HaloTag Engineering for Enhanced Fluorogenicity and Kinetics with a Styrylpyridium Dye

Carla Miró-Vinyals[+],[a, b] Alina Stein[+],[a] Sandro Fischer,[a] Thomas R. Ward,[a] and Alexandria Deliz Liang*[a, c]

HaloTag is a small self-labeling protein that is frequently used for creating fluorescent reporters in living cells. The small-molecule dyes used with HaloTag are almost exclusively based on rhodamine scaffolds, which are often expensive and challenging to synthesize. Herein, we report the engineering of HaloTag for use with a chemically accessible, inexpensive fluorophore based on the dimethylamino-styrylpyridium dye. Through directed evolution, the maximum fluorogenicity and the apparent second-order bioconjugation rate constants could be improved up to 4-fold and 42-fold, respectively. One of the top variants, **HT-SP5**, enabled reliable imaging in mammalian cells, with a 113-fold fluorescence enhancement over the parent protein. Additionally, crystallographic characterization of selected mutants suggests the chemical origin of the fluorescent enhancement. The improved dye system offers a valuable tool for imaging and illustrates the viability of engineering self-labeling proteins for alternative fluorophores.

Fluorescent reporters are a critical tool for examining cellular structure and processes. HaloTag is a self-labeling, non-fluorescent protein that can be covalently and selectively tagged with a fluorescent dye (Figure 1).[1,2] This reactivity was engineered by extensive mutation of a dehalogenase enzyme to yield HaloTag version 7 (hereafter HaloTag7), which reacts particularly well with rhodamine dyes.[2] Since then, tremendous effort has been placed on chemical modification of the fluorescent, small molecules to improve the photophysical properties of the dye-HaloTag7 constructs.[3,4] Within this context, the repertoire of small molecules used with HaloTag is still mostly limited to various rhodamine dyes.[3–9] Although these dyes are bright and bind rapidly to HaloTag7, they are also expensive, and their synthesis requires a sophisticated command of synthetic chemistry. Moreover, focusing on a single class of dyes limits the diversity of chemistry available to users. Different dyes that are also less synthetically demanding exist.[10–12] However, the HaloTag-fluorophore constructs that are formed with these simpler dyes often bind slowly to HaloTag, exhibit weak fluorescent enhancement upon HaloTag binding, or are not suitable for imaging in mammalian cells. An example is **F1** (Figure 1), which contains a dimethylamino-styrylpyridium core functionalized with a chloroalkane chain. Previous reports identified **F1** from a library of dimethylamino-styrylpyridium compounds functionalized with chloroalkane chains of varying lengths.[12] The authors report that **F1** exhibits a fluorescent enhancement of 27-fold when mixed with purified HaloTag7, indicating that **F1**-HaloTag7 is fluorogenic.[12] However, the accumulation of **F1** in the mitochondria prevented its use in mammalian cells.

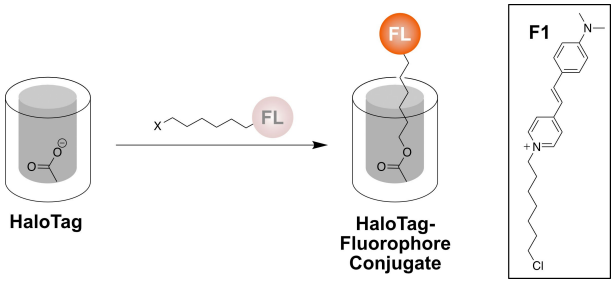

Figure 1. Bioconjugation of HaloTag with haloalkanes, such as **F1**.

Point mutations of HaloTag7 have shown improvements in the properties of the dye-HaloTag7 constructs.[10,13,14] Thus, we posited that the properties of **F1**-HaloTag7 might be improved by enzyme engineering. Such a construct would be an alternative to the typical rhodamine dyes and elucidate the potential fluorogenicity of HaloTag variants with other small molecules. To this end, we engineered several variants of HaloTag7 that rapidly bind and enhance the fluorogenicity with **F1**. The work herein demonstrates that mutations of HaloTag7 significantly improved the performance of the dye-HaloTag construct and allowed for reliable mammalian cell imaging.

