

October 2, 2007

BI/CH 421/621  
Biochemistry I

TOLAN

**Sample Protein Sequencing Problem  
ANSWER**

[NOTE: answer/logic is in BOLD. Newly placed sequences are underlined and those within parentheses or brackets means that the sequence is unclear, but that those amino acids are within the region of the parentheses.]

The amino acid composition of human  $\alpha$ -melanocyte-stimulating hormone is: Ala, 2 Gly, 3Glu, 2 Asp, 3Lys, 3 Pro, Trp, 2 Arg, Met, His, Phe, Tyr, and Ser. Edman degradation gave E1=Ala, E2=Glu.

**hormone is 22 amino acids long; first two are Ala-Glu**

A. Cyanogen Bromide treatment gave two peptides with the following amino acid composition and Edman degradation (E#):

Cn-1; Gly, Glu, Asp, Lys, 2 Pro, Arg, Ser, Trp, Phe, His  
E1=Glu, E2=His

**This is C-terminal because it does not have the product of CNBr treatment of Met. It has sequence beginning with Glu-His**

Cn-2; Gly, Ala, 2 Glu, 2 Lys, Arg, homoserine lactone, Tyr, Pro, Asp

**This is N-terminal due to homoserine lactone, which is the product of CNBr treatment of Met. From above we know the sequence so far as; Ala-Glu-X-X-X-X-X-X-Met.**

B. Trypsin hydrolysis gave six peptides with the following amino acid composition and Edman degradation:

T-1; Asp

**This must be THE C-terminus due to lack of Lys or Arg.**

T-2; Ala, Lys, Glu

**This is a peptide that ends in Lys, also from the fact that there is only one Ala and it is the N-terminal amino acid for the hormone, the sequence must be; Ala-Glu-Lys.**

T-3; Lys

**This means there is a Lys-Lys or Arg-Lys sequence somewhere in the protein.**

T-4; Phe, His, Glu, Arg, Met

**This means there is an X-X-X-X-Arg peptide, but from above we know that since the T4-peptide contains a Met and the sequence following the Met goes Glu-His that it must be either Met-Glu-His-Phe-Arg or Phe-Met-Glu-His-Arg. But, because the only Phe in the entire hormone is in the Cn-1 peptide, the Phe must be C-terminal to the Met, therefore, the peptide has the sequence Met-Glu-His-Phe-Arg.**

**So far we know this much sequence:**

**Ala-Glu-Lys-X-X-X-X-X-X-X-Met-Glu-His-Phe-Arg-X-X-X-X-X-(Lys/Arg)-Asp**

T-5; Gly, Tyr, Arg, Pro, Glu, Asp

E1=Asp, E2=Glu, E3=Gly

**This means there is an Asp-Glu-Gly-(Tyr-Pro)-Arg peptide with the order of Tyr-Pro or Pro-Tyr unclear. This is also a N-terminal peptide before the Met (Cn-2) due the presence of the Tyr. Because we know that there is the T-2 peptide with Ala-Glu-Lys, we know that it must be at the N-terminus and the next amino acid is either the single Lys (T-3) or this peptide (T-5). So far we know this much sequence:**

**Ala-Glu-Lys-[(Lys)-(Asp-Glu-Gly-(Tyr-Pro)-Arg)]-Met-Glu-His-Phe-Arg-X-X-X-X-X-(Lys/Arg)-Asp**

October 2, 2007

BI/CH 421/621  
Biochemistry I

TOLAN

T-6; Lys, Ser, 2Pro, Gly, Trp

This means there is X-X-X-X-X-Lys sequence in the whole peptide. And, because of the presence of the unique Ser, this also means that this is a C-terminal peptide after the Met (Cn-1). It must therefore fit in the unknown region above. Our sequence in then:

Ala-Glu-Lys-[(Lys)-(Asp-Glu-Gly-(Tyr-Pro)-Arg)]-Met-Glu-His-Phe-Arg-(Ser-Pro-Pro-Gly-Trp)-Lys-Asp

C. Chymotrypsin digestion gave four peptides with the following amino acid composition and Edman degradation:

C-1; Gly, Ser, 2Pro, Lys, Asp  
E1=Gly, E2=Pro, E3=Ser

This must be the C-terminal peptide due to the lack of any bulky hydrophobic amino acids (Tyr, Phe, Trp). The sequence must have an Lys-Asp at the C-terminal end and a Gly-Pro-Ser at the N-term. Since this is a six member peptide there is only one amino acid remaining, a Pro. ALSO, since we know from T-6 that there is a Trp in the T-6 peptide, it must be the chymotryptic cleavage point that created this C-1 peptide. Our sequence in then:

Ala-Glu-Lys-[(Lys)-(Asp-Glu-Gly-(Tyr-Pro)-Arg)]-Met-Glu-His-Phe-Arg-Trp-Gly-Pro-Ser-Pro-Lys-Asp

C-2; Phe, His, Glu, Arg, Met

This confirms that the placement of T-5 in the Cn-2 peptide is after the single Lys of T-3, otherwise it would have been a Lys in this peptide rather than the Arg. Our sequence in then:

Ala-Glu-Lys-Lys-Asp-Glu-Gly-(Tyr-Pro)-Arg-Met-Glu-His-Phe-Arg-Trp-Gly-Pro-Ser-Pro-Lys-Asp

C-3; Trp, Arg

This also confirms the T-4 to T-6 order. Our sequence in then:

Ala-Glu-Lys-Lys-Asp-Glu-Gly-(Tyr-Pro)-Arg-Met-Glu-His-Phe-Arg-Trp-Gly-Pro-Ser-Pro-Lys-Asp

C-4; 2Glu, Gly, Pro, Asp, 2Lys, Ala, Tyr

This clears up our last confusion surrounding the Pro-Tyr vs. Tyr-Pro. Since this chymotryptic peptide has a Pro, it must cleave after the Tyr and not between the Tyr-Pro, which would leave the Pro out of this peptide.

Our sequence in then COMPLETE:

Ala-Glu-Lys-Lys-Asp-Glu-Gly-Pro-Tyr-Arg-Met-Glu-His-Phe-Arg-Trp-Gly-Pro-Ser-Pro-Lys-Asp

Be aware, there are other methods and logic to solve this problem.