A peer-reviewed version of this preprint was published in PeerJ on 18 June 2018.

View the peer-reviewed version (peerj.com/articles/5043), which is the preferred citable publication unless you specifically need to cite this preprint.

Ye B, Shen W, Shi M, Zhang Y, Xu C, Zhao Z. (2018) Intein-mediated backbone cyclization of entolimod confers enhanced radioprotective activity in mouse models. PeerJ 6:e5043 https://doi.org/10.7717/peerj.5043
Intein-mediated backbone cyclization of entolimod confers enhanced radioprotective activity in mouse models

Bingyu Ye 1,2, Wenlong Shen 2, Minglei Shi 2, Yan Zhang 2, Cunshuan Xu Corresp., 1, Zhihu Zhao Corresp. 2

1 State Key Laboratory Cultivation Base for Cell Differentiation Regulation, College of Life Science, Henan Normal University, Xinxiang, China
2 Beijing Institute of Biotechnology, Beijing Institute of Biotechnology, Beijing, China

Corresponding Authors: Cunshuan Xu, Zhihu Zhao
Email address: 003009@htu.edu.cn, zhaozh@bmi.ac.cn

Background. Entolimod is a Salmonella enterica flagellin derivate. Previous work has demonstrated that entolimod effectively protects mice and non-human primates from ionizing radiation. However, it caused a “flu-like” syndrome after radioprotective and anticancer clinical application, indicating some type of immunogenicity and toxicity. Cyclization is commonly used to improve the in vivo stability and activity of peptides and proteins. Methods. We designed and constructed cyclic entolimod using split Npu DnaE intein with almost 100% cyclization efficiency. We adopted different strategies to purify the linear and circular entolimod due to their different topologies. Results. After Ni-chelating affinity purification, the linear and circular entolimod were purified by size-exclusion and ion-exchange chromatography, respectively. Compared with linear entolimod, the circular entolimod showed significantly increased both the in vitro NF-κB signaling and in vivo radioprotective activity in mice. Discussions/Conclusions. Our data indicates that circular entolimod might be a good candidate for further clinical investigation.
Intein-mediated backbone cyclization of entolimod confers enhanced radioprotective activity in mouse models

Bingyu Ye¹,², Wenlong Shen², Minglei Shi², Yan Zhang², Cunshuan Xu¹*, Zhihu Zhao²*

¹State Key Laboratory Cultivation Base for Cell Differentiation Regulation, College of Life Science, Henan Normal University, Xinxiang 453007, China.
²Beijing Institute of Biotechnology, 20 Dongdajie Street, Fengtai District, Beijing 100071, China.

Correspondence:
Cunshuan Xu, Tel: +86-373-3326001, Fax: +86-373-3326524, E-mail: 003009@htu.edu.cn;
Zhihu Zhao, Tel: +86-10-66948775, Fax: +86-10-63818253, E-mail: zhaozh@bmi.ac.cn

Email address:
Bingyu Ye: byy0804@gmail.com
Wenlong Shen: shenwl1988@gmail.com
Minglei Shi: shiml79@126.com
Yan Zhang: zany1983@163.com

Abstract

Background. Entolimod is a Salmonella enterica flagellin derivate. Previous work has demonstrated that entolimod effectively protects mice and non-human primates from ionizing radiation. However, it caused a “flu-like” syndrome after radioprotective and anticancer clinical application, indicating some type of immunogenicity and toxicity. Cyclization is commonly used to improve the in vivo stability and activity of peptides and proteins.

Methods. We designed and constructed cyclic entolimod using split Npu DnaE intein with almost 100% cyclization efficiency. We adopted different strategies to purify the linear and circular entolimod due to their different topologies.

Results. After Ni-chelating affinity purification, the linear and circular entolimod were purified...
by size-exclusion and ion-exchange chromatography, respectively. Compared with linear entolimod, the circular entolimod showed significantly increased both the in vitro NF-κB signaling and in vivo radioprotective activity in mice.

Discussions/Conclusions. Our data indicates that circular entolimod might be a good candidate for further clinical investigation.

