Nanoreactors for the Multi-Functionalization of Poly-Histidine Fragments

Marco Paolino,[a] Annalisa Reale,[a] Vincenzo Razzano,[a] Germano Giuliani,[a] Alessandro Donati,[a] Claudia Bonechi,[a] Gianfranco Caselli,[b] Michela Visintin,[b] Francesco Makovec,[b] Cinzia Scialabba,[c] Mariano Licciardi,[c] Eugenio Paccagnini,[d] Mariangela Gentile,[d] Laura Salvini,[e] Francesco Tavanti,[f] Maria Cristina Menziani,[f] Andrea Cappelli.[*,a]

[a]Dipartimento di Biotecnologie, Chimica e Farmacia (Dipartimento di Eccellenza 2018-2022), Università di Siena, Via A. Moro 2, 53100 Siena, Italy,

[b]Rottapharm Biotech S.p.A., Via Valosa di Sopra 9, 20900 Monza, Italy,

[c]Dipartimento di Scienze e Tecnologie Biologiche, Chimiche e Farmaceutiche (STEBICEF), Università degli Studi di Palermo, Via Archirafi 32, 90123 Palermo, Italy

[d]Dipartimento di Scienze della Vita, Università di Siena, Via A. Moro 2, 53100 Siena, Italy,

[e]Toscana Life Sciences Foundation, Via Fiorentina 1, 53100 Siena, Italy

[f]Dipartimento di Scienze Chimiche e Geologiche, Università di Modena e Reggio Emilia, Via Campi 103, 41121 Modena, Italy,

* Corresponding author. Tel: +39 0577 234320. E-mail: andrea.cappelli@unisi.it.

Abstract

Hexahistidine motif is often encoded at one extremity of amino acid sequences of recombinant proteins. The reaction of some water-soluble Morita-Baylis–Hillman adduct (MBHA) derivatives with equimolar amounts of *N*-acetylhexahistidine was investigated in order to evaluate their reactivity towards poly-histidine sequences. Very intriguingly, the results of kinetic studies along with the characterization of the reaction products strongly supported that multiple electrophilic attacks occurred to the same **Ac-His-6** molecule, as an apparently cooperative process. MBHA derivatives **2** and **3** were found to self assemble in water environment (as suggested by pyrene fluorescence analysis, DLS, CryoTEM, and high field NMR spectroscopy) to generate spherical micelles showing core-shell architectures. By virtue of their molecular and supermolecular organizations, the micelles appeared to be capable of performing massive attacks to the chemical species, which after an initial interaction with the solvated oligo(ethylene glycol) shell are able to reach the core. These results provide the bases for the development of a new approach in covalent protein functionalization.

Introduction

Hexahistidine motif is commonly encoded at the beginning or at the end of an amino acid sequence of a recombinant protein in order to allow its isolation and purification after its expression in bacteria from recombinant DNA vectors.[1-5] In this way, the protein of interest is easily separated from native proteins by affinity chromatography, and this technology has become the gold standard for preparation of recombinant proteins so that many engineered polypeptides contain in their sequences poly-histidine fragments.[6-10] Moreover, poly-histidine tags represent interesting opportunities for selective derivatization of these proteins.[11]

In previous works, some Morita-Baylis–Hillman adduct (MBHA) derivatives were prepared[12,13] and reacted with imidazole, *N*-acetylhistidine (**Ac-His-1**), and *N*-acetylhexahistidine (**Ac-His-6**) in order to evaluate their reactivity towards the imidazole fragment in different environment. Interestingly, naphtalene MBHA derivative **1** react with imidazole moiety providing biadduct compounds as the main reaction products (Figure 1).[13]

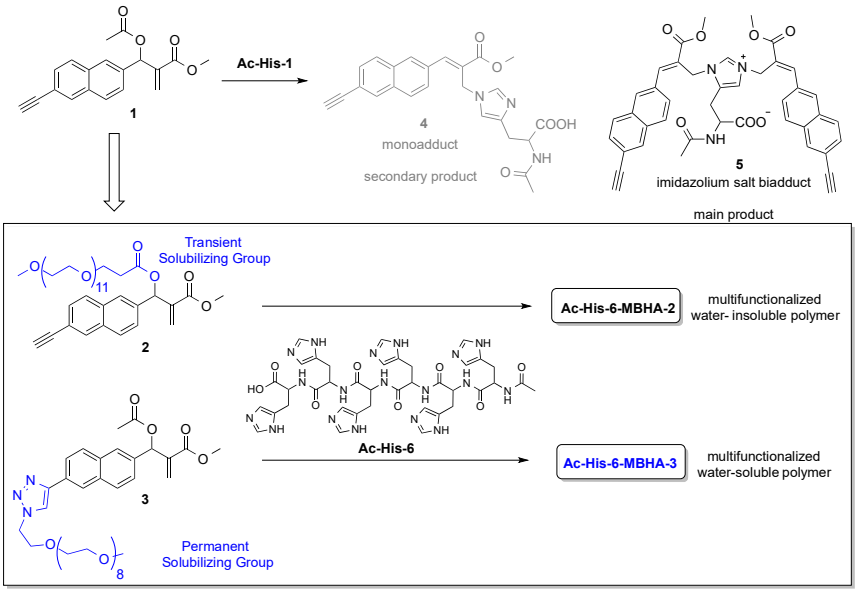

Figure 1. Reactivity of MBHA derivative **1** with *N*-acetylhistidine (**Ac-His-1**) and of the water-soluble MBHA **2** and **3** with *N*-acetylhexahistidine (**Ac-His-6**).

