EDGAR WINGENDER: DNA/PROTEIN SEQUENCE AND FUNCTION

MAKING COMPLEX DATA AVAILABLE FOR SCIENTIFIC ANALYSIS AND THE BEGINNINGS IN BRAUNSCHWEIG

JOHN COLLINS (HZI/GBF/TU BRAUNSCHWEIG, 1975-2010)

BIOINFORMATICS OF GENE REGULATION: 30 YEARS OF TRANSFAC (7-9 MARCH, 2018, GÖTTINGEN)

Edgar's time at the GBF:

01.1977 – 12.1977 Diplomarbeit with Professor Karl Wagner

05.12.1977 started Doktorarbeit, promoted 09.12.1980 (Karl Wagner, Jürgen Bode) till 30.04.1981 as Postdoc.

1981-1986 University of Marburg as Research Assistant (Lecturer) with Professor Klaus H. Seifart 1981 till 1986

Wingender, E. and Seifart, K. H. **Transcription in eukaryotes - the role of transcription complexes and their components** Angew. Chem. Int. Ed. Engl. 26, 218-227 (1987)

Jahn, D., Wingender, E. and Seifart, K. H. **Transcription complexes for various class III genes differ in parameters of formation and stability towards salt** J. Mol. Biol. 193, 303-313 (1987)

Wingender, E., Jahn, D. and Seifart, K. H. **Association of RNA polymerase III with transcription factors in the absence of DNA** J. Biol. Chem. 261, 1409-1413 (1986)

Wingender, E., Dilloo, D. and Seifart, K. H. **Zinc ions are differentially required for the transcription of ribosomal 5S RNA and tRNA in a HeLa-cell extract** Nucleic Acids Res. 12, 8971-8985 (1984)

Wingender, E., Shi, X. P., Houpert, A. and Seifart, K. H. **Isolation of a transcription complex for ribosomal 5S RNA** EMBO J. 3, 1761-1768 (1984)

Shi, X. P., Wingender, E., Böttrich, J. and Seifart, K. H. **Faithful transcription of ribosomal 5-S RNA in vitro depends on the presence of several factors** Eur. J. Biochem. 131, 189-194 (1983)

Returned to the GBF 01.02.1986 – 15.11.2002 as Research Assistant (Postdoc)

- founded Transfac 1987, spin-off from GBF ->-> founding of Biobase

1986 – 1990

Wingender, E. Compilation of transcription regulating proteins Nucleic Acids Res. 16, 1879-1902 (1988)

actin (cytoskeletal)	Xenopus laevis /HeLa	-94 to -75 (SRE)	1a, 4a, 4b	AAGATgcCCAtATtTGGcgATCTT	SRF (?)	4
Ad2MLP (adenovirus 2 major late prom.)	adenovirus/HeLa	-68 to -49	1a, 3, 4a, 4b	TGTAGGCCACGTGACCGG	UEF, USF, MLTF	5-10
		-63 to -52	1b	GGCCACGTGACC	USF	9
		-50 to -10	1a	TATAAAA	?	8
		-40 to +35	1a	TATAAAA	TFIID	9
		-34 to -22	1b	TATAAAA	TFIID	9
Adh (alcohol dehydrogenase) -distal prom. -proxim.prom.	Drosophila	- 85 to - 47	1a	GCTGCtGCTGCatcCGTCGaCGTCG	Adf-1	11
		-269 to -229	1a	TACTAA (4x)		11
		-151 to -105	1a	GCAGCGCTGCCGTCGccggctgaGCAGC	Adf-1 ?	11
		- 98 to - 77	1a	GAGATCGCGTAACGGTAGATAA		11
Adh1 (alcohol dehydrogenase)	maize	-190 to -186	4a (in vivo)	CCACG		12
		-145 to -138 (after ind.)	4a (in vivo)	CCCCGG		12
		-120 to -117 (after ind.)	4a (in vivo)	CGTGG		12
		-108 to -100	4a (in vivo)	CCCACAGGC		12
ADH2 (alcohol dehydrogenase)	yeast	-257 to -216	deletions	22 bp dyad symmetry	ADR1	13
albumin	rat/liver	-156 to -141	1a	GCAAGGGATTTAGTTA	NF-1	14
		-126 to -107	1a	TTTTTGGCAAGGAT		14
		-105 to - 89	1a	ATTTTGTAAT		14
		- 72 to - 35	1a	TGGTTAATGATCTACAGTTATTGGTTA		14
A-MuLV (amphotrop. murine leukemia virus)	/F9, PCC4	-87 to -59	1a, 3	CCAAT	EPBF	15
aP2 (adipocyte P2)	mouse/adipocytes	-124 to -108	1a, 3	AACATGACTCAGAGGAA	c-fos ?	16
BPV (bovine papil-loma virus-1)	/rec /HeLa	7613 to 7639 (enhancer)	1a	ATCGGTGCACCGAT	E2 ORF prod.	17
		7637 to 7667 (enhancer)	1a	TTGGGGctCCCCAA	AP-2	18
α2(I) collagen	mouse/3T3	- 95 to - 70	2a	CCAAT	NF-1	19, 20
		-315 to -295	1a, 2a, 3, 4b	TCGNNNNNGCCAA		19, 20
collagenase	human/HeLa,Hep82	-79 to -60	1a, 1b	ATGAGTCAGA	AP-1	21
CYC1	yeast	-293 to -278 (UAS 1, A)	1a, 3, 4b	CCGA	RAP	22
					HAP1	22

The meeting was designed to bring specialists from three traditionally diverse fields together, namely: human geneticists, molecular biologists and computer scientists. It will be increasingly necessary that these three groups interact in order to extract the most in terms of biological and medical information from the data being gathered on the various genomes being analysed (human, mouse, pin worm, fruit fly, yeast, bacteria etc.). This ongoing analysis will help focus on which methods are particularly efficient, where bottle-necks occur, in short, where the real emphasis in methodological developments, and where the real focus on biological (medical?) studies should be placed. Individual views

