ROLE OF TRYPTOPHANYL RESIDUES IN DRIVING MYOGLOBIN FOLDING

I. Sirangelo, M. Casillo, C. Malmo, G. Irace

Dipartimento di Biochimica e Biofisica, Seconda Università di Napoli

INTRODUCTION

The detailed mechanism by which the primary sequence of a protein encodes its tertiary structure, and by which a protein spontaneously acquires its native fold, is still unclear. Myoglobin provides a good versatile system to investigate the factors that govern the rate and mechanism of protein folding. At neutral pH, myoglobin has a well-defined globular structure, comprising eight helical segments (A-H) packed to form a compact hydrophobic core where the heme group is bound (1,2). On removal of the heme, the apoprotein remains well folded with extensive secondary and tertiary structure, only the F helix being dynamically disordered in the apomyoglobin.

The myoglobin folding proceeds through the formation of partly folded intermediate; the A, G and H helices and a part of the B helix are stabilized in the burst phase intermediate, while the helical structure in the remainder of the B helix and in C and E helices is stabilized during slower folding events (3).

In our previous work, we showed that tryptophanyl residues play a very important role in driving the early stages of the folding process (4). Molecular dynamics simulation studies revealed that the lack of some contacts between the tryptophanyl residues and the surrounding helices introduces structural distorsions which affect the next stages of the folding process (4). In this article, we provide further insights on the role played by these residues.

MATERIALS AND METHODS

Mutagenesis of the myoglobin gene - B.A. Springer and S.G. Sligar kindly provided plasmid pMb413 bearing the synthetic sperm whale myoglobin gene (5). DNA manipulations were essentially performed as described in Sambrook et al. (6). The tryptophanyl substitution was performed using the "Clontech Transformer site-directed mutagenesis" kit. Mutants were screened and confirmed by sequencing double-stranded DNA in the region of the mutation using the Sequenase kit purchased from United States Biochemical Corporation. Mutant myoglobins were expressed in Escherichia coli strain TB-1 [ara, Δ(lac-pro), strA, thi, T, Φ_{80} dLacZΔM15, r^-, m^+] (5).

Wild type myoglobin purification - Escherichia coli TB-1 harboring the pMb413 plasmid was grown at 37°C in Luria-Bertani in the presence of ampicillin (200 mg·L^{-1}). Protein was essentially purified as described by Springer and Sligar (5). Briefly, a 10-L culture of cells was harvested in late log phase, lysed overnight and sonicated. Cell debris was removed by centrifugation, and the supernatant was fractionated by ammonium sulfate precipitation. The 60-95% cut was centrifuged and suspended in 20 mM Tris, 1 mM EDTA, pH 8, and fractionated on a Sephadex G-50 (Pharmacia) gel filtration column (2.5 x 100 cm) equilibrated in the same buffer. Reddish-brown myoglobin containing fractions were collected and applied to a Whatman DEAE 52 ion exchange column (2.5 x 20 cm) equilibrated and resolved with 20 mM tris/HCl pH 8.4. Under these conditions, myoglobin did not stick on the column and was rapidly eluted.

RESULTS

Mammalian myoglobins contain two tryptophanyl residues at invariant positions 7 and 14 in the aminoacid sequence. Both residues reside on the A-helix, A-5 (W7) and A-12 (W14) (7,8). The replacement of one of the two tryptophanyl residues with a hydrophobic residue (phenylalanine) does not influence the overall fold of the protein, although the stability is significantly decreased (9,10). The simultaneous substitution of both tryptophanyl residues results in the expression of an unstable protein

that can not be purified in homogeneous form (4). The double W→F replacement is not tolerated probably because of distortions introduced in the compact AGH subdomain, which influence the subsequent steps of the folding pathway. In order to investigate the role played by the trytophanyl interactions on myoglobin folding and structure, we designed mutants in which one tryptophanyl residue was introduced in G, E and C helices after substituting both W7 and W14 with a phenylalanine residue. In particular, we replaced phenilalanine in the G helix (W7FW14F-F106W), leucine in C helix (W7FW14F-L40W) and leucine in E helix (W7FW14F-L72W). C and E helices were selected to verify the possibility that the introduction of a tryptophanyl residue might influence the subsequent steps of the folding process.

