoimmunity, making it highly relevant for the early detection of T1DM. The presence of ZnT8-specific CD8<sup>+</sup> T cells can be observed in the majority of T1DM patients, and these cells play a crucial role in the development of T1DM. As a target for immunotherapy, ZnT8 offers the potential to ameliorate beta-cell dysfunction in T1DM, presenting new avenues for the treatment of this condition (23).

It has been reported that anti-ZnT8 antibodies that target homologous membrane extension sequences exhibit cross-reactivity and have the ability to elicit robust immune responses, thus raising the possibility of a molecular mimicry mechanism that triggers T1DM (24). The literature suggests that infection with the *Mycobacterium avium* subspecies paratuberculosis is linked to T1DM by producing anti-ZnT8 antibodies via molecular homology (25). However, there are no studies available related to molecular mimicry between cow's milk proteins and ZnT8. Our study compared the amino acid sequences of BSA and ZnT8. The evaluation between these two proteins showed 54.87% sequence identity and 18.29% sequence similarity. When comparing the amino acid sequences of ZnT8 and BSA, we found a sequence identity of 41.00% and a sequence similarity of 27.66% between these two proteins.

Molecular modeling has enabled the comprehension of molecular mimicry caused by the cross-immune reaction to cow's milk protein-like antigens (BSA/BLG) with GAD-65/human insulin/ZnT8. The majority of molecular mimicry probably includes T-cell mediation, as T cells typically detect linear peptides that vary from 8 to 20 amino acids long (26).

Post-translational modifications refer to chemical alterations that occur in a protein after the translation of its genetic code. These modifications can involve the addition, removal, or alteration of functional groups in specific amino acid residues of the protein (27). Post-translational modifications play a crucial role in regulating the structure, function, and subcellular localization of proteins. They can influence enzymatic activity, protein stability, interaction with other proteins, and even cellular signaling (28). The analysis of post-translational modifications was not utilized in our study due to its focus on the primary sequence of the protein and its three-dimensional structure. In these cases, obtaining information about the specific composition of amino acids and their spatial organization is more important than investigating the chemical alterations that occur after genetic code translation.

A study *in silico*, similar to ours, demonstrated that BSA and BLG share sequence and structural homology with GAD-65 (16). However, our study is more comprehensive in terms of the larger number of proteins evaluated compared with the cited study. Investigations in molecular mimicry using bioinformatics are critically important in the understanding and management of autoimmune diseases. Bioinformatics approaches enable the identification and analysis of these mimetic epitopes, providing insights into the mechanisms underlying autoimmunity and aiding in the development of targeted therapies. Studies have shown the utility of bioinformatics in unraveling molecular mimicry in diseases (29). Thus, by elucidating the molecular basis of autoimmune reactions, these studies pave the way for precision medicine approaches and personalized therapeutic interventions.

Although similarities have been detected between BSA, BLG, GAD-65, human insulin, and ZnT8, a randomized clinical trial including 2,159 children from 15 countries explored the impact of excluding cow's milk proteins from infant formula on the development of T1DM in high-risk infants. The study found that removing cow's milk proteins from infant formulas did not prevent the development of T1DM in these high-risk children (30). This randomized international study included children with an HLA genotype associated with increased risk of T1DM and a first-degree relative affected by the disease were included. The majority of participants came from Canada, Finland, and the United States. The larger sample size of this study

provided greater statistical power compared with the pilot study (15), allowing for a more definitive conclusion on whether transitioning to an extensively hydrolyzed formula protects against diabetes. The study was designed to have two endpoints, specifically, the presence of two autoantibodies at age 6 and clinical T1DM at age 10. However, it is important to note that the results cannot be directly applied to the general population, as the participants were selected based on a positive family history and a specific HLA genotype associated with the risk of T1DM. Moreover, these findings may not be applicable to children with different HLA genotypes. However, it is worth noting the contrasting results obtained when compared with those of other studies (31-35).

In view of the above, we must consider the several limitations of *in silico* studies that focus on identifying how molecular mimicry between human pancreatic beta-cell antigens and cow's milk proteins may trigger T1DM. First, the molecular mimicry mechanism is not the only environmental trigger for T1DM development. Other factors, such as viral infections, gut microbiota, and dietary factors, may also contribute to T1DM pathogenesis (36). Second, the evidence for the molecular mimicry between cow's milk proteins and human insulin, GAD-65, and ZnT8 is mainly based on *in vitro* studies, and the clinical significance of these findings is still unclear.

In conclusion, the *in silico* analysis in the present study suggests that molecular mimicry mechanisms between proteins from cow's milk and human pancreatic beta-cell antigens may contribute to the autoimmune response that may trigger T1DM.

Disclosure: no potential conflict of interest relevant to this article was reported.

