



Found in Translation: Novel Insights Into Type 1 Diabetes and β -Cell Biology

Holger A. Russ and Howard W. Davidson

Diabetes 2021;70:2185–2186 | <https://doi.org/10.2337/dbi21-0031>

Although essential features in the natural history of type 1 diabetes are generally understood (1), many key questions remain. In particular, despite decades of research, the precise events that cause immune tolerance mechanisms to fail, triggering autoimmunity, remain unclear (2). One theory that has attracted considerable recent attention is that β -cell “stress” can trigger or potentiate autoimmunity by generating neoantigens (3–5), for example, by posttranslational modifications, hybrid peptide formation, or protein splicing (6). In this issue, Thomaidou et al. (7) significantly expand our knowledge of two other potential sources of “neo-antigens,” namely, the products of translation from noncanonical start sites, and alternatively spliced transcripts. Using a combination of ribosomal profiling (8) and long read RNA sequencing, they show that β -cells can use at least 15,014 unique translation start sites within transcripts from the 5,529 genes whose expression they detected. Ribosome profiling was initially developed by Ingolia, Weissman, and colleagues (8) and is based on next-generation sequencing of the \sim 30 base pairs fragments of mRNA that are protected from nuclease digestion by bound ribosomes. Combining this approach with pharmacological treatment to accumulate ribosomes at the sites of initiation allows accurate inference of translational reading frames. This powerful technique is now widely used to study many aspects of translation but had not previously been applied to human β -cells.

Surprisingly, only 17% of the start sites defined by Thomaidou et al. mapped to the reference for that transcript, thus revealing a wide range of previously unsuspected open reading frames (ORFs) and expanding the “translatome” to at least 22,670 distinct proteins. Unique ORFs included the products of upstream initiation (including some that would terminate before the annotated coding sequence and others that would extend it), translation in alternative reading frames, and downstream

initiation, including some ORFs originating from annotated 3' untranslated regions (7). Cytokine treatment resulted in initiation at 4,427 unique sites. As expected, many of these were from genes in pathways related to an inflammatory response, but importantly, some novel ORFs were in secretory granule proteins such as CHGA and so might be sources of “neoantigens,” though this remains to be confirmed.

A second surprising finding was that 20% of the start sites did not map to the reference transcriptome. While some of the discrepancies might possibly have arisen from sequencing or bioinformatics errors, long read sequencing revealed 6,892 new transcripts, suggesting that the β -cell transcriptome is also much more diverse than previously appreciated. Most (74%) of the new transcripts spanned a splicing junction, while others extended the annotated 3' untranslated region and so might represent novel polyadenylation sites. Importantly, 1,745 of the initiation sites that Thomaidou et al. (7) detected were located on the novel transcripts, including 499 that were specifically detected upon cytokine treatment, and so might also give rise to neoantigens. Preproinsulin is a major target of autoreactive T cells (9), and it is therefore of interest that Thomaidou et al. identified five novel transcripts from the *INS* gene. Four of these were of relatively low abundance, but the fifth was a novel 3' extension that comprised 3% of the total *INS* transcripts and would allow synthesis of a longer version of the defective ribosomal product that the authors have previously shown contains at least one major epitope recognized by diabetogenic T cells (10). Bioinformatics analyses indicate that the longer variant may contain additional T-cell epitopes, further highlighting the potential importance of the *INS-IGF2* locus as a source of neoantigens (11).

Beyond autoimmunity, the study by Thomaidou et al. may also provide new insight into basic β -cell biology. A

Barbara Davis Center for Diabetes, University of Colorado Anschutz Medical Campus, Aurora, CO

Corresponding author: Howard W. Davidson, howard.davidson@cuanschutz.edu

© 2021 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

See accompanying article, p. 2299.

considerable fraction (7%) of the novel start sites they found mapped to noncoding RNA, and particularly to long noncoding RNA, thereby revealing previously unsuspected polypeptides within them (7). This finding is consistent with many other related studies. Indeed, with the increasing use of ribosome profiling it has become apparent that there are a multitude of short ORF-encoded peptides and small proteins in eukaryotic cells whose existence and true significance are only gradually being revealed (12,13). Such “micropeptides” may play important cellular roles, for example, in regulating translation (14) and modulating the function of larger interaction partners (15). At present the functional significance of micropeptides in β -cell biology remains largely unknown. However, Thomaidou et al. have now cast new light on the “dark proteome” of the β -cell, which is a critical step toward gaining a complete understanding of its function in health and disease.