We first characterized the interaction between HaloTag7 and **F1** (Figures 2 and S1). We found this interaction to be very similar to the previously reported characteristics.[12] To improve

[a] C. Miró-Vinyals,[+] A. Stein,[+] S. Fischer, Prof. Dr. T. R. Ward, Prof. Dr. A. Deliz Liang
Chemistry Department, University of Basel
NCCR Molecular Systems Engineering, 4002 Basel (Switzerland)
E-mail: alexandriadeliz.liang@uzh.ch
[b] C. Miró-Vinyals[+]
Institute of Chemical Science and Engineering
EPF Lausanne, CH C2 425, Station 6, 1015 Lausanne (Switzerland)
[c] Prof. Dr. A. Deliz Liang
Department of Chemistry
University of Zurich, Winterthurerstrasse 190, 8057 Zurich (Switzerland)
[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202100424

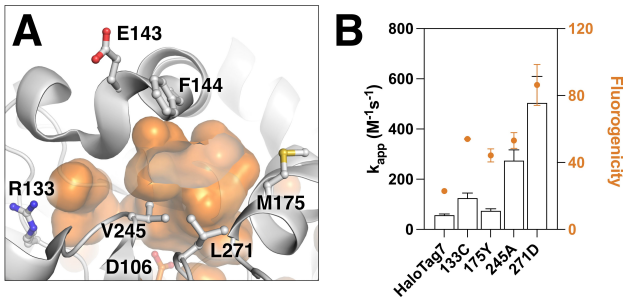

Figure 2. Single mutant libraries. A) The mutation sites were selected based on previous crystallography data (PDB: 5VNP).[15] The aspartic acid (D106) involved in bioconjugation and the residues mutated in this work are labeled and illustrated as sticks. The void for haloalkane binding is shown in orange. B) The apparent rate constants and fluorescence enhancement (λ_{ex} = 535 nm and λ_{em} = 610 nm) upon mixing 0.6–3.0 μM **F1** with 3.0 μM of selected single-mutant HaloTag7 variants. For each variant, the data are a composite of at least four biological replicates. The error bars represent the standard deviation from the mean.

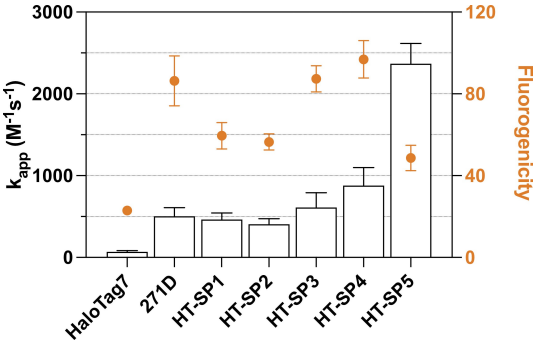

Figure 3. Top mutants from the second library. The apparent rate constants and fluorescence enhancement (λ_{ex} = 535 nm and λ_{em} = 610 nm) upon mixing 0.6–3.0 μM **F1** with 3.0 μM of HaloTag7, 271D, and variants from the second library. For each variant, the data are a composite of a minimum of three biological replicates. The error bars represent the standard deviation from the mean.

the properties of **F1**-HaloTag7, six positions within HaloTag7 were identified for mutagenesis based on crystallographic data (Figure 2A): E133, E143, F144, M175, V245, and L271.[10,11,15] Site-saturation mutagenesis at these six positions was conducted to yield an initial library of single mutants. These mutants were expressed and purified for functional screening. The mutants were not screened directly in *E. coli* cells to eliminate bias resulting from protein expression profiles in *E. coli*. The resulting protein libraries were screened for two factors, rate of fluorescence enhancement and maximum fluorogenicity.

Within this single mutant library, four mutations were particularly beneficial (Figures 2 and S2): 133C, 175Y, 245A, and 271D. The apparent second-order kinetics and the maximum fluorogenicity of these variants were determined (Figure 2B). The maximum fluorogenicity was improved for all four of the top single-mutant variants. The fluorescence response of mutant 245A was red-shifted from that of HaloTag7 and the other mutants (Figure S3). Mutation 271D demonstrated the most rapid and largest fluorogenic response upon binding to **F1**. Based on previous crystallographic data, we theorized that residue 271D might engage in an electrostatic interaction with the pyridinium of **F1**.