## REFERENCES

- Rojas M, Restrepo-Jiménez P, Monsalve DM, Pacheco Y, Acosta-Ampudia Y, Ramírez-Santana C, et al. Molecular mimicry and autoimmunity. *J Autoimmun.* 2018;95:100-23. doi: 10.1016/j.jaut.2018.10.012.
- Ilonen J, Lempainen J, Veijola R. The heterogeneous pathogenesis of type 1 diabetes mellitus. *Nat Rev Endocrinol.* 2019;15(11):635-50. doi: 10.1038/s41574-019-0254-y.
- Coppieters KT, Wiberg A, von Herrath MG. Viral infections and molecular mimicry in type 1 diabetes. *APMIS.* 2012;120(12):941-9. doi: 10.1111/apm.12011.
- Karissón MG, García J, Ludvigsson J. Cows' milk proteins cause similar Th1- and Th2-like immune response in diabetic and healthy children. *Diabetologia.* 2001;44(9):1140-7. doi: 10.1007/s001250100611.
- Persaud DR, Barranco-Mendoza A. Bovine serum albumin and insulin-dependent diabetes mellitus; is cow's milk still a possible toxicological causative agent of diabetes? *Food Chem Toxicol.* 2004;42(5):707-14. doi: 10.1016/j.fct.2004.01.002.
- Virtanen SM, Läärä E, Hyppönen E, Reijonen H, Räsänen L, Aro A, et al. Cow's milk consumption, HLA-DQB1 genotype, and type 1 diabetes: a nested case-control study of siblings of children with diabetes. Childhood diabetes in Finland study group. *Diabetes.* 2000;49(6):912-7. doi: 10.2337/diabetes.49.6.912.
- NetMHCpan-4.1 [internet]. Available from: <https://services.healthtech.dtu.dk/service.php?NetMHCpan-4.1>. Accessed on: Jun 24, 2023.
- Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol.* 2017;199(9):3360-8. doi: 10.4049/jimmunol.1700893.
- Damian RT. Molecular mimicry revisited. *Parasitol Today.* 1987;3(9):263-6. doi: 10.1016/0169-4758(87)90102-5.
- Cusick MF, Libbey JE, Fujinami RS. Molecular mimicry as a mechanism of autoimmune disease. *Clin Rev Allergy Immunol.* 2012;42(1):102-11. doi: 10.1007/s12016-011-8294-7.
- Wang J, Xue Y, Zhou L. Comparison of immune cells and diagnostic markers between spondyloarthritis and rheumatoid arthritis by bioinformatics analysis. *J Transl Med.* 2022;20(1):196. doi: 10.1186/s12967-022-03390-y.
- Zhang Y, Garofano F, Wu X, Schmid M, Krawitz P, Essler M, et al. Integrative analysis of key candidate genes and signaling pathways in autoimmune thyroid dysfunction related to anti-CTLA-4 therapy by bioinformatics. *Invest New Drugs.* 2020;38(6):1717-29. doi: 10.1007/s10637-020-00952-z.
- Wasmuth HE, Kolb H. Cow's milk and immune-mediated diabetes. *Proc Nutr Soc.* 2000;59(4):573-9. doi: 10.1017/s0029665100000811.
- Elliott RB, Harris DP, Hill JP, Bibby NJ, Wasmuth HE. Type I (insulin-dependent) diabetes mellitus and cow milk: casein variant consumption. *Diabetologia.* 1999;42(3):292-6. doi: 10.1007/s001250051153.
- Vaarala O, Knip M, Paronen J, Hämäläinen AM, Muona P, Väättäinen M, et al. Cow's milk formula feeding induces primary immunization to insulin in infants at genetic risk for type 1 diabetes. *Diabetes.* 1999;48(7):1389-94. doi: 10.2337/diabetes.48.7.1389.
- Benslama Y, Dennouni-Medjati N, Dali-Sahi M, Meziane FZ, Harek Y. Childhood type 1 diabetes mellitus and risk factor of interactions between dietary cow's milk intake and HLA-DR3/DR4 genotype. *J Biomol Struct Dyn.* 2022;40(21):10931-9. doi: 10.1080/07391102.2021.1953599.
- Miao D, Guyer KM, Dong F, Jiang L, Steck AK, Rewers M, et al. GAD65 autoantibodies detected by electrochemiluminescence assay identify high risk for type 1 diabetes. *Diabetes.* 2013;62(12):4174-8. doi: 10.2337/db13-0534.
- Guo L, Huang L, Liu J, Dou C, Song F, Wang Y, et al. Exploration of the structural basis for substrate specificity in glutamate decarboxylase. *Sci Rep.* 2018;8(1):2534.
- Dahlquist G, Savilahti E, Landin-Olsson M. An increased level of antibodies to beta-lactoglobulin is a risk determinant for early-onset type 1 (insulin-dependent) diabetes mellitus independent of islet cell antibodies and early introduction of cow's milk. *Diabetologia.* 1992;35(10):980-4. doi: 10.1007/BF00401429.
- Savilahti E, Saukkonen TT, Virtala ET, Tuomilehto J, Akerblom HK. Increased levels of cow's milk and beta-lactoglobulin antibodies in young children with newly diagnosed IDDM. The Childhood Diabetes in Finland Study Group. *Diabetes Care.* 1993;16(7):984-9. doi: 10.2337/diacare.16.7.984.
- Wenzlau JM, Juhl K, Yu L, Moua O, Sarkar SA, Gottlieb P, et al. The cation efflux transporter ZnT8 (Slc30A8) is a major autoantigen in human type 1 diabetes. *Proc Natl Acad Sci U S A.* 2007;104(43):17040-5. doi: 10.1073/pnas.0705894104.
- Kelleher SL, McCormick NH, Velasquez V, Lopez V. Zinc in specialized secretory tissues: roles in the pancreas, prostate, and mammary gland. *Adv Nutr.* 2011;2(2):101-11. doi: 10.3945/an.110.000232.