**Keywords:** entolimod; Npu DnaE intein; cyclization; radioprotective activity

1 Introduction

Entolimod (previously called CBLB502) is a truncated derivative of the *Salmonella* flagellin protein. It is substantially less immunogenic than flagellin but retains its TLR5-dependent NF-κB-inducing activity and radioprotective capability (Burdelya et al., 2008). Entolimod is currently under development as a medical radiation countermeasure (MRC) under the FDA’s Animal Efficacy Rule and has demonstrated efficacy for both reducing damage to radiosensitive hematopoietic (HP) and gastrointestinal (GI) tissues as well as in improving their regeneration (Krivokrysenko et al., 2012). Moreover, entolimod is also shown to be an effective antitumor agent in several in vivo models (Burdelya et al., 2012; Cai et al., 2011). Entolimod is now considered for clinical use because unlike activation of some other TLRs, the specific profile of cytokines induced following TLR5 stimulation by entolimod does not lead to septic shock-like syndrome or a “cytokine storm” of IL-1 and TNF (Akira et al., 2004; Carvalho et al., 2011; Vijay-Kumar et al., 2008). Entolimod thus has prospective clinical applications as a radioprotective and anticancer agent; however, being a flagellin variant, it can still cause a “flu-like” syndrome after injection, indicating some type of immunogenicity and toxicity (Ding et al., 2012; Burdelya et al., 2013; Hossain et al., 2014; Kojouharov et al., 2014; Yang et al., 2016; Brackett et al., 2016). Another concern is that the time frame for effective entolimod administration is relatively narrow, especially at very high doses of radiation exposure (10 Gy or 13 Gy, for instance) (Burdelya et al., 2008). Therefore, it is important to develop entolimod derivates with increased activity, stability, and more efficient recombinant production.
Intein-mediated backbone cyclization of proteins is a widely used approach for improving protein stability and biological activity (Tarasava et al., 2014; Iwai et al., 1999); therefore, it is now considered a powerful tool for enhancing the efficacy of protein-based therapeutics (Tavassoli, 2017). Studies have shown that relative to the linear protein, the cyclized protein is not susceptible to hydrolysis by exogenous proteases owing to its conformational rigidity arising from lack of both amino and carboxyl termini (Iwai et al., 1999; Evans et al., 1999; Horton et al., 2002). Both expressed protein ligation (EPL) (Camarero et al., 1999) and protein trans-splicing (PTS) (Scott et al., 1999) have been used to generate cyclic peptides and proteins. An alternative complementary method termed as split-intein mediated circular ligation of peptides and proteins (SICLOPPS) has also been developed to facilitate the convenient and efficient use of intein splicing (Scott et al., 1999). The SICLOPPS construct contains three parts: the C-terminal intein domain, target sequence, and N-terminal intein domain. In this construct, the target sequence can be head-to-tail cyclized. SICLOPPS originally used the naturally occurring split intein DnaE from *Synechocystis sp. PCC6803* (Ssp) (Wu et al., 1998), but recently this intein has been replaced with the faster splicing engineered intein from *Nostoc punctiforme (Npu)*, which is also significantly more tolerant of amino acid diversity in the extein sequence (Townend and Tavassoli, 2016). SICLOPPS is mainly used to generate cyclic peptide libraries in drug discovery. For example, researchers have discovered cyclic peptide inhibitors of DAM methyltransferase (Naumann et al., 2008), ClpXP protease (Cheng et al., 2007), hypoxia-inducible factor-1 (HIF-1), a variety of protein-protein interactions (Miranda et al., 2013), and inhibitors that reduce the toxicity of α-synuclein, a key protein in Parkinson’s disease (Kritzer et al., 2009). Cyclization of TEM-1β-lactamase, GFP and VP1 proteins has shown increased thermostability than the linear form and significantly more resistance to proteolysis by exopeptidase (Iwai et al., 1999; Zhao et al., 2010; Qi et al., 2017). However, the use of circular proteins as new and potential drug targets to prevent and treat diseases has not been explored.

In this work, based on SICLOPPS, we used the split intein DnaE from *Npu* to generate a split functional N- and C-terminal intein (I_C and I_N) to cyclize the entolimod protein in *E. coli*. The
cyclization reaction was achieved by sandwiching the entolimod with a His$_6$-tag between the C-terminal domain (Npu I$_C$) and N-terminal domain (Npu I$_N$) of the Npu DnaE intein in the order of Npu I$_C$-His$_6$/entolimod-Npu I$_N$. This design was successfully cloned, and expressed the entolimod protein in circular form, in which the N- and C-termini are covalently joined by an amide bond with approximately 100% cyclization efficiency. After expression, we found that the linear and circular form of entolimod required different purification strategies owing to their different topologies. After purification, we compared their biological activity both \textit{in vitro} and \textit{in vivo}; cyclic entolimod (hereafter referred as C-entolimod) showed significant activity in both \textit{in vitro} a dual-luciferase reporter assay and \textit{in vivo} radioprotective activity in C57BL/6 mouse models. Because of the excellent biological activity of C-entolimod, we could achieve the same effect as that of linear entolimod (or L-entolimod) by administering a much lower dose of C-entolimod, which could further reduce its potential immunogenicity and toxicity. Another advantage of C-entolimod is its increased stability, which could expand the time frame for effective administration.

2 Materials and methods

2.1 Cloning and construction of L- and C-entolimod expression plasmids

The full-length flagellin gene was amplified by forward primer 1 carrying the NdeI restriction site and His$_6$-tag (Table 1, bold font) and reverse primer 1 with the HindIII site from \textit{Salmonella enterica} serovar Dublin genome, as previously described (Burdelya et al., 2008). We used the flagellin gene as a template and amplified the entolimod N-terminal using forward primer 1 and reverse primer 2, and the C-terminal using forward primer 2 and reverse primer 1. The PCR products of the entolimod N-terminal and C-terminal were mixed and used as templates to amplify entolimod using forward primer 1 and reverse primer 1. The sequence 5′-TCCCCGGGAATTTCGCGGTGGGTGGTGGAATTCTAGACTCCATGGGT-3′ was the linker between entolimod N-terminal and C-terminal. The resulting DNA segment was digested with NdeI/HindIII and ligated into NdeI/HindIII digested pET-28a(+) (Novagen #69864-3),
resulting in the L-entolimod expression plasmid pET/entolimod-28a(+). We then used the same strategies to clone Ssp DnaE intein and Npu DnaE intein into the vector pET-28a(+). We amplified Ssp I_C using forward primer 3 carrying an NcoI restriction site and reverse primer 3 carrying a BamHI site from pSFBAD09 (Addgene # 11963), the product was digested with NcoI/BamHI and ligated into pET-28a(+) digested with the same two restriction enzymes, to generate pET/Ssp I_C-28a(+), the italicized base in reverse primer 3 contain a linker sequence with KpnI, BamHI, and NdeI cloning sites and three different termination codon frames TAA, TGA, and TAG. Ssp I_N was amplified using forward primer 4 with a BamHI site and reverse primer 4 with a HindIII site from pJJDuet30 (Addgene # 11962), and then ligated into pET/Ssp I_C-28a(+) with BamHI/HindIII digestion to produce pET/Ssp DnaE-28a(+) plasmid. We obtained pET/Npu I_C-28a(+) using forward primer 5 and reverse primer 5 with the same restriction enzymes as those used for pET/ Ssp I_C-28a(+) from pSKBAD02 (Addgene # 15335), and pET/ Npu DnaE-28a(+) was constructed using forward primer 6 and reverse primer 6 with the same restriction enzymes as pET/Ssp DnaE-28a(+) from pSKDuet01 (Addgene # 12172). Lastly, we used forward primer 7 carrying an NdeI site, and reverse primer 7 carrying a BamHI site, His6-tag (Table 1, bold font) to amplify entolimod from pET/entolimod-28a(+), digested the target sequence with NdeI/BamHI, and ligated it into NdeI/BamHI digested pET/Ssp DnaE-28a(+) and pET/Npu DnaE-28a(+) to generate the final C-entolimod expression plasmids pET/Ssp DnaE/entolimod-28a(+) and pET/Npu DnaE/entolimod-28a(+).