From the first DNA library, a recombination library was created comprising the top hits and some non-detrimental mutations at positions E143 and F144. This library contained double, triple, quadruple, and quintuple mutants. Screening of the second library was conducted on purified proteins, as reported for the first library. From this second round of engineering, several additional hits were identified: R133C/M175Y/V245A (**HT-SP1**), R133C/E143M/F144H/M175Y/V245A (**HT-SP2**), E143M/F144H/L271D (**HT-SP3**), M175Y/L271D (**HT-SP4**), and M175Y/V245A/L271D (**HT-SP5**). The bioconjugation kinetics and photophysical properties of these variants were characterized (Figure 3 and Table S4).

None of the variants from the second round were improved for maximum fluorescence enhancement over 271D, but several mutants displayed improved kinetics. Most of the improved variants contained mutation 271D (Figure S4), which seemed to increase the apparent second-order rate constant most dramatically. Interestingly, recombination of the most improved mutants, R133C/M175Y/V245A/L271D, resulted in complete loss of fluorogenicity (Figure S4). Additionally, two mutations that were not beneficial in the first rounds were found to be beneficial when combined with 271D: 143M and 144H (**HT-SP3**). **HT-SP5** was the fastest binding variant with an apparent second-order rate constant of $2.4 \cdot 10^3\ M^{-1}s^{-1}$. These results correspond to a 42-fold improvement in the rate constant. Nonetheless, the overall second-order rate constant for bioconjugation of **F1** with **HT-SP5** is lower than that for rhodamine dyes with HaloTag or HaloTag7 ($2.7 \cdot 10^6\ M^{-1}s^{-1}$ and $1.9 \cdot 10^7\ M^{-1}s^{-1}$, respectively).[1,2]

Based on the positive results with purified protein, we examined **F1** with the HaloTag7 variants expressed in *E. coli*. Each variant was expressed in BL21(DE3) cells under a T7 promoter. The cells were incubated with **F1**, and the fluorescence in each culture was monitored as a function of time (Figures 4 and S5). The maximum fluorescence

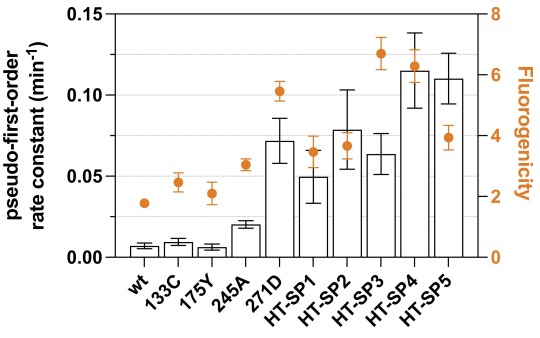

Figure 4. Whole-cell studies with *E. coli* BL21(DE3). To each *E. coli* culture, **F1** was added to a final concentration of 3.0 μM. The fluorescence was monitored as a function of time (λ_{ex} = 535 nm and λ_{em} = 610 nm). The pseudo-first-order rate constants were determined by a single exponential fit of the data. The studies were repeated in at least six biological replicates. The error bars represent the standard deviation from the mean.

enhancement for HaloTag7 was 1.77 ± 0.13. All of the mutants tested were improved over HaloTag7. The top variants were 271D, **HT-SP3**, and **HT-SP4**, which exhibited fluorescent enhancements (and standard deviation) of 5.46 ± 0.32, 6.70 ± 0.53, and 6.28 ± 0.54, respectively.

The kinetic traces were modeled to pseudo-first-order kinetics. Comparison between the apparent second-order rate constants in purified protein and the pseudo-first-order rate constants in *E. coli* displayed a similar trend. The notable difference was **HT-SP5**, which is one of the best mutants in the study with purified protein. **HT-SP5** expresses in *E. coli* much worse than HaloTag7 and other mutants (approximately 8-fold lower). The low expression of **HT-SP5** results in reduced concentration of the variant, which may prevent this experiment from being in the pseudo-first order range. This factor would explain why the apparent second-order rate constant appears lower in this *E. coli*-based assay.