23. Zhang X, Dong Y, Liu D, Yang L, Xu J, Wang Q. Antigen-specific immunotherapies in type 1 diabetes. *J Trace Elem Med Biol.* 2022;73:127040. doi: 10.1016/j.jtemb.2022.127040.
24. Masala S, Paccagnini D, Cossu D, Brezar V, Pacifico A, Ahmed N, et al. Antibodies recognizing *Mycobacterium avium* paratuberculosis epitopes cross-react with the beta-cell antigen ZnT8 in Sardinian type 1 diabetic patients. *PLoS One.* 2011;6(10):e26931. doi: 10.1371/journal.pone.0026931.
25. Sechi LA, Paccagnini D, Salza S, Pacifico A, Ahmed N, Zanetti S. *Mycobacterium avium* subspecies paratuberculosis bacteremia in type 1 diabetes mellitus: an infectious trigger? *Clin Infect Dis.* 2008;46(1):148-9. doi: 10.1086/524084.
26. Root-Bernstein R. Rethinking Molecular Mimicry in Rheumatic Heart Disease and Autoimmune Myocarditis: Laminin, Collagen IV, CAR, and B1AR as Initial Targets of Disease. *Front Pediatr.* 2014;2:85. doi: 10.3389/fped.2014.00085.
27. Emenike B, Nwajobi O, Raj M. Covalent Chemical Tools for Profiling Post-Translational Modifications. *Front Chem.* 2022;10:868773. doi: 10.3389/fchem.2022.868773.
28. Patwardhan A, Cheng N, Trejo J. Post-Translational Modifications of G Protein-Coupled Receptors Control Cellular Signaling Dynamics in Space and Time. *Pharmacol Rev.* 2021;73(1):120-51. doi: 10.1124/pharmrev.120.000082.
29. Feng D, Barnes BJ. Bioinformatics analysis of the factors controlling type I IFN gene expression in autoimmune disease and virus-induced immunity. *Front Immunol.* 2013;4:291. doi: 10.3389/fimmu.2013.00291.
30. Writing Group for the TRIGR Study Group; Knip M, Åkerblom HK, Al Taji E, Becker D, Bruining J, et al. Effect of Hydrolyzed Infant Formula vs Conventional Formula on Risk of Type 1 Diabetes: The TRIGR Randomized Clinical Trial. *JAMA.* 2018;319(1):38-48. doi: 10.1001/jama.2017.19826.
31. Patelarou E, Girvalaki C, Brokalaki H, Patelarou A, Androulaki Z, Vardavas C. Current evidence on the associations of breastfeeding, infant formula, and cow's milk introduction with type 1 diabetes mellitus: a systematic review. *Nutr Rev.* 2012;70(9):509-19. doi: 10.1111/j.1753-4887.2012.00513.x.
32. Wasmuth HE, Kolb H. Cow's milk and immune-mediated diabetes. *Proc Nutr Soc.* 2000;59(4):573-9. doi: 10.1017/s002966510000811.
33. Villagrán-García EF, Hurtado-López EF, Vasquez-Garibay EM, Troyo-Sanromán R, Aguirre-Salas LM, Larrosa-Haro A, et al. Introduction of pasteurized/raw cow's milk during the second semester of life as a risk factor of type 1 diabetes mellitus in school children and adolescents. *Nutr Hosp.* 2015;32(2):634-7. doi: 10.3305/nh.2015.32.2.9247.
34. Virtanen SM, Hyppönen E, Läärä E, Vähäsalo P, Kulmala P, Savola K, et al. Cow's milk consumption, disease-associated autoantibodies and type 1 diabetes mellitus: a follow-up study in siblings of diabetic children. *Childhood Diabetes in Finland Study Group. Diabet Med.* 1998;15(9):730-8. doi: 10.1002/(SICI)1096-9136(199809)15:9<730::AID-DIA646>3.0.CO;2-C.
35. Lamminsalo A, Lundqvist A, Virta LJ, Gissler M, Kaila M, Metsälä J, et al. Cow's milk allergy in infancy and later development of type 1 diabetes-nationwide case-cohort study. *Pediatr Diabetes.* 2021;22(3):400-6. doi: 10.1111/peidi.13181.
36. Xia Y, Xie Z, Huang G, Zhou Z. Incidence and trend of type 1 diabetes and the underlying environmental determinants. *Diabetes Metab Res Rev.* 2019;35(1):e3075. doi: 10.1002/dmrr.3075.