2.2 Expression of recombinant entolimod proteins and cell lysis

The expression plasmids pET/entolimod-28a(+), pET/Ssp DnaE/entolimod-28a(+) and pET/Npu DnaE/entolimod-28a(+) were transformed into E. coli BL21(DE3) (Tiangen Biotech, China). The cells were grown overnight at 37°C in LB medium supplemented with 50 μg/mL kanamycin. Protein overexpression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM in the bacterial culture (when optical density at 600 nm reached ~0.5) and the incubation was extended for an additional 3-5 h at 37°C. The cells containing the expression plasmids pET/entolimod-28a(+), pET/Ssp DnaE/entolimod-28a(+), and pET/Npu
DnaE/entolimod-28a(+) were subsequently harvested at 6000 × g for 25 min. After centrifugation, the three cell pellets were respectively sonicated (30% amplitude for 30 min, 5 s on, 5 s off) in lysis buffer: 20 mM Tris, 500 mM NaCl, 10 mM Imidazole, pH 8.0. The target proteins expressed as inclusion bodies (~80% of total) were confirmed by SDS-PAGE analysis.

2.3 Purification of L- and C-entolimod proteins

L-entolimod was purified described previously (inclusion bodies did not require refolding) (Burdelya et al., 2008). After purification by Ni-chelating affinity (GE Healthcare, UK) and size-exclusion chromatography (GE Healthcare, UK), the purity of L-entolimod was found to be > 90% after SDS-PAGE followed by Coomassie blue staining. However, the same purification methods were not suitable for C-entolimod. After purification by Ni chelating affinity chromatography, the total sample was desalted (20 mM phosphate buffer (PB), pH 7.4) using a G25 column (GE Healthcare, UK). After desalting, ion-exchange chromatography (IEC) was performed on a column of HiTrap Q HP (anion exchange) resin (GE Healthcare, UK). We used 20 mM PB, pH 7.4 as the ion-exchange binding buffer. C-entolimod was then eluted using an increasing salt (10 mM-1 M NaCl) gradient. Protein concentration was measured by a BCA Protein Assay Kit (Tiangen Biotech, China) and protein fractions were analyzed by SDS-PAGE stained with Coomassie blue. The buffers of purified L-entolimod and C-entolimod were changed to phosphate-buffered saline (PBS) and the proteins were stored at -80°C until used.

2.4 Q Exactive mass spectrometer (QE-MS)

The C-entolimod solution was adjusted to pH 8–9 with 100 mM NH₄HCO₃, and was then treated with 50 mM DTT, followed by 50 mM IAA. For peptides, the sample was equally divided into two parts, one for trypsin (T) digestion and the other for chymotrypsin (CT) digestion. The digested samples were diluted using solution A (2% ACN/98% H₂O/0.1% FA), centrifuged at 20,000 × g at 4°C for 30 min, and the supernatant was then transferred to the Dionex (U3000) LC-QE-MS system for C-entolimod sequence analysis. The entire volume of the digested C-entolimod solution was loaded onto a C18 trap column (C18 PepMap, 300 μm ID, 150 mm length, 300 Å pore size, 1.9 μm particle size). Mobile phase A: water with 0.1% FA, mobile
phase B: 80% ACN + 0.08% FA water solution. The flow rate was 0.2 ml/min. The QE-MS parameters were as follows: data collection time, 90 min; spray voltage, 1.9 kV; capillary temperature, 320°C; normalized collision energy, 27%; molecular weight range used for collection: 300–1400 Da. Data retrieval: Mascot search engine, self-built database, first level error 15 ppm, secondary error 20 mmu.

2.5 NF-κB-dependent dual-luciferase reporter assays

Human embryonic kidney cells 293(HEK293) (ATCC # CRL-1573) at 70–90% confluence were transfected with plasmids pNF-κB-Luc (Agilent #219078) and pRL-SV40 (Promega #E2231) using Lipofectamine™ 2000 (Thermo Fisher Scientific #11668019) as per the manufacturer's protocol. L- and C-entolimod were added to the cells at different concentrations (0.00001 μM, 0.001 μM, and 0.1 μM) at 24 h after transfection. Firefly and Renilla luciferase assays were performed 6 h after addition of L- and C-entolimod to the culture medium using the dual-luciferase reporter assay system (Promega #E1980).

2.6 Irradiation and radioprotection of mice

C57BL/6 mice were obtained from Vital River Laboratories (VRL) (Beijing, China). All animal experiments were performed according to the Guide for the Care and Use of Medical Laboratory Animals (Ministry of Health, China, 1998) and with the ethical approval of the Beijing Institute of Biotechnology. Male mice, 6-8 weeks old, weighing 18–22 g were used for the experiment (12 animals per group, 6 groups, a total of 72 animals). The mice were divided into 3 groups with similar weight distribution: PBS (control) group, L-entolimod treatment group, and C-entolimod treatment group. The mice were subjected to total body irradiation (TBI) with 60Co-gamma to a total dose of 9 Gy and 14 Gy at a dose rate of 153.33 cGy/min (Institute of Radiation Medicine, Beijing, China). Mice were irradiated on a rotating platform to ensure even dose delivery to all tissues. L- and C-entolimod were injected subcutaneously (s.c.) 30 min before TBI at a dose of 0.2 mg/kg. The PBS group was injected an equivalent volume of PBS instead. The survival rate of mice was observed 30 days after TBI.