Although providing important new information, the study by Thomaidou et al. had several limitations. First, for technical reasons, it focused exclusively on EndoC- β H1 cells (16). This line is generally accepted as the best available model for primary human β -cells and shows many functional similarities to isolated human islets. However, recent analyses have highlighted some key differences between EndoC- β H1 cells and islets (17,18), which likely reflect its fetal/embryonic origin and/or its transformed state. Indeed, ORFs for several key β -cell proteins such as IAPP and IGRP were not detected (7), indicating that the expanded proteome still remains incomplete. Thus, as acknowledged by the authors, this model exhibits an incomplete phenotype compared with bona fide β -cells but, nonetheless, represents a logical first step toward defining their transcriptome. Using primary tissue will be a critical next step, but the lower cell numbers and cellular heterogeneity of cadaveric islets present experimental challenges that will need to be overcome first. Pluripotent stem cell differentiation into β -cells is an alternative model that could provide large numbers of β -cells to further probe the transcriptome. However, meta-analysis of the transcriptomes of all available human β -cell models will likely provide the most thorough insights.

A second limitation was the lack of independent validation for the majority of the newly described ORFs. While comprehensive validation was beyond the scope of this study, it leaves open the possibility that some ORFs may represent bioinformatics artifacts. Thus, it will be important for future follow-up studies to be conducted to validate the newly identified polypeptides and determine their roles.

Funding. This work was supported by National Institutes of Health (NIH) grant R21 AI140044 (to H.A.R. and H.W.D.) and the Children’s Diabetes Foundation. H.W.D. also acknowledges the generous support of the Foundation for

Diabetes Research and Beatson Foundation. Work in the laboratory of H.A.R. is supported by NIH grants R01DK120444, R21AI140044, and DK1041162; the Culshaw Junior Investigator Award in Diabetes; a Gates Grubstake Award; and JDRF grant 2-SRA-2019-781-SB.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

References

- Battaglia M, Anderson MS, Buckner JH, et al. Understanding and preventing type 1 diabetes through the unique working model of TrialNet. *Diabetologia* 2017;60:2139–2147
- Erdem N, Montero E, Roep BO. Breaking and restoring immune tolerance to pancreatic beta-cells in type 1 diabetes. *Curr Opin Endocrinol Diabetes Obes* 2021;28:397–403
- Mallone R, Eizirik DL. Presumption of innocence for beta cells: why are they vulnerable autoimmune targets in type 1 diabetes? *Diabetologia* 2020;63:1999–2006
- Roep BO, Thomaidou S, van Tienhoven R, Zaldumbide A. Type 1 diabetes mellitus as a disease of the β -cell (do not blame the immune system?). *Nat Rev Endocrinol* 2021;17:150–161
- Piganelli JD, Mamula MJ, James EA. The role of β cell stress and neo-epitopes in the immunopathology of type 1 diabetes. *Front Endocrinol (Lausanne)* 2020;11:624590
- Nguyen H, Guyer P, Ettinger RA, James EA. Non-genetically encoded epitopes are relevant targets in autoimmune diabetes. *Biomedicines* 2021;9:202
- Thomaidou S, Sliker RC, van der Slik AR, et al. Long RNA sequencing and ribosome profiling of inflamed β -cells reveal an extensive transcriptome landscape. *Diabetes* 2021;70:2299–2312
- Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 2011;147:789–802
- Mannering SI, Pathiraja V, Kay TWH. The case for an autoimmune aetiology of type 1 diabetes. *Clin Exp Immunol* 2016;183:8–15
- Kracht MJ, van Lummel M, Nikolic T, et al. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat Med* 2017;23:501–507
- Kanatsuna N, Taneera J, Vaziri-Sani F, et al. Autoimmunity against INS-IGF2 protein expressed in human pancreatic islets. *J Biol Chem* 2013;288:29013–29023
- Delcourt V, Staskevicius A, Salzet M, Fournier I, Roucou X. Small proteins encoded by unannotated ORFs are rising stars of the proteome, confirming shortcomings in genome annotations and current vision of an mRNA. *Proteomics* 2018;18:e1700058
- Orr MW, Mao Y, Storz G, Qian S-B. Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Res* 2020;48:1029–1042
- Wu Q, Wright M, Gogol MM, Bradford WD, Zhang N, Bazzini AA. Translation of small downstream ORFs enhances translation of canonical main open reading frames. *EMBO J*. 2020;39:e104763
- Vitorino R, Guedes S, Amado F, Santos M, Akimitsu N. The role of micropeptides in biology. *Cell Mol Life Sci* 2021;78:3285–3298
- Ravassard P, Hazhouz Y, Pechberty S, et al. A genetically engineered human pancreatic β cell line exhibiting glucose-inducible insulin secretion. *J Clin Invest* 2011;121:3589–3597
- Lawlor N, Marquez EJ, Orchard P, et al. Multiomic profiling identifies cis-regulatory networks underlying human pancreatic β cell identity and function. *Cell Rep* 2019;26:788–801.e6
- Oleson BJ, McGraw JA, Broniowska KA, et al. Distinct differences in the responses of the human pancreatic β -cell line EndoC- β H1 and human islets to proinflammatory cytokines. *Am J Physiol Regul Integr Comp Physiol* 2015;309:R525–R534