Monatshefte für Chemie 131, 409–416 (2000)

Monatshefte für Chemie
Chemical Monthly

Proteolytic Preparation of a Heme Binding Fragment from Sperm Whale Myoglobin: Micro-Myoglobin

Stephan Schwarzinger[a], **Werner Ahrer,** and **Norbert Müller***

Institute of Chemistry, Johannes Kepler University, A-4040 Linz, Austria

Summary. Clostripain digestion of sperm whale apomyoglobin does not yield a heme binding fragment, contrary to horse heart apomyoglobin, from which mini-myoglobin has been obtained by this approach. However, in pepsin digests of sperm whale apomyoglobin we identified two fragments closely corresponding to the polypeptide encoded by the central exon of the myoglobin gene. One of these fragments consisting of 77 amino acid residues was purified. Spectroscopic data indicate that it has heme binding properties.

Keywords. Apomyoglobin; Proteolysis; Mass spectrometry; Pepsin; Clostripain.

Introduction

The oxygen storage protein myoglobin contains a single heme molecule as prosthetic group. Almost all contacts between the heme group and the polypeptide chain are located in that part of the protein which is encoded by the central exon of a total of three exons of the myoglobin gene [1]. The first heme binding globin fragments have been reported by *Craik et al*. [2, 3] who obtained a mixture of two polypeptide fragments (31–104 and 41–104) by proteolysis of human hemoglobin. Although the so-called intron-early theory [1] appears to be challenged by more recent results [4], these experiments and results reported by *De Sanctis* and coworkers [5–7] are in support of a correlation between protein functionality and the genes' exon-entities: Using the arginine specific protease clostripain to digest horse heart apomyoglobin they obtained the peptide fragment mini-myoglobin ranging from residue 32 to 139. Mini-myoglobin can be reconstituted with heme and Co-protoporphyrin IXa, and the reconstituted hemoprotein has the ability to bind oxygen reversibly, albeit with a lower binding constant than myoglobin. In more recent investigations this was attributed to the higher conformational flexibility of mini-myoglobin [8, 9]. These results corroborate the hypothesis that the polypeptide encoded by the central exon is responsible for the fundamental

* Corresponding author
[a] Present address: The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

functionality, whereas those parts of the sequence encoded by the outer exons serve to stabilize the protein in its native environment and to optimize the binding properties.

To our knowledge, so far most experiments to obtain functional myoglobin fragments have started from horse heart myoglobin (hh-Mb[1]). Obtaining functional fragments from sw-Mb is desirable for several reasons: Myoglobin from sperm whale (*Physeter catodon*) is currently the most extensively examined myoglobin with respect to crystal and solution structures at atomic resolution [10, 11]. Also, the binding and folding behaviour as well as other physico-chemical properties of both the holo- and the apoprotein have been investigated most thoroughly both by nuclear magnetic resonance and optical techniques, and are therefore exceptionally well characterized [12–19]. Due to that large amount of reference data available it seemed desirable to base an approach for the preparation and investigation of a minimal functional unit on this well known molecule. In the present communication we report the use of limited protease digestion of sperm whale apomyoglobin to obtain functional fragments corresponding even more closely to the inner exon (*i.e.* residues 31 to 105) of sperm whale Mb than mini-myoglobin does.

Results and Discussion

In contrast to horse heart myoglobin [1], sperm whale myoglobin contains two additional arginine residues in positions 45 and 118 (Fig. 1) providing additional cleavage sites for proteolysis with the highly arginine specific protease clostripain. Cleavage at Arg118, which is positioned close to the end of the G-helix, causes the loss of the entire H helix of Mb, which according to most recent results [20] is probably not essential for a functional heme binding protein. However, peptide bond cleavage at Arg45 would cause loss of essential residues involved in either forming the heme pocket or binding of the heme prosthetic (Leu32, Thr39, Lys42, Phe43, Arg45). The tested clostripain digestion of sperm whale apoMb under conditions where the apoMb exists in its stable, natively folded form assuming that the Arg45 site would be attacked slowly under these conditions, thus leading to intermediate products of the desired sequence length (Fig. 2). The experiments on digesting sw-apoMb using clostripain yielded more fragments than analogous experiments on horse heart myoglobin (Fig. 3). As can be seen from the mass spectrometric data in Table 1 no fragments with an intact Arg45–Leu46 peptide bond were found. From this we conclude that the 45–46 peptide bond is either sufficiently accessible in the native fold to be cleaved first or that any other cut at an Arg site makes possible a rapid attack by clostripain at the Arg45 site. This makes it impossible to obtain a heme binding fragment by clostripain digestion of sw-apoMb.