1st European Meeting
Human Genome Organisation (HUGO)

Hüthig

The long term aim of the human genome initiative is, with perhaps a total expenditure of some $ 2,000 to 3,000 million over the next 13 years, to yield a complete analysis of the genetic content of the human genome. Thus it is an important aim of this meeting to take

The meeting was designed to bring specialists from three traditionally diverse fields together, namely: human geneticists, molecular biologists and computer scientists. It will be increasingly necessary that these three groups interact in order to extract the most in terms of biological and medical information from the data being gathered on the various genomes being analysed (human, mouse, pin worm, fruit fly, yeast, bacteria etc.). This ongoing analysis will help focus on which methods are particularly efficient, where bottle-necks occur, in short, where the real emphasis in methodological developments, and where the real focus on biological (medical?) studies should be placed. Individual views on approach and emphasis vary greatly, and the opportunity to discuss this in a special

efforts so far and during the next two days, to wish you all an interesting, profitable and enjoyable meeting, and to open the meeting by calling for the first lecture on "Genome analysis and cancer" by the president of HUGO, Sir Walter.

1986 – 1990

Sömjen, D., Binderman, I., Schlüter, K.-D., Wingender, E., Mayer, H. and Kaye, A. M. **Stimulation by defined parathyroid hormone fragments of cell proliferation in skeletal derived cell cultures** Biochem. J. 272, 781-785 (1990)

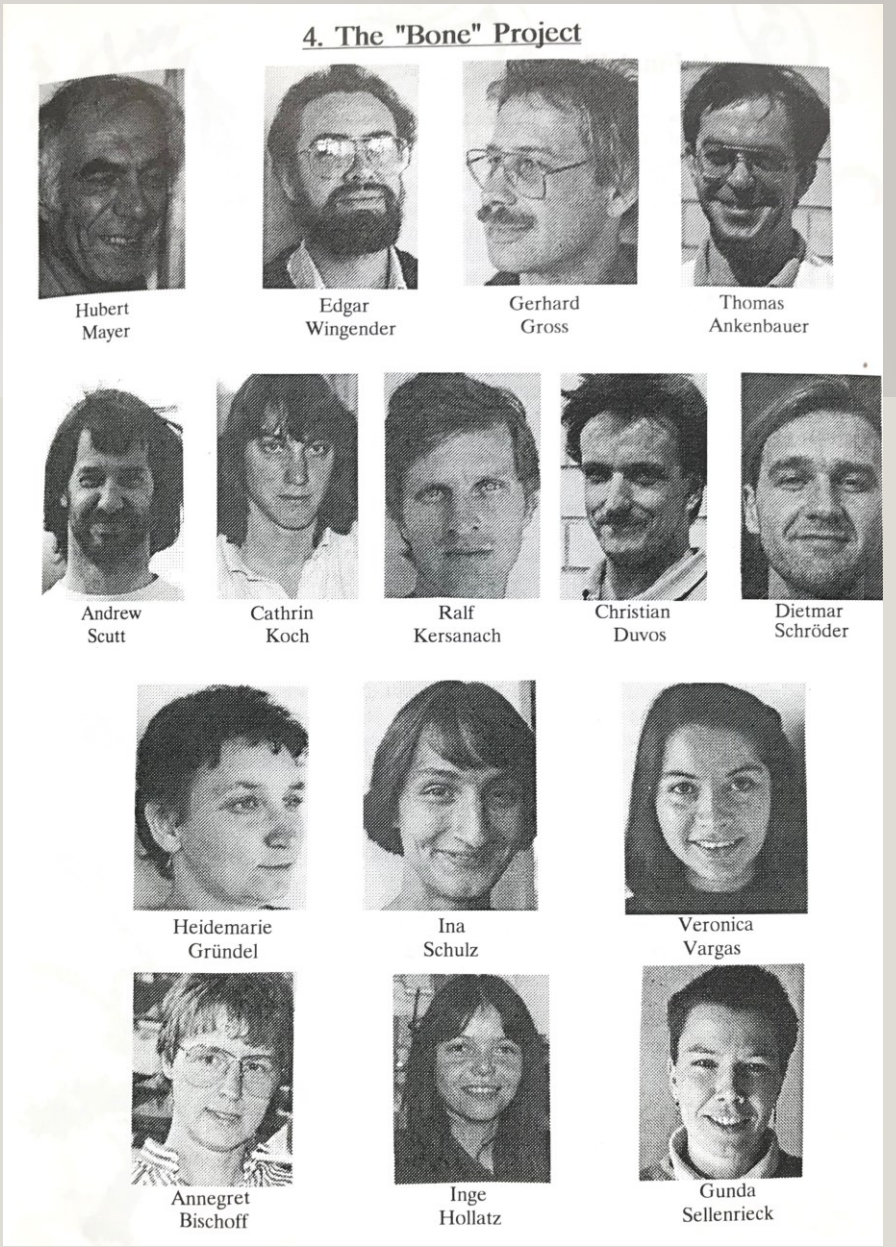
Rupp, E., Mayer, H. and Wingender, E. **The promoter of the human parathyroid hormone gene contains a functional cyclic AMP-response element** Nucleic Acids Res. 18, 5677-5683 (1990)

Wingender, E. **Transcription regulating proteins and their recognition sequences** CRC Crit. Rev. in Eukaryotic Gene Expression 1, 11-48 (1990)

Wingender, E., Bercz, G., Schlüter, K.-D. and Mayer, H. **Structure-Function Relationship in Parathyroid Hormone** Advances in Protein Design - International Workshop 1988 (H. Blöcker, J. Collins, R. D. Schmidt, eds.), GBF Monographs, Verlag Chemie, Weinheim 12, 167-176 (1989)