Figure 1 shows the western blotting analysis of the cell lysate of E. Coli TB-1 pMb413 mutants. A high level of myoglobin expression is evident only in the case of the wild type myoglobin (lane 2), while a weak band is visible for the mutant without tryptophanyl residues (lane 1) and for mutants having an indole residue in G, C and E helix (lanes 3, 4 and 5). The standard purification procedure failed when applied to mutant proteins. In fact, it was not possible to obtain stable proteins in homogeneous form. Using policlonal antibodies, these proteins were found in 60% ammonium sulfate precipitate instead of the precipitate obtained using 95% salt saturation. The next purification steps were also unsuccessful. The recombinant proteins were found in fractions eluted within the exclusion volume of a Sephadex G-50 column; moreover, these fractions showed extremely low ratios between the absorbance at 409 and 280 nm. These results indicate that the polypeptide chain does not fold correctly but might remain in a somewhat rod-like, flexible structure with an increased radius of gyration or might be aggregated. The myoglobin-containing fractions were collected and applied to a DEAE-Cellulose column. The recombinant proteins were not eluted with the equilibrating buffer, as happens with correctly folded myoglobins, but required a salt gradient from 0.1-0.3 M to be eluted (fig. 2). Moreover, protein was spread throughout the gradient as showed by western blot analysis (fig. 3).

