

Total Chemical Synthesis of a Functionalized GFP Nanobody

Yara Huppelschoten^{+, [a, b]} Angela F. Elhebieshy^{+, [a]} Dharjath S. Hameed, ^[a] Aysegul Sapmaz, ^[a] Jens Buchardt, ^[b] Thomas E. Nielsen, ^[b] Huib Ovaa, ^[a] and Gerbrand J. van der Heden van Noort^{*[a]}

In memory of Huib: "Thinking outside the box is a challenging but rewarding exercise".

Chemical protein synthesis has proven to be a powerful tool to access homogenously modified proteins. The chemical synthesis of nanobodies (Nb) would create possibilities to design tailored Nbs with a range of chemical modifications such as tags, linkers, reporter groups, and subsequently, Nb-drug conjugates. Herein, we describe the total chemical synthesis of a 123 amino-acid Nb against GFP. A native chemical ligationdesulfurization strategy was successfully applied for the syn-

Introduction

Camelid species produce unique heavy-chain IgG antibodies existing of a single antigen binding variable heavy-chain domain (V_HH) only, also referred to as nanobodies (Nbs).^[1,2] These Nbs have unique properties such as their small size (~15 kDa), robustness, high solubility and monomeric nature, all properties that have inspired many researchers to explore them as exquisite research tools in structural-, cell- and developmental biology ever since their serendipitous discovery.^[3-8] In addition, Nbs are seen as promising new therapeutics due to their high affinity (nM range) for their targets, easy tissue penetration, and low immunogenicity.^[9,10] Accordingly, a lot of interest has been raised for the functionalization of Nbs for various applications such as diagnostic tools, Nb-drug conjugates, and bivalent Nb conjugates.^[11] Conventionally, Nbs are produced via protein expression which provides a functional Nb with ease, albeit with limited modification possibilities such as epitope tagging and other naturally existing mutations. Chemical synthesis of proteins, however, expands the degree of freedom of modification with

- [b] Y. Huppelschoten,⁺ Dr. J. Buchardt, Prof. Dr. T. E. Nielsen Global Research Technologies, Novo Nordisk Novo Nordisk Park, 2760 Måløv (Denmark)
- [⁺] These authors contributed equally to this work.
- Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202200304
- © 2022 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

thesis of this GFP Nb, modified with a propargyl (PA) moiety for on-demand functionalization. Biophysical characterization indicated that the synthetic GFP Nb-PA was correctly folded after internal disulfide bond formation. The synthetic Nb-PA was functionalized with a biotin or a sulfo-cyanine5 dye by copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC), resulting in two distinct probes used for functional in vitro validation in pull-down and confocal microscopy settings.

both natural and unnatural amino acids.^[12] The accessibility of synthetic proteins has increased thanks to modern native chemical ligation (NCL) and desulfurization methods.^[12-14] The ability of a Nb to retain target affinity may be drastically compromised as a result of unselective chemical labeling, as is often observed with conventional NHS- or maleimide modification reagents. We imagined that, the chemical synthesis of a Nb would accelerate the process of generating homogeneous Nbconjugates. Many functional groups suitable for chemoselective labeling would be easily introduced at defined, non-interfering regions of the Nb through a synthetic approach. The structure of Nbs is highly conserved, making it an attractive protein for generic chemical synthesis, thus paving the way for a modular synthetic approach that, with minor customization, could be broadly applied for a multitude of nanobodies. The general Nb structure comprises nine β -strands organized in a four- and a five-stranded β -sheet forming the conserved framework regions (FRs), connected via the complementarity determining region (CDR) loops and a conserved disulfide bond.^[15,16] The specificity for its target is obtained through the three CDRs at the ends of the variable domains. The long CDR3 loop provides the most significant contribution to the specificity and affinity of the Nb. As a proof-of-concept, we selected a nanobody against GFP (referred to hereafter as GFP Nb) aiming to validate a synthetic approach that could prove useful within multiple applications (Figure 1).^[17]

