Engineered production of cancer targeting peptide (CTP)-containing C-1027 in *Streptomyces globisporus* and biological evaluation

Wenli Li[a] , Xiuling Li[b], Tingting Huang[c], Qihui Teng[c], Ivana Crnovcic[c], Christoph Rader[b,d] and Ben Shen*[a,c,d,e]

[a]Division of Pharmaceutical Sciences, University of Wisconsin, Madison, WI 53705, USA, and [b]Department of Cancer Biology, [c]Department of Chemistry, [d]Department of Molecular Therapeutics, and [e]Natural Products Library Initiative at The Scripps Research Institute, The Scripps Research Institute, Jupiter, FL 33458, USA.

**Correspondence to: E-mail: shenb@scripps.edu, Tel: (561) 228-2456, Fax: (561) 228-2472*

Supplementary Information

Table S1. Bacterial strains and plasmids used in this study

strains/ plasmids	genotype and description	reference
E. coli strains:		
DH5α	*E. coli* host for general subcloning	1
ET12567/pUZ8002	Methylation-deficient *E. coli* host for intergeneric conjugation; contains pUZ8002, a non-transmissible *oriT* mobilizing plasmid.	2
Streptomyces globisporus strains		
C-1027 (wild-type)	C-1027 wild type producer strain	3
SB1024	*cagA* mutant resulted from integration of pBS1138 into *S. globisporus* wild-type, *ermE* inserted in the middle of *cagA,* Erm^R, C-1027 nonproducing	This study
SB1025	*cagA* mutant resulted from integration of pBS1139 into *S. globisporus* wild-type, *ermE* inserted downstream of *cagA,* Erm^R, C-1027 nonproducing	This study
SB1026	recombinant strain resulted from integration of pBS1142 into *S. globisporus* SB1034, Apr^S and Erm^S, NGR-containing C-1027 producing	This study
Micrococcus luteus ATCC 9431	Enediyne bioassay strain to measure the antibacterial activity	4
Plasmids		
pGEM-T Easy	Vector in *E. coli* for sub-cloning	Promega
pUC18	Vector in *E. coli* for sub-cloning	Stratagene
pIJ4026	Vector containing the *ermE* cassette	5
pKC1139	*E. coli-Streptomyces* shuttle vector, temperature-sensitive replication origin, *aac(3)IV* resistance	6
pBS1005	pOJ446-derived *S. globisporus* genomic library cosmid	7
pBS1135	pUC18 containing 4.3-kb *Sac*I fragment from pBS1005	This study
pBS1136	*ermE* from pIJ4026 ligated into *Bst*XI site (downstream of *cagA*) of pBS1135	This study
pBS1137	*ermE* from pIJ4026 ligated into *Xho*I site (within the *cagA*) of pBS1035	This study
pBS1138	5.76-kb *Sac*I fragments from pBS1136 insert into *Eco*RV of pKC1139	This study
pBS1139	5.76-kb *Sac*I fragments from pBS1137 insert into *Eco*RV of pKC1139	This study
pBS1140	pGEM-T Easy containing PCR amplified *cagA* fragment encoding the NGR motif fused at its C-terminus	This study
pBS1141	pBS1135 containing the *Bcl*I/*Xho*I fragment from pBS1140	This study
pBS1142	4.3-kb *Eco*RI/*Hin*dIII fragment from pBS1141 inserted into pKC1139	This study

Figure S1: The C-1027 biosynthetic gene cluster from *S. globisporus* (accession no: AY048670) with the *cagA* gene highlighted with an arrow, and sequence alignment between CagA (accession no: BAA01609) and selected apo-proteins of enediyne chromoproteins NcsA for neocarzinostatin (accession no: BAA01764), KedA for kedarcidin (accession no: ADG01894), AxnA for actinoxanthin (accession no: BAA02014) and McmA for macromomycin (accession no: P01549). The leader peptides, which are cleaved after the transport of the chromoproteins outside of the cell, are indicated with a red box.

