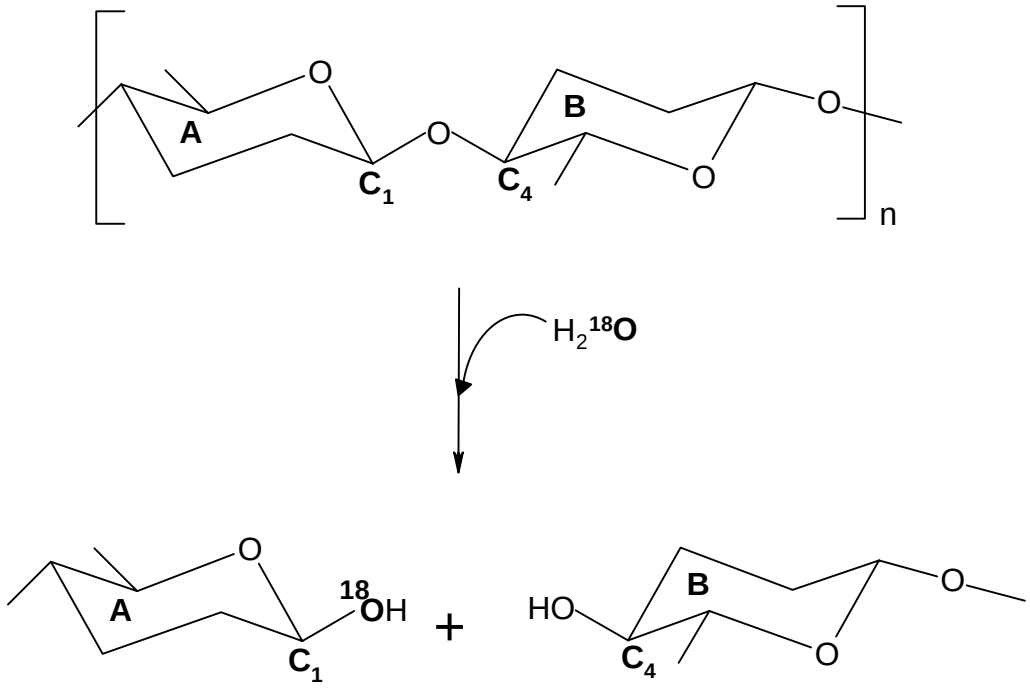
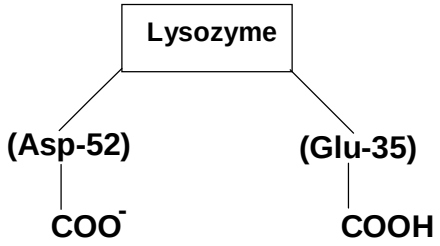
d. In the hydrolysis of the glycosidic bond, the glycosidic bridge oxygen goes with C_4 of the sugar **B**. On cleavage, ^{18}O from water will be found on C_1 of sugar **A**.



NOTE: The reaction proceeds with a carbonium ion stabilized on the C_1 of sugar **A**.

e. Most glycosidases contain two carboxylates at the active site that are catalytically important. Lysozyme is active only when one carboxylate is protonated and the other is deprotonated. A descending limb on the alkaline side of the pH profile is due to ionization of -COOH. An ascending limb on the acidic side is due to protonation of -COO^-. Thus the enzyme activity drops sharply on either side of the optimum pH. The ideal state of ionization at pH = 5 will be,



NOTE: It is desirable to study the amino acid side chains (R-groups) and their ionization properties. The pKa values of these groups significantly determine the pH dependence of enzyme activity.

f. Answers 2 and 4 are correct. Ionization of –COOH leads to generation of a negatively charged species, $–COO^-$. This charged species is poorly stabilized by diminished polarity and enhanced negative charge. Hence ionization of –COOH group is suppressed and the pKa is elevated.

g. The ratios of pseudo-first order rate constant (at 1M CH_3COO^-) in (a) to the first order rate constants in (b) and (c) provide the effective local concentrations.

