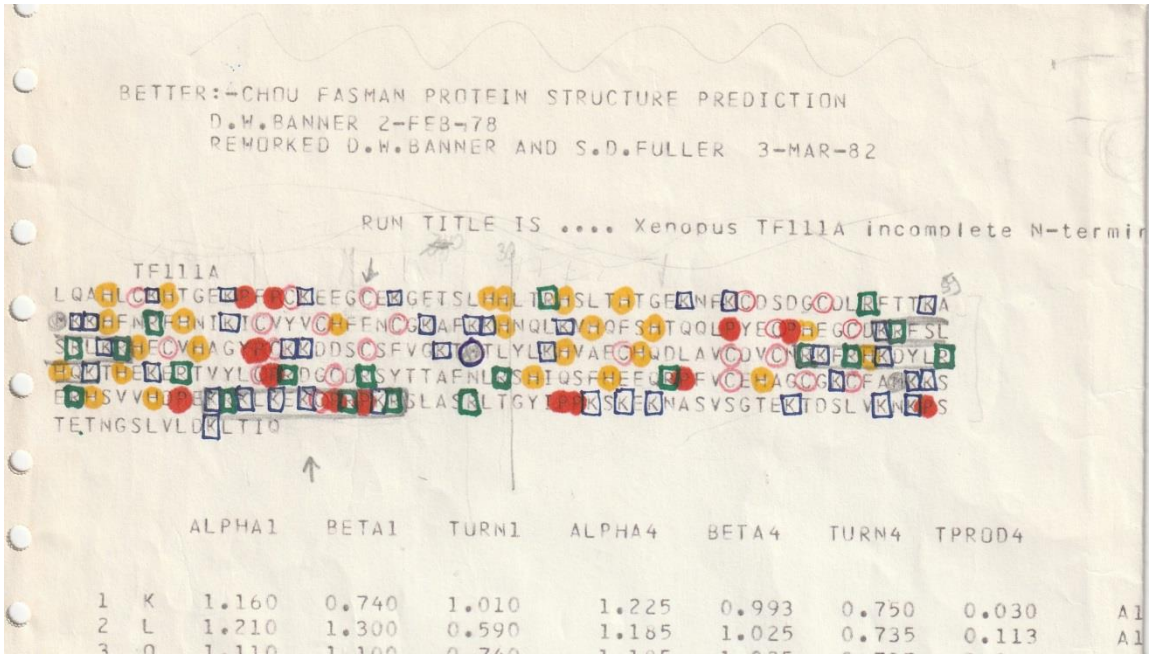
A Tale of Frogs, Fingers and Fruit Flies

Once upon a time there were white patches of snow still lying in the forest. These were the remnants of a late spring snowfall. All around the building the snow was piled up into frozen heaps. Towards the end of the afternoon, it was already getting dark. A freezing mist soon appeared above the trees.

On that day, in the last week of March 1984, I stepped out of the laboratory and walked down the corridor to fetch my mail. There I found a letter from Ian Jackson.[i] He had sent me the amino acid sequence of transcription factor IIIA (TFIIIA).[ii] A large amount of this protein, together with its partner 5S rRNA, was present in the eggs of immature frogs like Xenopus laevis [1]. I had chosen this ribonucleoprotein particle as the topic for my PhD at the European Molecular Biology Laboratory.[iii] My goal was to find its three-dimensional structure using X-ray crystallography. Primarily to show how RNA and protein molecules can specifically bind to each other. I was not to know then that I was embarking on an arduous 15-year odyssey.

Xenopus laevis - the African clawed frog. Photo by Brian Gratwicke licensed under CC BY-ND.

The amino acid sequence, which Ian had sent to me, was written out in the conventional three-letter code. The sequence of this protein was 316 amino acids long and was deduced from a cDNA.[iv] In 1984 the personal computer was just being invented, there were only primitive email networks and no World Wide Web with access to bioinformatics software. Success in scientific research required sharp eyes and a great deal of thought. Computer programs that existed were mainly in the hands of specialists. So, I asked a colleague, Christian Oefner, to run a secondary structure prediction using the Chou-Fasman method. I sat down to rewrite the TFIIIA sequence out in the single-letter code for amino acids.[v] Later that afternoon I went downstairs to our mainframe VAX 11/780 computer and typed it up. The computer program fortuitously printed out the sequence in lines 60 amino acids long.