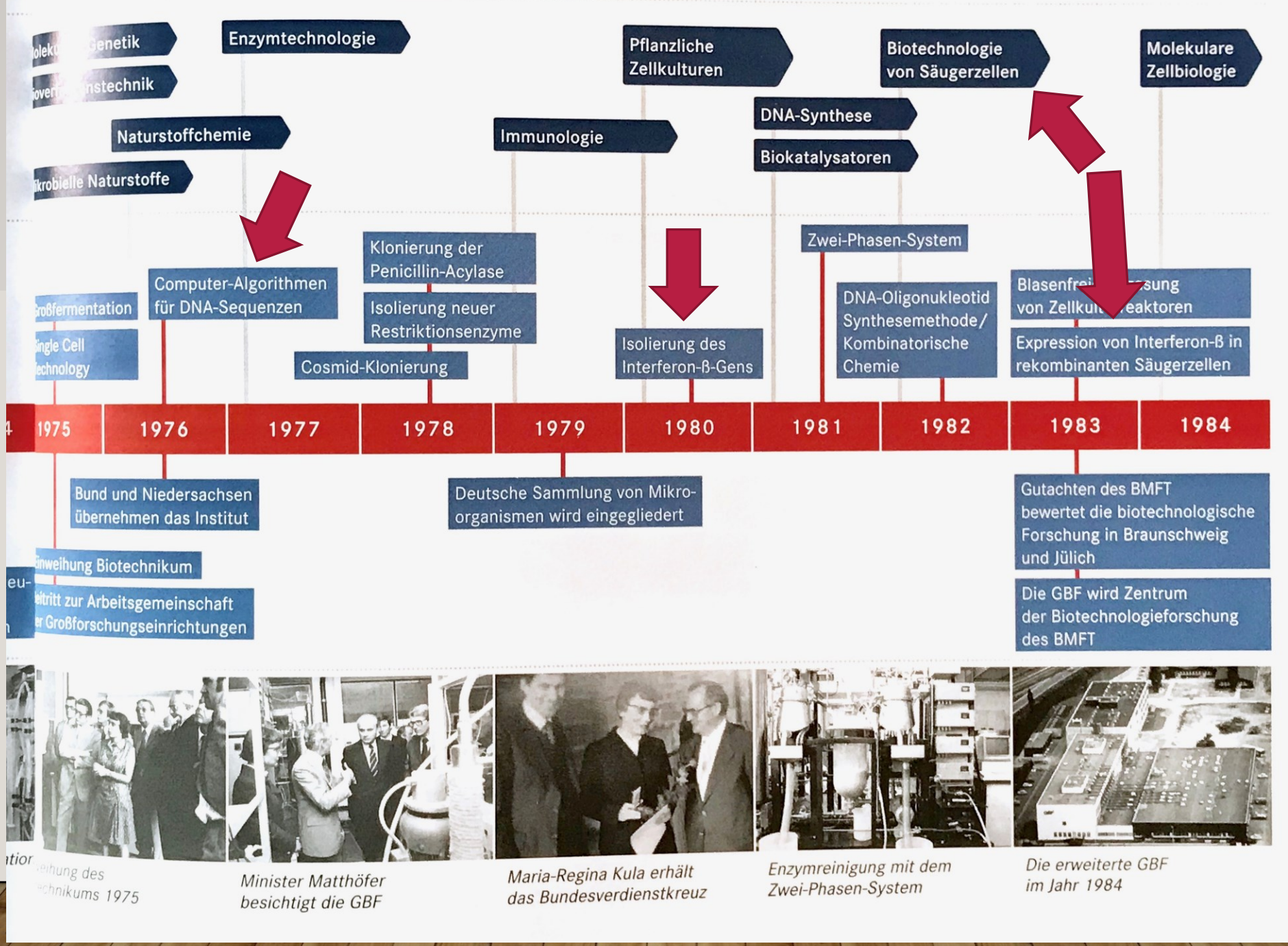
Schlüter, K.-D., Hellstern, H., Wingender, E. and Mayer, H. **The central part of parathyroid hormone stimulates thymidine incorporation of chondrocytes** J. Biol. Chem. 264, 11087-11092 (1989)

Wingender, E., Bercz, G., Blöcker, H., Frank, R. and Mayer, H. **Expression of human parathyroid hormone in *Escherichia coli*** J. Biol. Chem. 264, 4367-4373 (1989) Seifart, K. H.,



