Dan Tawfik

In Memoriam

Anthony Futerman

Dan Tawfik passed away in early May 2021 while on a climbing trip in Croatia. He was sixty-five and yet at the height of his powers. Dan was a faculty member at the Weizmann Institute of Science, and for the last twenty years or so, we were colleagues in the Department of Biomolecular Sciences. I knew him well and I admired his work. Dan was a second-generation Israeli, the son of Iraqi Jewish immigrants. As a young man, he served in the Israeli air force, reaching the rank of major before studying chemistry and biochemistry at the Hebrew University of Jerusalem. Dan completed his bachelor of science in 1988 and his master of science in 1990. He then enrolled in the PhD program at the Weizmann Institute, joining a research project supervised by Michael Sela and Zelig Eshhar, two of the Institute's most distinguished faculty members. Working in the new field of catalytic antibodies, Dan became fascinated by enzymes. After completing his doctorate, Dan joined Sir Alan Fersht's laboratory at Cambridge, where he effectively ran an independent research program. It was during this period that Dan developed an interest in directed evolution, and it was directed evolution that became the primary focus of his own research laboratory at the Weizmann Institute. He played such an important role in the development of the field that Frances Arnold made a point of mentioning his contributions as part of her 2018 Nobel Prize acceptance lecture. Dan received a number of awards for his work: the Weizmann Prize, the Teva Prize, and the EMET Prize, as well as the Enzyme Engineering Award from Engineering Conferences International.

Weizmann Institute faculty members, students, and colleagues gathered on Sunday, May 9 to remember Dan's life. All of the mourners commented on Dan's creativity and originality. We recalled his almost childlike joy in the day-to-day pursuit of his research and the great care he took of all the students and fellows who passed through his laboratory. One of his close colleagues mentioned Dan's love for hummus: he was, in words that he might well have appreciated, a great *hummus-fresser*; and he could often be found eating lunch with students and colleagues in one of the hummus restaurants near Rehovot. I remember sneaking out one lunchtime to a well-known restaurant with some of my students and finding Dan regaling his colleagues at the next table, a great hummus mound shrinking on his plate.[1]

On May 9, we all knew that we would miss him, and we were right.

Dan's scientific research focused on proteins and, in particular, on enzyme evolution. The majority of his work was concerned with how new protein functions evolved from existing functional proteins. More recently he had begun working on the most difficult challenge in biochemical evolution: reconstructing the metabolic pathways that may have led to the emergence of the first functional proteins. While he made great progress in the first area, the extent of his progress in the second was more limited. The problem is very difficult; it may be intractable.

I heard Dan lecture twice during the past year or so. The first lecture was part of a departmental seminar series that was held during October 2020. The second was delivered in April 2021 at a conference organized by the Israel Society for Astrobiology and the Study of the Origin of Life. Dan's presentation was entitled "From So Simple a Beginning—How Did the First Proteins Evolve?" I can think of no better way to honor Dan's memory than by reviewing the ideas that he discussed in these lectures, and by highlighting not only the progress made, but also the open questions that remained.

Dan's work on the evolution of new functions in existing enzymes is both well established and relatively uncontroversial. The fundamental experimental basis of this approach, known as directed evolution, is to take a preexisting gene, subject it to iterative rounds of mutagenesis, and then select and isolate proteins with the desired or unexpected new function. With this classical approach, Arnold was able to convert an enzyme that normally only cleaves peptide bonds in water to one that cleaves peptide bonds in organic solvents. Similarly, Dan was able to increase the catalytic efficiency of an existing enzyme, serum paraoxonase 1, by 10^5-fold, and to evolve a protein that became able to use substrates to which the parental protein showed no detectable activity. In a 2012 paper,

he suggested that "new functions evolve by augmenting weak, promiscuous functions in existing proteins."[2] There can be no doubt that random mutagenesis in the laboratory can lead to the generation of better enzymes or enzymes with novel functions. As Dan remarked in a recent lecture, "Teaching old dogs new tricks is actually not so difficult." Far more challenging is the question of how the first dog learned *his* trick—that is, how to explain the emergence of functional proteins or protein domains in abiotic scenarios. Dan was focused on this question during the last years of his life.