Considering the importance of the N-terminal clustering of the CDR loops in the affinity of Nbs towards their targets, we envisioned that incorporating a propargyl moiety at the Cterminus would be ideal for later modifications using copper mediated azide alkyne cycloaddition (CuAAC) chemistry which uses mild, near physiological reaction conditions.^[18] We here present a native chemical ligation-based synthesis for the generation of a functionalizable GFP Nb. On-demand conjugation of the synthetic Nb to either an affinity tag or a fluorescent

 [[]a] Y. Huppelschoten,⁺ A. F. Elhebieshy,⁺ Dr. D. S. Hameed, Dr. A. Sapmaz, Prof. Dr. H. Ovaa, Dr. G. J. van der Heden van Noort Oncode Institute and Dept. Cell and Chemical Biology Leiden University Medical Centre Einthovenweg 2, 2333 ZC, Leiden (The Netherlands) E-mail: gvanderheden@lumc.nl

ChemBioChem 2022, 23, e202200304 (1 of 6)





Figure 1. Structure of GFP Nb (PDB: 30GO). Indicated are the CDR domains in purple, the conserved disulfide bridge in yellow and the C-terminus as point of modification.

moiety results in easy access to Nb-conjugates in a divergent manner which can be used as tools in e.g. pull-down experiments and confocal microscopy.

Results and Discussion

Although the GFP Nb is relatively small, containing 123 amino acids, the β -sheet rich structure is known to increase the likelihood of aggregation, on-resin and in-solution, during Fmoc-based solid-phase peptide synthesis (SPPS).^[19] Initial investigations proved that a three-segment NCL approach was necessary for the synthesis of the GFP Nb.[20] The GFP Nb contains two native Cys residues, of which only one, however, is located at an appropriate potential ligation position (Cys₉₇) (Figure 2A). Therefore, NCL-desulfurization chemistry was chosen to assemble the Nb, as an Ala-to-Cys mutation could facilitate a NCL position that could be converted back to the native Ala using radical desulfurization post NCL. Accordingly, we envisioned the use of the acetamidomethyl (Acm) group to protect the other native Cys and prevent unwanted thioesterification or desulfurization during the construction of the GFP Nb. Our strategy for the synthesis of GFP Nb is outlined in



Figure 2. A. Sequence of GFP Nb, with underlined pseudo-proline dipeptides and iso-acyl dipeptides used in SPPS. B. Synthetic strategy for obtaining GFP Nb. C. UPLC analysis of the NCL of 1 and 2, the NCL of 3 and 4, purification of 5 and folding of 5. ' indicates MPAA thioesters, * indicates dibenzofulvene adduct, # indicates MPAA disulfide adduct.

ChemBioChem 2022, 23, e202200304 (2 of 6)



Figure 2, where we divided the polypeptide sequence into three fragments. Thioester fragment GFP Nb₁₋₄₉ (1), hydrazide fragment Cys-GFP Nb₅₀₋₉₆ (2), and Cys-GFP Nb₉₇₋₁₂₃ (4) were all prepared according to Fmoc-SPPS strategy on hydrazide or 2chlorotrityl resins.^[21] Peptide 1 and 2A were both prepared as hydrazides for subsequent (in-situ) activation and thiolysis. Peptide 1 was synthesized uneventfully with a final yield of 9% (Figure S1). During the initial attempt, the synthesis and purification of peptide 2 proved to be challenging. In order to increase yield and crude purity, four pseudo-proline building blocks were incorporated during SPPS (underscored in Figure 2A and Table S2). In addition, to overcome purification difficulties, the Fmoc was retained at the N-terminus of the peptide, increasing the retention time of the product significantly, compared to the capped deletion sequences formed during synthesis. This resulted in a significant increase in yield (3.6% yield in the initial attempt for 2 towards 13.2% yield for 2A) and enhanced purification efficiency (Figure S2). Furthermore, we experienced a difficult purification of peptide 4 using reversed phase-HPLC due to low solubility of the peptide most likely caused by the hydrophobic stretch at the N-terminal end of the sequence. The addition of a GT iso-acyl dipeptide to 4 (underscored in Figure 2A) increased the solubility and improved the purification process, resulting in pure peptide (7% yield) (Figure S3).