Rudi Balling, GBF/HZI
Director 2001-2009

Leopold Flohé,
GBF, Director 1991-1994



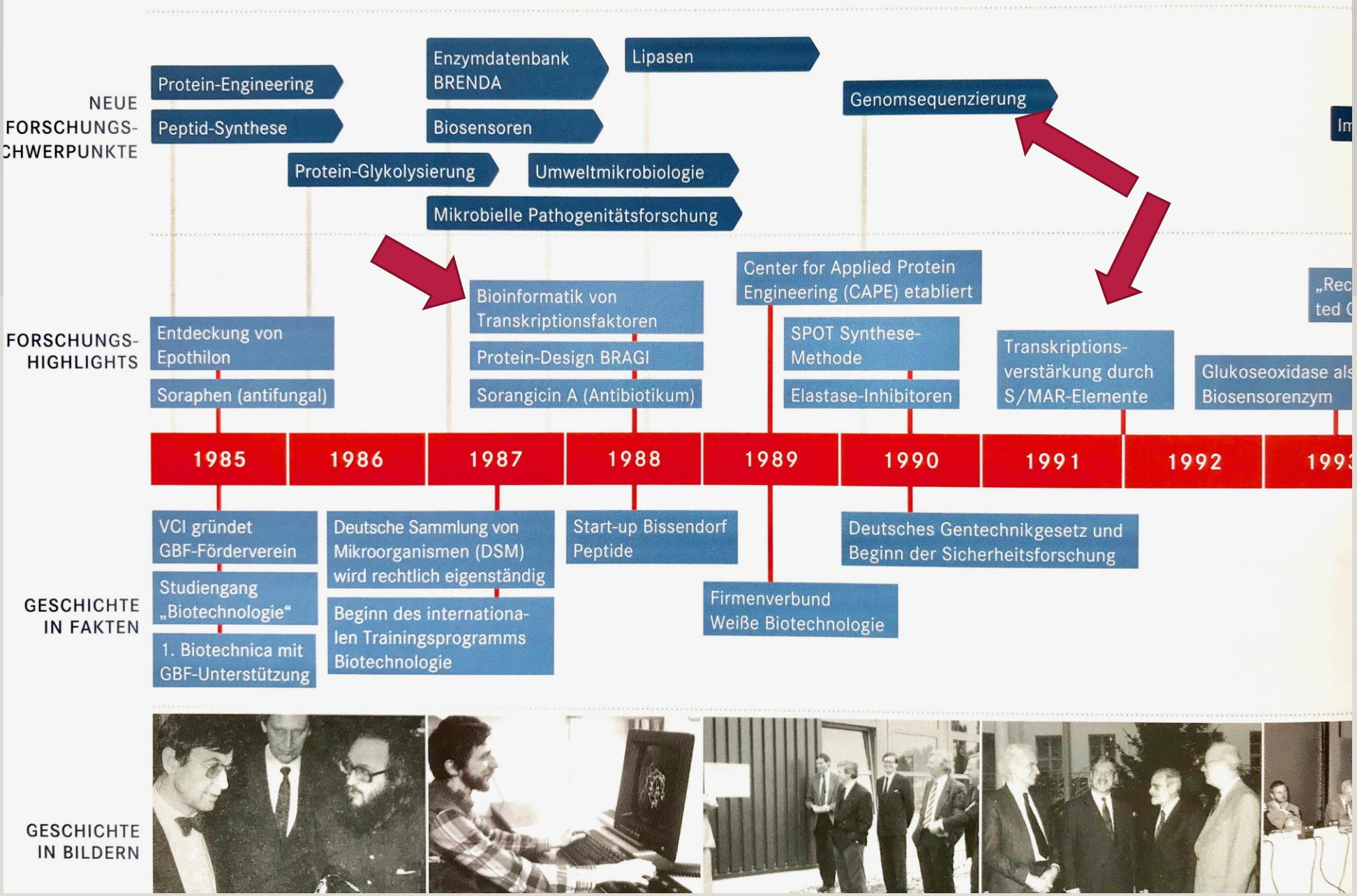
...hung des ...chnikums 1975

Minister Matthöfer besichtigt die GBF

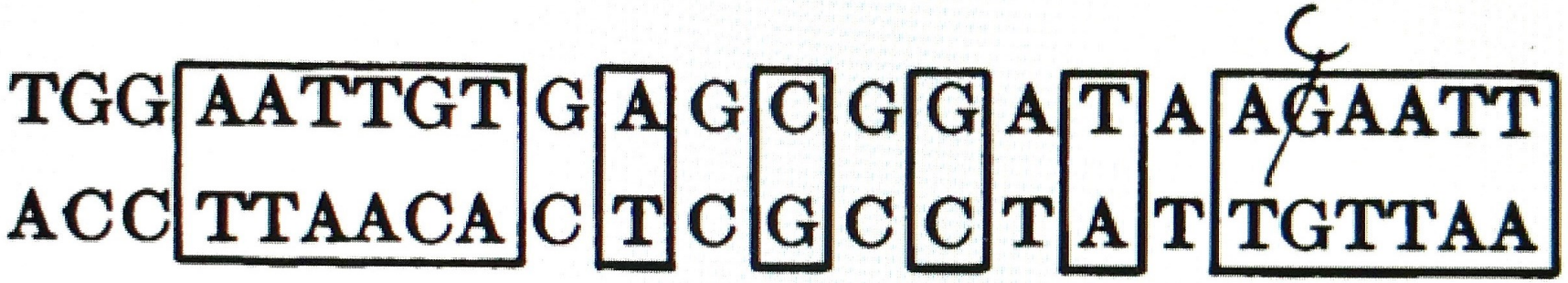
Maria-Regina Kula erhält das Bundesverdienstkreuz

Enzymreinigung mit dem Zwei-Phasen-System

Die erweiterte GBF im Jahr 1984


NEUE FORSCHUNGS-SCHWERPUNKTE
Protein-Engineering
Peptid-Synthese
Enzymdatenbank BRENDA
Biosensoren
Lipasen
Protein-Glykolysierung
Umweltmikrobiologie
Mikrobielle Pathogenitätsforschung
Genomsequenzierung
FORSCHUNGS-HIGHLIGHTS
Entdeckung von Epothilon
Soraphen (antifungal)
Bioinformatik von Transkriptionsfaktoren
Protein-Design BRAGI
Sorangicin A (Antibiotikum)
Center for Applied Protein Engineering (CAPE) etabliert
SPOT Synthese-Methode
Elastase-Inhibitoren
Transkriptions-verstärkung durch S/MAR-Elemente
Glukoseoxidase als Biosensorenzym
1985
1986
1987
1988
1989
1990
1991
1992
GESCHICHTE IN FAKTEN
VCI gründet GBF-Förderverein
Studiengang „Biotechnologie“
1. Biotechnica mit GBF-Unterstützung
Deutsche Sammlung von Mikroorganismen (DSM) wird rechtlich eigenständig
Beginn des internationalen Trainingsprogramms Biotechnologie
Start-up Bissendorf Peptide
Firmenverbund Weiße Biotechnologie
Deutsches Gentechnikgesetz und Beginn der Sicherheitsforschung
GESCHICHTE IN BILDERN