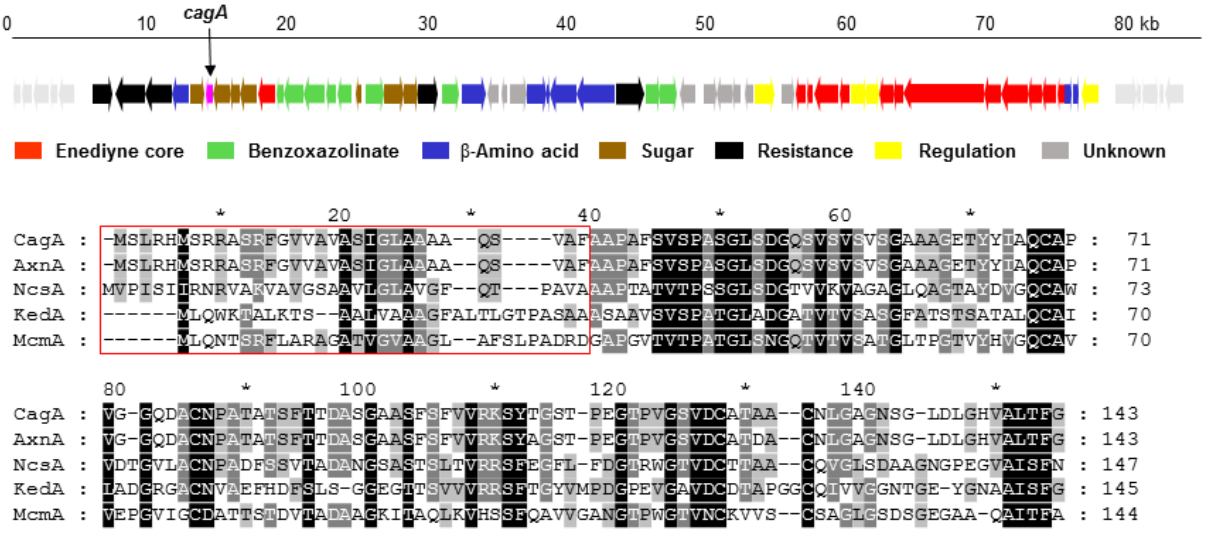


Figure S2. Electrospray mass spectra for native and NGR-containing CagAs. The predicted molecular weights of native and NGR-containing CagAs are 10,501.5 and 11,149.2 Da, respectively.

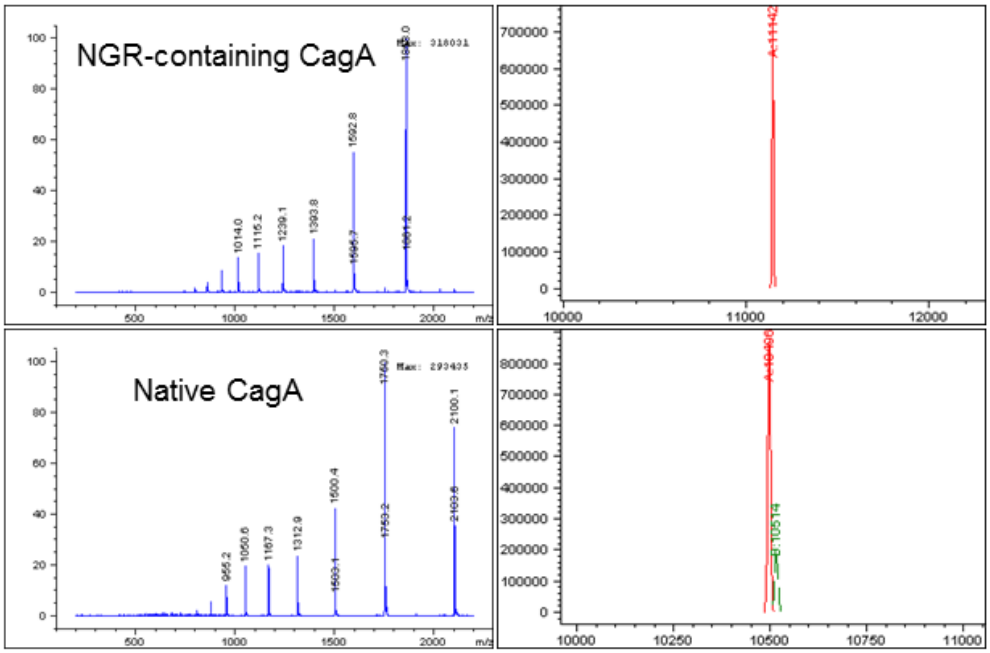


Figure S3. Flow cytometric pattern of CD13 expression in selected breast cancer cell lines. The expression levels of CD13 in breast cancer cell lines were analyzed by flow cytometry. Cells were stained with PE-conjugated anti-CD13 antibody.