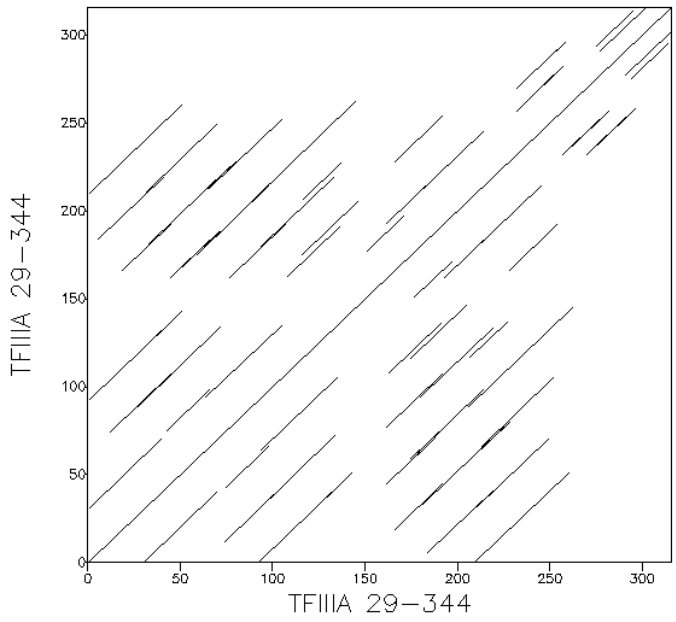


The original printer output of the TFIIIA sequence showing my amino acid markings (March 1984).

The biological activity of TFIIIA posed something of a structural paradox. This protein, which forms a complex with 5S rRNA called the 7S ribonucleoprotein storage particle, was found to be a transcription factor. It could bind to double-stranded DNA and activate the 5S DNA gene promoter. With this in mind, I drew colored circles around those amino acids that might interact with nucleic acids. These were the positively charged basic amino acids; lysine (K) marked in blue, arginine (R) in green and histidine (H) in yellow. I saw that the 40 lysines and 19 arginines were scattered along the entire length of this sequence. A span of basic amino acids, from positions 280 to 293, stood out.[vi] The 22 cysteines (C) were marked with pink open circles. 14 prolines (P) were circled in orange. The prolines often show where there are disruptions in the secondary structure.

I noticed that pairs of histidines marked in orange, were bunched at the beginning and also in the middle of the sequence segments.[vii] These were separated by approximately the same distance apart. I also noticed that each pair of histidines was followed closely by a pair of cysteines marked in pink. In total there were eight pairs of cysteines and nine pairs of histidines with a fixed separation of 12 residues between them. It was obvious that there was a pattern made up of at least eight similar repeats throughout this sequence. I discussed this with Dimitri Tsernoglou, a protein crystallographer. Dimitri said that he had not seen this pattern before in any known protein structure.

A few days later at the beginning of April, Christian Oefner drove on his motorcycle to the Max Planck Institute for Medical Research. I followed him in my car down to Heidelberg. We went to visit Chris Sander who had expertise in secondary structure analysis and protein folding. Chris Sander computed a so-called Dot Plot to look for any similar or repeated features in the protein sequence.[viii] He displayed the resulting Dot Plot of TFIIIA on the green screen of his Tektronix terminal.



A self-comparison Dot Plot of the TFIIIA sequence [2]. Each position on the horizontal axis was compared pairwise with all of the positions on the vertical axis. A match found between the two copies of TFIIIA was plotted as a dot. Two closely similar amino acid sequences will produce a diagonal line of dots. Positions 1 to 316 on each of the axes in the Dot Plot shown above, correspond to positions 29 to 344 of the TFIIIA amino acid sequence. The extent of the matching was revealed by scoring over a 20-residue sliding window.[ix]

Sequence repeats stood out in eight parallel diagonal lines spanning positions 29-254. The display confirmed the pattern I had already noticed earlier by eye. Obviously TFIIIA was almost entirely made of sequence repeats. The shortest diagonal consisted of 25 residues. Longer diagonals showed that the repeats followed one-after-another. The distance between diagonal lines was 28 ±2 residues. This meant that there was a similar feature, in the amino acid sequence, repeating itself at a regular interval.

GEKNFT**C**DSDG**C**DLRFTTKANMKK**H**FNRF**H**NIKICVYV**C**HFEN**C**GKAFK

QD-LAV**C**DV--**C**NRKFRHKDYLRD**H**-QKT**H**EKERTVYL**C**PRDG**C**DRSYT

KHNQLKV**H**-QFS**H**TQQLPYE**C**PHEG**C**DKRFSLPSRLKR**H**EKV**H**AGYPCKK

TAFNLRS**H**IQSF**H**EEQRPFV**C**EHAG**C**GKCFAMKKSLER**H**SVV**H**D--PEKR

A typical long diagonal from the Dot Plot. This results from the comparison of positions 41-138 (top and third lines) with positions 161-254 (second and fourth lines). Spanning over three Cys2-His2 repeats with few insertions or deletions.

What new facts had emerged in April 1984 from our analysis of the TFIIIA sequence?

First, that TFIIIA had at least eight similar segments. An extra pair of histidines implied that more repeats were present.

Second, the eight repeats were arranged in tandem without intervening sequences.