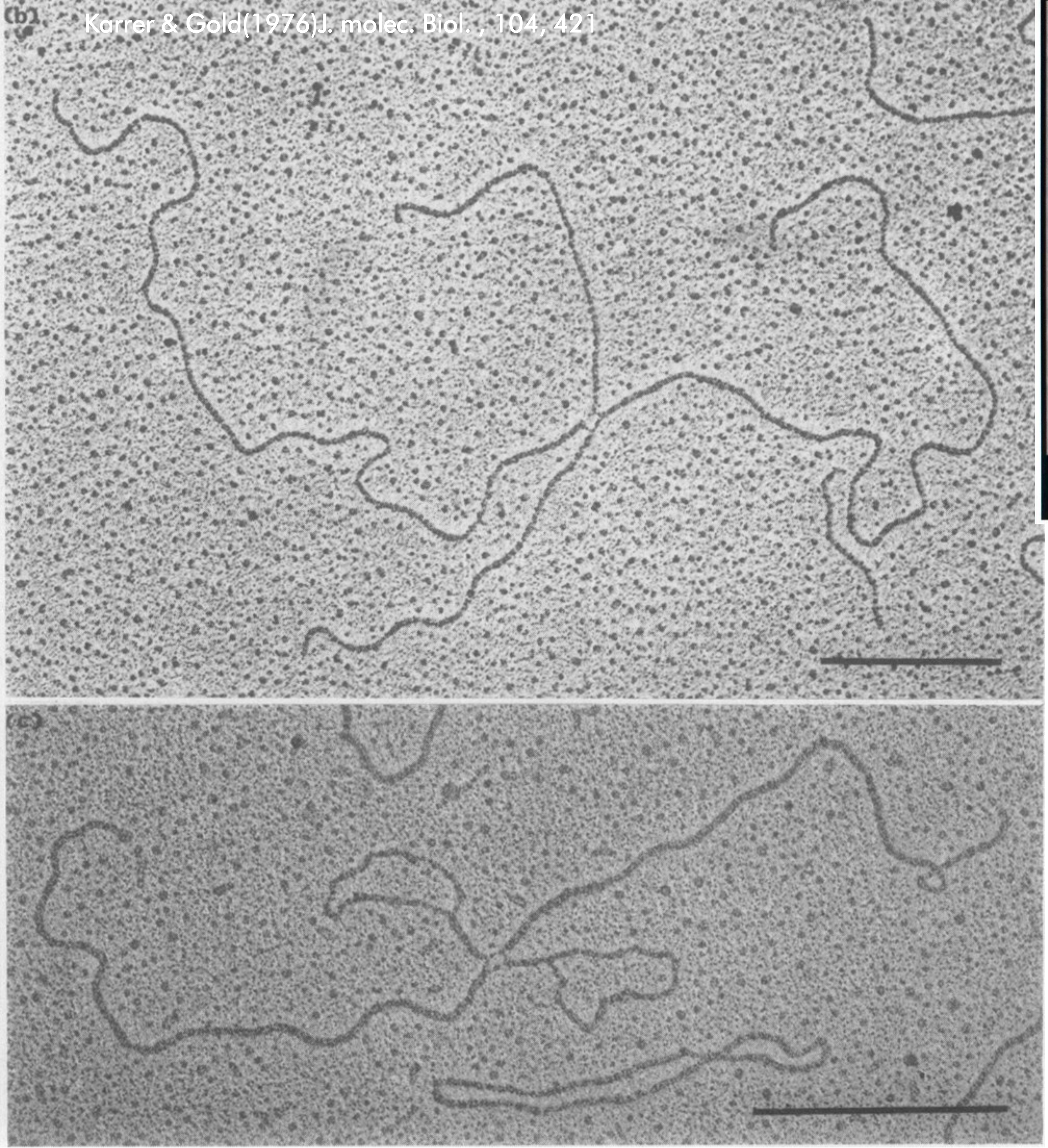
below:

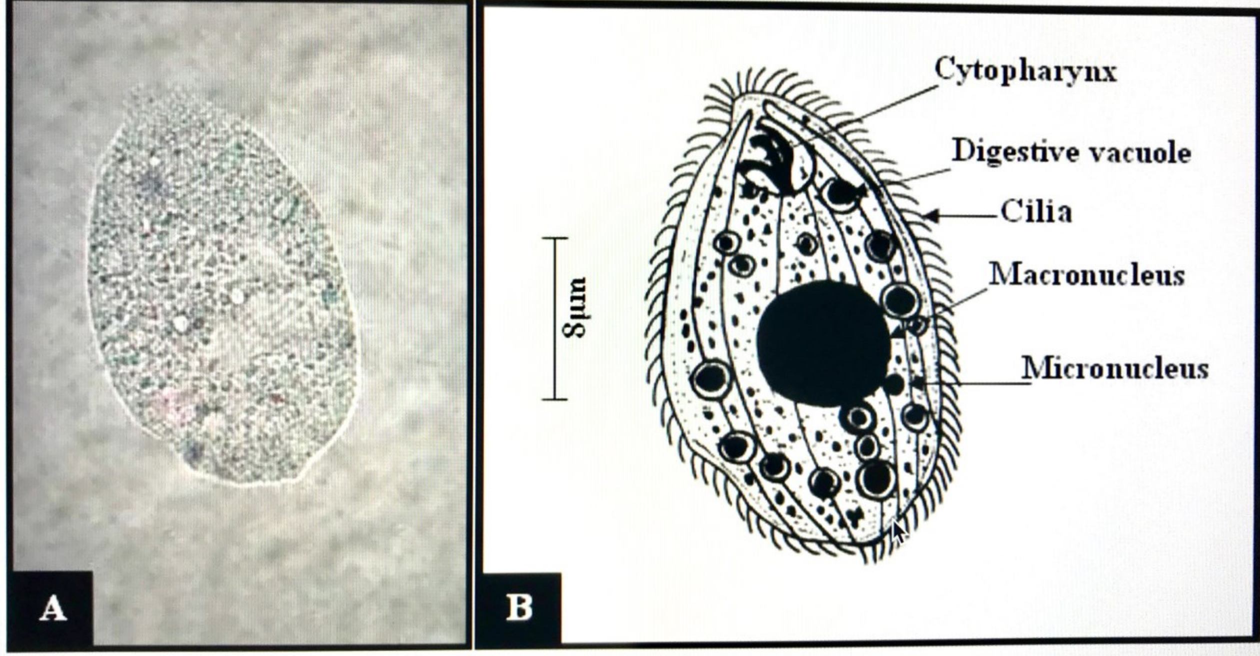


(However, two to three of these matched base pairs might be symmetrical by chance.) This symmetry could permit the repressor protein to interact with DNA on a 2-fold symmetry axis: either two subunits could interact, each with the appropriate half of the symmetry region; or (as suggested by Dr. Thomas Steitz) all four subunits could interact, two with the left-hand region, the other two, related by a 2-fold axis, with the right-hand region. A 2-fold symmetry was suggested on