Thus, the structures of these compounds were manipulated by introducing oligo(ethylene glycol) chains in the aim of obtaining water-soluble derivatives **2** and **3**, which were characterized from the point of view of their reactivity towards *N*-acetylhexahistidine in PBS (Figure 1).

Very interestingly, mass spectrometry, NMR spectroscopy, and photophysical studies confirmed the presence of considerable amounts of biadduct residues in both the polymeric materials **Ac-His-6-MBHA-2** and **Ac-His-6-MBHA-3** obtained with ten-fold excess of compounds **2** and **3,** respectively.[13]

In the present article we investigate the macromolecular mechanisms at the basis of the multi-functionalization of *N*-acetylhexahistidine using amphiphilic molecules **2** and **3** in water solutions.

Intrigued by the reactive double addition phenomenon, in order to evaluate the relative reactivity of **2** and **3**[13] in the same reaction conditions, kinetic experiments were performed in NMR tubes containing **Ac-His-6**, and only one equivalent of these MBHA derivatives (concentration around 1 mg/mL) in PBS medium prepared with deuterium oxide using DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid sodium salt) as an internal standard (Figures SI1 and SI2).

The ^{1}H NMR experiments demonstrated that after 48 h at room temperature the initial amounts of MBHA derivatives were almost completely converted, whereas substantial amounts of **Ac-His-6** appeared to be present in the reaction mixture. The residual amounts of unreacted **Ac-His-6** could be estimated with some uncertainness to be in the range of 60-80%, suggesting that multiple electrophilic attacks occurred to the same **Ac-His-6** molecule.

Surprised by this intriguing results, further kinetic experiments were performed with MBHA derivative **3**, which showed the highest relative reactivity and produced water-soluble polymeric materials, as suggested by both the clear reaction mixtures and the presence of broad signals in the spectra in Figure SI2. In particular, the kinetic experiments were performed in NMR tubes containing **Ac-His-6**, only one equivalent of the MBHA derivative at the initial concentrations of 3 and 5 mg/mL in PBS medium prepared with deuterium oxide without the internal standard DSS,

which was excluded from the reaction mixture in order to eliminate its potential interaction with the reagents.

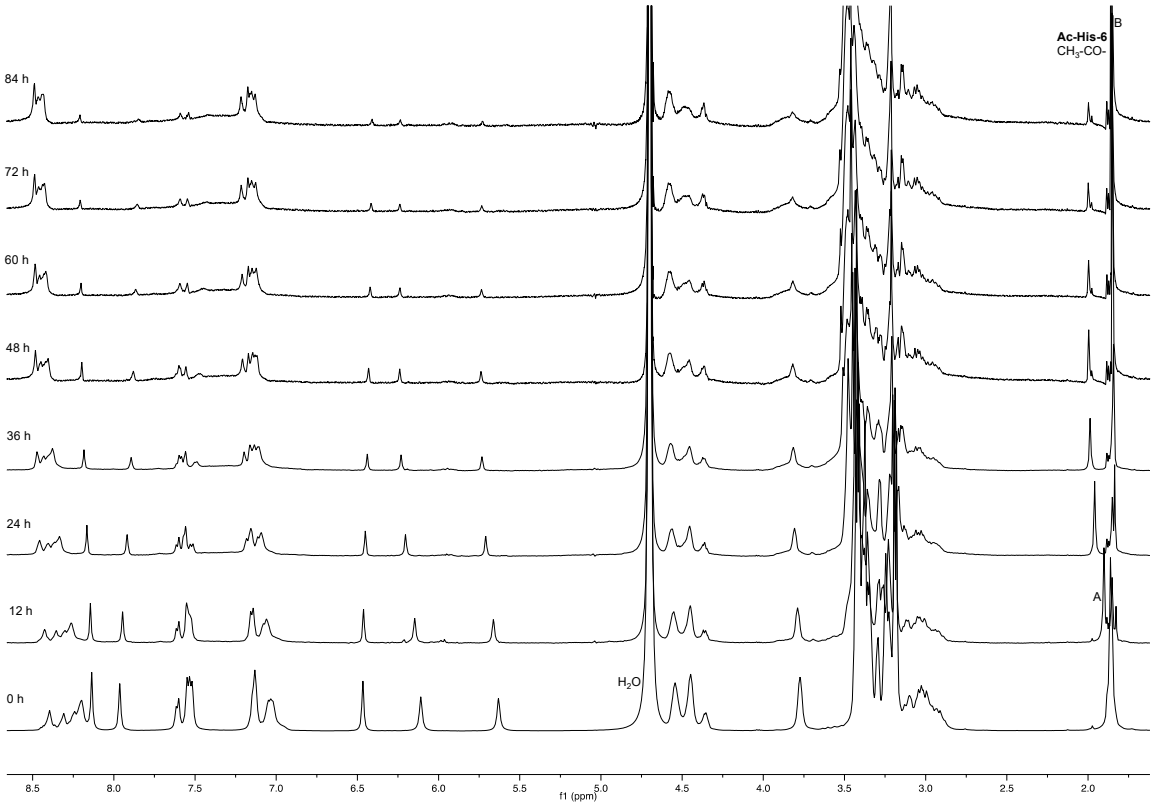

Figure 2. ^{1}H NMR spectra (500 MHz, PBS-D_2O) of the mixtures obtained by reaction of **Ac-His-6** with an equimolar amount of compound **3** (initial concentration 5 mg/mL). The peaks marked with A (attributed to the acetyl moiety of **3**) and B (attributed to acetate produced by the concerted addition–elimination process) were used to follow the reaction progress.