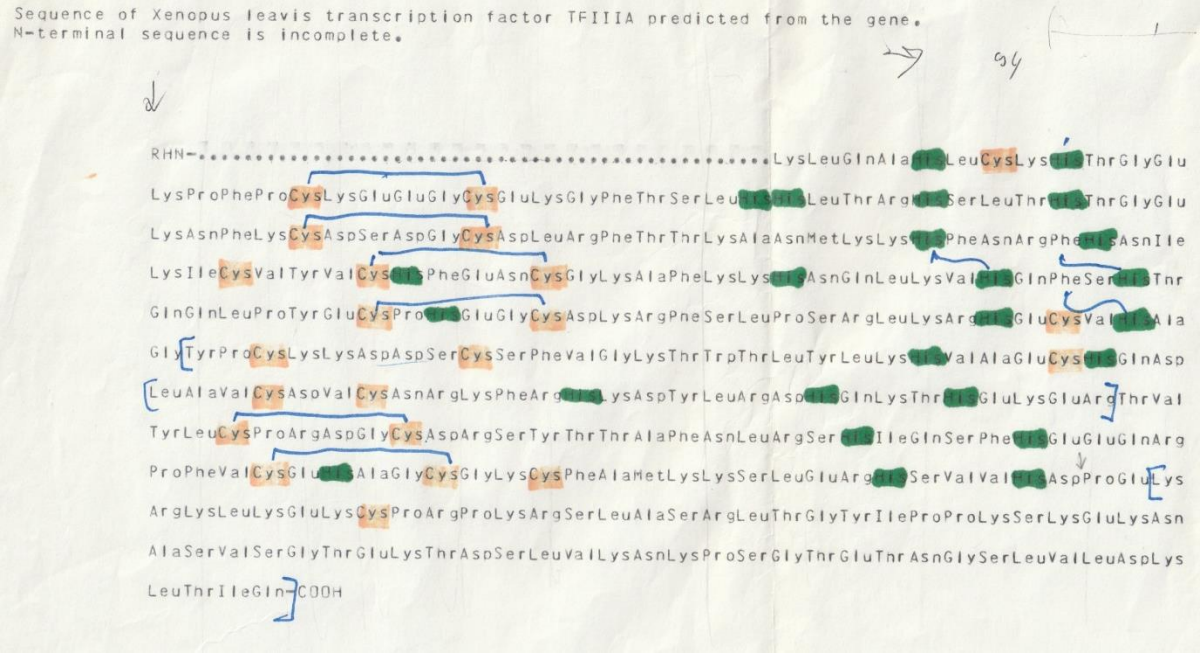
Third, the average length of a repeat was 28 ±2 residues.

Fourth, there was a Cys2X12His2 motif in each sequence repeat. X12 represents the 12 positions where a variety of amino acids were present.

Through that summer of 1984 I developed a method for the large-scale preparation of the 7S particle. Milligram quantities were needed for crystallization trials. Electron microscopy of the 7S particle was undertaken by Kevin Leonard. Surprisingly the 7S particle had a star shape in negative stain. This suggested to us that the protein was wrapped around the outside of the RNA. I also purified TFIIIA protein. Some was given to Akira Tsugita for protein sequencing to confirm the translated cDNA sequence. My supervisor Fritz Winkler and I had already carried out circular dichroism studies of TFIIIA and of the 7S particle at the Max Planck Institute for Biophysics in Frankfurt. These circular dichroism measurements were continued with Michelle Hollecker at the Institut des Sciences de l'Ingenieur in Nancy, and then with Jennifer Reed at the Deutsches Krebsforschungszentrum, Heidelberg. We obtained a secondary structure content of 17.5 ±4.6 % alpha helix and 33.5 ±4.8 % beta sheet [3]. There was a surprisingly low amount of alpha helix in TFIIIA. This prompted Jennifer Reed to propose that TFIIIA was a protein with several very short alpha helices.

In July, Tony Crowther was visiting the EMBL. I gave him the amino acid sequence and some purified TFIIIA protein to take to a former colleague, Daniella Rhodes. My intention was to set a precedent. Sending a signal that I knew of the TFIIIA sequence and testing the waters. This failed completely on both accounts. I expected to get a swift response. But no word or acknowledgment came from Cambridge.

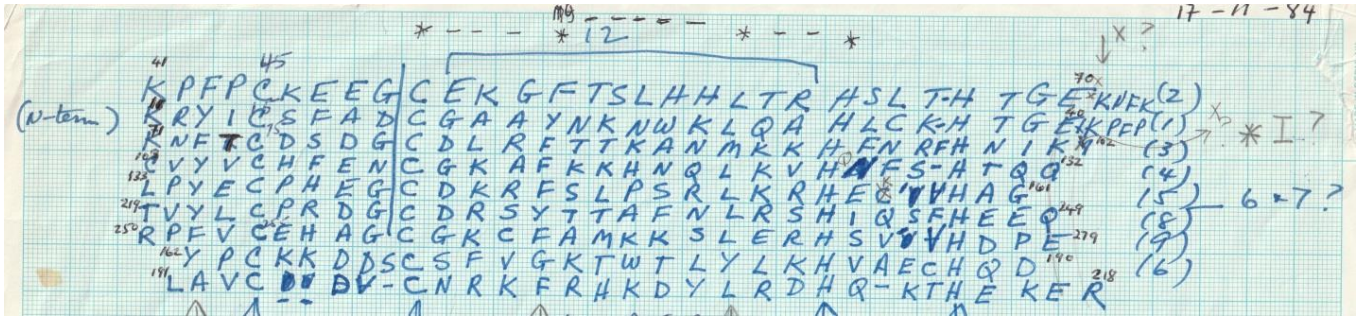
I made a phone call to Aaron Klug in September. We discussed the results of our zinc and lead cleavage experiments with phenylalanine tRNA. During this conversation Aaron told me that they had looked for metals in the 7S particle. He asked me to take a guess at their result. I replied that the metal was probably zinc. I pointed out that RNA might also bind metal ions. Aaron did not disclose to me precisely how much zinc had been found in the 7S particle. At that time I was not aware that Jay Hanas had already discovered zinc in the TFIIIA protein a year earlier in 1983 [4].