For example, (2) (0.4) / (0.002) = 200

i.e the effective concentration = 200 M

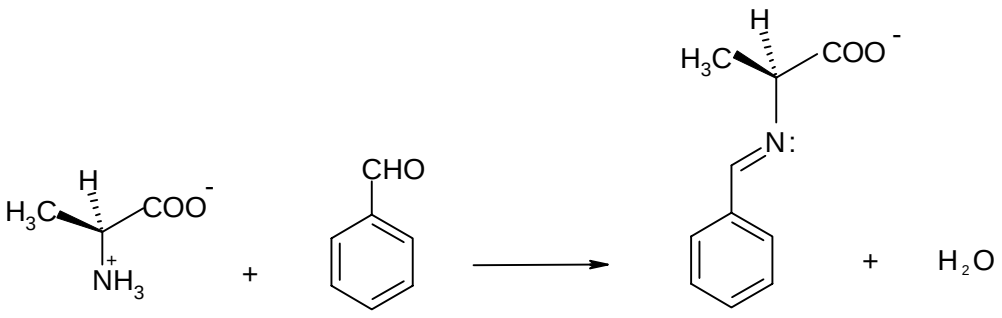
(3) (20) / (0.002) = 10,000

i.e. the effective concentration = 10,000 M

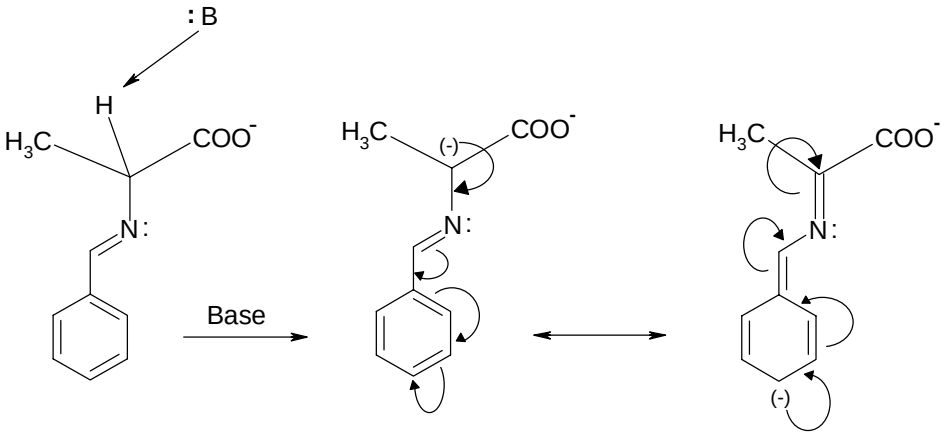
h. In addition to the enhanced local concentration effect, the COO^- group in (3) is better oriented to act in catalysis. The double bond restricts the motion of COO^- and thus reduces the number of unsuitable orientation of $–COO^-$, thereby enhancing the reaction rate.

22. Coenzyme chemistry

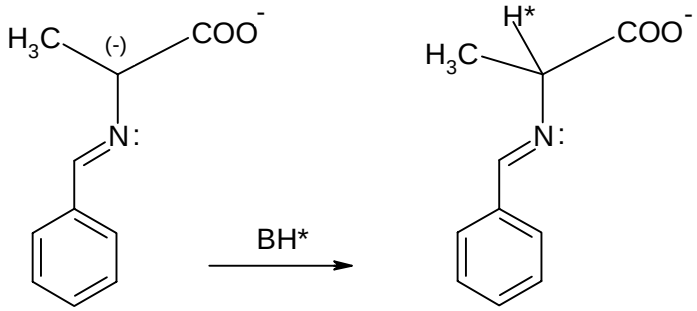
a. Step 1: Schiff base formation



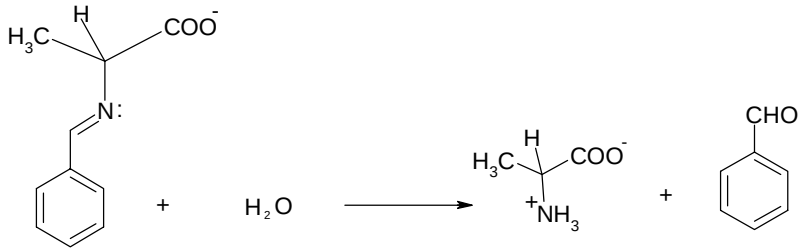
Step 2: Proton abstraction



Step 3: Reprotonation



Step 4: Hydrolysis



b. From the information stated in the problem, the following conclusions can be drawn:

Structure 2: Removal of the phosphate group does not hamper the activity. This indicates that the phosphate is not critical for the activity of PLP.