In 2020, Kar et al.^[22] described a one-pot Fmoc deprotection and NCL strategy which provided an excellent strategy for the first NCL of peptide 1 with peptide 2A, via *in-situ* preparation of 2. Peptide 2A was first subjected to piperidine treatment at pH 10.7 to furnish the deprotected peptide 2 within 10 minutes (Figure S4). After pH adjustment to pH 7.0, peptide thioester 1 was added to the reaction mixture containing 2 to afford ligated product 3 A after 16 hours at 37 °C (Figure 2C). Due to the low reactivity of the Val thioester, 4-mercaptophenylacetic acid (MPAA)^[23] was used as additive in this ligation and afterwards removed using a 3 kDa molecular weight cutoff spin filter, monitoring the MPAA removal by LC–MS/UV chromatography (Figure S5). Subsequently, the Cys residue was desulfurized by applying radical desulfurization conditions (TCEP, VA-044 and GSH) leading to **3B** after 16 hours (Figure S6).^[14]

The integrity of the propargyl moiety on peptide 4 is not compatible with Pd chemistry, therefore we opted to conduct the Acm removal in 3B before the final ligation step between peptide **3B** and **4**. We envisioned that the use of PdCl₂ for the Acm removal of protected Cys₂₃ in **3B** would be a good option due to its compatibility with the denaturing 6 M Gdn buffer.^[24] All chemicals from the desulfurization reaction were removed using a 3 kDa cutoff spin filter before the addition of PdCl₂ and the reaction was completed within 1 hour and subsequently quenched with DTT, prior to purification of the peptide using size-exclusion chromatography (SEC), to obtain product 3 in a 54% (Figure S5). We used SEC as an alternative purification method using the high chaotropic NCL buffer, 6 M Gdn buffer, as the purification of fragment 3 proved to be troublesome using standard reversed-phase HPLC due to its hydrophobic nature. Before we continued with the second ligation, peptide 4A was incubated at pH 7-8 for 10 minutes to obtain the native peptide bond after an $O \rightarrow N$ acyl shift (Figure S7). Followed by the one-pot thioesterification of **3** and ligation to **4**, final product **5** was obtained in 36% yield (Figure 2C, Figure S8–9).

With the full-length GFP Nb₁₋₁₂₃-linker-PA in hand, we continued to the folding step including disulfide bond formation. The folding was carried out by stepwise dialysis in phosphate buffered saline (PBS). As spontaneous disulfide formation did not occur within 48 hours in PBS we added 2,2'-Dithio-dipyridin (DTP),^[25,26] a known disulfide bond formation accelerator. Indeed, after the addition of 1 mM DTP to the folding buffer, initiation of disulfide bond formation was observed after 1 hour and completed within 16 hours as was clearly observed in the high-resolution mass spectrum by the loss of 2 Da (Figure 3B).

To confirm the proper folding of the chemically synthesized and folded GFP Nb (5), we conducted circular dichroism (CD) experiments. The synthetic GFP Nb exhibited absorptions of βsheet structures similar to the recombinantly expressed GFP Nb,^[17] indicating that their folding is comparable (Figure 3C). In addition, biolayer interferometry (BLI) experiments were performed to determine the affinity of the synthetic GFP Nb 5 to its target protein GFP and compared to that of the recombinantly expressed GFP Nb. The C-terminal His-tag on both the expressed and synthetic Nb was used to immobilize the Nbs on Ni-NTA biosensor tips and untagged GFP was used as the analyte at concentrations ranging from 1 nM to 100 nM. With this setup, the expressed and synthetic GFP Nb showed similar binding profiles (Figure 3D) and affinities of 1.12 and 1.11 nM, respectively. These values are in agreement with previously reported data in literature (1.4 nM).^[17]