2.7 Statistical analysis
Each column in Fig. 5 is presented as means ± SD of three independent experiments. The results were statistically evaluated for significance using the Student’s $t$-test. The survival rate was statistically evaluated for significance using the log-rank test. $P < 0.05$ was considered statistically significant.

### 3 Results

#### 3.1 Npu DnaE split intein-mediated entolimod cyclization

According to the principle of SICLOPPS, the two fragments of intein, $I_N$ and $I_C$, interact with each other to form an active intein that splices to cyclize the proteins or peptides placed in between. In this case, the entire entolimod gene was sandwiched between the $I_C$ and $I_N$ of $Ssp$ or $Npu$ intein with CFN residues at the C terminus from the native C-terminal extein sequence of DnaE intein, and HM and GS residues at its C terminus and N terminus due to cloning (Fig. 1A, B, C).

C-entolimod was readily apparent by SDS-PAGE upon IPTG induction under the expression plasmids pET/$Npu$ DnaE-28a(+). Fig. 2A shows the expression and the expected splice product C-entolimod (~33.7kDa); the intein ($I_N$ and $I_C$) band was clearly observed in two fragments, $I_N$ (~11.9 kDa) and $I_C$ (~4.1 kDa), after splicing. C-entolimod migrated more rapidly in SDS-PAGE analyses than did L-entolimod (Fig.3A, ~32 kDa), implying an additional topological constraint (Iwai et al., 1999; Zhao et al., 2010; Qi et al., 2017). We then tested whether entolimod cyclization could be mediated by $Ssp$ intein. Unfortunately, the results indicated that pET/$Ssp$ DnaE/entolimod-28a(+) failed to induce the expression of C-entolimod, and the precursor ($Ssp$ DnaE- entolimod) was clearly observed in the gel (Fig. 2B). Therefore, $Npu$ intein was considered a better choice for entolimod cyclization. Previous studies have shown that $Npu$ intein is more tolerant of amino acid substitutions in the C-terminal extein sequence than $Ssp$ intein (Townend and Tavassoli, 2016), and also has exhibited the highest efficiency for the protein trans-splicing reaction so far ($t_{1/2}$ of ~60 s), which was 33–170-fold higher than that of $Ssp$ intein (Zettler et al., 2009). As reported previously, L-entolimod is mainly expressed in the form of
inclusion bodies (~80%), but is resistant to thermal denaturation at 90°C for 20 min and does not need refolding (Burdelya et al., 2008). In our study, after sonication, C-entolimod also predominantly formed inclusion bodies (Fig. 2A). Thus, we could successfully express C-entolimod by Npu intein-mediated its cyclization.

3.2 Purification of L- and C-entolimod

In this step, we chose different strategies to purify L- and C-entolimod, respectively. First, the L- and C-entolimod-expressing cells were suspended in lysis buffer comprising 20 mM Tris, 500 mM NaCl, and 10 mM Imidazole, pH 8.0, followed by sonication in an ice water bath to minimize thermal damage to proteins. The supernatant and pellet were then collected. The pellet was used for further purification. The pellet could be completely solubilized in Buffer A (20 mM Tris, 500 mM NaCl, 2 M Urea, pH 8.0). Then, the L-entolimod was further purified by Ni-chelating affinity column chromatography and could be readily eluted in Buffer B (20 mM Tris, 500 mM NaCl, 250 mM Imidazole, pH 8.0). After analyzing the purity of L-entolimod by SDS-PAGE, we only found a few bands below L-entolimod (Fig. 3A). However, the one-step elution was not suitable for C-entolimod purification. Therefore, we increased the concentration of imidazole (0-500 mM) in a step-wise manner to elute C-entolimod. Fig. 3B shows the result of using three imidazole gradients (50 mM, 200 mM, and 500 mM) to elute C-entolimod. Unfortunately, it could not yield the same purity as L-entolimod. We further purified L-entolimod according to the procedure described previously, and performed desalting followed by size-exclusion chromatography (Burdelya et al., 2008). Through these two purification steps, L-entolimod was purified to > 95% (Fig. 3C). However, size-exclusion chromatography to purify C-entolimod yielded results similar to those obtained using Ni-chelating affinity purification (data not shown). As predicted by an online protein isoelectric point calculator tool, the isoelectric point of C-entolimod was approximately 5.6. We, therefore, decided to purify C-entolimod by anion exchange chromatography using a HiTrap Q HP resin. After desalting (20 mM PB, pH 7.4) by a G25 column, we used an elution buffer (20 mM PB, 1 M NaCl, pH 7.4) in a step-wise manner to collect the product, as shown in Fig. 3D. C-entolimod with high purity
was obtained with the elution buffer containing 10 and 20 mM NaCl (lane 3 and 4). These two were then combined as the final C-entolimod purification (Fig. 3E). To further confirm the purified product was C-entolimod, we used MALDI mass spectrometry and Q Exactive mass spectrometer (QE-MS) to analyze its molecular weight and sequence, the results indicated that both size and sequence were right as expected (Fig. 4 and Fig. S1).

3.3 Effect of C-entolimod on NF-κB pathway activation

To investigate whether C-entolimod has in vitro biological activity, we added both the L-entolimod and C-entolimod to the dual-luciferase reporter plasmids transfected HEK293 cells. Then the luciferase activity was measured to check/determine the in vitro activity that the both products activate the NF-κB-dependent pathway by inducing production of numerous bioactive factors including anti-apoptotic proteins, ROS scavengers, cytokines, and anti-inflammatory agents (Burdelya et al., 2013; Lockless et al., 2009).