Similarly,

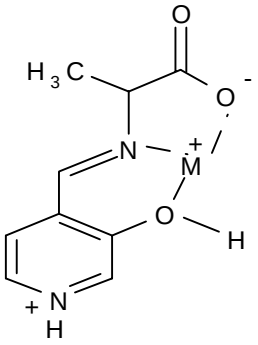
Structure 3: CH_2-OH is not critical.

Structure 4: Phenolic OH is needed in the free form.

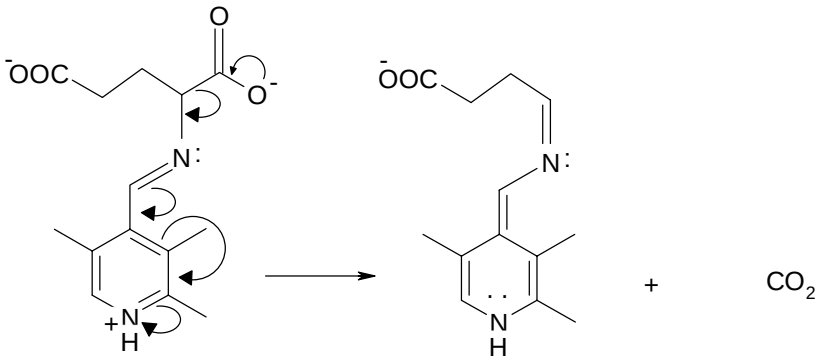
Structure 5: NO_2, a well-known electron withdrawing group, causes benzaldehyde to become activated. Hence positively charged nitrogen in structure 3 must be also important for its electron withdrawing effect.

Structure 6: Electron withdrawing effect of NO_2 is only effective from the *para* position. Introduction of this group at *meta* position leads to an inactive analog.

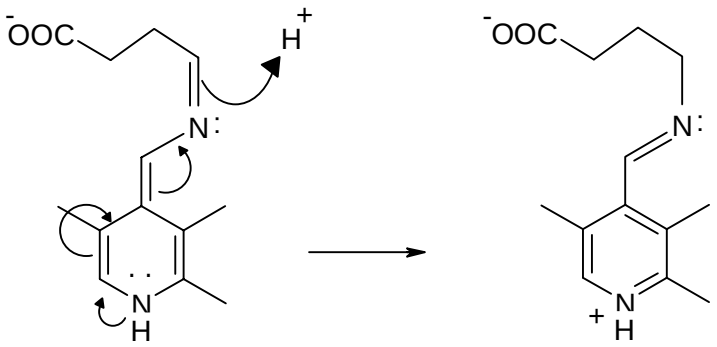
c. **Role of metal ion**: The metal ion is involved in a chelation, as shown below, and provides an explanation for the critical role of the phenolic OH. The planar structure formed due to chelation assists in the electron flow.



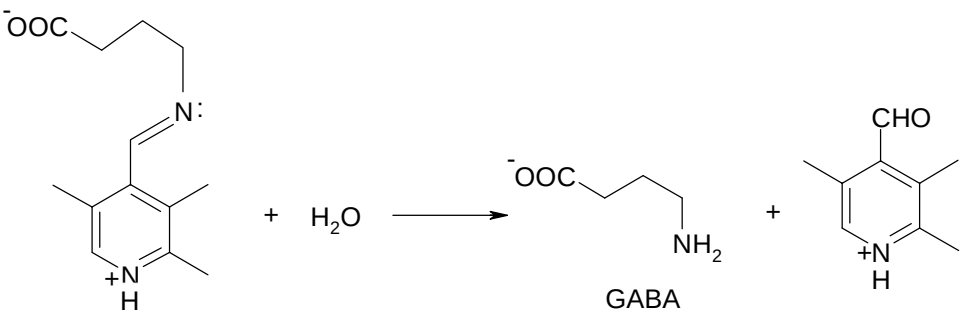
d. Step 1: Schiff base formation and decarboxlyation