Next, we wanted to show the adaptability of the fully synthetic Nb as a chemical tool by modifying the C-terminal propargyl moiety using a bio-orthogonal labeling strategy. Accordingly, we used copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) chemistry to functionalize the synthetic Nb with an azide-functionalized biotin molecule for the purpose of pull-down experiments (Figure 4A). Unfolded, purified 5 was reacted with biotin-azide using mild reaction conditions (3 mM CuSO₄, 10 mM sodium ascorbate, and 2 mM tris-(hydroxypropyltriazolylmethyl)amine (THPTA)) to form the GFP Nb-biotin conjugate (6) (Figure S11). After the CuAAC, the synthetic Nb was folded as described previously for compound 5, removing all additives from the CuAAC reaction during the dialysis step. In addition, we envisioned that modification of 5 with sulfo-Cyanine5-azide (Cy5) would lead to the opportunity to validate the proper functioning and target binding of our Nb by co-localization of the Nb with GFP tagged proteins in cells using confocal microscopy. Hence, we used the same procedure to synthesize a GFP Nb-Cy5 conjugate (7) (Figure S12). Next, we measured CD to warrant correct folding of the functionalized Nbs (Figure 4B) and continued with CD denaturing experiments to investigate the stability of the synthetic Nb-conjugates compared to the expressed GFP Nb. Both Nb-conjugates showed a similar denaturing pattern as the expressed GFP Nb indicating that introduction of C-terminal cargo onto the Nb did not alter its biophysical properties (Figures S14-S16).

ChemBioChem 2022, 23, e202200304 (3 of 6)





Figure 3. A. Deconvoluted mass spectrum and ESI Mass spectrum (inset) of unfolded 5. B. Deconvoluted mass spectrum and ESI Mass spectrum (inset) of folded 5. C. Circular dichroism spectra comparing recombinant and synthetic GFP Nb. D. Bio Layer Interferometry analysis of folded 5 in comparison to expressed GFP Nb.



Figure 4. A. Functionalization of synthetic Nb 5. Nanobody structure PDB ID: 30GO.^[17] B. circular dichroism of expressed GFP Nb, synthetic GFP Nb conjugates 6 and 7.

We wanted to test Nb conjugates **6** and **7** in *in vitro* assays for recognition of GFP labeled proteins. First, we set out to validate the binding of the synthetic GFP Nb to GFP-fusion proteins in a complex protein mixture by performing a pulldown assay. For this purpose, we decided to use the MelJuSo cell line established in our lab^[27] that stably expresses GFPtagged small GTPase Rab7 that is a central regulator of membrane trafficking in multiple directions.^[28] Hence, we incubated the GFP Nb-biotin conjugate (**6**) with MelJuSo cell lysate expressing GFP-Rab7 to perform a pull-down assay. As negative control to ensure GFP-specific binding, we used WT MelJuSo cells that did not express GFP-Rab7. After two hours of incubating conjugate **6** with the cell lysates of both GFP-Rab7 MelJuSo or WT MelJuSo, we were able to selectively pull-down the GFP-Rab7 protein from the cell lysate showing a signal around 55 kDa, equal to the molecular weight of GFP-Rab7 (Figure 5A). Negligible background signal confirms the selectivity of the synthetic GFP Nb for GFP over other proteins present in the cell lysate (Figure S18).

Next, we continued with confocal microscopy validation, and used GFP Nb-Cy5 conjugate **7** to examine co-localization with GFP tagged proteins located in different cellular compartments (cytoskeleton, endoplasmic reticulum (ER), peroxisomes, endosomes, mitochondria and Golgi). MelJuSo cells were hence transfected with six different plasmids encoding for GFP tagged Actin, VAMP-Associated Protein A (VAPA), Peroxisomal bio-





Figure 5. A. Western blot analysis of the pull-down of GFP-Rab7 from cell lysate using GFP Nb-biotin conjugate 6. The signal around 25 kDa is GFP as a result of protein degradation. B. Confocal images of MelJuSo cells expressing GFP-Rab7 in the presence of GFP Nb-Cy5 7. C. Illustration of a cell highlighting the cellular compartments visualized in B.

genesis factor 3 (PEX3), Rab7, Protein tyrosine phosphatase interacting protein 51 (PTPIP51) and Oxysterol binding protein (OSBP-PH) proteins (Figure 5B and C). The cells were fixed and the membranes were permeabilized before incubation with the conjugate **7** for 1 hour. The incubation of the transfected MelJuSo cells with **7** resulted in a complete overlap between the GFP signal and the Cy5 signal, indicating the co-localization of the synthetic Cy5-Nb with the GFP tagged proteins (Figure 5B). The staining with conjugate **7** resulted in a strong signal in each of the tested cell compartments with minimal background, indicating full target engagement of our synthetic Nb (Figure 5B). The ability to visualize GFP fusion proteins at various cellular locations further efficiently showcases the broad applicability of the synthetic Nb.

