Site-Specific and Trigger-Activated Modification of Proteins by Means of Catalytic Hemin/G-quadruplex (hGQ) DNAzyme Nanostructures

Jordi F. Keijzer and Bauke Albada*

Laboratory of Organic Chemistry, Wageningen University & Research, Stippeneng 4, 6807 WE, Wageningen, the Netherlands.

ABSTRACT: Tyrosine residues of proteins can be modified using nanostructured DNA-based catalysts. One such catalyst is the complex between hemin and G-quadruplex DNA, efficiently catalyzes the modification of proteins with N-methyl luminol derivatives in the presence of hydrogen peroxide. Time-resolved analysis of the modification reaction mixture shows that final conversions are reached within 15–30 min, and LC-MS analysis of tryptic digests of the proteins shows that the reaction proceeds with chemoselectivity for electron-rich aromatic residues (Tyr and Trp), and that the site-specificity of the modification depends on the sequence and secondary structure folding of the G-quadruplex topology. The modification can be applied on proteins with different biomedical functions, and can be regulated by an external stimulus.

The hemin/G-quadruplex (hGQ) nanostructure is a DNA-based catalyst that can mimic reactions of peroxidase enzymes. The activity of these so-called hGQ DNAzymes is determined by the sequence and secondary structure formed by the layers of guanine tetrads. Further enhancement of the catalytic activity \((k_{cat})\) can be achieved by nucleotide supplements, or by conjugation of the hGQ DNAzyme to an aptamer sequence that binds to the substrate. Such so-called nucleoapzymes can be subjected to rational design or incorporation into supramolecular assemblies. The predictable formation of the catalytically active DNAzyme nanostructure has resulted in its incorporation into complex oligonucleotide assemblies of which the designed activity depends on an external trigger. Apart from the oxidation of chemical substrates in sensor-type setups, some conversions mimic biological processes such as the oxidation of dopamine to aminochrome or of \(N\)-hydroxy-L-arginine to nitric oxide and L-citrulline. We now establish hGQ DNAzymes as potent catalysts for site-specific oxidative protein modification.

Protein modification is ideally performed in a rapid, efficient and site-specific manner. The latest methods apply bio-orthogonal click-chemistry reactions are superior in rate and selectivity. Alternatively, synthetic catalysts or enzymes have been applied for the modification of native proteins or for the conversion of (a) genetically encoded handle(s). However, when it comes to the application of biomimetic catalytic species, only a few methods exist.

In order to test if the hGQ DNAzyme nanostructure would allow us to design catalyst that display variations in protein modification ability, and respond to external triggers, we used hGQ DNAzymes based on various topologies (Figure 1A) as catalysts for the oxidative modification of proteins using N-methyl luminol derivatives and hydrogen peroxide \((H_2O_2)\) (Figure 1B). We used lysozyme, thrombin, bovine serum albumin (BSA), and the therapeutically relevant immunoglobulin trastuzumab as a representative set that covers a large range of protein sizes, i.e. 14 kDa for lysozyme to 150 kDa for trastuzumab.

Lysozyme (14.3 kDa) contains three tyrosine (Tyr) residues with different solvent accessible area (SAA): Tyr23 (SAA: 32%), Tyr20 (SAA: 28%), Tyr53 (SAA: 15%). The presence of potentially competing aromatic amino acid residues tryptophan (Trp), phenylalanine, and histidine, allowed assessment of the chemoselectivity of the reaction.