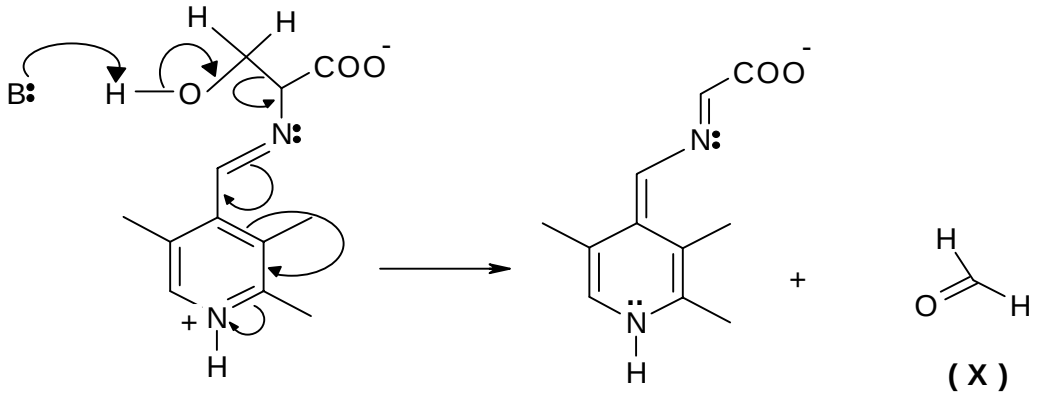
Step 2: Tautomerization



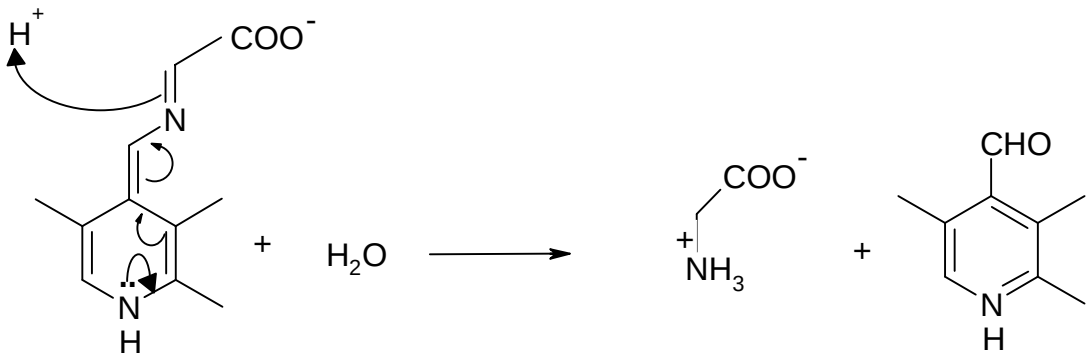
Step 3: Hydrolysis



e. Step 1: Schiff base formation followed by carbon- carbon bond scission.

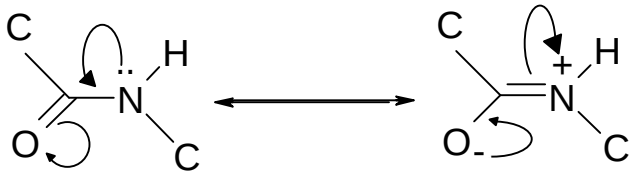


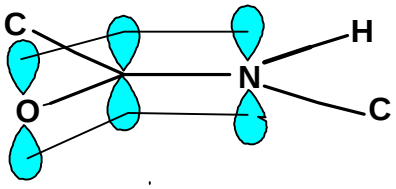
Step 2: Tautomerization followed by hydrolysis