Conclusion

In conclusion, we have developed a practical native chemical ligation-based synthetic approach for the generation of a synthetic GFP Nb ready for on-demand functionalization. With this method, we can obtain homogenous batches of labeled Nb by selectively labeling of the Nb using CuAAC without altering the properties of the Nb, such as folding or thermo-stability. This technology was successfully applied to modify the GFP-Nb with both a biotin or a fluorophore which were used in pull-down and confocal microscopy experiments, respectively. The Nb labeling is performed without compromising the antigenbinding site, making this method also applicable for Nbs containing a Cys residue in their CDR domains. In addition, easy



modification of the CDR3 domain can be obtained because it is introduced in one of the final synthesis steps. We envision that with this protocol in hand, Nbs against other targets might be synthesized, using a similar strategy, due to the high sequence and structure similarities between Nbs. This methodology could potentially also be applied to the streamlined preparation of Nb drug-conjugates or (heterogeneous) Nb multimers. Moreover, unnatural amino acids can be easily introduced through the SPPS protocol, e.g. to confer stability of the Nb against degradation *in vivo* and improve the applicability of this protein scaffold.^[29,30]

Experimental Section

The data that support the findings of this study are available in the Supporting Information of this article.

Acknowledgements

We would like to thank Dr. Birol Cabukusta at the LUMC for plasmids used for confocal microscopy experiments. Y. H. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 765445 (ITN UbiCODE), A. F. E. acknowledges funding by ICI (grant No ICl00013) and G. J. vdH. vN. acknowledges funding by NWO (VIDI grant VI.192.011).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: chemical protein synthesis · click chemistry · nanobodies · protein modifications · solid phase peptide synthesis

- M. M. Harmsen, H. J. De Haard, Appl. Microbiol. Biotechnol. 2007, 77, 13– 22.
- [2] M. Dumoulin, K. Conrath, A. Van Meirhaeghe, F. Meersman, K. Heremans, L. G. J. Frenken, S. Muyldermans, L. Wyns, A. Matagne, *Protein Sci.* 2009, 11, 500–515.
- [3] H. D. Herce, D. Schumacher, A. F. L. Schneider, A. K. Ludwig, F. A. Mann, M. Fillies, M. A. Kasper, S. Reinke, E. Krause, H. Leonhardt, M. C. Cardoso, C. P. R. Hackenberger, *Nat. Chem.* 2017, *9*, 762–771.