Dan was well aware that determining how the first functional proteins were generated would be far from easy. In a 2013 article entitled "Close to a Miracle," he was quoted as saying that,

> What we lack is a hypothesis for the earlier stages, where you don't have this spectrum of enzymatic activities, active sites and folds from which selection can identify starting points. Evolution has this catch-22: nothing evolves unless it already exists.[3]

Dan's approach to this "daunting challenge," as he described it,[4] can be gleaned from the topics raised in these two lectures and from some of the recent publications from his laboratory.[5] In one of these papers, he discussed three critical issues concerning the pathways of de novo protein evolution in some detail. First, "de novo emergence demands a transition from disordered polypeptides into structured proteins with well-defined functions."[6] Under abiotic conditions, polypeptides must have been generated randomly and thereafter "happen[ed] to provide some benefit," although the precise benefit is never clearly defined in the paper. Second, such polypeptides must have exhibited functions of "evolutionary relevance," with such polypeptides evolving into modern proteins, perhaps by a "series of duplications and fusions and mergings with other peptide fragments." Finally, "the earliest proteins ... were likely based on abiotic, spontaneously synthesized amino acids." As a result, the amino-acid composition of prebiotic proteins was unlikely to be similar to the amino acids found in modern proteins.

Dan postulated that the binding of nucleic acids by peptides was likely a critical event in the emergence of life. For this reason he focused on peptides that might exhibit nucleic acid–binding activity. Based on this assumption, he attempted to determine a putative gradual evolutionary trajectory leading from a polypeptide to a ubiquitous nucleic acid–binding protein. Dan proposed that the first primordial enzymes bound and utilized ligands that contain a phosphate moiety.[7] Such proteins, known as phosphate-binding loop (P-loop) NTPases, apparently comprise up to 40% of all domains with known structure in bacteria. Dan and colleagues were able to generate small functional P-loop proteins that emerged via extensions of the same seed element. This putative ancestral P-loop was also able to bind ATP/GTP and ssDNA/RNA. These early nucleic acid–binding proteins may have contained ornithine, a basic amino acid that can directly bind nucleic acids. Although ornithine is not found in today's proteins, Dan suggested that an abiotic chemical reaction was able to convert ornithine into arginine, which improves both the structure and function of the P-loops. This led him to conclude that, "[F]unctional proteins may arise from short and simple sequences that included ornithine."[8]

I greatly admire scientists engaged in a good-faith effort to understand the pathways that may have led to the emergence of life on earth. Studying some of the possible pathways is better by far than not studying any of them. But what is possible and what is real are two different matters, and if certain pathways are possible, there remains the question whether they work in the real world. Is there a statistically viable chance that random processes could have led to the emergence of a relatively simply peptide, such as small, functional phosphate-loop proteins containing about forty amino acid residues? Modern proteins are made up from twenty amino acids. But more than eighty amino acids, including ornithine, have been generated in abiotic scenarios, and around 1,000 can be generated by natural pathways. If ornithine is indeed involved in abiotic synthetic reactions, then it would be disingenuous to exclude the possibility that other abiotic amino acids might also have been involved in the generation of the first simple peptides. If each residue added to a putative ancestral P-loop is randomly selected from a pool of approximately one hundred candidate amino acids, the probability of randomly inserting any one particular amino acid is 1 in 100. The abiotic probability of randomly generating a single 40-residue polypeptide sequence is 1 in 100^{40}, or 1 in 10^{80}. Although each amino acid can exist in D or L stereoconformations, only the latter is incorporated into modern proteins. The probability of randomly inserting a D-amino acid or an L-amino acid is 1 in 2. Since this probability is the same for every amino acid in the polypeptide, the probability of randomly assembling an all-L 40-residue polypeptide is 1 in 2^{40}, or 1 in 10^{12}. In a 40-residue peptide, 39 peptide bonds need to be formed, with the probability of any one peptide bond forming along the backbone, rather than between amino acid side chains, being 1 in 2. Assuming a pool of about 100 abiotic amino acids, the probability of randomly generating one specific 40-residue, all-L, putative ancestral P-loop peptide is 1 in $100^{40} \times 2^{40} \times 2^{39} = 1 \text{ in } 10^{80} \times 10^{12} \times 10^{11} = 1 \text{ in } 10^{103}$. In principle, anything is possible, but the probability of such a scenario occurring randomly appears to be unlikely. And this is before evolution needs to convert such simple beginnings into more complex proteins.