We compared the biological activity of L-entolimod and C-entolimod at three different concentrations (0.00001, 0.001, and 0.1 μM). Three independent replicates were set up for each concentration and PBS control. The results shown that C-entolimod exhibited an excellent performance at all three concentration levels. The differences between the corresponding L-entolimod and C-entolimod groups were statistically significant (Fig. 5, **P < 0.01).

Additionally, even at the very low concentration (0.00001 μM), C-entolimod still activated NF-κB signaling at approximately 40% higher than that by L-entolimod (Fig 5). Cyclization of proteins has been previously reported to enhance biological activity (Iwai et al., 1999; Zhao et al., 2010; Qi et al., 2017). In the present work, our novel protein C-entolimod also showed a similar outcome. These results thus indicate that C-entolimod has a significant biological activity in NF-κB pathway activation.

3.4 Effect of C-entolimod on mouse survival after irradiation

As an anti-radiation drug, L-entolimod showed a significant radioprotective effect on mouse and primate models. To address whether C-entolimod has a in vivo protective effect same as or even better than that of L-entolimod, we injected both agents subcutaneously (s.c.) into C57BL/6 mice
at 0.2 mg/kg of body weight. We administered L-entolimod and C-entolimod 30 min prior to irradiation at two different doses (9 Gy and 14 Gy), and compared them with control mice that received PBS (12 animals per group, in total 72 mice). We determined the survival rate of mice at 30 days after TBI at radiation doses of 9 Gy and 14 Gy. At the lower radiation dose of 9 Gy, which killed all the PBS-treated mice, there was no significant difference in survival rate between C-entolimod and L-entolimod-treated mice (Fig. 6A, \(P = 0.3173\)). However, in the higher radiation dose of 14 Gy group, 40% of the C-entolimod-treated mice survived whereas all L-entolimod-treated mice were killed by the ninth day (Fig. 6B, \(* * P < 0.01\)). These results indicate that C-entolimod has a better radioprotective effect than L-entolimod at higher radiation doses. Thus, C-entolimod is a potential novel and effective drug with radioprotective application.

### 4 Discussion

In this study, we demonstrated that \(Npu\) DnaE split intein possesses robust trans-splicing activity for circularization entolimod in our SICLOPPS construct. After expression and purification, the cyclized entolimod was clearly observed upon SDS-PAGE analysis and no precursor protein was observed, suggesting almost 100% splicing efficiency. To investigate whether C-entolimod has better \textit{in vitro} biological activity and \textit{in vivo} radioprotective activity than L-entolimod, we used the luciferase reporter assay and TBI, respectively. It was found C-entolimod could dramatically augment NF-κB signaling activity in HEK293 cells and increase the survival rate of mice at 30 days after TBI.

In our study, the first requirement was to find a split DnaE intein with the highest splicing activity for cyclization of the entolimod. We chose \(Npu\) DnaE split intein for our experiment, which is significantly more tolerant of amino acid diversity in the extein sequence, as reported previously (Townend and Tavassoli, 2016). In intein-mediated protein, many important factors can influence the intein splicing efficiency. Previous studies indicated that a C-terminal Cys-Phe-Asn sequence was necessary to achieve highly efficient and rapid splicing (Lockless et al., 2009; Shah et al., 2012), but replacement of Phe with Trp or Met in the \(Npu\) intein also resulted in a
splicing efficiency similar to that of the native extein, and the natural extein was not found to be the fastest splicing substrate (Cheriyan et al., 2013). Iwai et al. compared the co-expression of the Npu I\textsubscript{N}/Ssp I\textsubscript{C} with that of Ssp I\textsubscript{N}/Ssp I\textsubscript{C} constructs. They found that the extein sequence at the splicing junction has little importance for the splicing activity (Iwai et al., 2006). Another important reason for the higher activity of Npu intein is possibly the substitution of less conserved residues. The key residues in the active sites of both inteins may slightly alter their own geometry or polarization, which might influence the splicing reaction (Zettler et al., 2009). Together, we speculated that the requirements for trans-splicing seem to be specific to the inteins themselves, and that the difference in splicing efficiency between Npu and Ssp inteins may arise from the proper formation of their three-dimensional structures.

Besides, as the L- and C-entolimod have different topological structure, the purification strategies for L- and C-entolimod are different. After purification by Ni-chelating affinity chromatography, we found that the distribution of E. coli host proteins was quite different between the one-step purified products of L-entolimod and C-entolimod. As shown in Fig. 3A and B, the extra non-target bands mainly migrated below L-entolimod contrasted to above C-entolimod. We assumed that the process of intein-mediated entolimod cyclization might have caused this different, because the circularized C-entolimod product might migrate faster relative to the liner one. This expectation was supported by the following purification procedures where we could obtain high purity for C-entolimod by ion-exchange chromatography but not by size-exclusion chromatography.

The biological safety of L-entolimod has been tested in 150 healthy human subjects. A phase I study of L-entolimod in patients with advanced solid tumors has been completed and a second one is now undergoing (http://www.cbiolabs.com/entolimod). Additionally, a phase I clinical trial of L-entolimod in patients with metastatic liver disease is under preparation (Brackett et al., 2016). However, because entolimod is a flagellin variant, it can still cause a “flu-like” syndrome after administration. Therefore, it is essential to further reduce its potential immunogenicity and toxicity. The circular protein C-entolimod has shown significant much better in vitro and in vivo
biological activity than the linear one. Thus, the same treatment effect could be obtained by using reduced doses of C-entolimod than the linear one, which may reduce the potential immunogenicity and toxicity. Moreover, protein head-to-tail cyclization is not susceptible to hydrolysis by exogenous proteases, thus might be a general approach to increase the stability of linear proteins (Iwai et al., 1999; Zhao et al., 2010; Qi et al., 2017). Following the above principle to C-entolimod, the time frame for effective administration would be widened and could improve its therapeutic effect.