Late 1975 Jan Engberg (Panum Medical School, Copenhagen) and myself writing on our work and ideas on the role of symmetry in DNA regulation in *Tetrahymena pyriformis*

Karrer & Gold(1976)J. molec. Biol. , 104, 421





Free Ribosomal DNA Molecules from *Tetrahymena pyriformis* GL are Giant Palindromes

JAN ENGBERG, POUL ANDERSSON, VAGN LEICK

Biochemical Institute B
University of Copenhagen
Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

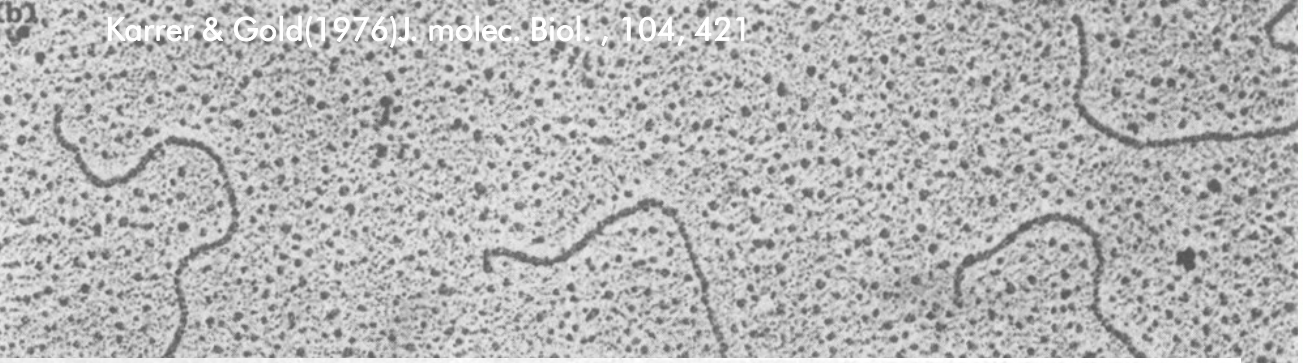
AND

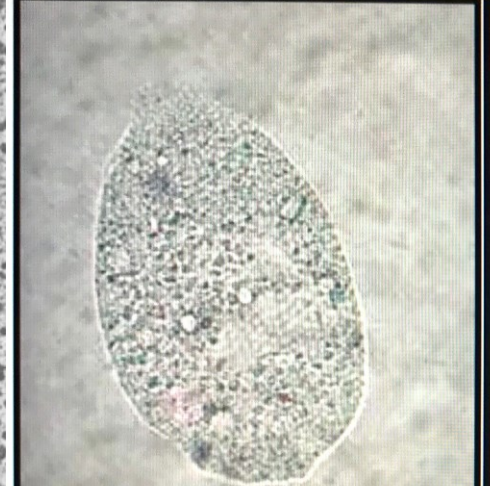
JOHN COLLINS†

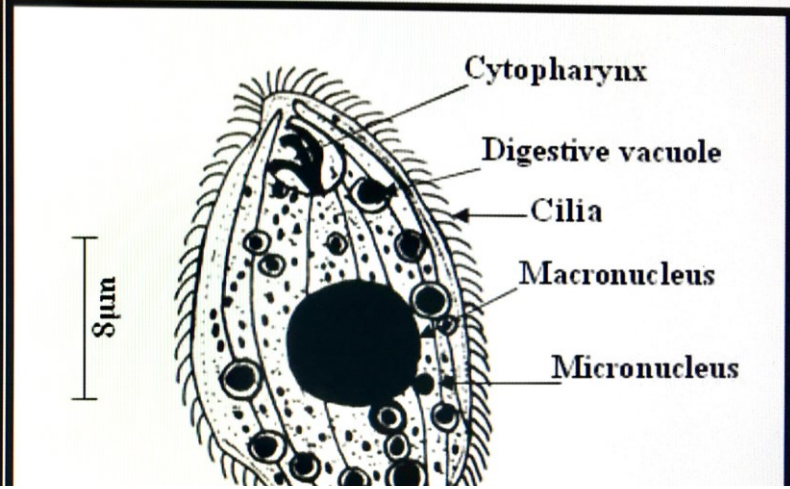
Institute of Microbiology
University of Copenhagen, Øster Farimagsgade 2A, DK-1353 Copenhagen K
Denmark

(Received 29 December 1975, and in revised form 16 February 1976)

Restriction endonuclease *Eco*RI was used to study the structure of the free ribosomal DNA molecules from *Tetrahymena pyriformis*, strain GL. From the







(7)

J. theor. Biol. (1977) **66**, 573–582

A Model for Switching on Ribosomal RNA Synthesis by Creating a Palindromic DNA Sequence in the Promoter Region of the Ribosomal RNA Cistron: the "Structon"

JOHN COLLINS

Gesellschaft fur Biotechnologische Forschung mbH, D-3300 Braunschweig-Stockheim, Mascheroder Weg 1, *Federal Republic of Germany*

AND

JAN ENGBERG

Biochemical Institute B, University of Copenhagen, Blegdamsveg 3, 2200 *Copenhagen, Denmark*

(*Received* 21 *October* 1976)