I divided the TFIIIA sequence up into segments of 30 amino acids in length. This was printed out in the conventional three-letter amino acid code. Eight of the segments contain Cys2 colored orange and His2 colored in green. A computer printout dated 2 October 1984 is shown above.

At the end of October, I attended a symposium at the neighboring Max Planck Institute for Nuclear Physics. There I met with Robert Roeder, Daniella Rhodes and Akira Tsugita in the main hall. We stood next to the wall of fossils. Akira spoke about his protein sequencing work. He had identified the last three amino acids of TFIIIA to be ThrIleu/LeuGln. But Akira was unsuccessful with sequencing in from the N-terminus. The first amino acid was probably modified and therefore resistant to chemical release. To my consternation, Akira passed around my computer printout depicting the eight Cys2-His2 segments, as shown in the figure above. However, this produced no visible reaction from any of the others present. Robert Roeder said he had recently finished the TFIIIA cDNA sequence. I asked him for the complete sequence and he generously agreed to send it to me.

Early In November, I brought the amino acid sequence of TFIIIA to Patrick Argos. He had recently arrived to take up a post at EMBL. He was an expert in the computer analysis of protein structures. That afternoon Pat lined up the invariant pairs of cysteines (C) and histidines (H). Where necessary gaps were inserted into the sequence alignment to maximize any similarity. The amino acids were positioned according to their chemical and physical characteristics. The resulting aligned nine segments contained between 26 and 32 residues.



Pat Argos's alignment of the nine repeats in Xenopus TFIIIA dated 17 November 1984.

k p Y/f v **C** - - d/e G/d **C** d/g k/r - F/y t t k - - L k/r r/k **H** - - s/k **H** t/e g e/q

The 30-residue consensus sequence derived from nine TFIIIA repeats. Shown in the single-letter code for amino acids. Invariant amino acids are shown in capital letters. The alignment revealed three more highly conserved residues in each repeat; tyrosine (Y), phenylalanine (F) and leucine (L). They might be involved in a hydrophobic cluster or core.

Pat Argos computed a secondary structure averaged over the nine aligned Cys2-His2 segments.[x] The result predicted some beta strand, a Cys-Cys loop, followed closely by a short alpha helix containing eight residues [2]. It was Pat Argos who pointed out that two cysteines and a neighboring pair of histidines could bind to a zinc ion.[xi] He had already published in 1978, a study of tetrahedral zinc coordination in the active sites of several proteins [5]. Pat's suggestion seemed to conflict with the zinc content reported by Jay Hanas in 1983 [4]. Jay had found only three zinc ions rather than the nine we expected to be present in TFIIIA.[xii]

Sitting at the desk in Pat Argos's office, an image from the past suddenly flashed through my mind. I saw the caterpillar from Walt Disney's film "Alice in Wonderland".[xiii] So that was what TFIIIA looked like. This was my "Aha" moment. The structure of TFIIIA protein was made up of small folded units. Each unit possessed a two-cysteine loop and a pair of histidines that could coordinate a zinc ion. I imagined a caterpillar-like protein made up of identical segments. That afternoon in November I had learned from Pat Argos about sequence alignment, how to make a consensus sequence, and about tetrahedral coordination of zinc in proteins. Pat had also computed a plausible secondary structure for the repeats. We guessed that they were likely to be structural units. At the end of my session with Pat Argos, I went back into the laboratory and told Fritz Winker about the predicted Cys-Cys loops and the alpha helices.

The presence of short alpha helices in TFIIIA was supported by some experimental data. Our prediction of 21% alpha helix was consistent with the 17.5 ±4.6 % alpha helix measured by circular dichroism. Earlier in 1983, Geoff Kneale and I had previously observed the reversible quenching of tryptophan fluorescence in the 7S particle [6]. Two tryptophans in the predicted alpha helices, W28 and W177, were the only candidates. This quenching meant that they had to be in close proximity to 5S rRNA. It was widely accepted in 1984 that a short alpha helix could bind to the major groove of DNA. This was the current model for protein recognition of specific

DNA base pairs [7]. We realized that the short alpha helices of TFIIIA may bind in a similar way to the double helix. Such an elongated protein could easily stretch over the 50 base pairs of the 5S gene promoter. However, the exact mode of binding to 5S rRNA remained a mystery.

What additional facts came to light in November 1984?

First, there were nine consecutive copies of an amino acid sequence motif in TFIIIA.

Second, we had derived a 30-amino acid consensus sequence: X2YXCX3GCX3FX5LX2HX3HX3. A variety of different amino acids occur at positions marked X. C and H were always present, and Y, G, F and L are highly conserved.