![Figure 1. A: G-quadruplex topologies used to construct the DNAzymes in this study (for G6 and G8 only three G tetrads are shown). B: hemin/G-quadruplex (hGQ) DNAzyme catalyzed modification of a protein-based tyrosine residue with N-methyl luminol derivative I in the presence of H_2O_2.](image)

Much to our delight, LC-MS analysis of the reaction mixtures revealed substantial levels (32–96%) of lysozyme modification by various hGQ DNAzyme nanostructures (Table 1). Whereas only 2% of lysozyme modification was observed for hemin alone (Table 1) or unstructured ssDNA or dsDNA (see Supplementary Table S2), the presence of G-quadruplex structures led to higher amounts of modified protein. Notable differences were observed for different G-quadruplex topologies: intramolecular parallel GQs formed the most active complexes, followed by the intramolecular mixed/hybrid GQs and intermolecular parallel GQs, and with the intramolecular anti-parallel GQs generating the least active hGQ DNAzymes.

Time-resolved HPLC-analysis of the reaction mixtures revealed that the modification was nearly complete after 15 minutes (Supplementary Fig. S1).
Interestingly, conjugation of a lysozyme-binding aptamer (LBA) to G-quadruplex structures affected the ability of the hGQ DNAzyme to modify the protein (Supplementary Table S2). In general, higher H$_2$O$_2$ concentrations led to faster conversion, whereas higher NML and/or DNAzyme concentrations increased the number of modifications (Supplementary Table S3–S4). Furthermore, hGQ DNAzyme-induced modifications resulted in a decrease in the glycan-hydrolase activity (Supplementary Table S5).

The effect of G-quadruplex topology on the residues that were modified was studied by tryptic digestion in combination with LC-MS/MS analysis of several mixtures. As expected, the obtained fragments revealed that modification of Tyr was preferred over the modification of Trp. Singly modified lysozyme occurred on either Tyr23 or Tyr20 (Fig. 2A, Table 1), which are the most exposed residues with SAA of 32% and 28%, respectively. That site-specificity is not merely dictated by the amount of modification is apparent from the results obtained for G8 and PW17, which show very similar modification ability. Whereas the G8-based hGQ globally modifies lysozyme, the DNAzyme based on PW17 restricts its modification to one side of the protein (Figure 2A). Interestingly, when the PW17 sequence is conjugated to a lysozyme-binding aptamer (LBA) an additional modification of Tyr53 is observed (Supplementary Table S6), which shows the potential influence of an aptamer on the modification ability of hGQ DNAzymes. Importantly, LBA itself was not able to enhance the background activity of hemin.

Following these encouraging results, we subjected human alpha-thrombin (33.6 kDa) to the same hGQ DNAzyme sequences.

Table 1. Details of the hGQ DNAzyme catalysed modification of lysozyme or thrombin with N-methyl luminol derivative 1.

| code | GQ type | lysozyme modification* (%) | thrombin modification* (%) |
|------|---------|----------------------------|---------------------------|
|      |         | total 1 2 >2 | total 1 2 >2 |
| h    | hemin   | 2 2  - - | 5 5  - - |
|      | alone   |             |                          |
| G6   | interm. | 36 36 - - | 81 58 20 3 |
|      | parallel|             |                          |
| TBA  | intramol. | 35 35 - - | 29 25 4 - |
| HT   | antiparallel| 32 32 - - | 44 32 2 - |
| Bcl2 | intramol. | 72 62 10 - | 96 69 27 - |
| PW17 | mixed/hybrid| 74 65 8 1 | 97 41 39 17 |
| EA2  | intramol. | 91 70 18 3 | 96 41 40 15 |
| cMyc | parallel |             |                          |

*Conditions: 10 µM hemin, 10 µM DNA, 140 µM lysozyme, 700 µM NML-N$_2$ (1), and 2800 µM H$_2$O$_2$ (reaction time: 30 min). The effect of modification of Tyr was preferred over the modification of Trp. Interestingly, when the PW17 sequence is conjugated to a lysozyme-binding aptamer (LBA) an additional modification of Tyr53 is observed (Supplementary Table S6), which shows the potential influence of an aptamer on the modification ability of hGQ DNAzymes. Importantly, LBA itself was not able to enhance the background activity of hemin.

Following these encouraging results, we subjected human alpha-thrombin (33.6 kDa) to the same hGQ DNAzyme sequences.