**5 Conclusions**

The above results indicate that C-entolimod might be a good candidate for further clinical investigation.

**Acknowledgements**

We thank members of our laboratory for critical discussions.

**References**

Akira, S., and Takeda, K. (2004). Functions of toll-like receptors: Lessons from KO mice. *C. R. Biol.* 327, 581–589.

Brackett, C.M., Kojouharov, B., Veith, J., Greene, K.F., Burdelya, L.G., Gollnick, S.O., Abrams, S.I., and Gudkov, A.V. (2016). Toll-like receptor-5 agonist, entolimod, suppresses metastasis and induces immunity by stimulating an NK-dendritic-CD8+ T-cell axis. *Proc. Natl. Acad. Sci. USA.* 113, E874–E883.

Burdelya, L.G., Krivokrysenko, V.I., Tallant, T.C., Strom, E., Gleiberman, A.S., Gupta, D., Kurnasov, O.V., Fort, F.L., Osterman, A.L., Didonato, J.A., Feinstein, E., and Gudkov, A.V. (2008). An agonist of Toll-like receptor 5 has radioprotective activity in mouse and primate models. *Science* 320, 226–230.
Toll-like receptor 5 agonist protects mice from dermatitis and oral mucositis caused by local radiation: implications for head and neck cancer radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 83, 228–234.

Burdelya, L.G., Brackett, C.M., Kojouharov, B., Gitlin, I.I., Leonova, K.I., Gleberman, A.S., Aygun-Sunar, S., Veith, J., Johnson, C., Haderski, G.J., Stanhope-Baker, P., Allamaneni, S., Skitzki, J., Zeng, M., Martsen, E., Medvedev, A., Scheblyakov, D., Artemicheva, N.M., Logunov, D.Y., Gintsburg, A.L., Naroditsky, B.S., Makarov, S.S., and Gudkov, A.V. (2013). Central role of liver in anticancer and radioprotective activities of Toll-like receptor 5 agonist. *Proc. Natl. Acad. Sci. USA* 110, E1857–1866.

Cai, Z., Sanchez, A., Shi, Z., Zhang, T., Liu, M., and Zhang, D. (2011). Toll-like receptor 5 on breast cancer cells by flagellin suppresses cell proliferation and tumor growth. *Cancer Res.* 71, 2466–2475.

Camarero, J.A., and Muir, T.W. (1999). Biosynthesis of a head-to-tail cyclized protein with improved biological activity. *J. Am. Chem. Soc.* 121, 5597–5598.

Carvalho, F.A., Aitken, J.D., Gewirtz, A.T., and Vijay-Kumar, M. (2011). TLR5 activation induces secretory interleukin-1 receptor antagonist (sIL-1Ra) and reduces inflammasome-associated tissue damage. *Mucosal. Immunol.* 4, 102–111.

Cheng, L., Naumann, T.A., Horswill, A.R., Hong, S.J., Venters, B.J., Tomsho, J.W., Benkovic, S.J., and Keiler, K.C. (2007). Discovery of antibacterial cyclic peptides that inhibit the ClpXP protease. *Protein Sci.* 16, 1535–1542.

Cheriyan, M., Pedamallu, C.S., Tori, K., and Perler, F. (2013). Faster protein splicing with the Nostoc punctiforme DnaE intein using non-native extein residues. *J. Biol. Chem.* 288, 6202–6211.

Ding, X., Bian, G., Leigh, N.D., Qiu, J., McCarthy, P.L., Liu, H., Aygun-Sunar, S., Burdelya, L.G., Gudkov, A.V., and Cao, X. (2012). A TLR5 agonist enhances CD8(+) T cell-mediated graft-versus-tumor effect without exacerbating graft-versus-host disease. *J. Immunol.* 189, 4719–4727.
Evans, T.C. Jr., Benner, J. and Xu, M.Q. (1999). The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins. *J. Biol. Chem.* 274, 18359–18363.

Horton, D.A., Bourne, G.T., and Smythe, M.L. (2002). Exploring privileged structures: the combinatorial synthesis of cyclic peptides. *Mol. Divers.* 5, 289–304.

Hossain, M.S., Ramachandiran, S., Gewirtz, A.T., and Waller, E.K. (2014). Recombinant TLR5 Agonist CBLB502 Promotes NK Cell-Mediated Anti-CMV Immunity in Mice. *PLoS One* 9, e96165.

Iwai, H, and Plückthun, A. (1999). Circular beta-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett.* 459, 166–172.

Iwai, H., Züger, S., Jin, J., and Tam, P.H. (2006). Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. *FEBS Lett.* 580, 1853–1858.

Krivokrysenko, V.I., Shakhov, A.N., Singh, V.K., Bone, F., Kononov, Y., Shyshynova, I., Cheney, A., Maitra, R.K., Purmal, A., Whitnall, M.H., Gudkov, A.V., and Feinstein, E. (2012). Identification of granulocyte colony-stimulating factor and interleukin-6 as candidate biomarkers of CBLB502 efficacy as a medical radiation countermeasure. *J. Pharmacol. Exp. Ther.* 343, 497–508.

Kritzer, J.A., Hamamichi, S., McCaffery, J.M., Santagata, S., Naumann, T.A., Caldwell, K.A., Caldwell, G.A., and Lindquist, S. (2009). Rapid selection of cyclic peptides that reduce α-synuclein toxicity in yeast and animal models. *Nat. Chem. Biol.* 5, 655–663.

