The Amino Acid Sequence of Cytochrome b_{562} of Escherichia coli*

Eiji Itagaki** and Lowell P. Hager
Biochemistry Division, Department of Chemistry
and Chemical Engineering, University of Illinois
Urbana, Illinois 61801

Received August 12, 1968

Cytochrome b_{562} has been purified and crystallized from Escherichia coli (Itagaki and Hager, 1966). It has been established that this hemoprotein binds iron protoporphyrin IX as prosthetic group. In this respect, b_{562} resembles hemoglobin and myoglobin. However, the enzymatic function of b_{562} classifies it as a soluble electron carrier in E. coli. In this respect, b_{562} is similar to mammalian cytochrome c. Since b_{562} has a low molecular weight and contains the minimum number of histidine residues required to fill both the fifth and sixth heme ligand positions, we undertook the amino acid sequence determination of b_{562} as a contribution to the study of structure-function relationships in b-type cytochromes. We approached the determination of the amino acid sequence of cytochrome b_{562} using the overlap method employing chymotryptic, tryptic, and cyanogen bromide peptides.

In this communication, we describe the amino acid sequences of the cyanogen bromide, chymotryptic and tryptic peptides, and the amino acid sequence of cytochrome b_{562} as deduced from these peptides. The details of this study will be published elsewhere.

The amino terminal sequence of cytochrome b_{562} was established by a three step Edman degradation (Doolittle, 1965) and by dinitrophenylation (Frankel-Conrat et al, 1955) of the denatured apoprotein. The amino terminal sequence

* This investigation was supported by a research grant (RG 7768) from the Institute of General Medical Science, National Institutes of Health.

** Present address: Faculty of Pharmaceutical Sciences, Kanazawa University, Kanazawa, Japan.

is NH_2-Ala-Asp-Leu-Glu. The carboxyl terminal sequence of b_{562} is Ala-Tyr-Lys-His-Gln-Lys-Tyr-Arg. This sequence was established by combining the results obtained from carboxypeptidase B digestions with sequence studies on chymotryptic and tryptic peptides.

Two chymotryptic digests of b_{562} were prepared. Apocytochrome b_{562} (9 μmoles), denatured by HCl-acetone treatment at room temperature, was digested with 3% (w/w) α-chymotrypsin (Worthington) for 160 minutes, and 20 μmoles of the apoprotein were also digested under the same condition for 24 hours. A trypsin digest of denatured apocytochrome b_{562} (22.5 μmoles) was prepared by incubation with 5% (w/w) trypsin (Calbiochem, B grade) for 19 hours at 38°. These digests were charged on Bio-Rad AG-50W-X2 columns equilibrated with 0.1 M pyridine-acetate buffer (pH 3.1). Chymotryptic peptides and tryptic peptides were eluted with a pyridine-acetate linear gradient buffer system (Margoliash and Smith, 1962). Each peptide fraction was further purified by paper chromatography and/or paper electrophoresis at pH 6.5 or 4.6. Twelve chymotryptic peptides and nineteen tryptic peptides were obtained. The amino acid sequences of the individual peptides were studied using a variety of standard techniques, including the Edman degradation method, leucine aminopeptidase (Light, 1967), carboxypeptidase A or B digestion, dinitrophenylation, and papain or trypsin digestion. Table I summarizes the sequence studies on the twelve chymotryptic peptides, and Table II shows the sequences of the nineteen tryptic peptides.

Cyanogen bromide treatment of denatured apo b_{562} (1 μmole treated with excess cyanogen bromide in 60% formic acid at 30° for 40 hours, Steevs, et al, 1965) yielded 4 peptides (CNBr I, II, III, and IV) which were separated and purified on Sephadex G-50 columns and silicic acid thin layer chromatograms (Willand and Georgopoulos, 1964). The amino terminal sequence of CNBr I was established to be Ala-Asp-Leu- by a three step Edman degradation. CNBr II was digested with α-chymotrypsin, and two peptides (CNBr-II-C-1 and C-2) were purified. The sequence of CNBr-II-C-1 was established to be Glu-Thr-Leu-Asn-