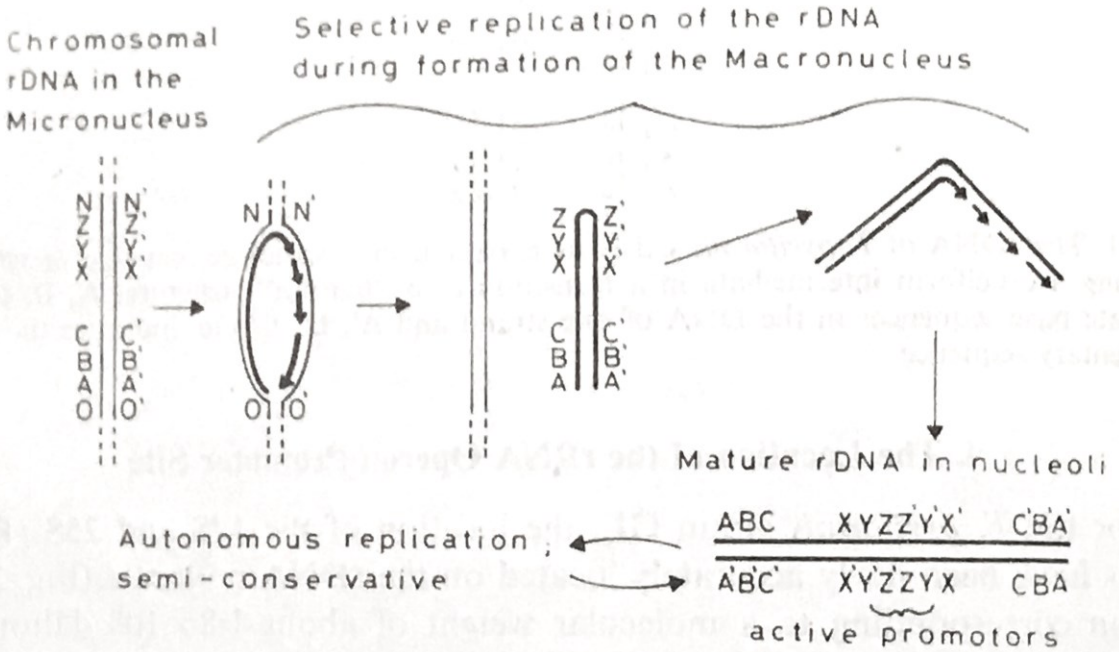
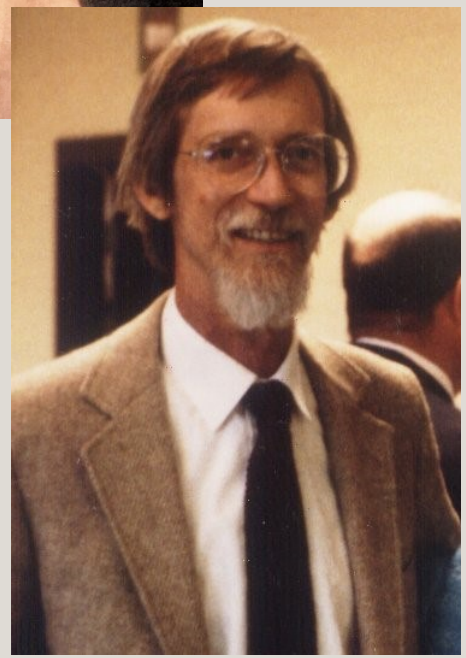


FIG. 2. The "structon"; a segment of the micronuclear chromosomal DNA equivalent to just over one ribosomal RNA cistron is selectively replicated through a "hairpin" intermediate, during the development of the macronucleus. The "hairpin" is then replicated to the open linear form which is found in the nucleoli. The model further postulates that NZYX alone is not sufficient to act as a promoter, but some element in Y′Z′ZYX is capable of functioning as a promoter leading the efficient rRNA synthesis only in the nucleolus.

It therefore **seemed imperative** to us to have a tool to analyse DNA sequences for various possible structures including initially **symmetric regions** (*palindromic sequences*) ***even though we could analyze ALL the sequences that were available in 1975/76 by hand in a short time.***

Generally useable sequencing technology *didn‘t come into being until the* ***end of 1977****: Maxam & Gilbert and the Sanger method which required either isolated end-labelled (P^{32}) DNA fragments or single-stranded DNA (e.g. from an M13-cloned fragement) and a synthetic DNA primer*



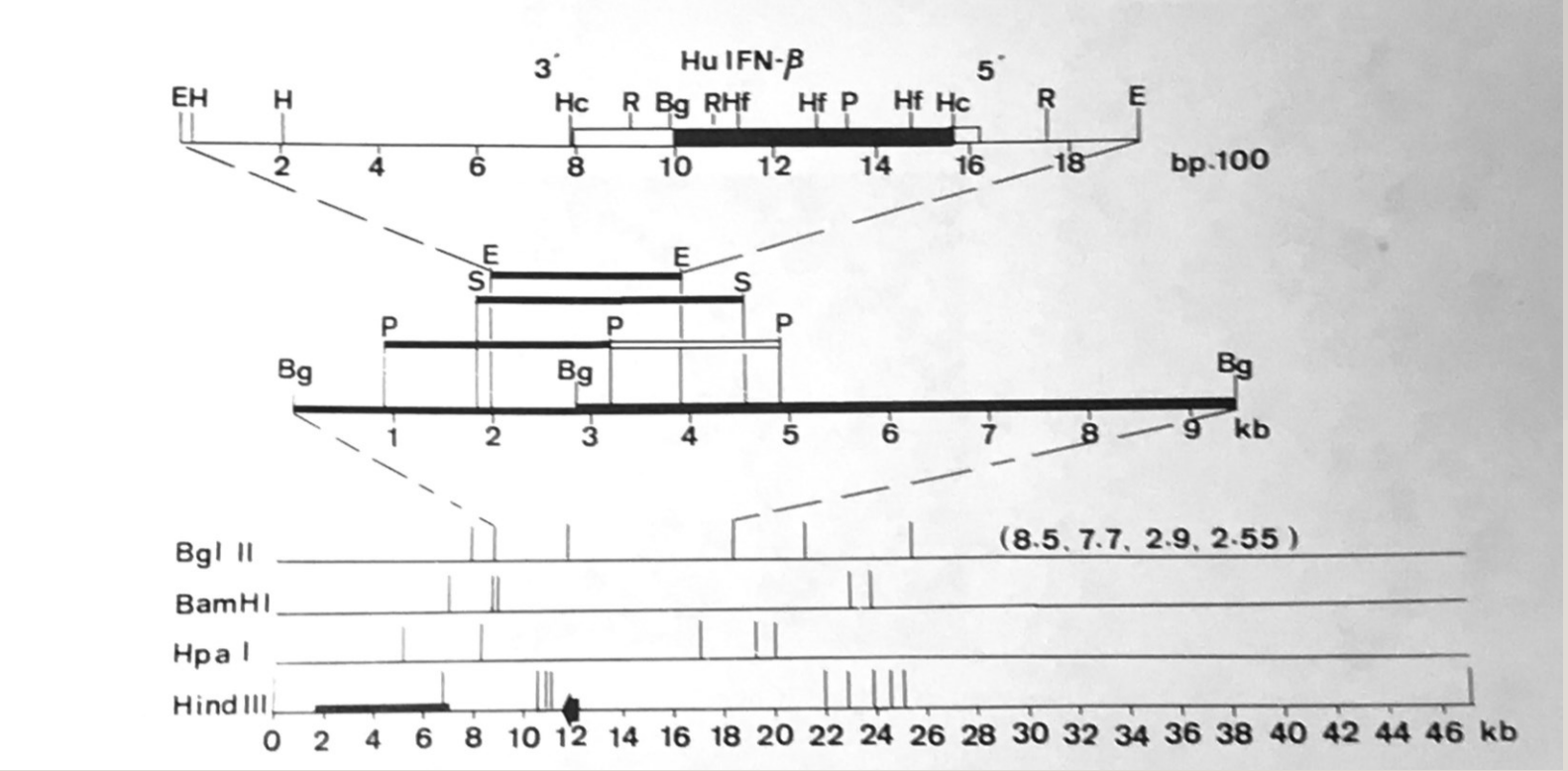
First computer algorithms searching for inverse symmetry (palindromic sequences)
Dave Lincoln & John Collins, 1975