Pepsin is a less specific protease than clostripain which, however, does not cleave the Arg peptide bonds. The amino acid residues in the neighborhood of the start (31) and the end (105) of the central exon are the same in sperm whale and horse heart myoglobin and, as can be seen from Fig. 2, these residues are also in the neighborhood of pepsin cleavage sites. Pepsin digests of sperm whale

[1] For abbreviations used, see Experimental

Fig. 1. Amino acid sequences of sperm whale (*Physeter catodon*, Swissprot file MYG PHYCA) myoglobin and horse (*Equus cabalus*, Swissprot file MYG HORSE) myoglobin. The differences in the sequence are indicated by inverse print, arginine residues (the target site for clostripain) in both sequences are highlighted as bold print in gray boxes. The recognition sites and cuts for a pepsin digest are indicated below the SW-Mb sequence by arrows: ↑ . . .pepsin specific site where proteolysis occurred in our experiments, ⇞ . . .pepsin specific site where proteolysis was not observed in our experiments, ⇑ . . . additional sites (unusual for pepsin) at which cleavage occured

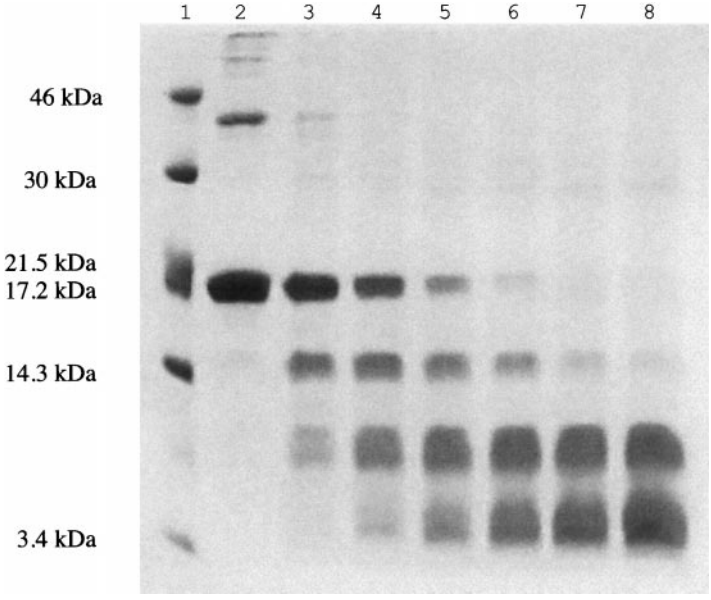

Fig. 2. SDS-PAGE monitoring of digestion progress; the time course of a pepsin digest is monitored by SDS-PAGE (E/S = 1/1000); lane 1: marker, lane 2:0 min, lane 3: 5 min, lane 4:10 min, lane 5:15 min, lane 6:20 min, lane 7:25 min, lane 8:30 min

myoglobin contain more fragments than their clostripain counterparts. From the pepsin lysates obtained under the conditions given in the experimental part we isolated fragments larger than 4 kDa by chromatography on a C-4 RP-HPLC column (Fig. 4). The major fragments were analyzed by mass spectrometry. The results are summarized in Table 2. In addition to proteolysis at the expected sites,