Table 2. Details of the hGQ DNAzyme catalysed modification of lysozyme or thrombin with N-methyl luminol derivative 1.

| code | GQ type | lysozyme modification* (%) | thrombin modification* (%) |
|------|---------|----------------------------|---------------------------|
|      |         | total 1 2 >2 | total 1 2 >2 |
| h    | hemin   | 2 2  - - | 5 5  - - |
|      | alone   |             |                          |
| G6   | interm. | 36 36 - - | 81 58 20 3 |
|      | parallel|             |                          |
| TBA  | intramol. | 35 35 - - | 29 25 4 - |
| HT   | antiparallel| 32 32 - - | 44 32 2 - |
| Bcl2 | intramol. | 72 62 10 - | 96 69 27 - |
| PW17 | mixed/hybrid| 74 65 8 1 | 97 41 39 17 |
| EA2  | intramol. | 91 70 18 3 | 96 41 40 15 |
| cMyc | parallel |             |                          |

*Conditions: 10 µM hemin, 10 µM DNA, 140 µM lysozyme, 700 µM NML-N$_2$ (1), and 2800 µM H$_2$O$_2$ (reaction time: 30 min). The effect of modification of Tyr was preferred over the modification of Trp. Interestingly, when the PW17 sequence is conjugated to a lysozyme-binding aptamer (LBA) an additional modification of Tyr53 is observed (Supplementary Table S6), which shows the potential influence of an aptamer on the modification ability of hGQ DNAzymes. Importantly, LBA itself was not able to enhance the background activity of hemin.

Following these encouraging results, we subjected human alpha-thrombin (33.6 kDa) to the same hGQ DNAzyme sequences.

Table 3. Details of the hGQ DNAzyme catalysed modification of lysozyme or thrombin with N-methyl luminol derivative 1.

| code | GQ type | lysozyme modification* (%) | thrombin modification* (%) |
|------|---------|----------------------------|---------------------------|
|      |         | total 1 2 >2 | total 1 2 >2 |
| h    | hemin   | 2 2  - - | 5 5  - - |
|      | alone   |             |                          |
| G6   | interm. | 36 36 - - | 81 58 20 3 |
|      | parallel|             |                          |
| TBA  | intramol. | 35 35 - - | 29 25 4 - |
| HT   | antiparallel| 32 32 - - | 44 32 2 - |
| Bcl2 | intramol. | 72 62 10 - | 96 69 27 - |
| PW17 | mixed/hybrid| 74 65 8 1 | 97 41 39 17 |
| EA2  | intramol. | 91 70 18 3 | 96 41 40 15 |
| cMyc | parallel |             |                          |

*Conditions: 10 µM hemin, 10 µM DNA, 140 µM lysozyme, 700 µM NML-N$_2$ (1), and 2800 µM H$_2$O$_2$ (reaction time: 30 min). The effect of modification of Tyr was preferred over the modification of Trp. Interestingly, when the PW17 sequence is conjugated to a lysozyme-binding aptamer (LBA) an additional modification of Tyr53 is observed (Supplementary Table S6), which shows the potential influence of an aptamer on the modification ability of hGQ DNAzymes. Importantly, LBA itself was not able to enhance the background activity of hemin.

Following these encouraging results, we subjected human alpha-thrombin (33.6 kDa) to the same hGQ DNAzyme sequences.