These experiments confirm the presence in solution of a high amount of unreacted **Ac-His-6** despite the virtual disappearance of the MBHA **3** in the reaction mixture.

The composition of the reaction mixture of the kinetic experiments was also studied by ^{13}C NMR analysis (Figure SI4) of the mixture obtained by reaction of **Ac-His-6** with an equimolar amount of compound **3** (initial concentration 5 mg/mL) in order to get information about the structure of the species present in the reaction mixture (i. e. the polymeric material).

The comparison shown in Figure SI4 confirmed the presence in the reaction mixture of substantial amounts of unreacted **Ac-His-6**, after the almost complete disappearance of **3** (after 3 weeks at room temperature), and of a polymeric material showing spectroscopic features very similar to those shown by of previously published polymeric material **Ac-His-6-MBHA-3** obtained by reaction of **Ac-His-6** with a ten-fold excess of **3**. This similarity was particularly evident when the addition of $CDCl_3$ in the NMR tube containing the reaction mixture allowed the polymeric material to be extracted selectively (Figure SI4B), with the unreacted **Ac-His-6** molecules and mono-substituted species [**Ac-His-6-MBHA-3(1)**] remaining in the PBS solution (as confirmed by MALDI-TOF analysis, data not shown).

Moreover, the composition of the reaction mixtures of the kinetic experiments performed at 3 mg/mL and 5 mg/mL was also studied by MALDI-TOF-MS measurements (Figures SI5 and SI6). These experiments suggested that the reaction mixtures showed a quite heterogeneous composition. In fact, the exploration of the spot in the MALDI target showed that different compositions were associated to different areas in the spot. In particular, the edges were rich in both unreacted **Ac-His-6** molecules and mono-substituted species [**Ac-His-6-MBHA-3(1)**] (see Figure SI5), whereas the center of the spot was particularly rich in the polymeric material showing a MALDI spectrum (see Figure SI6) very similar to that of the material previously obtained by reaction of **Ac-His-6** with a ten-fold excess of **3** (i. e. **Ac-His-6-MBHA-3**),[13] as suggested by ^{13}C NMR spectra of the reaction mixture.

Curiously, the comparison of the results of these kinetic experiments (shown in Figures SI2, SI3 and 2) suggested that the initial concentration of MBHA derivative **3** could affect the apparent reaction rate and the shape of the peaks attributed to the unreacted **Ac-His-6**. Moreover, significant variation in the chemical shift values were observed for the signals (at 5.63 and 6.10 ppm in Figure 2) attributed to the reactive methylene moiety of MBHA **3** with the progress of the reaction and in agreement with the decrease of the amount of **3**. Thus, the effects of compound **3** concentration on the chemical shift values was ascertained by high field NMR spectroscopy studies. The comparison

of the ^{1}H NMR spectra obtained with **3** solutions in PBS showing increasing concentrations (i. e. 0.1, 0.3, 0.5, 1.0, 3.0, and 5.0 mg/mL; Figure 3) showed a clear dependence of the chemical shift values on the concentration. In particular, significant down-field shifts were observed in the aromatic/vinylic region of the spectra with the decrease of the concentration up to chemical shift values observed in the spectrum recorded with a solution in deuterated chloroform showing a concentration value of 3 mg/mL.