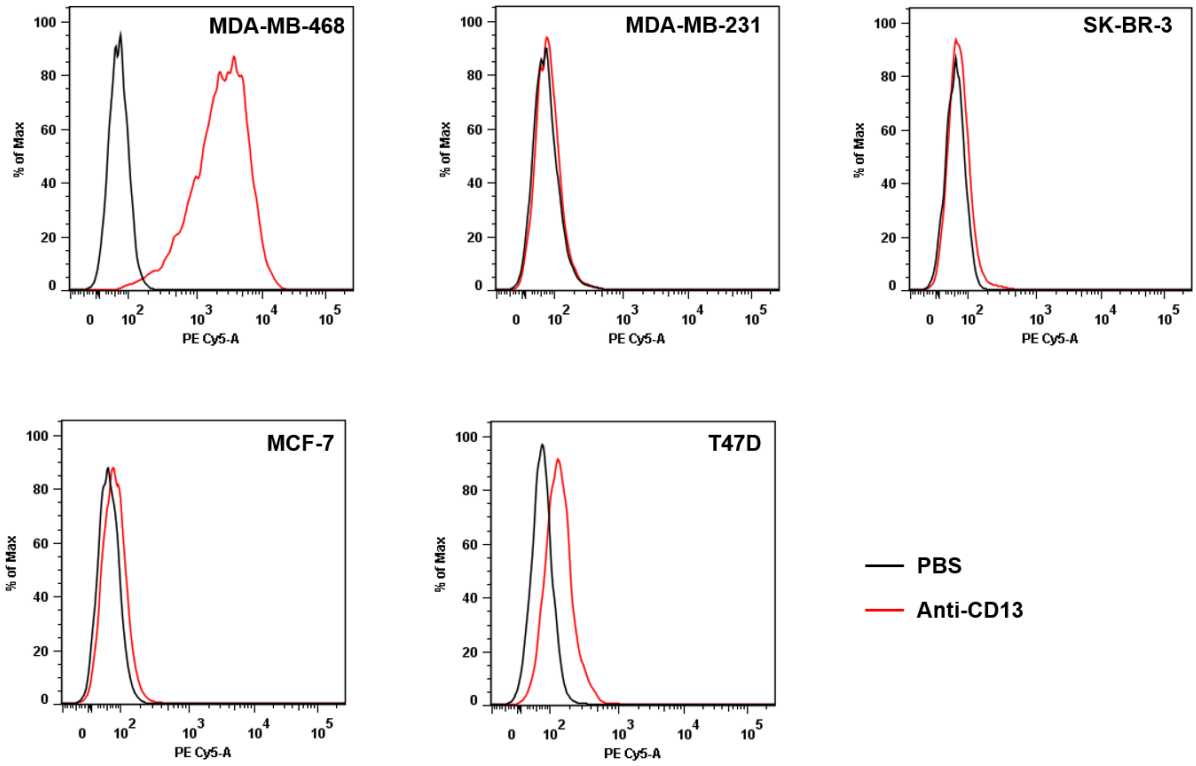
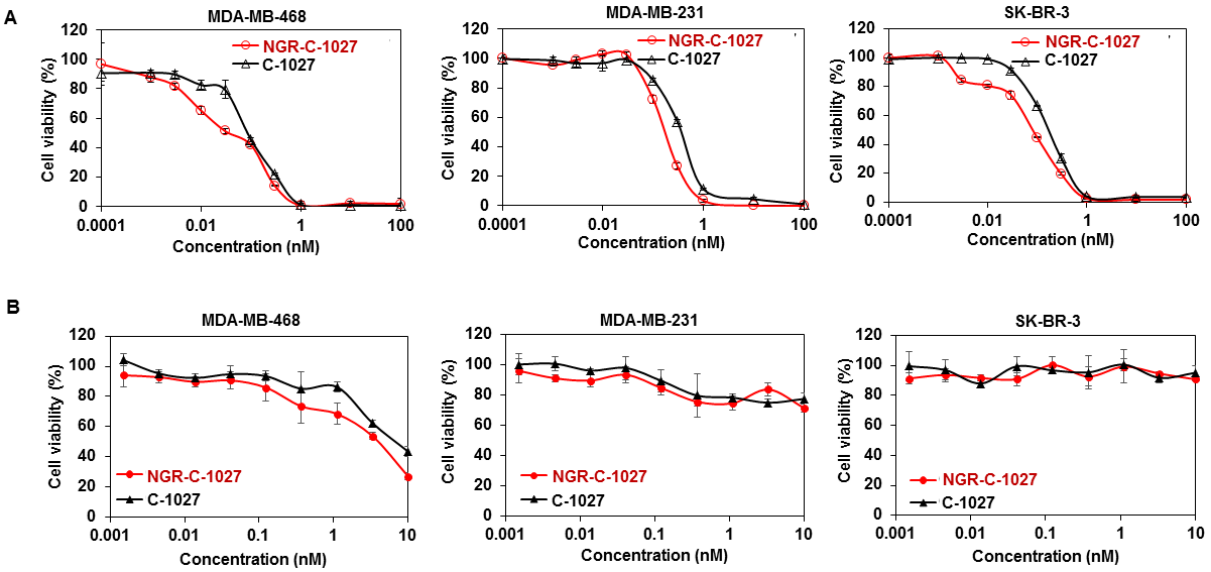