Based on the dramatic improvements in both reaction rate and fluorescence enhancement, we tested HaloTag7 and two mutants-**HT-SP4** (the highest fluorogenicity) and **HT-SP5** (the fastest mutant for bioconjugation)-in HeLa cells (Figure 5). The results were compared with the commercial rhodamine-based dye Janelia Fluor® 549 HaloTag® (hereafter **JF549**, Promega, Scheme S1). Previous work and chemical intuition would suggest that the cationic **F1** might accumulate in the mitochondria.[12] However, we were hopeful that the improved reaction rate and fluorogenicity with the HaloTag7 variants might make this system suitable for imaging in other cellular locations. Both HaloTag7 and the variants were tagged with a localization target for actin (Lifeact).[16] The cells transfected with HaloTag7, **HT-SP4**, and **HT-SP5** were compared upon incubation with **JF549** (Figures 5 and S7) or **F1** (Figures 5 and S8). For all **F1** experiments, a washing step was applied prior to imaging.

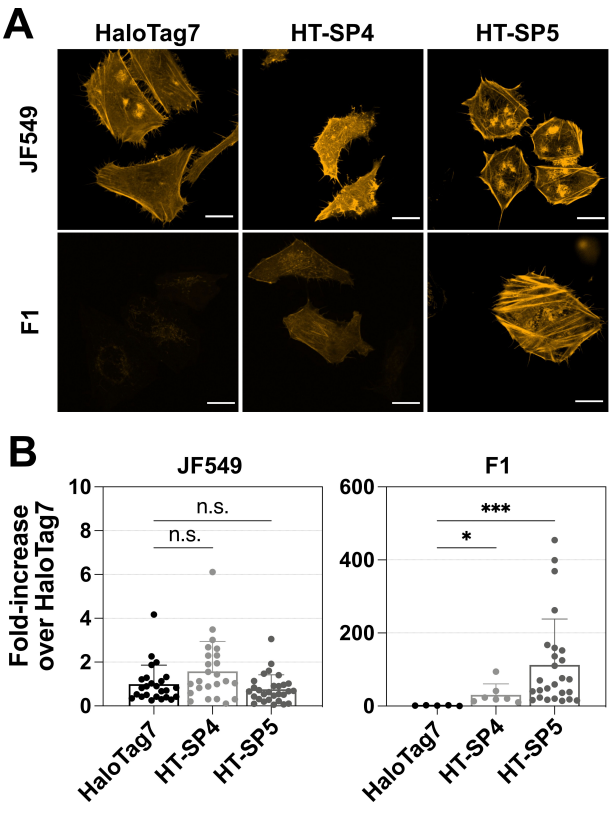

Figure 5. Live cell imaging with HaloTag7 variants conjugated to **F1** or **JF549** in HeLa cells. A) HeLa cells were transiently transfected with HaloTag7, **HT-SP4**, or **HT-SP5**. The cells were incubated with **JF549** (200 nM, 15 min incubation, no wash; top row) or **F1** (500 nM, 30 min, 16 h wash; bottom row). Scale bar = 20 μm. λ_{exc} = 561 nm, λ_{em} = 600/52 nm. B) Comparison of the fluorescence intensity of cells transiently expressing HaloTag7, **HT-SP4**, or **HT-SP5** incubated with the commercial dye **JF549** or with **F1**. The data were normalized to the mean of HaloTag7. *** = 0.0001, * = 0.0341.

The fluorescence was high for **JF549** with HaloTag7 and the two variants. No significant difference in the fluorescence was observed for the different protein conjugates with **JF549** (Figure 5B, left). These results indicate that all of these constructs are capable of stabilizing the zwitterionic form of the rhodamine dye. The minimal of change in the fluorescence response is not surprising, given that the linker length of **JF549** is much longer than **F1** (Scheme S1).[17] In contrast, dramatically different fluorogenicities are observed for the HaloTag7 variants conjugated to **F1** (Figures 5 and S8).