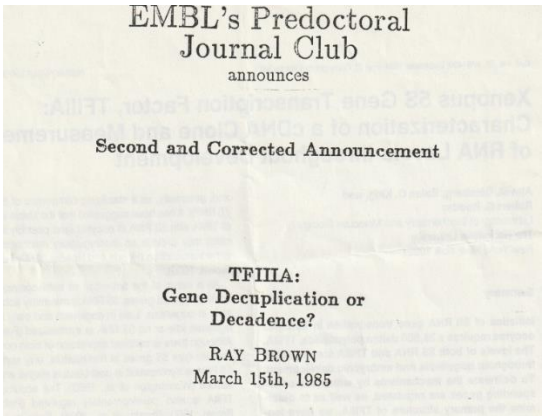
Third, it had been shown that TFIIIA required zinc for its biological activity [4]. The most obvious candidates to bind zinc were the nine adjacent pairs of cysteines and histidines. But we knew of no experimental data which identified the ligands or showed the folding of the zinc binding sites. As TFIIIA was present only inside cells, its pairs of cysteines were not expected to form disulphide bonds.

Fourth, the predicted amount of alpha helix agreed well with circular dichroism measurements. The short alpha helices of the repeats might bind to specific DNA sequences.

Fifth, we envisioned an elongated, modular protein structure. The periodicity plot suggested a linear arrangement of small units probably created by gene duplication. I had measured a Stokes radius of 34 Angstroms that was consistent with an extended protein molecule tumbling inside a sphere [3].

The Sequel

A complete amino acid sequence of Xenopus TFIIIA from Robert Roeder's laboratory was eventually published in the journal Cell [8]. This journal appeared on library shelves in the last week of January 1985. I was very surprised to find that there was no mention of the Cys2-His2 pattern. Now the way was clear for us to report our sequence analysis. At the time it was important that I should publish some of my PhD work. There had been no effort to keep any of our ideas secret. These were freely shared with others and were also presented at laboratory talks during 1984-85.


EMBL's Predoctoral
Journal Club
announces

Second and Corrected Announcement

TFIIIA:
Gene Decuplication or
Decadence?

RAY BROWN
March 15th, 1985

A notice of a seminar on 15 March 1985 at EMBL.

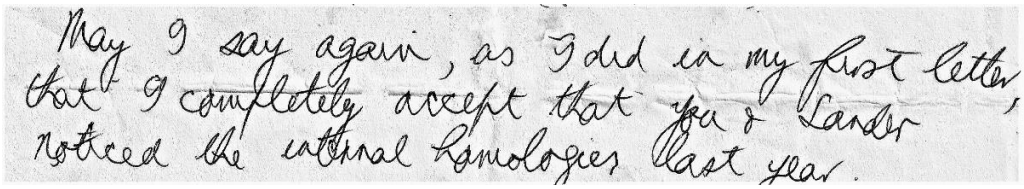
Lennart Philipson, Riccardo Cortese, Dimitri Tsernoglou and others were present at an EMBL seminar in mid-March 1985. There I described our sequence analysis, the 30-amino acid motif and its predicted secondary structure. I had noticed that the most similarity was shared by repeats 2, 5 and 9. But I was not able to assign them to DNA sequences within the 5S gene promoter. Finally, I speculated on a gene-duplication origin and whether the sequence homologies near the C-terminus were the remnants of primordial repeats.

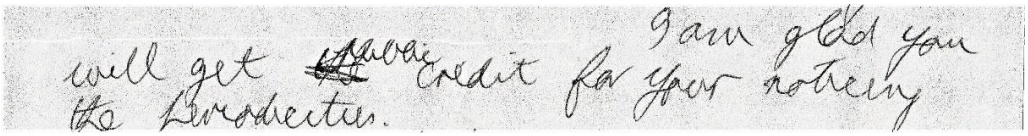
My family spent the Easter of 1985 at the house of a colleague, Jean-Claude Jesior, in the French Alps near Grenoble. I had brought with me a computer search of 2675 protein sequences in the newly-formed National Biomedical Research Foundation PIR data bank. Short spans, which contained closely spaced cysteines and histidines, were identified in protein sequences. This search proved to be unfruitful and I found no patterns that closely matched the Cys2X12His2 motif of TFIIIA.

In my absence from EMBL a manuscript was received on the 4th of April at the office of the EMBO journal [9]. The authors had requested an in-house referee for their research paper. I believe that they probably thought I would be chosen as an in-house referee. In fact, I was neither asked nor selected to be the referee. I was not shown their manuscript and I neither read nor saw any of its contents. Later I learnt that the Director-General of EMBL, Lennart Philipson, had made a phone call on my behalf to Aaron Klug.

Attempts were made to contact me at the EMBL outstation in Grenoble. But nothing reached me during our Easter vacation in nearby Laval. On my return to Heidelberg, I found out that we had acquired rivals. I hastily completed the manuscript describing our sequence analysis of TFIIIA and structure predictions. I suggested we might include a schematic drawing but Pat Argos disagreed. One morning soon after, Chris Sander appeared in the laboratory and asked to be included as an author. In February Chris Sander had attended a symposium at Churchill College, Cambridge. I have no knowledge of what topics were discussed there.