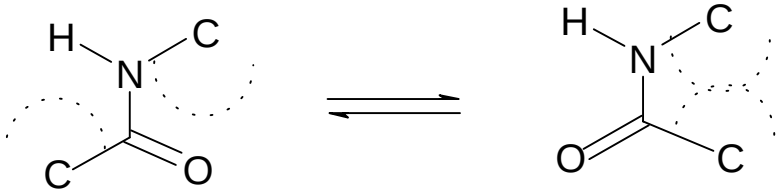
23. Protein folding

a. The planar amide group, that is, C_α, O, H and the next C_α are in a single plane - is stabilized by resonance. The C-N bond of the amide assumes partial double bond character and the overlap between p orbitals of O, C and N is maximized. The C_α's across this partial double bond can assume *cis* or *trans* arrangement.

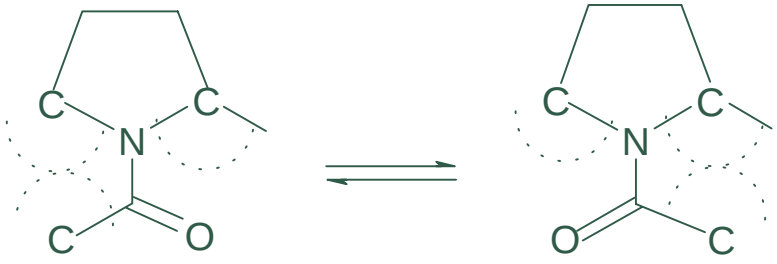




b. With nineteen of the amino acids, the *trans* arrangement is sterically favoured (i. e. it is comparatively less crowded). In the case of <u>proline</u>, *cis* and *trans* arrangements are almost equally crowded.



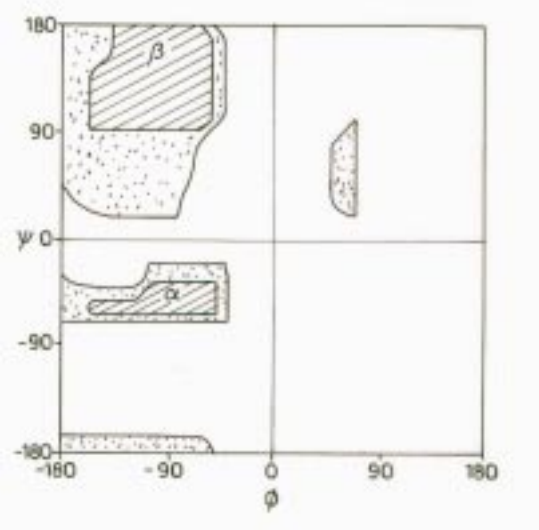
trans – amide *cis* - amide



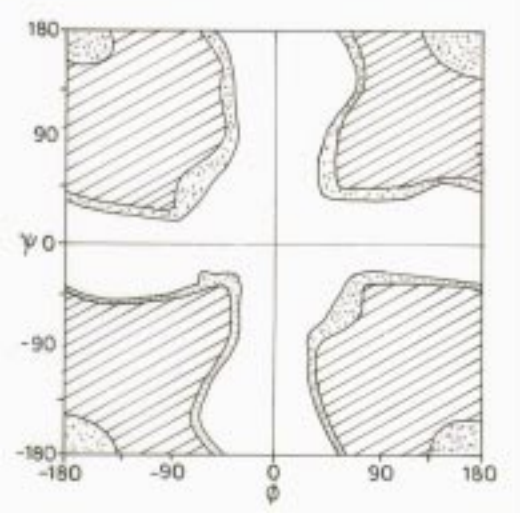
trans – amide *cis* - amide

c. **Note about Ramchandran diagram**: In a polypeptide, the amide units are planar (partial double bond character across the N-C bond) but the bonds connecting N and C_α, and the carbonyl carbon and C_α are free to rotate. These rotational angles are defined as ϕ and ψ, respectively. The conformation of the main chain is completely defined by these angles. Only some combinations of these angles are allowed while others are disallowed due to steric hindrance. The allowed range of ϕ and ψ angles are visualised as a steric contour diagram, shown below, known as the Ramachandran diagram.