Figure 1. Amino Acid Sequence of *E. coli* Cytochrome b_{562}

```
 1                               10
Ala-Asp-Leu-Glu-Asp-Asp-Met-Gln-Thr-Leu-Asn-Asp-Asn-Leu-Lys-Val-

                20                                        30
Ile-Glu-Lys-Ala-(Asx,Asx,Glx)-Lys-Ala-Asn-Asp-Ala-Ala-Gln-Val-

                                    40
Lys-Leu-Lys-Met-Arg-Ala-Ala-Ala-Leu-Asn-Ala-Gln-Lys-Lys-Ala-Thr-

             50                                      60
Pro-Pro-Lys-Leu-Glu-Asp-Lys-Ser-Pro-Asn-Ser-Gln-Pro-Met-Lys-Asp-

                                 70
Phe-Arg-His-Gly-Phe-Asp-Ile-Leu-Val-Gly-Glu-Ile-Asp-Asp-Ala-Leu-

 80                                      90
Lys-Leu-Ala-Asn-Glu-Gly-Lys-Val-Lys-Glu-Ala-Gln-Ala-Ala-Glu-Ala-

              100                                          110
Gln-Leu-Lys-Thr-Thr-Arg-Asn-Ala-Tyr-Lys-His-Gln-Lys-Tyr-Arg.
```

Figure 2. Comparison of Cytochrome b_{562}, Sperm Whale Myoglobin and Horse Hemoglobin

```
                              b562
 -Arg-His-Gly-Phe-Asp-Ile-Leu-Val-Gly-Glu-Ile-Asp-Asp-Ala-Leu-Lys-
  |65 |   |           |70 |       |       75                 |80
  |   |   |           |   |Myoglobin                         |
  |   |   |           |   |       |                          |
 -Lys-His-Gly-Val-Thr-Val-Leu-Thr-Ala-Leu-Gly-Ala-Ile-Leu-Lys-Lys-
   63 |   |           |68 |       |       73                 |78
      |   |           |   |Hemoglobin                        |
      |   |           |   |       |                          |
α-Ala-His-Gly-Lys-Lys-Val-Ala-Asp-Gly-Leu-Thr-Leu-Ala-Val-Gly-His-
   62 |   |           |67 |       |       72                 |77
      |   |           |   |       |                          |
β-Ala-His-Gly-Lys-Lys-Val-Leu-His-Ser-Phe-Gly-Glu-Gly-Val-His-His-
   62                  67                 72                  77
```

Asp-Asn-Leu by a four step Edman degradation plus carboxypeptidase A digestion. A five step Edman degradation of CNBr III gave the amino terminal sequence Arg-Ala-Ala-Ala-Leu. The amino terminal sequence of CNBr IV was shown to be

Lys-Asp-Phe-Arg-His by a five step Edman degradation.

The amino acid sequence of cytochrome b_{562} of *E. coli* could be deduced

Table I. Amino Acid Sequence of Chymotryptic Peptides Obtained from *E. coli* Cytochrome b_{562}

Peptide No.	Amino Acid Sequence*
C-1	Ala-Asp-Leu-Glu-Asp-Asp-Met-Gln-Thr-Leu
C-2	Asn-Asp-Asn-Leu
C-3	Val-Ile-Glu-Lys-Ala-(Asx,Asx,Glx,Lys) (Ala,Asn,Asp,Ala)- Ala-Glu-Val-Lys-Leu
C-4	(Lys)-Met
C-5	(Arg)-Ala-Ala-Ala-Leu
C-6	Asn-Ala-Gln-Lys-Lys-Ala-Thr-(Pro,Pro,Lys) (Leu,Glu,Asp,Lys) (Ser,Pro,Asp,Ser,Glu,Pro,Met,Lys)-Asp-Phe
C-7	(Arg)-His-Gly-Phe-Asp-Ile-Leu
C-7-a	(His,Gly)-Phe
C-7-b	Asp-Ile-Leu
C-8	Val-Gly-Glu-Ile-Asp-Asp-Ala-Leu-Lys-Leu
C-8-a	Val-Gly-Glu-Ile-Asp-Asp-Ala-Leu
C-8-b	Lys-Leu
C-9	Ala-Asn-Glu-Gly-Lys-Val-Lys-Glu-Ala-Glu-Ala-Ala-Glu-Ala-Gln-Leu
C-10	Lys-Thr-Thr-Arg-Asn-Ala-Tyr
C-10-a	Lys-Thr-Thr
C-10-b	Arg-Asn-Ala-Tyr
C-11	Lys-His-Gln-Lys-Tyr
C-11-a	(His)-Glu-Lys-Tyr
C-12	Arg

* See the legend of Table II for notations.

from the above information (see Fig. 1). Some interesting points emerge from this sequence. There is a characteristic pattern for the distribution of

Table II. Amino Acid Sequences of Tryptic Peptides Obtained from *E. coli* Cytochrome b_{562}