No less intractable is the question of how evolution chose which amino acids to use in modern proteins from

among the large set that can be generated by abiotic synthesis. In one of the two lectures I mentioned, Dan was asked a similar question: Why is ornithine not used in today's proteins? It is instructive to consider his reply:

> [I]n general this question is hard to answer because we forget that chance plays a huge role in evolution ... but the reason traditionally noted is that ornithine is less stable [when] incorporated into tRNAs.

This argument *presupposes* the existence of tRNAs (transfer RNAs) and in the end appeals to the existence of proteins (acting as enzymes) that it was designed to explain.

To Dan's credit, he was always mindful of the open questions and ambiguities in the study of de novo protein generation. I remember asking him a question at one of his lectures about how nature selected for L- rather than D-amino acids. Dan was gracious enough to acknowledge that he had no good answer. Open questions drive science forward, but surely there must be a point when our understanding of the complexity of living systems prompts researchers to seek alternative theories to explain the origins and evolution of proteins, and the origins of life in general. Has this time already come? Was Darwin correct when he suggested that, "[F]rom so simple a beginning, endless forms most beautiful and most wonderful have been, and are being evolved"?[9] Although Darwin was fully aware of the "endless forms," he could not have known anything about their molecular details. Dan may have thought that only a "small" shopping list would be needed to account for those molecular details: a self-containing functional motif N-helix phosphate binding site; a minimal, abiotic amino-acid alphabet, including nonproteogenic amino acids such as ornithine; the oligomerization of simpler forms of self-assembly; multifunctionality, the binding of ssDNA, ATP, and polyphosphate. A detailed understanding of the pathways needed to generate such molecules suggests, at least to me, that random events were unlikely to have been the sole cause behind their generation.

Dan was a giant in his field. Only time will tell whether his research actually supports the notion that random abiotic pathways can lead to the de novo generation of proteins, or whether an entirely different approach is needed.

May Dan's memory be blessed.

Anthony Futerman is the Joseph Meyerhoff Professor of Biochemistry in the Department of Biomolecular Sciences at the Weizmann Institute of Science in Rehovot, Israel.

1. More personal tributes to Dan can be found in the obituary published on the website of the Weizmann Institute. See also a video in which Dan discusses his award of the prestigious EMET prize: "הכוז , קיפות חלס וד 'פורפ 2020 היימיכויב – סייחה יעדמב תמא סרפ," *YouTube* video, uploaded October 14, 2020.
2. Moshe Goldsmith and Dan Tawfik, "Directed Enzyme Evolution: Beyond the Low-Hanging Fruit," *Current Opinion in Structural Biology* 22, no. 4 (2012): 407, doi:10.1016/j.sbi.2012.03.010.
3. Rajendrani Mukhopadhyay, "Close to a Miracle," *ASBMB* Today, September 23, 2013.
4. Liam Longo et al., "Primordial Emergence of a Nucleic Acid–Binding Protein via Phase Separation and Statistical Ornithine-to-Arginine Conversion," *Proceedings of the National Academy of Sciences* 117, no. 27 (2020): 15,731–39, doi:10.1073/pnas.2001989117.
5. For a review of the work undertaken by the Dan's laboratory and some notable publications, see Tyler Hampton, "Dan S. Tawfik Group: The New View of Proteins," *Inference: International Review of Science* 1, no. 1 (2014), doi:10.37282/991819.14.8.
6. Liam Longo et al., "Primordial Emergence," 15,731.
7. Liam Longo et al., "Short and Simple Sequences Favored the Emergence of N-Helix Phospho-Ligand Binding Sites in the First Enzymes," *Proceedings of the National Academy of Sciences* 117, no. 10 (2020): 5,310–18, doi:10.1073/pnas.1911742117.
8. Liam Longo et al., "Primordial Emergence," 15,731.
9. Charles Darwin, *On the Origin of Species* (London: John Murray, 1872), 429.

DOI: 10.37282/991819.21.39

Published on August 24, 2021

https://inference-review.com/article/in-memoriam-dan-tawfik