For nineteen amino acids, the conformational choice is largely restricted to the so-called α and β regions on left half of the Ramachandran diagram (Panel A). This is due to the L - chiral nature of amino acids and the steric effects of their R groups. Glycine is an achiral residue with H as the R group. Therefore, much larger conformational regions on both left and right halves of Ramachandran diagram are accessible to this residue (Panel B).



Panel A



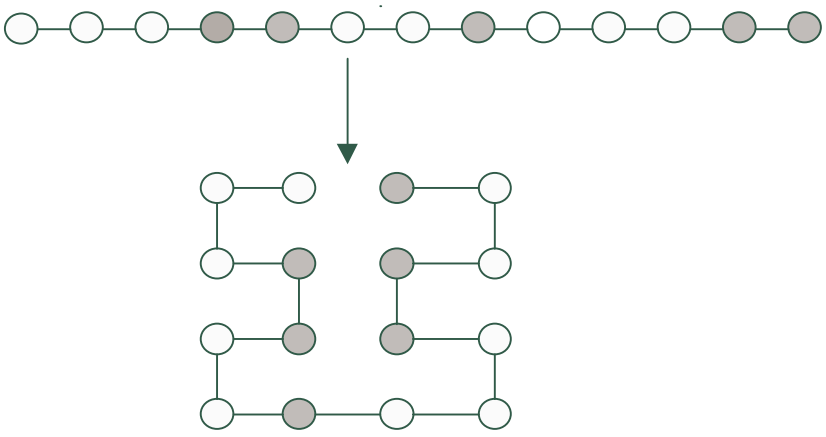
Panel B

d. Consecutive residues in α conformation form the α-helix. Similarly, consecutive residues in β conformation form the β-sheet. Both α-helix and β-sheet structures feature extensive networks of hydrogen bonds which stabilise them. Thus random combinations of α and β conformations are rarely found.