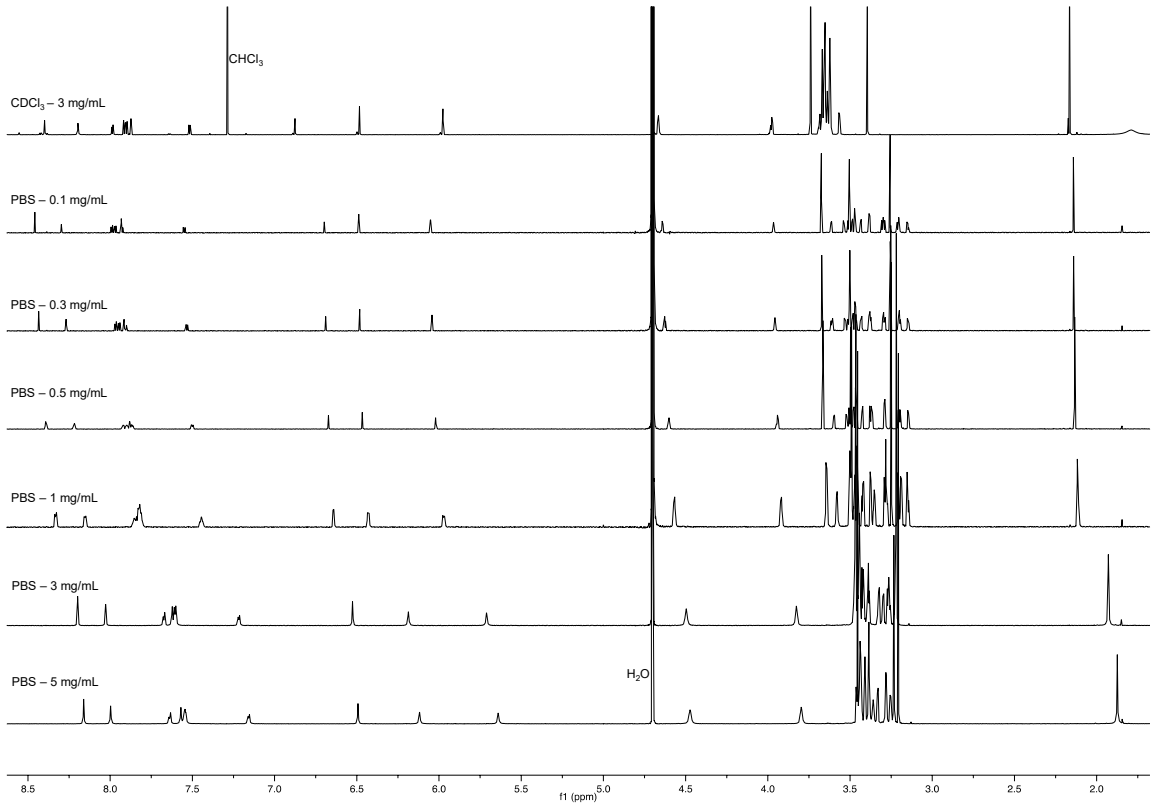

Figure 3. Comparison of the ^{1}H NMR spectra (950 MHz) obtained with **3** solutions in PBS showing increasing concentrations with that obtained in $CDCl_3$.

This behavior suggested the presence of an aggregation phenomenon, which could be easily inferred from the amphiphilic nature of these compounds.

Therefore, the aggregation behavior of the water-soluble amphiphilic MBHA derivatives **2** and **3** was investigated by using four different techniques: (a) pyrene fluorescence analysis,[14] (b) dynamic

light scattering (DLS), (c) cryo transmission electron microscopy (CryoTEM) techniques, and (d) bidimensional high field (i. e. 950 MHz) NMR spectroscopy.

The critical aggregation concentration (CAC) values of **2** and **3** were determined by fluorescence analysis using pyrene as probe and the results were shown in Figure 4A and 4B. The CAC value was taken from the intersection of the regression lines obtained before and after the slope change (low and high concentration regions). MBHA derivative **2** showed a CAC value (i.e. 0.85 mg/mL, $1.0 \cdot 10^{-3}$ mmol/mL), whereas compound **3** showed a CAC value of 0.68 mg/mL ($8.9 \cdot 10^{-4}$ mmol/mL). Interestingly, the variations in the chemical shift values observed in the ^{1}H NMR spectra with MBHA **3** were consistent with the CAC measured by fluorescence analysis using pyrene as probe (i. e. 0.68 mg/mL). In fact, the decrease of the concentration below 0.5 mg/mL produced only minor changes, with respect to the variations observed above 1 mg/mL. Overall, the results of these measurements confirmed our assumption on the relationships between the amphiphilic structure of MBHA derivatives **2** and **3** and their aggregation liability.