Just before it was sent off to the journal FEBS Letters I changed the title of the manuscript. The title was altered so as to include all of the sequence homologies. Showing that the aligned segments were in agreement with the periodicity plot [2]. The new title proved to be an historic mistake and led to some unfounded and unkind criticism. I was told, for example, that our sequence motif was too long and that we did not understand the role played by zinc. In fact we were the first to set eyes on the Cys2 and His2 pattern and the sequence repeats in TFIIIA, almost a year earlier.


May I say again, as I did in my first letter, that I completely accept that you & Sander noticed the internal homologies last year.


I am glad you will get ~~the~~ credit for your noticing the homologies.

Excerpts from a letter sent by Aaron Klug dated 26 July 1985.

Finally, our luck had completely ran out. The referee at FEBS Letters was taken seriously ill. I believe this was because of a heart related issue. Unfortunately, this resulted in a six-week delay in the publication of our manuscript in FEBS Letters. Hearing of this setback, Pat Argos remarked to me that they, our rival group in Cambridge, will get all of the credit for figuring out the Cys2-His2 motifs in TFIIIA. Although both research groups had presented identical conclusions [2, 9], our insights into the structure and function of zinc fingers have been largely unrecognized, disregarded and mostly forgotten.

Paucis verbis

The C2H2-type zinc finger motif was discovered in 1984 at EMBL in the course of a secondary structure computer analysis. Short alpha helices were identified in TFIIIA that could recognize DNA.

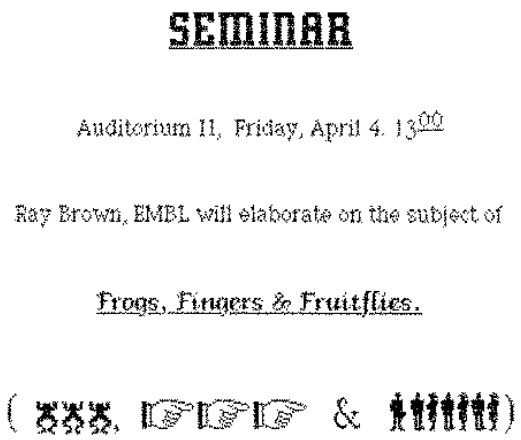
A Postscript

In mid-1985 two reports were published that identified multiple Cys2-His2 sequence repeats in Xenopus TFIIIA [2, 9]. The pivotal idea that "2 cysteines and 2 histidines may coordinate one zinc ion" was proposed by both research groups. One of these reports described some additional analysis and biochemical experiments.[xiv] These results from the 7S ribonucleoprotein storage particle were extrapolated to the TFIIIA protein.[xv] They predicted that the DNA-binding region was "a long, twisted ribbon of beta sheet" [9]. This model of the "zinc-binding domains" proved to be wrong. Our alternative proposal of short alpha helices, which "could interact with DNA in an extended mode", turned out to be correct [2, 6]. The Cys2-His2 motif, discovered in TFIIIA, was soon to acquire the name of zinc finger.[xvi]

The sequences of Serendipity *sry beta* and *sry delta* genes soon followed in November of 1985 [10]. These encoded two zinc finger proteins found in fruit flies. Six zinc fingers were present in the sry beta protein and seven zinc fingers in the sry delta protein.

Drosophila melanogaster - the fruit fly. Images from https://toppng.com/john3.

Eventually there was compelling evidence that the Cys2-His2 repeats were likely independent structural units. The first six zinc fingers of the TFIIIA protein were found to correspond exactly to six exons in the *gtf3a* gene [11]. I presented the state of knowledge as of spring 1986 on April 4 at the University of Aarhus, Denmark.


SEMINAR

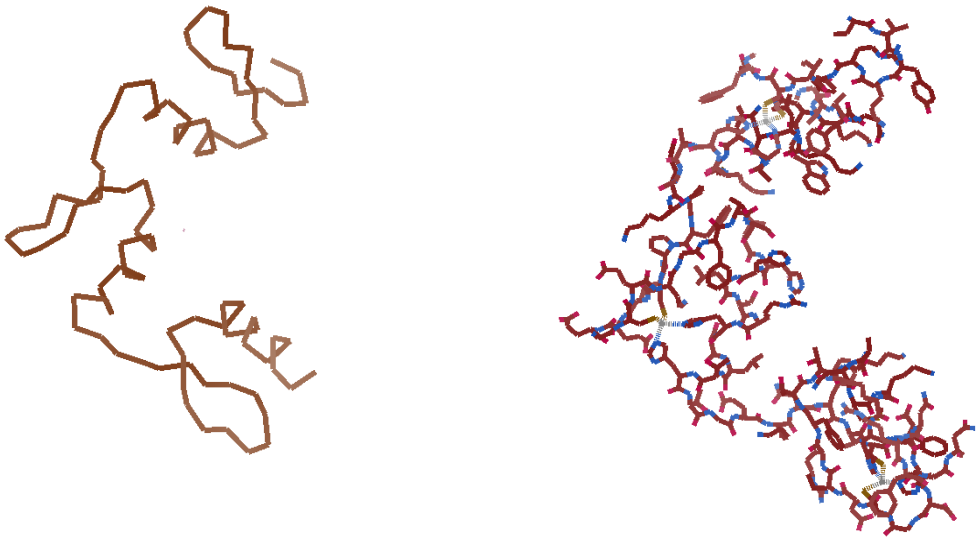
Auditorium II, Friday, April 4. 13[00]

Ray Brown, EMBL will elaborate on the subject of

Frogs, Fingers & Fruitflies.