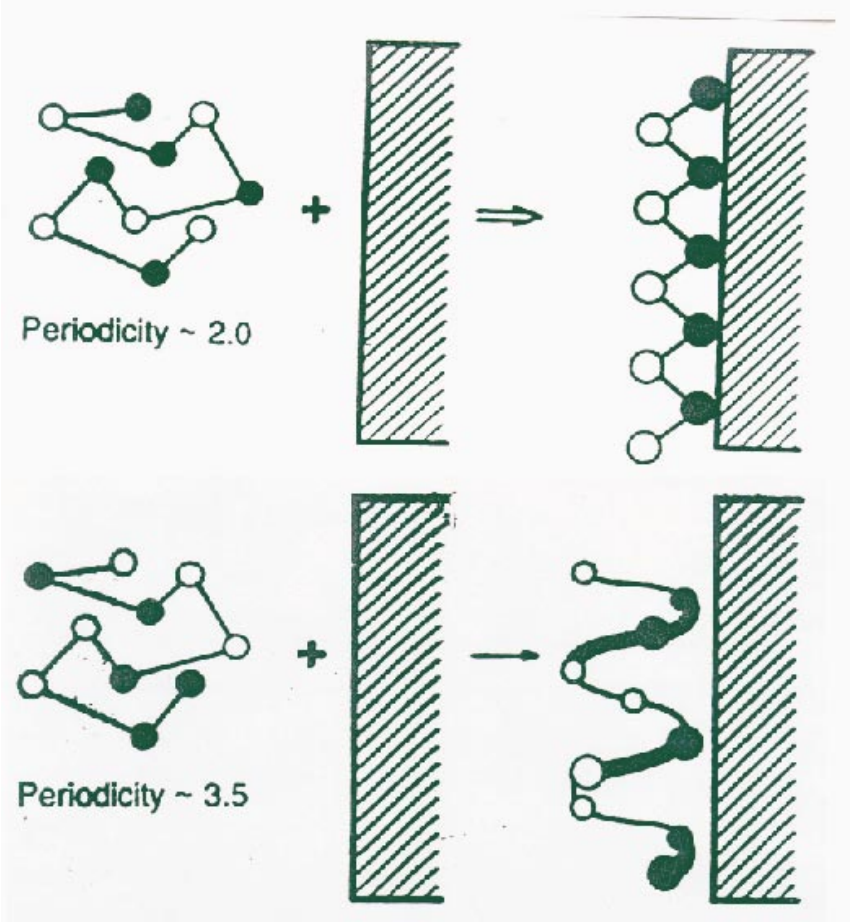
α - helix

β - sheet

e. For a polypeptide to fold in an aqueous environment, nearly half the R groups should be nonpolar (water hating) and the other half polar (water loving). Upon folding to form a globular protein, the nonpolar R groups are packed inside (away from water) while the polar groups are positioned on the surface (in contact with water). The phenomenon is similar to the hydrophobic aggregation of a micellar structure in water. If all the R groups are either polar or non-polar, no hydrophobic segregation is possible, and no folding will occur.



f. Alternating polar/nonpolar periodicity of R groups favors β-sheets. All the non-polar groups will face the apolar surface while the polar groups will be exposed to water. So the net folding will be like a β-sheet. On the other hand, a complex periodic pattern of R group polarities is needed in forming the α-helix.



Hydrophobic periodicity (primary sequence) + **Apolar surface (interface)** ⇒ **Induced (secondary structure)**

24. Protein sequencing

The sequence of amino acids in a protein or polypeptide is expressed starting from the N-terminal amino acid. From Edman degradation method the N-terminal amino acid is Asp. In the N-terminal fragment generated by trypsin or CNBr this amino acid should, therefore, be in position1. All other peptides generated by CNBr cleavage will

be preceded by Met on their N-terminal side. Likewise, all peptides generated by trypsin should be preceded by Arg or Lys. As we proceed from N-terminal amino acid to C-terminal amino acid, we carefully examine the different amino acids in each position shown in Table1(a) and 1(b)

For the first fragment starting from N-terminal Asp in position 1, we look for residues common in each position to CNBr and trypsin cleaved peptides. This gives

Position 1 2 3 4 5 6

Residue Asp -Pro/Tyr - Tyr -Val -Ile/Leu -Arg .....(1)

At position 6 Arg will render the polypeptide susceptible to trypsin. Therefore, 7th residue of this CNBr fragment (Table1a) should be same as residue1 in another peptide generated by trypsin and 8th residue of this CNBr fragment will be same as residue 2 in Table 1(b). Therefore we get

7 8

Gly/Phe - Tyr .....(2)

Since 8 will be Tyr, Pro will be assigned to position 2 of the polypeptide .....(3)

Residue 9 in the polypeptide should be at position 3 in the Table1(b) and residues 10,11,12,13 and 14 should be at positions 4,5,6,7 and 8 respectively in Table1(b). The same residues should be in positions 1 onwards in Table1(a).

None of the residues in position 3 (Table1b) is same as in position 1 in Table 1(a). However, positions 4 to 8 in Table 1(b) have residues common with positions 1 to 5 in Table 1(a). Further Glu in position 1 (Table 1a) will be preceded by Met (since it is a part of CNBr cleaved peptide). And position 3 in Table 1(b) has Met. Therefore, we get

9 10 11 12 13 14

Met- Glu - Thr - Ser - Ilu - Leu .....(4)

Position 5 in the polypeptide can now be firmly assigned to Ilu .....(5)

Positions 15 and 16 in the polypeptide will be beyond residue 8 in the trypsin cleaved peptide (not shown here). We now attempt to construct the remaining trypsin or CNBr fragments.

Table 1 (a) shows Arg in position 1. This will be preceded by a Met. Matching of the unassigned residues in position 2 in Table 1(a) with those in position 1 in Table 1(b) and for subsequent positions by the procedure demonstrated earlier that will give.