- [4] A. Manglik, B. K. Kobilka, J. Steyaert, Annu. Rev. Pharmacol. Toxicol. 2017, 57, 19–37.
- [5] R. W. Cheloha, T. J. Harmand, C. Wijne, T. U. Schwartz, H. L. Ploegh, J. Biol. Chem. 2020, 295, 15307–15327.
- [6] G. Bao, M. Tang, J. Zhao, X. Zhu, *EJNMMI Res.* 2021, 11, https://doi.org/ 10.1186/s13550-021-00750-5.
- [7] J. Wesolowski, V. Alzogaray, J. Reyelt, M. Unger, K. Juarez, M. Urrutia, A. Cauerhff, W. Danquah, B. Rissiek, F. Scheuplein, N. Schwarz, S. Adriouch, O. Boyer, M. Seman, A. Licea, D. V. Serreze, F. A. Goldbaum, F. Haag, F. Koch-Nolte, *Med. Microbiol. Immunol.* **2009**, *198*, 157–174.
- [8] H. Zhang, Y. Han, Y. Yang, F. Lin, K. Li, L. Kong, H. Liu, Y. Dang, J. Lin, P. R. Chen, J. Am. Chem. Soc. 2021, 143, 16377–16382.
- [9] C. Ackaert, N. Smiejkowska, C. Xavier, Y. G. J. Sterckx, S. Denies, B. Stijlemans, Y. Elkrim, N. Devoogdt, V. Caveliers, T. Lahoutte, S. Muyldermans, K. Breckpot, M. Keyaerts, *Front. Immunol.* **2021**, *12*, 632687.
- [10] C. Vincke, R. Loris, D. Saerens, S. Martinez-Rodriguez, S. Muyldermans, K. Conrath, J. Biol. Chem. 2009, 284, 3273–3284.
- [11] E. Y. Yang, K. Shah, Front. Oncol. 2020, 10, 1182.
- [12] Y. Tan, H. Wu, T. Wei, X. Li, J. Am. Chem. Soc. 2020, 20288–20298.
- [13] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, Science 1994, 266, 776–779.
- [14] Q. Wan, S. J. Danishefsky, Angew. Chem. Int. Ed. 2007, 46, 9248–9252; Angew. Chem. 2007, 119, 9408–9412.
- [15] S. Muyldermans, Annu. Rev. Biochem. 2013, 82, 775-797.
- [16] L. S. Mitchell, L. J. Colwell, Proteins Struct. Funct. Bioinf. 2018, 86, 697– 706.
- [17] M. H. Kubala, O. Kovtun, K. Alexandrov, B. M. Collins, Protein Sci. 2010, 19, 2389–2401.
- [18] C. W. Tornøe, C. Christensen, M. Meldal, J. Org. Chem. 2002, 67, 3057– 3064.
- [19] M. Paradís-Bas, J. Tulla-Puche, F. Albericio, Chem. Soc. Rev. 2016, 45, 631–654.
- [20] L. Hartmann, T. Botzanowski, M. Galibert, M. Jullian, E. Chabrol, G. Zeder-Lutz, V. Kugler, J. Stojko, J. M. Strub, G. Ferry, L. Frankiewicz, K. Puget, R. Wagner, S. Cianférani, J. A. Boutin, *Protein Sci.* 2019, 28, 1865–1879.
- [21] G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. Bin Li, Y. Lin, H. K. Cui, L. Liu, Angew. Chem. Int. Ed. 2011, 50, 7645–7649; Angew. Chem. 2011, 123, 7787–7791.
- [22] A. Kar, J. Mannuthodikayil, S. Singh, A. Biswas, P. Dubey, A. Das, K. Mandal, Angew. Chem. Int. Ed. 2020, 59, 14796–14801; Angew. Chem. 2020, 132, 14906–14911.
- [23] T. M. Hackeng, J. H. Griffin, P. E. Dawson, Proc. Natl. Acad. Sci. USA 1999, 96, 10068–10073.
- [24] S. K. Maity, M. Jbara, S. Laps, A. Brik, Angew. Chem. Int. Ed. 2016, 55, 8108–8112; Angew. Chem. 2016, 128, 8240–8244.
- [25] J. Giribaldi, D. Wilson, A. Nicke, Y. El Hamdaoui, G. Laconde, A. Faucherre, H. M. O. Maati, N. L. Daly, C. Enjalbal, S. Dutertre, *Toxin Rev.* 2018, 10, 222.
- [26] K. Maruyama, H. Nagasawa, A. Suzuki, *Peptides* **1999**, *20*, 881–884.
- [27] A. Sapmaz, I. Berlin, E. Bos, R. H. Wijdeven, H. Janssen, R. Konietzny, J. J. Akkermans, A. E. Erson-Bensan, R. I. Koning, B. M. Kessler, J. Neefjes, H. Ovaa, *Nat. Commun.* **2019**, *10*, 1454.
- [28] F. Guerra, C. Bucci, Cell Sci. 2016, 5, 34-62.
- [29] V. De Filippis, F. De Antoni, M. Frigo, P. P. De Laureto, A. Fontana, Biochem. 1998, 37, 1686–1696.
- [30] W. Cabri, P. Cantelmi, D. Corbisiero, T. Fantoni, L. Ferrazzano, G. Martelli, A. Mattellone, A. Tolomelli, *Front. Mol. Biosci.* 2021, 8, 697586.

Manuscript received: June 7, 2022 Accepted manuscript online: August 3, 2022 Version of record online: August 23, 2022