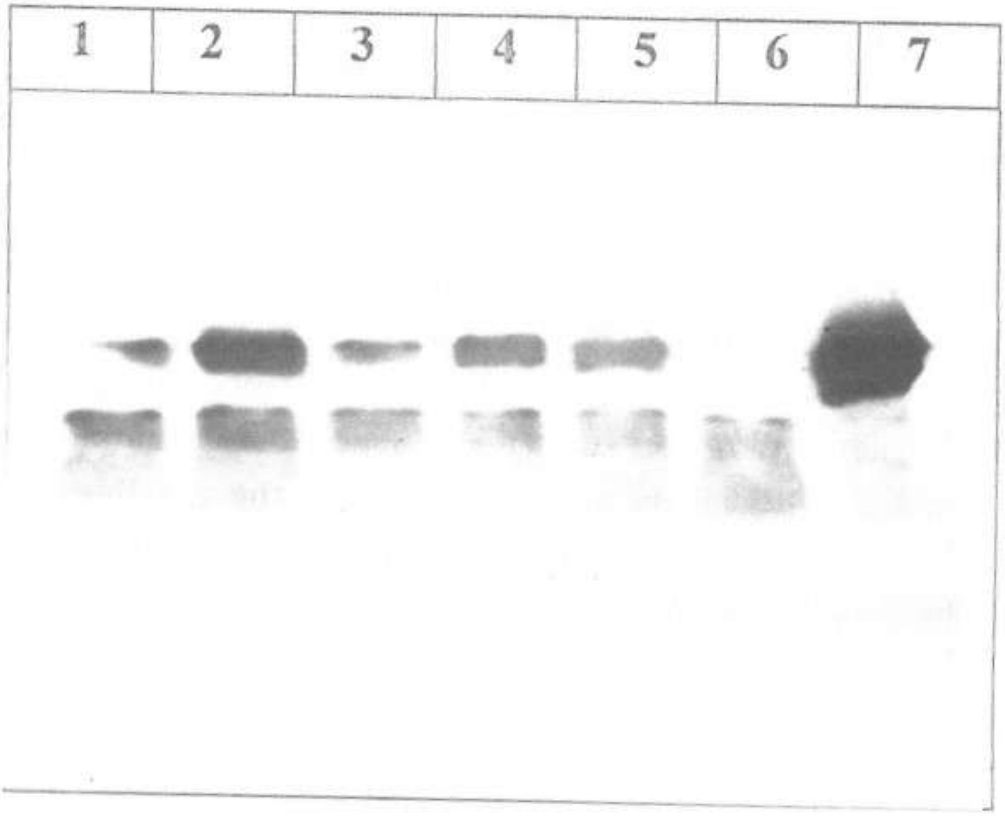

Fig. 1 - Western blot analysis of E. coli TB-1 pMb413 cell extracts. Lane 1, W7FW14F; lane 2, wild type; lane 3, W7FW14F-F106W; lane 4, W7FW14F-L40W; lane 5, W7FW14F-L72W; E. coli TB-1 pUC19 cell extract; lane 6, sperm whale myoglobin (Sigma). Each lane was loaded with 7.5 μg of cell extract.

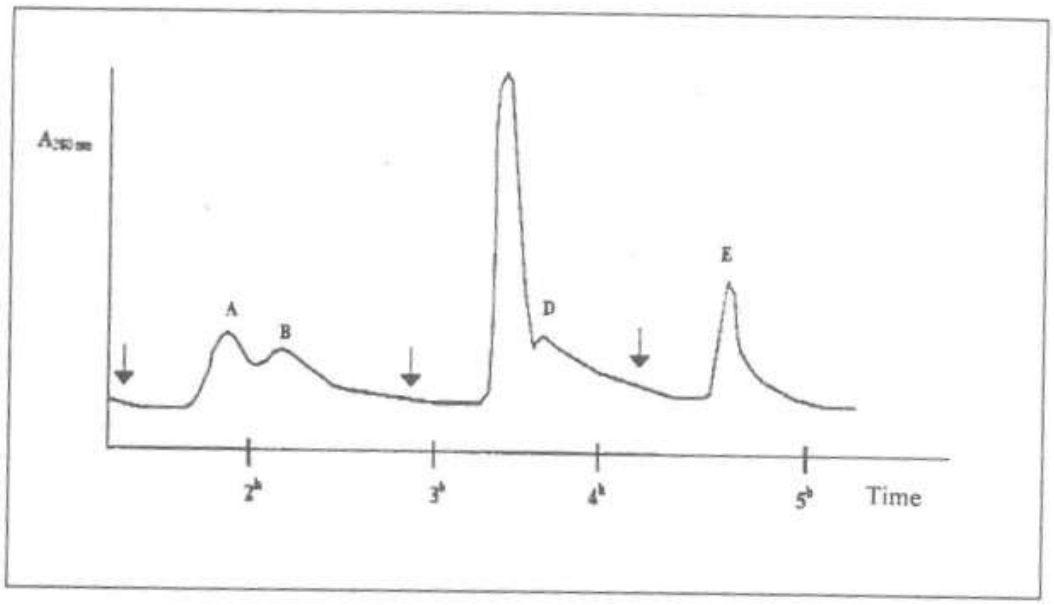

Fig. 2 - Elution diagram obtained after sample application on DEAE-Cellulose. Equilibration buffer: 0.01 M Tris HCl- 0.001 M EDTA, pH 8.4; the arrows indicate the application of saline gradient: 0.1, 0.2 and 0.3 M NaCl, respectively.

This observation further corroborates the idea that the polypeptide chain does not fold correctly but may fluctuate from one to another partly folded state.

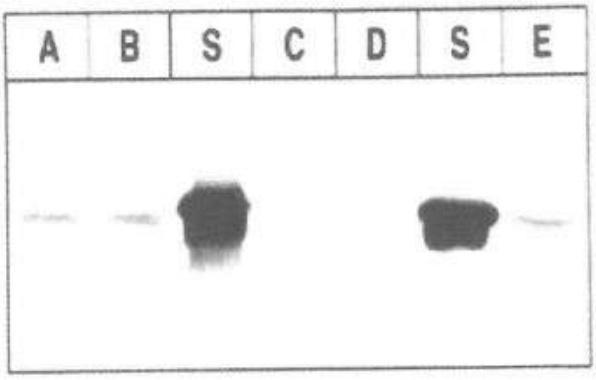

Fig. 3 - Western blot analysis of the fractions eluted from DEAE-Cellulose chromatography. A, B, C, D, and E correspond to elution diagram peaks of fig. 2; S, sperm whale myoglobin standard.