Figure 2. A–B: Relevant sides of lysozyme (A) and human alpha-thrombin (B) with the position and solvent accessible area (SAA) of the residues that were modified with the respective hGQ DNAzymes (given between the square brackets; an ‘h’ indicates that this residue is also modified by hemin alone), or the Tyr/Trp residues in thrombin that have a high SAA but that are not modified (in red) [based on PDB-codes 3JIV (lysozyme) and SEW2 (thrombin)]. Modified residues are displayed in ball-and-stick, unmodified residues as sticks; active site residues are shown in green ball display. C: List of all Tyr and Trp residues in thrombin in decreasing SAA percentage, with the residues that are modified in bold (LC refers to the thrombin light chain, the other residues are on its heavy chain).
Thrombin contains various Tyr and Trp residues that would be available for modification judging from their SAA (Fig. 2B–C). As expected, hGQ DNAzymes displayed differences in their activity and site-specificity to modify thrombin, following comparable topology-related trends as was found for lysozyme (Supplementary Tables S7–S9). For both lysozyme and thrombin we found that crosslinking of the proteins did not occur in the absence of NML-derived 1 (Supplementary Fig. S2).

LC-MS-MS analysis of tryptic digests of the reaction mixtures with thrombin revealed that in the presence of any of the hGQ DNAzymes, NML-derivative 1 and H2O2, modification took place on Tyr88 (SAA: 24%) and Trp148 (SAA: 79%) (Table 1, and Supplementary Table S10). Depending on the G-quadruplex structure, additional modifications were detected (Fig. 2B). Interestingly, two of the more exposed Tyr residues were not modified: Tyr71 (SAA: 62%) and Tyr47 (SAA: 28%) (Fig. 2C). Tyr71 is located at the anion-binding exosite I of thrombin, which is also the known binding site of the G-quadruplex thrombin binding aptamer (TBA).20 Similarly, Tyr47 is located at the periphery of cationic exosite II, which is the binding site for thrombin binding aptamer HD22. Based on this we propose that the hGQ DNAzymes interact at these sites, thereby blocking modification of these specific residues.

Apparently, modification can be limited to only a few exposed residues. To examine if this also applied to larger proteins that potentially contain many more exposed reactive residues, we investigated modification of bovine serum albumin (BSA, 66 kDa) and the monoclonal therapeutic antibody trastuzumab (150 kDa). For these proteins we used SDS-PAGE analysis and visualized the modification using a two-step labelling approach in which NML-Ni-modified protein was derivatized by means of a strain-promoted alkyme-azide cycloaddition (SPAAC) reaction to a 4 kDa BCN-functionalized PEG unit. As expected, both proteins were modified in the presence of NML-Ni (1), H2O2, and hGQ DNAzymes (see Supplementary Figs. S3–S4). DNAzyme activities appear to be similarly related to the different topologies as was the case for lysozyme and thrombin. Specifically, whereas BSA was modified once by hemin alone,14 in the presence of GQ sequences higher numbers of modification were observed. The therapeutically relevant antibody trastuzumab was primarily modified on the heavy chain with up to three modifications for the most active hGQ DNAzymes (i.e. PW17 and cMyc) and higher concentrations of reagents (see Supplementary Figs. S5). As was observed for thrombin, the number of modifications decreased when the amount of DNAzyme was reduced 2- or 4-fold (i.e. from 0.9 equiv. with respect to the protein, to 0.45 and 0.225).

Now that we established that our protein modifying catalysts display features usually associated with enzymes (i.e. high rate, chemoselective, and site-specific), we designed a system that allowed regulation of the hGQ catalysed protein modification reaction by means of a switchable element (Fig. 3A). Upon addition of an ssDNA sequence that is complementary to the PW17 sequence (i.e. an OFF switch), a DNA duplex would be formed that does not have the ability to activate hemin. Indeed, we were able to switch the activity of the DNAzyme between its active ("ON") and inactive ("OFF") form by means of an external stimulus. Specifically, in the ON state, the DNAzyme modifies approximately 80% of lysozyme with NML derivative 1, whereas in the OFF state the modification conversion drops to ~5%, which is similar to hemin alone (see HPLC traces in the Supplementary Figs. S6–S11). Importantly, the hGQDNAzyme was formed again after addition of an activating ssDNA strand that was complementary to the ssDNA OFF strand. Since the activating strand contains a high number of guanine bases, we designed a strand that in itself does not form an active hGQ DNAzyme (see Supplementary Fig. S11). Indeed, the reformed hGQ DNAzyme complex regained its original protein modifying ability, also when a DNAzyme-aptamer conjugate was applied (Supplementary Figs. S6–S9). Using the larger and fluorescent lissamine-NML conjugate 2 the switchable character of the hGQ DNAzyme is visualized by SDS-PAGE analysis for both lysozyme and thrombin (Fig. 3B–C).