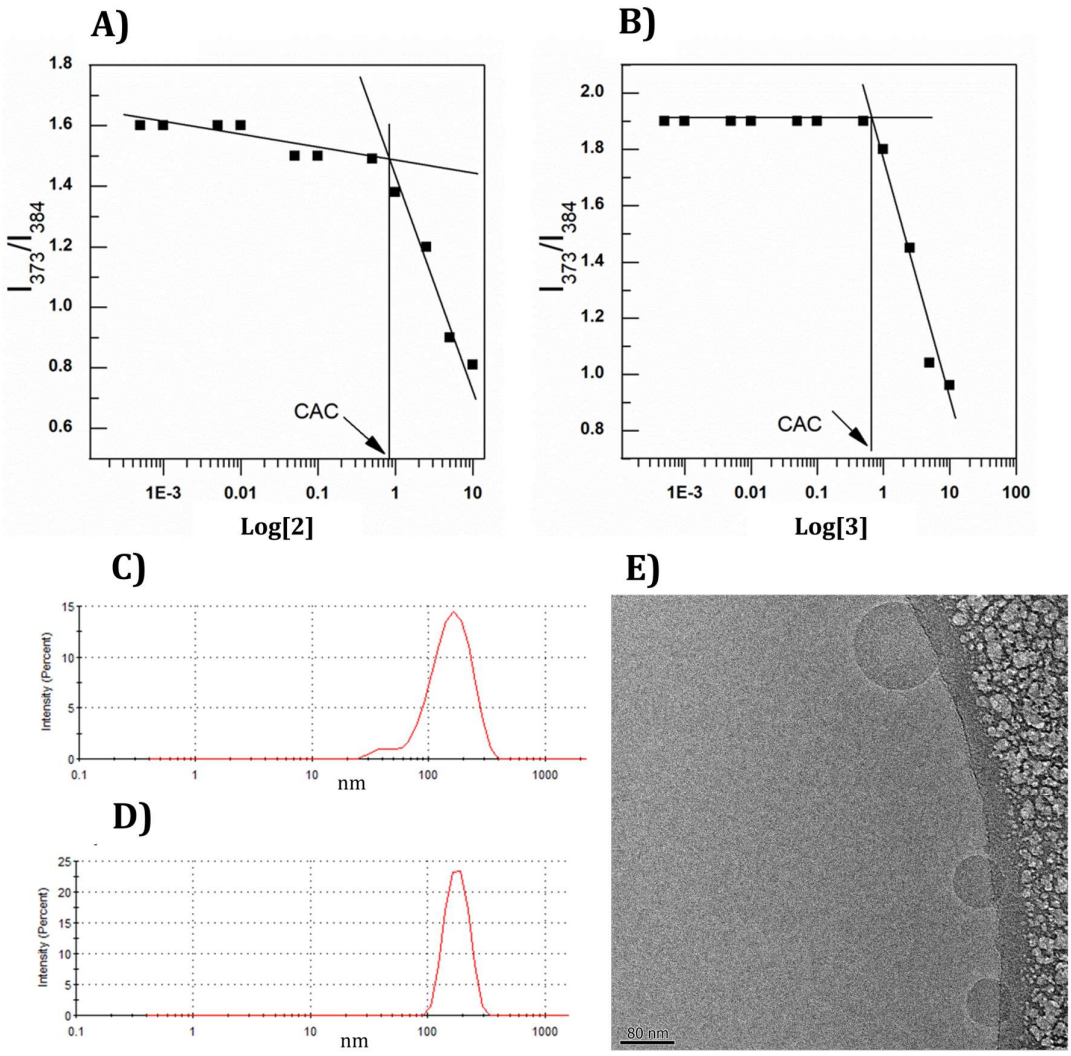

Figure 4. Aggregation studies: I_{373}/I_{384} intensity ratio obtained from pyrene emission spectra in the presence of **2** (A) and **3** (B) as a function of the logarithm of their concentrations; DLS size distribution histograms of **2** (C) and **3** (D) dispersions in PBS; CryoTEM image obtained with a water solution of **3** at a concentration (i. e. 4.5 mg/mL) well above its CAC. The three spherical micelles interact with the edge of the QUANTIFOIL® R 2/1 300-mesh copper grids.

The solutions of water-soluble derivatives **2** and **3** in PBS were completely clear. However, DLS analyses of the solutions confirmed the tendency of these MBHA derivatives to form nanoaggregates in PBS solution at the concentration above their CAC (i. e. 1.0 mg/mL). DLS measurements were performed at 25 °C using a Malvern Zetasizer NanoZS instrument; the results

obtained with dispersions of **2** and **3** in PBS are summarized in Table S1 and the most significant histograms are shown in Figures 4C and 4D. The results of the DLS analyses confirmed again the self-assembling features of the naphthalene MBHA derivatives. In particular, the results suggested that the dispersions of compound **2** in PBS appeared to contain two heterogeneous populations: one showed a small mean diameter (around 40 nm) and was assumed to be constituted small aggregates composed by few molecules. The other population showed greater dimensions (around 170 nm) and were obviously constituted by a large number of molecules probably organized in micelle-like structures. On the contrary, only one population (diameter around 200 nm) was observed with **3**. Thus, this latter MBHA derivative was selected for further characterization in cryoTEM studies. In order to overcome possible problems due to the presence of salts in the amorphous water, we decided of performing cryoTEM studies on MBHA derivative **3** by using a solution in water at a concentration well above its CAC (i. e. 4.5 mg/mL). Very interestingly, this electron microscopy technique showed the presence of spherical objects showing dimensions around 80-100 nm (Figure 4E).

The apparent discrepancy between DLS and cryoTEM analysis could be explained by the architecture of the aggregates in relation to the features of the analytic techniques and/or to the different composition of the medium used for the analysis.

In order to collect other information on the self-assembling architecture, NOESY experiments were performed in PBS and in $CDCl_3$ on MBHA **3** (Figure 5). These studies strongly supported the aggregation in the aqueous solvent and the molecularly dissolved state in the organic solvent. In fact, the cross peaks in the NOESY spectrum recorded in PBS (Figure 5A) showed the same sign of the diagonal peaks, whereas they showed the opposite one in the NOESY recorded in $CDCl_3$ (Figure 5B). For molecules with short correlation times τc the cross peaks have the opposite sign of the diagonal peaks, while larger molecules with long correlation times have the same sign for all peaks.[15] Furthermore, the cross peaks in the NOESY spectrum recorded in PBS are intense when involved signals in the aromatic/vinylic region, suggesting that the aromatic regions of different

molecules are close each other in the space. On the other hand, the cross peaks involving the signals of the nona(ethylene glycol) chains were weak, supporting a micelle architecture showing the side chains fluctuating in the water environment around the micelle core.