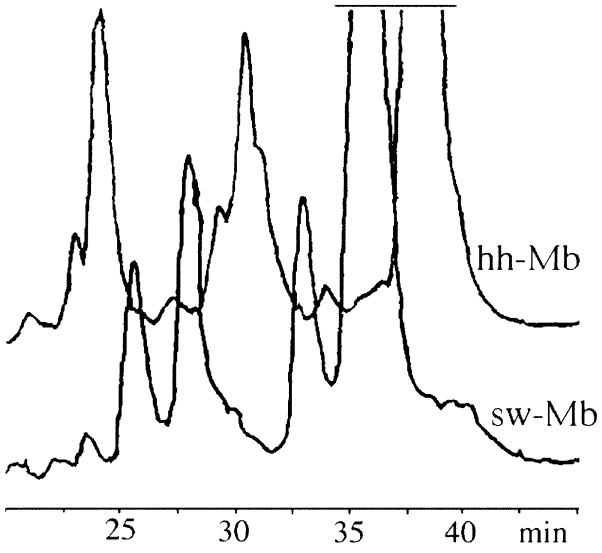

Fig. 3. Comparison of HPLC separations of clostripain digests of horse heart apoMb (upper trace, 2 major fragment peaks, native protein at 38 min.) and sperm whale apoMb (lower trace, 3 major fragment peaks, native protein at 36 min.); the maximum peak height corresponds to 0.12 absorption units

Table 1. Results from mass spectroscopic analysis (mass accuracy: 0.02%) of HPLC fractions of apoMb – clostripain digests

sw-apoMb (native)			hh-apoMb		
Fragment	Calc. mass/Da	Found mass/Da	Fragment	Calc. mass/Da	Found mass/Da
1–153	17199.9	17199	1–153	16951.5	16951
1–31	3468.9	3467.5	1–31	3404.8	3408.0
1–45	5197.9	5196.7	1–139	15416.7	not detected
46–118	8226.6	8224.7	32–139	12028.9	12035
46–139	10454.1	10451	32–153	13564.9	13562
46–153	12019.9	12019	140–153	1553.7	not detected
119–153	3811.3	3810.5	–	–	–

non-specific proteolytic cleavage was found between residues Leu69/Thr70 and Thr70/Ala71 to which all complimentary fragments were also found (*e.g.* 1–69/70–106/107–153 or 30–70/71–153). None of the major fragments has apparently been cleaved at Leu72 and Leu89. It should be noted that there may well be fragments smaller than 4 kD in the lysate that stem from cleavage at their sites, but these fragments have not been analyzed due to overlaps caused by the large number of similar fragments.

Two fragments were found that correspond closely to the polypeptide encoded by the central exon. These fragments were named micro-myoglobin A (μMb A, Ile 30-Phe106, 77 residues), which is the prevailing fragment according to chromatography and mass spectra, and micro-myoglobin B (μMb B, Phe33-Phe106, 74 residues). Whereas the HPLC fraction containing μMb A was essentially pure, μMb B was found in two fractions together with other lower