Met - Arg - Tyr - Pro - His - Asn - Trp - Phe - Lys - Gly - Cys .....(6)

(The last two residues are the unassigned residues in position 1 and 2 in Table 1b) Considering (2), (5) and (6) together it is now possible to firmly assign position 7 on the polypeptide to Gly .....(7)

a. The amino acid sequence common to the first fragments (N-terminal) obtained by CNBr and trypsin treatments is

1 2 3 4 5

Asp - Pro - Tyr - Val - Ile

b. The sequence of the first fragment generated by CNBr treatment is

1 2 3 4 5 6 7 8

Asp- Pro - Tyr- Val- Ile -Arg -Gly -Tyr

To complete the sequence of the polypeptide we need to construct the sequence of another trypsin fragment. Starting from position 4-(Arg) in Table 1(a) we get the sequence,

Arg-Phe-His-Thr-Ala ..... (8)

At this stage, we again examine the unassigned residues. The Arg in (8) will have to be serially preceded by Asn, Gln, Gly and Met (these are the unassigned residues in respective positions in Table 1(a). We then get the sequence,

Met-Gly-Gln-Asn-Arg-Phe-His-Thr-Ala .....(9)

And following the Ala in (9)

Leu-Ser-Cys-Glu … ..(10)

From (9) and (10), we get the sequence

Met-Gly-Gln-Asn-Arg-Phe-His-Thr-Ala-Leu-Ser-Cys-Glu …..(11)

Since the smallest fragment is a dipeptide (Table 1b) and (6) shows that it follows Lys, it follows that this will be at the C-terminal end. Therefore, the partial sequence shown in (6) will follow the partial sequence shown in (11).Thus, we get

Met-Gly-Gln-Asn-Arg-Phe-His-Thr-Ala-Leu-Ser-Cys-Glu-Met-Arg-Tyr-Pro-His-Asn-Trp-Phe-Lys-Gly-Cys …..(12)

There is already a Met in position 9 of the polypeptide. The next Met can only come earliest at position 17 since CNBr fragment have at least 8 amino acids. Therefore, the starting residues of (12) can be assigned position 17.

This leaves positions 15 and 16 which will be filled by the unassigned residues Val and Ala in the CNBr fragment at positions 6 and 7 (Table 1a).

c. The final sequence, therefore, will be

						Trypsin							**CNBr**									
1	2	3	4	5	6	↓	7	8	9	↓	10	11										
Asp	-	Pro	-	Tyr	-	Val	-	Ile	-	Arg	-	Gly	-	Tyr	-	Met	-	Glu	-	Thr		
										CNBr										**Trypsin**		
12	13	14	15	16	17	↓	18	19	20	21	↓	22										
Ser	-	Ile	-	Leu	-	Val	-	Ala	-	Met	-	Gly	-	Gln	-	Asn	-	Arg	-	Phe		
													CNBr	**Trypsin**								
23	24	25	26	27	28	29	30	↓	31	↓	32	33										
His	-	Thr	-	Ala	-	Leu	-	Ser	-	Cys	-	Glu	-	Met	-	Arg	-	Tyr	-	Pro		
									Trypsin													
34	35	36	37	38	↓	39	40															
His	-	Asn	-	Trp	-	Phe	-	Lys	-	Gly	-	Cys										

Arrows (↓) indicate the CNBr and trypsin-labile sites.

d. There are 6 basic amino acid residues in the polypeptide. 6/40 = 15%

e. An α helix has 3.6 amino acid residues per turn of 5.4Å.

Thus, the length of the polypeptide in α helical conformation will be :

$40/3.6 \times 5.4 = 59.4$ Å.

e. The polypeptide has 40 amino acids. Since each amino acid is coded for by a triplet of nucleotides, the total number of nucleotide pairs in the double stranded DNA of the exon will be

40 x 3 = 120 base pairs.

The molecular weight of the DNA making the exon

= 330 x 2 x 120

= 79200 Da

g. If the exon contains 120 base pairs and A and C are in equal numbers, there will be 60 A-T pairs and 60 G-C pairs. Each A-T pair is held by two H-bonds and each G-C pair is held by three H-bonds. Hence the total number of H-bonds holding this double helix is :

(60 x 2) + (60 x 3) = 300