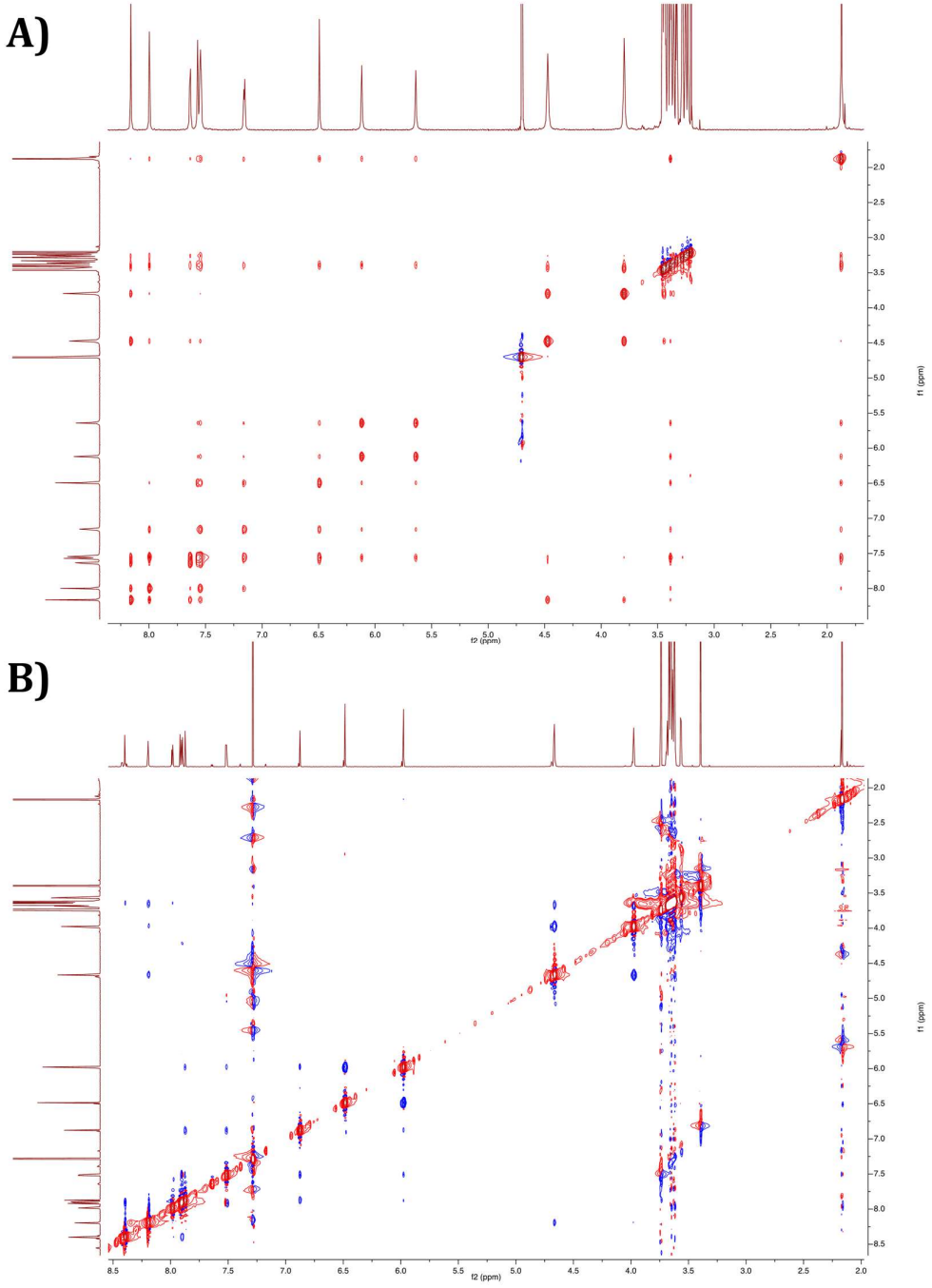

Figure 5. NOESY spectra (950 MHz) obtained with a **3** solution in PBS-D_2O at the concentration of 5 mg/mL (A), and in $CDCl_3$ at the concentration of 3 mg/mL (B).

.

Thus by virtue of its hydrophobic nature, the naphthalene moiety was assumed to drive the aggregation and to localize into a micelle core, whereas the amphiphilic oligo(ethylene glycol) chains-tails are located in the micelle edge and exposed to the water environment. In this organization, the reactive moiety of MBHA derivatives **2** and **3** is likely to be located inside the core of the micelle, which can therefore behave as a functional nanoreactor towards the chemical species capable of interacting with the micelle core.[16,17]

These results appeared to easily explain the simultaneous massive attack to the same **Ac-His-6** molecule. In particular, the **Ac-His-6** molecule can be assumed to interact with the MBHA micelle in two different steps. In the first one, it could interact with the shell formed by the solvated OEG chains-tails up to reaching the core where are located the reactive groups. Thus, in the second step **Ac-His-6** molecule interacts with the surface of the core, where the massive attack can occur. Imaginatively, as in an aquarium, the poly-histidine sequence (the clownfish) interacts with the shell formed by the solvated nona(ethylene glycol) chains (the tentacles of the sea anemone) up to reaching the core where are located the reactive groups and the massive attack can occur to generate the multifunctionalized fishbone-like poly-histidine.