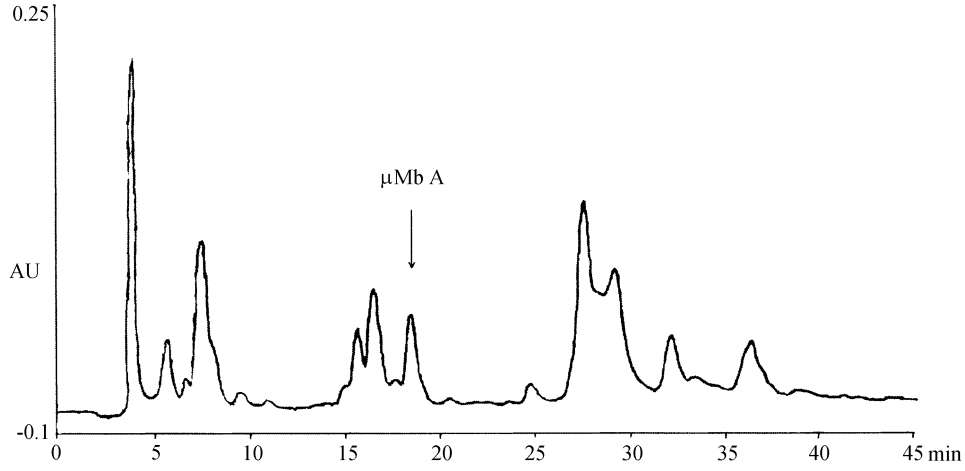

Fig. 4. RP-HPLC chromatogram (approximately 1.5 mg of total protein were applied to a 10×250 mm C4 column) monitored at 214 nm of a pepsin digest of sperm whale apomyoglobin. μMb A elutes as the peak at 18.5 min indicated by the arrow, whereas μMb B was found by mass spectrometry as a minor component in the two peaks at 16.5 and 17.5 min immediately preceding it

Table 2. Pepsin proteolysis fragments from sperm whale apoMb and their calculated and measured masses (mass accuracy 0.02%)

sw-apoMb fragment (residues)	Calculated mass Da	Found mass Da	Polypeptide
1–153	17199.9	17199	sw-apoMb
1–106	12050.0	12047	
1–69	7920.1	7919.4	
30–153	14018.3	14015	
30–106	8868.4	8865.5	μMb A
30–70	4839.6	4838.7	
30–69	4738.8	4739.6	
33–153	13635.8	13635	
33–106	8485.9	8485.0	μMb B
33–69	4356.03	4355.6	
70–153	9297.8	9293.7	
70–137	7426.8	7426.8	
70–106	4147.9	4146.3	
71–153	9196.7	9196.2	
107–149	4762.5	4766.6	
107–153	5167.9	5167.1	

molecular mass peptides. Since the yield of μMb B as estimated from the HPLC peak integrals and relative ion abundance in the mass spectra was far below 1%, no further purification was attempted.

Structure and heme binding investigations are limited by the small sample amounts remaining after HPLC (estimated 5% yield of μMb A starting from a total of 1 mg raw protein lysate) and the restricted supply of native sw-Mb. Evidence for

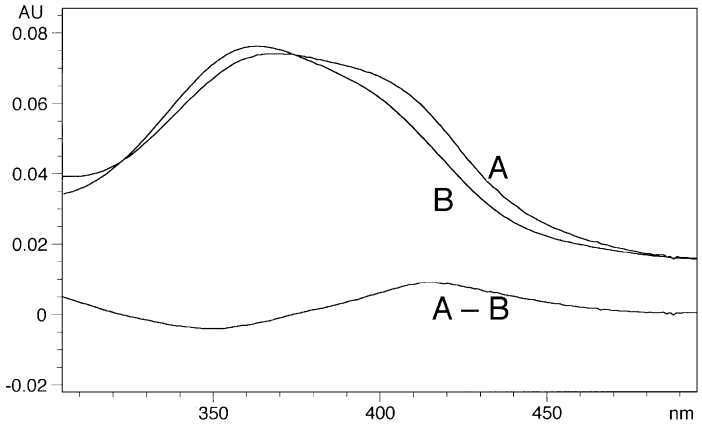

Fig. 5. *Soret* region of UV/Vis spectra of a freshly prepared solution of hemin and μMb A (A), reference spectrum of a hemin solution (B), and their difference spectrum (A–B) in the *Soret* region show the changes characteristic of heme binding by μMb A; spectra were recorded on a HP 8453 diode array UV/Vis spectrometer at room temperature

heme-binding by μMb A was found using UV/Vis spectroscopy (Fig. 5). When μMb A is added to a solution of hemin, the peak maximum shifts from 385 to 404 nm. Due to the excess of heme the change is small but can be seen clearly in the difference spectrum. This shift is characteristic of heme binding as observed from native myoglobins where the 385 nm *Soret* peak decreases in amplitude, whereas a shoulder at 404 nm transforms into a narrow and more intense peak at longer wavelength, a behaviour which is attributed to coordination of the fifth and sixth coordination site of the iron. In accordance with the trend observed in the heme binding studies of horse heart mini-myoglobin [6], the effect is even smaller for μMb A, apparently due to the lower number of structural constraints in the shorter polypeptide. Preliminary CD experiments (not shown) also indicate very little helical structure, which increases upon addition of ethanol or trifluoroethanol.