Figure S4. Cytotoxicity of native and NGR-containing C-1027 (NGR-C-1027) chromoproteins toward MDA-MB-468 (CD13-positive), MDA-MB-231 (CD13-negative), and SK-BR-3 (CD13-negative) cell lines. (A) Cells were incubated with the indicated concentration of C-1027 chromoproteins for 24 h before the MTS assay was developed. (B) Cells were incubated with indicated concentration of drugs for 15 min before being washed with PBS. Cells were further incubated with drug-free complete medium for 72 h before the MTS assay was developed. Cell viability was calculated as percentage of untreated control. Each point represents the mean ± SD of 3 replicates, and the IC_{50}s were determined by a computerized curve fitting using GraphPad Prism as summarized in Table 1.



References

1. Sambrook, J.; Fritsch, E, F.; Maniatis, T. Molecular Cloning: A Laboratory Manual, Second Edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; **1989**.
2. Kieser, T., Bibb, M. J.; Chater, K. F.; Buttner, M. J.; Hopwood, D. Practical *Streptomyces* genetics: a laboratory manual. Norwich: John Innes Foundation, Norwich, UK; **2000**.
3. Liu, W.; Christenson, S. D.; Standage, S.; Shen, B. Biosynthesis of the enediyne antitumor antibiotic C-1027. *Science.* **2002**, *297*, 1170-1173.
4. Hu, J. L. Xue, Y. C. Xie, M. Y. Zhang, R. Otani, T. Minami, Y. Yamada, Y. Marunaka, T. A new macromolecular antitumor antibiotic, C-1027. I. Discovery, taxonomy of producing organism, fermentation and biological activity. *J Antibiot*. **1988**, *41*, 1575-1579.
5. Bibb, M. J.; Janssen, G. R.; Ward, J. M. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. *Gene*, **1985**, *38*, 215–226.
6. Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene*. **1992**, *116*, 43-69.
7. Liu, W.; Shen, B. Genes for production of the enediyne antitumor antibiotic C-1027 in *Streptomyces globisporus* are clustered with the *cagA* gene that encodes the C-1027 apoprotein. *Antimicrob Agents Chemother*. **2000**, *44*, 382-392.
8. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. **2011**, 7, 539.
9. Nicholas, K. B., Nicholas, H.B. Jr., Deerfield, D.W. II. GeneDoc: Analysis and Visualization of Genetic Variation, EMBNEW.NEWS. **1997**, *4*, 14.