November 15, 1984

Dear Dan:

I have almost completed the sequencing of the mPRL, mGH, and mPLC cDNA clones, but I am now at the point that I can predict the complete amino acid sequence for each protein. A few ambiguities remain in two of the sequences in the 3'-untranslated region and part of these sequences have been determined on only one strand (due to the limitations of reading through the G-C tails). I need to verify the mPRL sequence because of two odd findings. First, the published protein sequence for mPRL begins with Leu-Pro-Ile-etc. The cDNA sequence predicts that the amino acid at the -1 position is a proline, and no protein reported in Von Heijne's review paper contains a proline residue at this position. Second, the initiator ATG for mPRL, based on homology to other prolactins and on the length of the leader sequence, is just downstream of another ATG, and this other ATG appears to obey the rules for a strong translation start site. If I can verify these findings by sequencing independent mPRL cDNA clones, they may add some twists to the signal sequence cleavage and translation initiation dogmas.

I am enclosing the amino acid sequences for mPRL, mGH, and mPLC (the placental cDNA clone that hybridizes weakly to proliferin). They are given with the homology of mPLC to each of mPRL, mGH, and mPLF. mPLC has several interesting features: (1) as predicted from the larger mRNA, mPLC encodes a protein larger than any of the other three - 244 amino acids in the precursor; (2) mPLC lacks the amino terminal two Cys residues (the same as mGH) and has the remaining 4 Cys residues common to the other three proteins, but it has an extra Cys as well (circled in green); (3) the protein extends beyond the final Cys, again similar to GH; (4) this stronger structural homology to GH is not repeated in comparisons of sequence homology; (5) the first 10 amino acids are identical to those predicted from the sequence of PLF-1 (this is extremely surprising since the Leu at position 5 in PLF-1 is predicted to be a Ser from PLF-2 and from Krebs RNA); (6) this region of homology is reflected in the nucleotide sequence - see the other enclosure - where mPLF and mPLC are nearly identical throughout the 5'-untranslated region and into the coding region (this homology is suggestive of an exon shuffling event in the not too distant past; I do not know the exon boundary, but my guess is that it will come between the codons for amino acids 10 and 11; undoubtedly, this region explains the cross-hybridization); (7) certain regions in mPLC are more homologous to bPRL than to mPRL, suggesting that mPLC and mPRL have drifted apart in those regions following gene duplication faster than a mouse and a bovine gene have drifted apart from a common gene in some mammalian ancestor; for example:

	92										135							
mPLC	I	T	K	A	F	N	S	C	H	T	P	L	K	H	L	V	T	E
mPRL	M	V	K	V	I	N	D	C	P	T	P	L	F	Q	L	I	T	G
bPRL	I	T	M	A	L	N	S	C	H	T	P	L	Y	H	L	V	T	E

(8) mPLC has 3 potential glycosylation sites, all in the amino terminal half of the protein, and 2 Arg-Lys sequences. I am sure that you will find some other points of interest that I have neglected. I have not yet put together the sequence comparisons between the mammalian GH genes and between the mammalian PRL genes; that will be done in the near future.

I am curious to hear how the proliferin story is developing outside of Illinois. Although I am quite busy here and am enjoying my interactions with the other labs in the department, I do feel a bit out of touch. Finally, I have another favor to ask of you. I have applied for a Basil O'Conner Starter Grant from the March of Dimes, and they would like a letter from you describing my potential to become an independent scientist. Would you be willing to write on my behalf to

Samuel J. Ajl, Ph.D.
Vice President for Research
Basil O'Conner Starter Research Grant
March of Dimes Birth Defects Foundation
1275 Mamaroneck Avenue
White Plains, New York 10605

The application and supporting letters are due by January 1, 1985. Thank you in advance.

Sincerely,

Dan

P.S. 1) Did you take care of the respone to Philipson, or should I?
2) When I write up the PLC sequence, do we want to send a preprint to our patent attornies?
3) I read Nilsen-Hamilton's papers after learning of them from Se-Jin — is anything happening on that front?