The proposed mechanism led us to assume the existence of a certain degree of disorder in the poly-alkylated poly-histidine species, which could, however, organize their structures by virtue of the propensity to the formation of biadducts featuring the characteristics of imidazolium salts. Thus, in order to evaluate the potential stability of the poly-histidine sequences containing biadducts residues of molecule **3**, molecular dynamics simulations have been performed considering a three-fold biadduct substitution.

The total energies averaged over the 100 ns production run for each **Ac-His-6-MBHA-3** conjugate are reported in Table S2. The energy values span a very narrow range meaning that the formation of these **Ac-His-6-MBHA-3** conjugates is almost equally probable. Coulomb and π-π interactions between the rings of **MBHA-3** and the imidazole ring of histidines are mainly

responsible for the stability of these compounds (Figure 6).

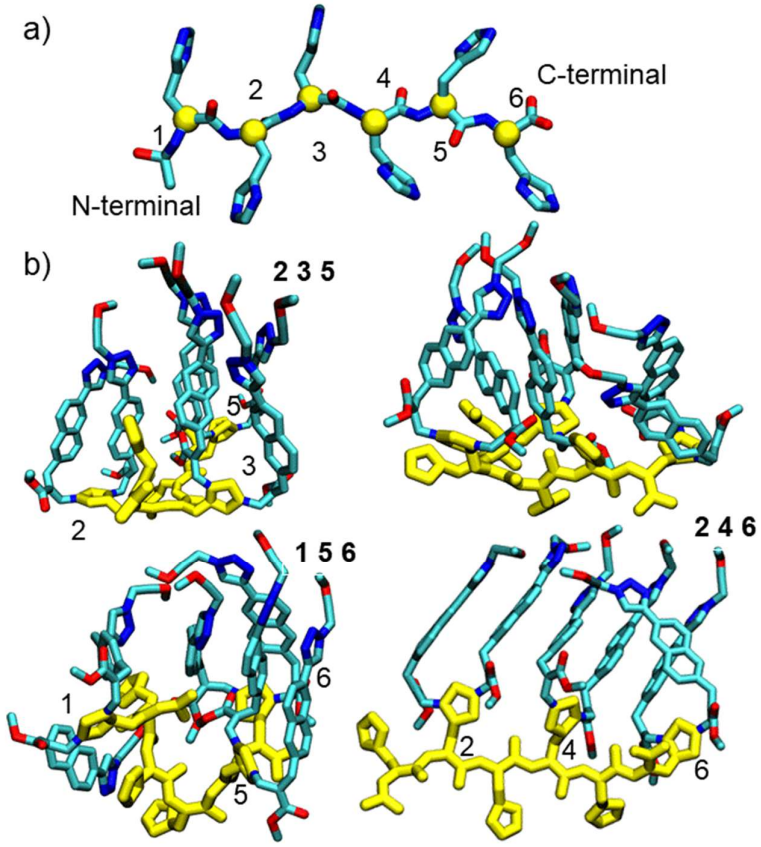

Figure 6. a) The three-dimensional model of **Ac-His-6** with numbering, b) four representative **Ac-His-MBHA-3** conjugates conformations obtained from the 100 ns production run. The **Ac-His-6** is colored in yellow.

Then, the oligo(ethylene glycol) chains have been added to the most stable "1 5 6" **Ac-His-6-MBHA-3** conjugate and the dynamics behavior of the final compound has been studied. The full-length **Ac-His-6-MBHA-3** structure shows a similar folding characterized by a close packing of the **MBHA-3** aromatic rings and the imidazole ring of the **Ac-His-6** histidine residues, whereas the oligo(ethylene glycol) flexible tails fold in a helix conformations, show higher mobility and tend to wrap themselves over the **Ac-His-6-MBHA-3** conjugate, as shown in the movie (see **ESI**).

Thus, the results of the preliminary modeling studies suggested that stable covalent core-shell architectures could be created in the **Ac-His-6** as a consequence of subsequent attacks, providing a quite complex dynamic picture of the events. Aware of the level of the approximation, we have

tried to develop a rational mechanism as it follows. When the concentration is above its CAC, **3** exists in PBS solution in the molecularly dissolved form **3**(**m**) in equilibrium with the micellar form **3**(**M**) showing different concentrations, dynamics, and reactivity. **Ac-His-6** molecules could be attacked rapidly by **3**(**m**) to give mono-substituted species **Ac-His-6-MBHA-3(1)** or more slowly, but massively by **3**(**M**) with the formation of polymeric species **Ac-His-6-MBHA-3**. The presence of substituents could increase the susceptibility towards subsequent attacks as a consequence of the creation of a hydrophobic local region, so that the reactivity increases after the first attacks, but decreases with the following ones for steric reasons. This assumption appeared to agree with the observation that multi-substituted polymeric species were observed also in kinetics experiments performed below the CAC (data not shown).

These results could be the basis of a potential selectivity, which could be exploited in the selective derivatization of histidine residues located in precise sequences of complex peptides such as poly-histidine tags and can pave the way for the development of a new biochemical technology.

ASSOCIATED CONTENTS

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI:

ACKNOWLEDGMENTS

The authors are grateful to dr. Fabio Calogiuri (CERM, Firenze, Italy) for the technical assistance in performing the high-field NMR experiments and CREA scarl (Colle di Val d'Elsa, Italy).