(🐸, ☞☞☞ & 🧍🧍🧍🧍🧍🧍)

Efforts to prove the zinc finger hypothesis became a self-fulfilling prophecy. The first three-dimensional models of a zinc finger were built in 1987 by Jeremy Berg [12] and by Toby Gibson [13, 14]. Structural studies soon followed. Synthetic peptides, designer-constructs and protein fragments confirmed folding of the polypeptide chain. What came out was more or less equal to what was put into these projects. Importantly, however, our prediction of the short DNA-binding alpha helices in zinc fingers was vindicated. Zinc and its ligands in six of the TFIIIA fingers were finally revealed in 1998.[xvii] It was not a surprise that the N-terminal zinc fingers 1, 2 and 3 of TFIIIA wrapped around the DNA double helix [15]. Each zinc finger spreads over five or six base pairs in the major groove.[xviii] In contrast, zinc fingers 4, 5 and 6 were stretched out along the DNA double helix. TFIIIA is not a typical zinc finger protein. In this particular case the zinc fingers and their linkers have evolved so as to fit to both DNA and RNA structures [16]



The beta hairpin-alpha helix topology (on the left) and three-dimensional structure (on the right) of zinc fingers 1, 2 and 3 of Xenopus TFIIIA. Figure created from https://www.rcsb.org/structure/1TF6.

A Post Postscript

There are 1854 genes containing at least one C2H2 zinc finger in the human genome [17]. It is striking that these make up the second-largest class of human genes.

Endnotes

i A member of Donald Brown's laboratory at the Carnegie Institute of Washington in Baltimore, USA.

ii I am completely indebted to Riccardo Cortese who proposed this project in the summer of 1982. Riccardo Cortese, Fritz Winkler and I agreed on a research plan at a dinner party in my apartment in Boxbergring. That evening we dined on pigeon salmis and drank several bottles of Margaux wine. Later In February 1984 Riccardo suggested that I should write to Don Brown and ask for his unpublished cDNA sequence of TFIIIA. In my letter I stated that my interest was in the structure of TFIIIA.

iii This project was carried out at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany. https://www.embl.org.

iv This sequence of the polypeptide chain deduced from cDNA was 316 amino acids long. The first 28 amino acids of TFIIIA were missing from the sequence [8].

v Three to one letter amino acid code at https://www.bioline.com/media/calculator/01_17.html.

vi A basic sequence, KRKLKEKCPRPKRS spanning positions 280 to 293. The arrangement of lysines suggests it may be a nuclear localization signal. The C-terminus of TFIIIA has a putative nuclear export signal sequence, LVLDKLIQ at positions 337-344.

vii In the summer of 1983 I visited the Medical Research Council Laboratory of Molecular Biology in Cambridge, UK. Hans Christian Thøgersen and I pondered the unusual high percentage of histidine (7%) in the amino acid composition of p45 (TFIIIA) [1].

viii Chris Sander also noticed two sequences TGEKN and TGEKP in TFIIIA. These are examples of the universal TGERP "linker" sequence found in Cys2-His2-type zinc finger proteins.

ix A Dot Plot server is available at https://www.ebi.ac.uk/Tools/seqstats/emboss_dotmatcher/.

x Secondary structure prediction of multiple aligned sequences https://www.compbio.dundee.ac.uk/jpred/.

xi Other combinations of the Cys and His ligands could not be ruled out.

xii The discrepancy arose from an overestimation of the TFIIIA protein concentration. A factor of 0.62 was missing from their calculation using the Bio-Rad protein assay because of the unusual amount of color [3].

xiii See https://www.alice-in-wonderland.net/resources/pictures/caterpillar/.

xiv Digestion with trypsin cuts the TFIIIA protein 58 times while chymotrypsin cuts at 77 possible sites (https://web.expasy.org/peptide_cutter). Protection from cutting may be provided by 5S rRNA. Folded regions of the protein could also be resistant. Gel bands separated by 1, 2 and 9 kD were observed for trypsin and 1, 2, 1.5 and 2 kD for chymotrypsin digests [8]. Evidence for tightly-folded zinc fingers would be a "ladder" of gel bands with a spacing of 3.5 kD. In addition, I had found the 7S particle to be unaffected by as much as 20 mM dithiothreitol.

xv Neither of the research groups prepared any "zinc-binding domains" from TFIIIA for an analysis and characterization of their structure, composition or contents.