In conclusion, we showed that clostripain digestion is not a suitable approach to prepare heme binding fragments from sw-Mb. However, we were able to prepare micro-myoglobin, a polypeptide corresponding to the inner exon of sw-Mb by pepsin digestion of sperm whale apo-myoglobin. Due to the whale protection sperm whale myoglobin is of limited availability. Therefore it appears worthwhile to find alternate routes for the preparation of larger amounts of heme and potentially oxygen binding myoglobin fragments of less than 80 amino acids, in particular for structural studies. Such studies are currently in progress and will be reported elsewhere [21].

Experimental

Abbreviations: Mb, myoglobin; apoMb, apomyoglobin; sw, sperm whale; hh, horse heart; HPLC, high performance liquid chromatography; RP, reversed phase; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; E/S, enzyme/substrate; μMb, micro-myoglobin.

Preparation of the apoprotein

Sperm whale and horse heart apoMb were prepared according to a modified version of the method of *Taele* [22]. 10 mg of the respective myoglobin (obtained from Sigma)[2] were dissolved at 0°C in 1 cm^3 20 m*M* acetate buffer (*pH* 3.8) containing 0.1 *M* NaF. The resulting solution was extracted three times with three volumes of chilled 2-butanone saturated with water. The remaining aqueous phase was rapidly diluted into 9 volumes of chilled 10 m*M* ammonium acetate buffer at *pH* 6 and dialyzed twice against approximately 20 volumes of 10 m*M* ammonium acetate buffer at *pH* 6 and 4°C. After dialysis the protein solution was lyophilized.

Clostripain proteolysis

2 mg of sw-apoMb or hh-apoMb were dissolved in 0.2 cm^3 of 20 m*M* *pH* 8 phosphate buffer containing 2 m*M* $CaCl_2$ and 2 m*M* 2-mercaptoethanol and centrifuged to remove insoluble material. The solution was incubated with clostripain at an E/S ratio of 1/100 at room temperature for 30 min. The reaction was either quenched by rapid freezing in liquid nitrogen followed by lyophilization or by addition of SDS to give a final concentration of 0.5%.

Pepsin proteolysis

Digests employing pepsin were carried out at *pH* 2 in 20 m*M* citrate buffer containing 100 m*M* sodium chloride by dissolving 1 mg of sw-apoMb per 100 mm^3 buffer. The experimental conditions for preparative digests were optimized with respect to the E/S ratio as well as to the incubation time (Fig. 1). The progress of proteolysis was monitored by SDS-PAGE according to the protocol of *Laemmeli* [23]. For the preparative run an E/S ratio of 1/1000 was applied, and the reaction was allowed to proceed for 15 min at 22°C. The reactions were quenched as described above.

The digests were analyzed by reversed phase HPLC using a Vydac C4 column (214TP54, 2.1 × 2.50 mm, 0.2 cm^3/min flow rate. The samples were loaded at 30% acetonitrile in H_2O containing 0.1% trifluoroacetic acid and eluted with a linear gradient to 45% acetonitrile over 40 min. For the larger scale preparation starting from 1.5 mg of apoMb we used a semi-preparative Vydac C4 column (214TP1010, 10 × 250 mm preceded with a Vydac guard column cartridge, flow rate 4.0 cm^3/min). Fractions were collected manually and subsequently lyophilized and analyzed by mass spectrometry.

Mass spectrometry was carried out using a HP 5989B quadrupole mass analyzer equipped with a pneumatically assisted electrospray interface (HP59987A). The drying gas (99.999% N_2) was delivered at 150°C and a flow of 7 dm^3/min; nebulizing gas pressure: 60 psi. Samples were dissolved in 150 to 250 mm^3 of a 1/1 mixture of H_2O and acetonitrile containing 1 to 5% formic acid. Continuous flow injection at a flow rate of 5 mm^3/min was applied using a syringe pump. The mass spectrometer was tuned with horse heart cytochrome c; sw-Mb and hh-Mb were used as external mass reference standards. The calculated masses of the standards and the fragments were computed using tabulated values for the average masses of the individual amino acids [24]. Mass accuracy was 0.02–0.05%.