As expected with HaloTag7, **F1** accumulates significantly in mitochondria and displays a weak fluorescent signal, excluding this construct as a viable fluorescent reporter (Figures 5 and S8). **HT-SP4**, the most fluorogenic variant in purified protein, exhibited a 31-fold fluorescence enhancement over HaloTag7 (Figure 5B). Although actin labeling can be observed with **HT-SP4**, the signal is weak (Figures 5A and S8). **HT-SP5**, the most rapidly binding variant, was 113-fold more fluorogenic with **F1** than the parent HaloTag7 (Figure 5B). This strong fluorogenicity is particularly remarkable given that, unlike rhodamine-based dyes, **F1** does not undergo a spirocyclization equilibrium that can be exploited to induce large fluorogenicities. For **HT-SP5**, cell images could be obtained that showed clear actin fibers (Figure 5A). These results suggest that the rate of bioconjugation may be the primary factor for improvement of cell imaging. Notably, it was still necessary to wash the cells with **HT-SP5** and **F1** (Figure S9).

To elucidate the chemical origin of the improved fluorescence response, we attempted to obtain crystallographic data of the variants bound to **F1**. Crystallographic data with well-resolved **F1** were obtained for HaloTag7 and **HT-SP2** (Figure 6). In the crystal structure of **F1**-HaloTag7, the dimethylamino-styrylpyridium core is loosely packed by the F144, A145, T148, T172, and M175. In contrast, mutations in **HT-SP2** reposition the dimethylamino-styrylpyridium core within the wide opening of HaloTag. In **HT-SP2**, the primary packing interactions with **F1** are formed by W141, P142, 144H, A145, P243, 245A, and L271. Residue 144H of **F1**-**HT-SP2** adopts two conformations. Notably, in both structures, no hydrogen bonding or electrostatic interactions were observed between the protein and **F1**. These structural data highlight that HaloTag7 can be mutated to favor large changes in the orientation of **F1**, underscoring the plasticity of the scaffold.

Although a crystal structure could not be obtained for **F1**-**HT-SP5**, chemical intuition suggests that an electrostatic interaction could be formed between 271D and the pyridinium

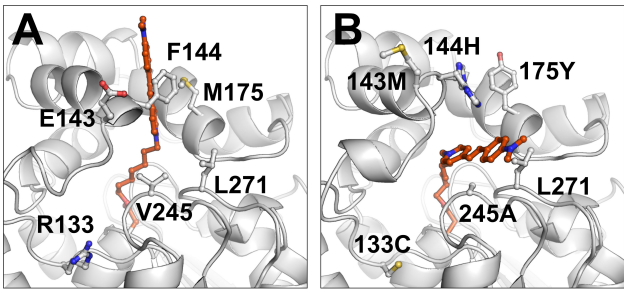

Figure 6. Crystallographic data. Crystal structures were obtained for **F1**-HaloTag7 (A, PDB: 7OO4) and **F1**-**HT-SP2** (B, PDB: 7OND). In each panel, **F1** (orange) and the residues target in this work (light grey) are shown as sticks.

of **F1**. This interaction may be key in improving two factors: i) rate of bioconjugation and ii) the rigidity of the dimethylamino-styrylpyridium, which improves intramolecular charge transfer necessary for fluorescence with this small molecule.

The work herein presents an evolved protein for fluorescent labeling in mammalian cells with a chemically accessible dye. Additionally, we demonstrate that an easily synthesized probe can be made useful through the careful engineering of HaloTag. Although much work in the field has focused on fluorescence optimization through synthetic chemistry,[3,4] this work suggests that an equally suitable route is the genetic optimization of the scaffold proteins (such as HaloTag7) for use with alternative fluorophores beyond the rhodamine family. We expect that this strategy could be extended to additional fluorophore families enabling the use of a wider range of tools for live-cell imaging.

Acknowledgements

T.R.W. acknowledges the NCCR Molecular Systems Engineering for its generous financial support. A.D.L. thanks the EU for Marie Skłodowska-Curie fellowship (H2020-MSCA-IF-2017) for funding. Open access funding provided by Universitat Basel.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: directed evolution · HaloTag · live-cell imaging · protein engineering