DISCUSSION

Constitutive expression of myoglobin in E. coli is related to the efficiency of the promoter, codon usage, rate of hemin synthesis, and intrinsic stability of apoproteins and holoproteins. Because the same expression vector and host strain were used, differences in the expression of myoglobin mutants must be due to changes in the overall stability and/or to incorrect protein folding. The instability of mutant proteins could be responsible for an extremely poor constitutive expression in E. coli TB-1 cells presumably because of denaturation processes rapidly occurring in the bacterial cytoplasm. In the case of double tryptophanyl substitution, drastic alterations in the overall fold on the polypeptide chain may be responsible not only for the observed structural heterogeneity but also for the loss of the ability to bind the prosthetic group. The introduction of a tryptophanyl residue in the G, E, and C helix is not able to restablish a correct protein folding, thus indicating that the tryptophanyl residues must be correctly positioned and that small variations are not tolerated. In conclusion, the data reported in this work indicate that the presence of a tryptophanyl residue in a region different from A-helix is not able to restore the formation of tertiary key interactions necessary at various stages of myoglobin folding.

Mammalian myoglobins contain two tryptophanyl residues at the invariant positions 7 (A-5) and 14 (A-12) in the N-terminal region (A helix) of the protein molecule. The simultaneous substitution of both tryptophanyl residues causes an incorrect folding with subsequent loss of heme binding. The introduction of a indolic residue in different molecular regions, i.e. G, E, and C helix resulted in a not correctly folded protein, suggesting that the tryptophanyl residues are strong structural determinants.

1) ELIEZER D., WRIGHT P., J. Mol. Biol., 1996, 263, 531-538.
2) LECOMTE J.T., SUKITS S.F., BHATTACHARJYA S., FALZONE C.J., Protein Sci., 1999, 8, 1484-1491.
3) GARCIA C., NISHIMURA C., CAVAGNERO S., DYSON H.J., WRIGHT P.E., Biochemistry, 2000, 39, 11227-11237.
4) SIRANGELO I., TAVASSI S., MARTELLI P.L., CASADIO R., IRACE G., Eur. J. Biochem., 2000, 267, 3937-3945.
5) SPRINGER A., SLIGAR S.G., Proc. Natl. Acad. Sci. USA, 1987, 84, 8961-8965.
6) SAMBROOK J., FRITSCH F., MANIATIS T., Molecular Cloning, Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.
7) BALESTRIERI C., COLONNA G., GIOVANE A., IRACE G., SERVILLO L., Comp. Biochem. Physiol., 1978, 60B, 195-199.
8) BISMUTO E., COLONNA G., SAVY F., IRACE G., Int. J. Peptide Protein Res., 1985, 26, 195-207.
9) KAY M.S., BALDWIN R.L, Nat. Struct. Biol., 1996, 3, 439-445.
10) SIRANGELO I., TAVASSI S., IRACE G., Biochim. Biophys. Acta, 2000, 1476, 173-180.

KEY WORDS: apomyoglobin folding, tryptophanyl replacement, apomyoglobin structure.

This work has been supported by Progetto Finalizzato Biotecnologie and by a MURST Program of Relevant National Interest (PRIN).

Lectured at the meeting held in Naples on November 24, 2000.
Received: November 27, 2000; accepted: December 12, 2000.

Address reprint requests/correspondence to Prof. G. Irace, Dipartimento di Biochimica e Biofisica, Via Costantinopoli 16, I-80138 Napoli.
E-mail: gaetano.irace@unina2.it