Kojouharov, B.M., Brackett, C.M., Veith, J.M., Johnson, C.P., Gitlin, I.I., Toshkov, I.A., Gleichman, A.S., Gudkov, A.V., and Burdelya, L.G. (2014). Toll-like receptor-5 agonist Entolimod broadens the therapeutic window of 5-fluorouracil by reducing its toxicity to normal tissues in mice. *Oncotarget* 5, 802–814.

Lockless, S.W. and Muir, T.W. (2009). Traceless protein splicing utilizing evolved split inteins. *Proc. Natl. Acad. Sci. USA.* 106, 10999–11004.

Miranda. E., Nordgren, I.K., Male, A.L., Lawrence, C.E., Hoakwie, F., Cuda, F., Court, W., Fox,
K.R., Townsend, P.A., Packham, G.K. Eccles, S.A., and Tavassoli, A. (2013). A cyclic peptide inhibitor of HIF-1 heterodimerization that inhibits hypoxia signaling in cancer cells. J. Am. Chem. Soc. 135, 10418–10425.

Naumann, T.A., Tavassoli, A., and Benkovic, S.J. (2008). Genetic selection of cyclic peptide Dam methyltransferase inhibitors. ChemBioChem. 9, 194–197.

Qi, X., and Xiong, S. (2017). Intein-mediated backbone cyclization of VP1 protein enhanced protection of CVB3-induced viral myocarditis. Sci. Rep. 7, 41485.

Scott, C.P., Abel-Santos, E., Wall, M., Wahnon, D.C., and Benkovic, S.J. (1999). Production of cyclic peptides and proteins in vivo. Proc. Natl. Acad. Sci. USA. 96, 13638–13643.

Shah, N.H., Dann, G.P., Vila-Perelló, M., Liu, Z., and Muir, T.W. (2012). Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering. J. Am. Chem. Soc. 134, 11338–11341.

Tarasava, K., and Freisinger, E. (2014). An optimized intein-mediated protein ligation approach for the efficient cyclization of cysteine-rich proteins. Protein Eng. Des. Sel. 27, 481–488.

Tavassoli, A. (2017). SICLOPPS cyclic peptide libraries in drug discovery. Curr. Opin. Chem. Biol. 38, 30–35.

Townsend, J.E., and Tavassoli, A. (2016). Traceless production of cyclic peptide libraries in E. coli. ACS. Chem. Biol. 11, 1624–1630.

Vijay-Kumar, M., Aitken, J.D., Kumar, A., Neish, A.S., Uematsu, S., Akira, S., and Gewirtz, A.T. (2008). Toll-like receptor 5-deficient mice have dysregulated intestinal gene expression and nonspecific resistance to Salmonella-induced typhoid-like disease. Infect. Immun. 76, 1276–1281.

Wu, H., Hu, Z., and Liu, X.Q. (1998). Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803. Proc. Natl. Acad. Sci. USA. 95, 9226–9231.

Yang, H., Brackett, C.M., Morales-Tirado, V.M., Li, Z., Zhang, Q., Wilson, M.W., Benjamin, C., Harris, W., Waller, E.K., and Gudkov, A.V. (2016). The toll-like receptor 5 agonist entolimod suppresses hepatic metastases in a murine model of ocular melanoma via an NK cell-dependent mechanism. Oncotarget 7, 2936–2950.
Zettler, J., Schütz, V., and Mootz, H.D. (2009). The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett.* 583, 909–914.

Zhao, Z., Ma, X., Li, L., Zhang, W., Ping, S., Xu, M.Q., and Lin, M. (2010). Protein cyclization enhanced thermostability and exopeptidase-resistance of green fluorescent protein. *J. Microbiol. Biotechnol.* 20, 460–466.
Figure 1 (on next page)

The mechanism of SICLOPPS intein splicing and entolimod cyclization in vivo using Npu DnaE intein

(A) The two fragments of intein, I_N and I_C, interact to form an active intein that splices to cyclize entolimod placed in between. Splicing mediates the ligation of the N and C termini of entolimod through a native peptide bond. (B) Schematic representation of expression plasmids pET/Ssp DnaE-28a(+) or pET/Npu DnaE-28a(+) and pET/Ssp DnaE/entolimod-28a(+) or pET/Npu DnaE/entolimod-28a(+). (C) The fusion protein sequence of Npu I_C - entolimod - Npu I_N. The C-terminal 39-residue segment and the N-terminal 102-residue segment of the Npu DnaE intein are enclosed within rectangles. The linker sequence of CFNHM and H_6GS is shown in bold font.
A. SICLOPPS

B. Diagram showing the structure of pT7 and its components:

- pT7
- I_c
- linker
- I_n
- T7 ter

C. Npu I_c and Npu I_n sequences:

- Npu I_c: MIKIATRKYLGKQNVYGVERDHNFALKNGFIASNCFNHMWS
- Npu I_n: HPQFEKAQVINTNSLSLLTQNNLNSQSSLSSAIERLSSGLRINSA
- KDDAAAGQAIAFRTSNIKGLTQASRNANDGISIAQTEGALNEIN
- NNLQRVRELVQATNGNDSDDLKSIQDEIQQRLEIEDRVSNQT
- QFNGVKVLSQDNQMKIQVQGANDGETITDLQKIIVKSLGLDG
- NVNSPGISGGGIGILDMSGMTLINEDAAKSTANPLASIDSAQL
- KVDAVRSSLGAIQNRFDSACTIONLTGNLTNLNSARSRIEDADYATE
- VSNMSQAQILQQAGTSVLAAQNPVPNVLSSLRHHHHHGS
- CLSYETEILTVEYGLPIG KvEKT VYSVDNNGNITQPVAQ
- WHDRGEQEVFEYCELDGLRATKDHKMFVTDGQMLPIDEIFER
- ELDLMRVDNLNP
- Npu I_n: H_e GS
Figure 2 (on next page)