xvi In April 1986 at a symposium at the University of Kent, Louise Fairall told me that it was Horace Drew who coined the name "fingers" for the DNA-binding domains of TFIIIA.

xvii From 1985-1987 I studied crystals of 5S rRNA-TFIIIA and 5S rRNA-TFIIIA zinc fingers 1-9. These crystals diffracted to 8 Angstroms resolution. From 1988 to 1995 various DNA complexes containing zinc fingers 1-6, 1-7 and 7-9 of TFIIIA were crystallized. It became possible to collect X-ray data when in 1991 a way was found to freeze crystals in liquid nitrogen. Zinc was identified from its anomalous X-ray scattering.

xviii Some or all of the amino acid side chains at alpha helix positions -1,+3 and +6, bind to three consecutive DNA bases (in the 3'-5' direction). A "simple" triplet code does not include other contacts to DNA base pairs, for example, made by the amino acid at position +2 (https://dnaprodb.usc.edu).

Bibliography

[1] Isolation of a 7S particle from Xenopus laevis oocytes: a 5S RNA-protein complex. Picard B, Wegnez M. (1979) Proc Natl Acad Sci USA. 76, 241-245. [web site]

[2] The primary structure of transcription factor TFIIIA has 12 consecutive repeats. Brown RS, Sander C, Argos P. (1985) FEBS Letters 186, 271-274. [web site]

[3] Structural investigation of the Xenopus 7S ribonucleoprotein complex. Brown RS. (1989) PhD Thesis. Birkbeck College, University of London. [web site]

[4] Xenopus transcription factor A requires zinc for binding to the 5S RNA gene. Hanas JS, Hazuda DJ, Bogenhagen DF, Wu FYH, Wu CW. (1983) J Biol Chem. 258, 14120-14125. [web site]

[5] Similarities in active center geometries of zinc-containing enzymes, proteases and dehydrogenases. Argos P, Garavito RM, Eventoff W, Rossmann MG, Brändén CI. (1978) J Mol Biol. 126, 141-158. [abstract]

[6] Fingers and helices. Brown RS, Argos P. (1986) Nature 324, 215. [web site]

[7] Protein-DNA recognition. Pabo CO, Sauer RT. (1984) Annu Rev Biochem. 53, 293-321. [abstract]

[8] Xenopus 5S gene transcription factor, TFIIIA: Characterization of a cDNA clone and measurement of RNA levels throughout development. Ginsberg AM, King BO, Roeder RG. (1984) Cell 39, 479-489. [abstract]

[9] Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. Miller J, McLachlan AD, Klug A. (1985) EMBO J. 4, 1609-1614. [web site]

[10] Sequence and structure of the Serendipity locus of *Drosophila Melanogaster*. A densely transcribed region including a blastoderm-specific gene. Vincent A, Colot HV, Rosbash M. (1985) J Mol Biol. 186, 149-166. [abstract]

[11] Structure of the gene for Xenopus transcription factor TFIIIA. Tso JY, Van Den Berg DJ, Korn LJ. (1986) Nucleic Acids Res. 14, 2187-2200. [web site]

[12] Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Berg JM. (1988) Proc Natl Acad Sci USA. 85**,** 99-102. [web site]

[13] European Molecular Biology Laboratory Research Reports 1987, pages 149-150. [web site]

[14] A model for the tertiary structure of the 28 residue DNA-binding motif (‘zinc finger’) common to many eukaryotic transcriptional regulatory proteins. Gibson TJ, Postma JPM, Brown RS, Argos P. (1988) Protein Engineering 2, 209-218. [abstract]

[15] Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. Nolte RT, Conlin RM, Harrison SC, Brown RS. (1998) Proc Natl Acad Sci USA. 95, 2938-2943. [web site]

[16] Molecule of the Month-Zinc Fingers. Goodsell D. [March 2007] RCSB Protein Data Bank. [web site]

[17] The diversity of zinc finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. Lukic S, Nicolas J-C, Levine AJ. (2014) Cell Death Differ. 21, 381-387. [web site]

Caveat lector

The history of science relies upon those events that can be reconstructed from the published literature. What was envisioned, discussed or shared is completely missing from the historical record. Contents of a scientific report will likely favor the writer. The inevitable race to publish distorts the timeline of key events. A breakthrough in science is not always a sudden revelation but rather the ending of a long investgation. Happening when enough facts have accumulated from different sources. Many people could have been involved in setting the stage and preparing the ground. It is a great loss to the history of science that their contributions go unrecognized. Participants and events should not be left out simply because they detract from a good story. When the story is retold, there is a tendency to adapt and improve the narrative. In this way, what probably started merely as an educated guess, becomes transformed into an important “discovery”. This article is my personal memoir describing the sequence of steps, the insights, predictions and events that took place in 1984.

About the Author

Ray Brown worked at the European Molecular Biology Laboratory from 1978 to 1987. He received a PhD in Biophysics in 1989 from the University of London. His research career in molecular and structural biology spanned more than 45 years. Spending periods at the Medical Research Council Laboratory of Molecular Biology, Harvard University and the Universities of Aarhus and Connecticut and XTAL Biostructures Inc. His other interests are genealogy and French food.