We note that with derivative 2 the modification efficiency was reduced; with the smaller derivative 1 conversions of 80% were achieved.

![Image](Image 300x356 to 767x659)

Figure 3. A: Design of the switchable hGQ DNAzyme system. B: SDS-PAGE analysis of the trigger-regulated modification of lysozyme (left) and thrombin (right) by means of a fluorescent N-methyl luminol derivative 2. C: Structure of lissamine-NML conjugate 2.

In conclusion, we describe how the hemin/G-quadruplex (hGQ) DNAzyme nanostructure can be used for the oxidative modification of Tyr residues, and to a lesser extent Trp residues, with N-methyl luminol derivatives in the presence of H2O2. In the absence of NML-tag, protein-protein crosslinking was not observed. We observed a correlation of the differences in protein modification and the GQ- folding conformation, where the parallel GQ sequences are more active than antiparallel GQs. Furthermore, we found preferences for specific residues that are modified by the different GQ topologies. This suggests differences in interaction between the various hGQ DNAzyme with the different target proteins, a process that can further be affected by the application of protein-binding aptamers. Lastly, we show that the catalytic nanostructure can be inactivated by the application of an external trigger, thereby lowering protein modification to the background activity of <5% that we observed for hemin alone. The observed hGQ DNAzyme-catalysed modification of proteins is novel, adding a yet unknown C–N bond forming reaction to the hGQ DNAzyme catalytic repertoire. Furthermore, the
rapid rate of modification, its chemoselectivity, site-specificity (which is potentially influenced by the presence of a protein-binding aptamer), and ability to respond to an external trigger, make this biomimetic protein modification process not that dissimilar from biological enzymatic processes. In view of the many GQ structures that can bind to proteins, the growing applications for DNA nanotechnology and the importance of modified proteins for many lines of research in many scientific disciplines, we expect that our approach will uncover novel catalysts with specific protein modification abilities.

ASSOCIATED CONTENT
Supporting Information describes the synthesis of the two NML derivatives, contains two full sets of conversion results for 22 hGQ DNAzymes for lysozyme and thrombin (including HPLC traces of all reaction mixtures), additional experiments that determine the effect of variables, and SDS-PAGE gel images of BSA and trastuzumab. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
bauke.albada@wur.nl

Author Contributions
All authors have given approval to the final version of the manuscript.

Funding Sources
Funded by the ECHO grant from the Dutch Organization for Scientific research (NWO) (project number 711.017.004).

Notes
Any additional relevant notes should be placed here.

ORCID:
Jordi F. Keijzer: 0000-0001-7124-7500
Bauke Albada: 0000-0003-3659-2434

ACKNOWLEDGMENT
We thank Jorick Bruins for supplying BCN-lissamine derivative that was used for lissamine-NML derivative 2, and Frank Claassen for his assistance on the LC-MS analysis.

ABBREVIATIONS
NML, N-methyl luminol; LBA, lysozyme-binding aptamer; (h)GQ, (hemin/)G-quadruplex; SAA, solvent accessible area.
Table of Contents artwork:
(a) Chidchob, P.; Sleiman, H. Recent advances in DNA nanotechnology. *Curr. Opin. Chem. Biol.* **2018**, *46*, 63–70. (b) Bujold, K.; Lacroix, A.; Sleiman, H. DNA nanostructures at the interface with biology. *Chem.* **2018**, *4*, 495–521.