AUTHOR INFORMATION

Corresponding Authors

*E-mail: Andrea Cappelli: andrea.cappelli@unisi.it, tel: +39 0577 234320.

ORCID:

Marco Paolino: 0000-0003-1387-7875

Annalisa Reale: 0000-0001-6687-186X

Vincenzo Razzano: 0000-0001-6299-4504

Germano Giuliani:

Alessandro Donati:

Claudia Bonechi: 0000-0002-4328-3253

Gianfranco Caselli: 0000-0002-1075-7408

Michela Visintin: 0000-0003-3490-6807

Francesco Makovec:

Cinzia Scialabba: 0000-0002-5209-9319

Mariano Licciardi: 0000-0003-4539-9337

Eugenio Paccagnini: 0000-0002-9784-2057

Mariangela Gentile:

Laura Salvini:

Francesco Tavanti: 0000-0002-0594-5308

Maria Cristina Menziani: 0000-0003-3428-5297

Andrea Cappelli: 0000-0003-4140-3028

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Hochuli, E.; Bannwarth, W.; Döbeli, H.; Gentz, R.; Stüber D. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technology.* **1988**, *6*, 1321-1325.

(2) Lilius, G.; Persson, M.; Bulow, L.; Mosbach, K. Metal affinity precipitation of proteins carrying genetically attached polyhistidine affinity tails. *Eur. J. Biochem.* **1991**, *198*, 499–504.

(3) Janknecht, R.; Martynoff, G. D.; Lou, J.; Hipskind, R.; Nordheim, A.; Stunnenberg, H. Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proc. Natl. Acad. Sci.* **1991**, *88*, 8972–8976.

(4) Pina, A. S.; Lowe, C. R.; Roque, A. C. Challenges and opportunities in the purification of recombinant tagged proteins. *Biotechnol. Adv.* **2014,** *32*, 366-381.

(5) Porath, J.; Carlsson, J. A. N.; Olsson, I.; Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **1975**, *258*, 598–599.

(6) Hochuli, E.; Döbeli, H.; Schacher, A. New metal chelate adsorbent selective for proteins and peptide containing neighbouring histidine residues. *J. Chromatogr.* **1987**, *411*, 177-184.

(7) Ljungquist, C.; Breitholtz, A.; Brink-Nilsson, H,; Moks, T.; Uhlen, M.; Nilsson, B. Immobilization and affinity purification of recombinant proteins using histidine peptide fusions. *Eur. J. Biochem.* **1989**, *186*, 563–569.

(8) Spriestersbach, A; Kubicek, J; Schäfer, F; Block, H; Maertens, B. Purification of His-tagged proteins. *Methods Enzymol*. **2015**, 559, 1-15.

(9) Bornhorst, J. A.; Falke, J. J. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol.* **2000**, *326*, 245–254.

(10) Lindner, P.; Guth, B.; Wülfing, C.; Krebber, C.; Steipe, B.; Müller, F.; Plückthun, A. Purification of native proteins form the cytoplasm and periplasm of Escherichia coli using IMAC and histidine tails: a comparison of proteins and protocols. *Methods: A Companion to Methods in Enzymology*. **1992**, *4*, 41–55.

(11) Kepanidis, A. N.; Ebright, Y. W.; Ebright, R. H.; Site-specific incorporation of fluorescent probes into protein: hexahistidine-tag-mediated fluorescent labeling with (Ni^{2+}:nitrilotriacetic acid)$_n$-fluorochrome conjugates. *J. Am. Chem. Soc*. **2001**, *123*, 12123-12125.

(12) Razzano, V.; Paolino, M.; Reale, A.; Giuliani, G.; Artusi, R.; Caselli, G.; Visintin, M.; Makovec, F.; Donati, A.; Villafiorita-Monteleone, F.; Botta, C.; Cappelli, A. Development of imidazole-reactive molecules leading to a new aggregation-induced emission fluorophore based on the cinnamic scaffold. *ACS Omega*, **2017**, *2*, 5453-5459.

(13) Razzano, V.; Paolino, M.; Reale, A.; Giuliani, G.; Donati, A.; Giorgi, G.; Artusi, R.; Caselli, G.; Visintin, M.; Makovec, F.; Battiato, S.; Samperi, F.; Villafiorita-Monteleone, F.; Botta, C.; Cappelli, A. Poly-histidine grafting leading to fishbone-like architectures. *RSC Advances*, **2018**, *8*, 8638-8656.

(14) F. Bongiovì; G. Di Prima; F. S. Palumbo; M. Licciardi; G. Pitarresi; G. Giammona, Hyaluronic acid-based micelles as ocular platform to modulate the loading, release and corneal permeation of corticosteroids, *Macromol. Biosci*. **2017**, *17*, 1700261

(15) Levitt, M. H. Spin Dynamics: Basics of Nuclear Magnetic Resonance, John Wiley & Sons, 2001

(16) Vriezema, D. M.; Aragonès M. C.; Elemans, J. A. A. W.; Cornelissen, J. J. L.; Rowan, A. E.; Nolte, R. J. M. *Chem. Rev*. **2005**, *105*, 1445.

(17) Yoshizawa, M.; Klosterman, J. K.; Fujita, M. *Angew. Chem. Int. Ed*. **2009**, *48*, 3418.

TOC graphic