[1] G. V. Los, L. P. Encell, M. G. Mcdougall, D. D. Hartzell, N. Karassina, C. Zimprich, M. G. Wood, R. Learish, R. F. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. H. Klaubert, R. F. Bulleit, K. V Wood, *ACS Chem. Biol.* **2008**, *3*, 373–382.
[2] L. P. Encell, O. R. Friedman, K. Zimmerman, P. Otto, G. Vidugiris, M. G. Wood, G. V Los, M. G. Mcdougall, C. Zimprich, N. Karassina, R. D. Learish, R. Hurst, H. James, S. Wheeler, P. Stecha, J. English, K. Zhao, J. Mendez, H. A. Benink, N. Murphy, D. L. Daniels, M. R. Slater, M. Urh, A. Darzins, D. H. Klaubert, R. F. Bulleit, K. V Wood, *Curr. Chem. Genomics* **2012**, *6*, 55–71.
[3] J. B. Grimm, B. P. English, J. Chen, J. P. Slaughter, Z. Zhang, A. Revyakin, R. Patel, J. J. Macklin, D. Normanno, R. H. Singer, T. Lionnet, L. D. Lavis, *Nat. Methods* **2015**, *12*, 244–250.
[4] L. Wang, M. Tran, E. D'Este, J. Roberti, B. Koch, L. Xue, K. Johnsson, *Nat. Chem.* **2020**, *12*, 165–172.
[5] P. E. Deal, P. Liu, S. H. Al-Abdullatif, V. R. Muller, K. Shamardani, H. Adesnik, E. W. Miller, *J. Am. Chem. Soc.* **2020**, *142*, 614–622.
[6] C. Deo, A. S. Abdelfattah, H. K. Bhargava, A. J. Berro, N. Falco, H. Farrants, B. Moeyaert, M. Chupanova, L. D. Lavis, E. R. Schreiter, *Nat. Chem. Biol.* **2021**, *17*, 718–723.
[7] R. S. Erdmann, S. W. Baguley, J. H. Richens, R. F. Wissner, Z. Xi, E. S. Allgeyer, S. Zhong, A. D. Thompson, N. Lowe, R. Butler, J. Bewersdorf, J. E. Rothman, D. St Johnston, A. Schepartz, D. Toomre, *Cell Chem. Biol.* **2019**, *26*, 584–592.e6.
[8] J. B. Grimm, A. N. Tkachuk, L. Xie, H. Choi, B. Mohar, N. Falco, K. Schaefer, R. Patel, Q. Zheng, Z. Liu, J. Lippincott-Schwartz, T. A. Brown, L. D. Lavis, *Nat. Methods* **2020**, *17*, 815–821.
[9] O. M. Thirukkumaran, C. Wang, N. J. Asouzu, E. Fron, S. Rocha, J. Hofkens, L. D. Lavis, H. Mizuno, *Front. Chem.* **2020**, *7*, 938.
[10] M. Kang, H. Lee, H. Kim, Y. Dunbayev, J. K. Seo, C. Lee, H.-W. Rhee, *Chem. Commun.* **2017**, *53*, 9226–9229.
[11] Y. Liu, K. Miao, N. P. Dunham, H. Liu, M. Fares, A. K. Boal, X. Li, X. Zhang, *Biochemistry* **2017**, *56*, 1585–1595.
[12] S. A. Clark, V. Singh, D. V. Mendoza, W. Margolin, E. T. Kool, *Bioconjugate Chem.* **2016**, *27*, 2839–2843.
[13] K. Deprey, J. A. Kritzer, *Bioconjugate Chem.* **2021**, *32*, 964–970.
[14] M. S. Frei, M. Tarnawski, J. Roberti, B. Koch, J. Hiblot, K. Johnsson, *bioRxiv* **2021**, https://doi.org/10.1101/2021.04.01.438024.
[15] Y. Liu, M. Fares, N. P. Dunham, Z. Gao, K. Miao, X. Jiang, S. S. Bollinger, A. K. Boal, X. Zhang, *Angew. Chem. Int. Ed.* **2017**, *56*, 8672–8676.
[16] M. S. Frei, P. Hoess, M. Lampe, B. Nijmeijer, M. Kueblbeck, J. Ellenberg, H. Wadepohl, J. Ries, S. Pitsch, L. Reymond, K. Johnsson, *Nat. Commun.* **2019**, *10*, 4580.
[17] J. B. Grimm, A. K. Muthusamy, Y. Liang, T. A. Brown, W. C. Lemon, R. Patel, R. Lu, J. J. Macklin, P. J. Keller, N. Ji, L. D. Lavis, *Nat. Methods* **2017**, *14*, 987–994.

Manuscript received: August 17, 2021
Revised manuscript received: October 4, 2021
Accepted manuscript online: October 5, 2021
Version of record online: October 19, 2021