SDS-PAGE analysis of C-entolimod expression

(A) C-entolimod is mainly expressed as inclusion bodies. Lane 1: Before induction; Lane 2: C-entolimod expression after induction with IPTG; Lane 3: Supernatant proteins containing C-entolimod after sonication; Lane 4: Precipitated proteins containing C-entolimod after sonication. The two fragments (Iₐ and Iₖ) of Npu intein are clearly seen in Lane 2 and 3. (B) Ssp intein in mediating entolimod cyclization. Lane 1: Before induction; Lane 2: Precursor protein Ssp intein-entolimod after induction with IPTG.
Figure 3 (on next page)

SDS-PAGE analysis of L- and C-entolimod purification

(A) Purification of L-entolimod by Ni-chelating affinity chromatography. Lane 1: Before induction; Lane 2: L-entolimod expression after induction with IPTG; Lane 3: Elute with 250 mM Imidazole. (B) Purification of C-entolimod by Ni chelating affinity chromatography. Lane 1: Loaded sample; Lane 2: Flow-through; Lane 3: Elute with 50 mM Imidazole; Lane 4: Elute with 200 mM Imidazole; Lane 5: Elute with 500 mM Imidazole; Lane 6: Elute with 50 mM Imidazole (Non-reduced); Lane 7: Elute with 200 mM Imidazole (Non-reduced); Lane 8: Elute with 500 mM Imidazole (Non-reduced). (C) Purification of L-entolimod after a second step by size-exclusion chromatography. Lane 1: Before induction; Lane 2: L-entolimod expression after induction with IPTG; Lane 3: Supernatant proteins containing L-entolimod after sonication; Lane 4: Precipitated proteins containing L-entolimod after sonication; Lane 5: Final protein of L-entolimod. (D) Purification of C-entolimod by anion exchange chromatography using HiTrap Q HP resin. Lane 1: Loaded sample; Lane 2: Flow-through; Lane 3: Elute with 10 mM NaCl; Lane 4: Elute with 20 mM NaCl; Lane 5: Elute with 30 mM NaCl; Lane 6: Elute with 40 mM NaCl; Lane 7: Elute with 50 mM NaCl; Lane 8: Elute with 100 mM NaCl; Lane 9: Elute with 1 M NaCl. Lane 3 and 4 are combined for the final product. (E) Final protein of C-entolimod. Lane 1: C-entolimod (non-reduced); Lane 2: C-entolimod.
The molecular weight of C-entolimod is analyzed by MALDI mass spectrometry. The data analysis are the means ± SD of three independent experiments. The size is ~33.7 kDa.
**Figure 5** (on next page)

Induction of NF-κB-responsive transcription by L-entolimod and C-entolimod

HEK293 cells carrying an NF-κB-dependent luciferase reporter construct were incubated with the indicated concentrations (0.00001 μM, 0.001 μM and 0.1 μM) of L-entolimod or C-entolimod. Luciferase activity was measured after 6 h. Each column represents the average of three independent experiments and error bars indicate standard deviations (S.D.). P values were determined by Student’s t-test, **P < 0.01.
Protection of mice from lethal irradiation by C-entolimod compared to L-entolimod

(A) C57BL/6 mice were injected s.c. with C-entolimod (0.2 mg/kg), L-entolimod (0.2 mg/kg), or PBS 30 min prior to receiving 9 Gy TBI. P values were determined by the log-rank test. ****P < 0.0001 for comparison of survival in C-entolimod- and PBS-treated groups or L-entolimod- and PBS-treated groups. P = 0.3173 for comparison of survival in C-entolimod- and L-entolimod-treated groups. (B) C57BL/6 mice were injected s.c. with C-entolimod (0.2 mg/kg), L-entolimod (0.2 mg/kg) or PBS 30 min prior to receiving 14 Gy TBI. P values were determined by the log-rank test. ****P < 0.0001 for comparison of survival in C-entolimod- and PBS-treated groups or L-entolimod- and PBS-treated groups. **P < 0.01 for comparison of survival in C-entolimod- and L-entolimod-treated groups.
**Table 1** (on next page)

Primer sequences used in this study
| Primer           | Sequence (5’ to 3’)                                                                 |
|-----------------|--------------------------------------------------------------------------------------|
| forward primer 1| AAACATATGCATCATCATCATCATCACGCACAAGTCATTAA TACAAACAG                                 |
| reverse primer 1| GCAGAAGCTTTTATTAACGCAGTAAAGAGAGGACGTT                                               |
| forward primer 2| TAGAAACGCGATCGATTTTCTCCAGACGTTGCTGAATTTTATCC TGGATAG                                |
| reverse primer 2| GGTGGTGTTGGTGGGAATTTCTAGACTCCATGCGGTACATTAATCAA TGAAGAC                             |
| forward primer 3| TAAACATGGGCATGAAAAATATCGGTC                                                       |
| reverse primer 3| TAAAGGATCCGGTACCTACGTACGTGATTGATTGAAACATTTG                                       |
| forward primer 4| TAAACATGGGCATGAAAAATATCGGTC                                                       |
| reverse primer 4| TAAAGGATCCGGTACCTACGTGATTGATTGAAACATTTG                                       |
| forward primer 5| TAAACATGGGCATGAAAAATATCGGTC                                                       |
| reverse primer 5| TAAACATGGGCATGAAAAATATCGGTC                                                       |
| forward primer 6| TAAACATGGGCATGAAAAATATCGGTC                                                       |
| reverse primer 6| TAAACATGGGCATGAAAAATATCGGTC                                                       |
| forward primer 7| TAAACATGGGCATGAAAAATATCGGTC                                                       |
| reverse primer 7| TAAACATGGGCATGAAAAATATCGGTC                                                       |