Acknowledgements

The authors thank Prof. Dr. *Wolfgang Buchberger* for generous access to the mass spectrometer and Dr. *Rita Grandori* for constructive criticism of the manuscript. *S.S.* acknowledges inspiring

[2] The stock of sperm whale myoglobin at Johannes Kepler University was purchased from Sigma before the global ban on whale products.

discussions with Prof. Dr. *Angelo Fontana*, University of Padua, Italy, in the initial phase of this work. Financial support was obtained through a research grant (206–510-A) from the City of Linz to *S.S.*, from the province government of Upper Austria (*Forschungsförderung des Landes Oberösterreich*), and the *Allgemeine Sparkasse Oberösterreich* (Sparkasse-Forschungsförderung 1997), and by the *Fonds zur Förderung der wissenschaftlichen Forschung* (project P12696 CHE). *N.M.* also acknowledges a research grant from the *Jubiläumsfonds of the Austrian National Bank* (project no. 6808).

References

[1] Gō M (1981) Nature **291**: 90
[2] Craik CS, Buchman SR, Beychok S (1980) Proc Natl Acad Sci USA **77**: 1384
[3] Craik CS, Buchman SR, Beychok S (1981) Nature **291**: 87
[4] Cho G, Doolittle RF (1997) J Mol Evol **44**: 573
[5] De Sanctis G, Falcioni G, Giardina B, Ascoli F, Brunori M (1986) J Mol Biol **188**: 73
[6] De Sanctis G, Falcioni G, Giardina B, Ascoli F, Brunori M (1986) J Mol Biol **200**: 725
[7] De Sanctis G, Falcioni G, Grelloni F, Desideri A, Polizio F, Giardina B, Ascoli F, Brunori M (1991) J Mol Biol **222**: 637
[8] Di Iorio EE, Yu W, Calonder C, Winterhalter KH, De Sanctis G, Falcioni G, Ascoli F, Giardina B, Brunori M (1993) Proc Natl Acad Sci USA **90**: 2025
[9] De Sanctis G, Ascoli F, Brunori M (1994) Proc Natl Acad Sci USA **91**: 11507
[10] Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR, Phillips DD, Shore VC (1954) Nature (London) **185**: 422
[11] Jennings PA, Stone MJ, Wright PE (1995) J Biomol NMR **6**: 271
[12] Jennings PA, Wright PE (1993) Science **262**: 892
[13] Eliezer D, Yao J, Dyson HJ, Wright PE (1998) Nature Structural Biology **5**: 148
[14] Griko YV, Privalov PL, Venyaminov SY, Kutyshenko VP (1988) J Mol Biol **202**: 127
[15] Hughson FM, Wright PE, Baldwin RL (1990) Science **249**: 1544
[16] Hughson FM, Barrick D, Baldwin RL (1991) Biochemistry **30**: 4114
[17] Barrick D, Hughson FM, Baldwin RL (1994) J Mol Biol **237**: 588
[18] Marko H, Müller N, Falk H (1990) Eur J Biochem **193**: 573
[19] Wagner UG, Müller N, Schmitzberger W, Falk H, Kratky C (1995) J Mol Biol **247**: 326
[20] Cavagnero S, Dyson HJ, Wright PE (1999) J Mol Biol **285**: 269
[21] Grandori R, Schwarzinger S, Müller N (2000) Eur J Biochem (in press)
[22] Taele FWJ (1959) Biochem Biophys Acta **35**: 543
[23] Laemmeli UK (1970) Nature **227**: 680
[24] Current Protocols in Protein Science on CD-ROM (1998) Wiley, New York

Received October 28, 1999. Accepted November 23, 1999