	binary		binary
A= 3	011	G= 6	110
T= 4	100	C= 1	001
A+T = 7	0111	C+G = 7	0111
A+A= 6	0010	C+T= 5	0101
T+T= 8	1000	G+T= 10	1010
G+G= 12	1100	A+G= 9	1001
C+C= 2	0010	C+A = 4	0100

****

Choose length of scan for symmetry e.g. 6 bp:
DNA sequence
.....G G A A T T C T A A.....
a b c d e f ->
[(a+f)-7] if SUM= 0 -> [(b+e)-7,
if SUM =0 -> [(c+d)-7] ->
if SUM =0 -> write out sequence and position (first steps to GENMON and computi
on small computers compared to mainframe).

Later modified for partial homology (consensus) and partial symmetry comparisons.



Collins J. and Hohn B. (1978) Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage Lambda heads. Proc. Natl. Acad. Sci. USA 75, 4242-4246

Gross G., Mayr U., Bruns W., Grosveld F., Dahl H. H. M. and Collins J. (1981) The structure of a thirty-six kilobase region of the human chromosome including the fibroblast interferon gene IFN-ß. Nucl. Acids Res. 9, 2495-2507

Hauser H., Gross G., Bruns W., Hochkeppel H.-K., Mayr U. and Collins J. (1982) Inducibility of human ß-interferon gene in mouse L-cell clones. Nature (Lond.) 297, 650- 655

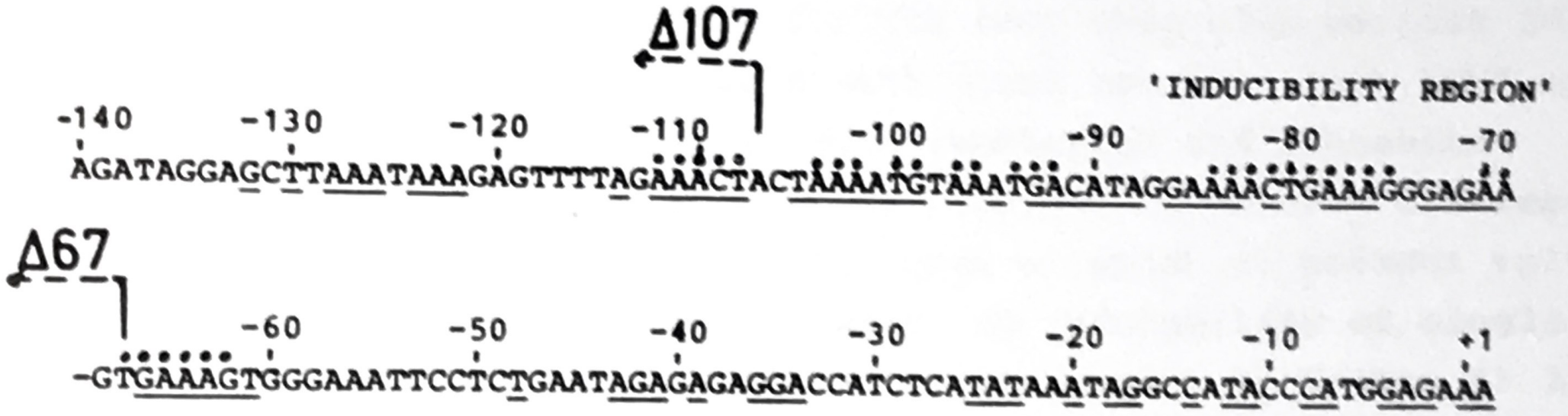


Fig. 4. The immediate 5'distal sequence to the transcription start (designated +1) of the IFNß gene are represented as well as the limits of the deletions Δ-107 and Δ-67 which define the borders of the "inducibility region". Homology of this sequence to the equivalent huIFNα1 sequence is underlined. Short sequence runs having high homology to the concensus sequence AAACTGAAAG are dotted.

Edgar, thanks to you, and all those who worked with you, to facilitate interpretation of the wealth of data that accumulated on gene regulation, structure and function.

Wishing you a happy and healthy retirement!