Peptide No.	Amino Acid Sequence*
T-1	Ala-Asp-Leu-(Glu,Asp,Asp,Met,Glu,Thr,Leu,Asn,Asp)Asn-Leu-Lys
T-2	Val-Ile-Glu-Lys
T-3	Ala-Asn-Asp-Ala-Ala-Gln-Val-Lys
T-4	Met
T-5	Arg
T-6	Ala-Ala-Ala-Leu-Asn-Ala-Gln-Lys
T-7	Lys
T-8	Ala-Thr-Pro-Pro-Lys
T-9	Leu-Glu-Asp-Lys
T-10	Ser-Pro-Asn-Ser-Glu-Pro-Met-Lys
T-11	Asp-Phe-Arg
T-12	Gly-Phe-(Asp,Ile,Leu,Val,Gly,Ile,Asp,Asp,Ala,Leu)Lys
T-13	Leu-Ala-Asn-Glu-Gly-Lys
T-14	(Glu,Ala,Glu,Ala,Ala,Glu,Ala,Glu,Leu,Lys)
T-15	Thr-Thr-Arg
T-16	Asn-Ala-Tyr
T-17	Lys
T-18	(His)-Gln-Lys
T-19	Arg

*Solid arrows pointing to the right represent sequences determined by the Edman degradation method. Dashed arrows pointing to the right indicate sequences determined by leucine aminopeptidase hydrolysis. Parentheses indicate sequences determined by dinitrophenylation. Arrows pointing to the left represent sequences determined by carboxypeptidase A or B. Tryptic peptides and papain peptides are designated by —T— and —P—, respectively.

proline residues and basic and acidic amino acids in the protein. A preponderance of acidic residues is found in the amino terminal end, all four proline residues are located in the middle of the molecule, and six basic amino acids are located in the carboxyl terminal region between residues 98 and 110. The single pair of histidine residues in b_{562} , which may coordinate to the iron atom in the heme group as fifth and sixth ligands, are in positions 66 and 106. It is interesting to note that one of the potential histidine ligands in both cytochrome b_5 (Ozols and Strittmatter, 1968) and b_{562}, as well as the known histidine ligands in myoglobin and hemoglobin (Perutz, 1965), are located near the center of the amino acid sequence. In contrast, the histidine residue of cytochrome c which serves as a heme ligand is situated near the amino terminal end at residue 18 (Margoliash and Smith, 1961). The spacing between the two histidines in b_{562} is 40 residues compared to a 31 and 29 residue spacing between the histidine ligands in hemoglobin and myoglobin.

A comparison of the amino acid sequence in the His_{66} region of cytochrome b_{562} with that around one of the histidine-heme ligands in sperm whale myoglobin and horse hemoglobin shows an interesting correlation. The amino acid sequences from Arg_{65} to Lys_{80} in cytochrome b_{562}, from Ala_{62} to His_{77} in hemoglobin, and from Lys_{63} to Lys_{78} in myoglobin show a striking homology (see Fig. 2). Sperm whale myoglobin and cytochrome b_{562} have the sequence -His-Gly- and a basic amino acid, Lys or Arg, on the amino terminal side of histidine (Fig. 2). The α and β chains of horse hemoglobin also have the -His-Gly- sequence. Val_{68} of myoglobin corresponds to Val_{67} in hemoglobin and Ile_{70} in cytochrome b_{562}. Leu_{69} of myoglobin corresponds to Ala_{68} in the α chain, Leu_{68} in the β chain of hemoglobin and Leu_{71} in cytochrome b_{562}. Another region of homology between myoglobin and b_{562} occurs further down the chain in the area between the two histidine ligands. The sequence Glu_{83}-Ala-Glu-Leu-Lys_{87} in myoglobin is duplicated in b_{562} between residues 94 and 98. Whether these similarities represent sequence fragments left over from evolutionary divergence or result from functional convergence in heme proteins poses an interesting problem.

The authors wish to acknowledge the excellent technical assistance of Mrs. Marie Kindt and Miss Helga Nagler.

References

1. Doolittle, R. F., Biochem. J., 94, 742 (1965).

2. Frankel-Conrat, H., Harris, J. L., and Levy, A. L., Methods Biochem. Anal., 2, 359 (1955).

3. Itagaki, E. and Hager, L. P., J. Biol. Chem. 241, 3687 (1966).

4. Light, L., Methods in Enzymology, Vol. XI., 29 (1967).

5. Margoliash, E. and Smith, E. L., J. Biol. Chem., 237, 2151 (1962).

6. Margoliash, E. and Smith, E. L., Nature 192, 1121 (1961).

7. Ozols, J. and Strittmatter, P., J. Biol. Chem., 243, 3376 (1968).

8. Perutz, M. F., J. Mol. Biol., 13, 646 (1965).

9. Steevs, E., Jr., Craven., R. G., Anfinsen, C. B., and Bethune, J. L., J. Biol. Chem., 240, 2478 (1965).

10. Wieland, Th. and Georgopoulos, D., Biochem. Z., 340, 476 (1964).