



# GMEzy: A Genetically Modified Enzybiotic Database

Hongyu Wu<sup>1,3\*</sup>, Jinjiang Huang<sup>1\*</sup>, Hairong Lu<sup>1,3</sup>, Guodong Li<sup>2,3</sup>, Qingshan Huang<sup>1,2,3\*</sup>

**1** State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Sciences, Fudan University, Shanghai, China, **2** Shanghai Engineering Research Center Of Industrial Microorganisms, Shanghai, China, **3** Shanghai High-Tech United Bio-Technological R&D Co., Ltd., Shanghai, China

## Abstract

GMEs are genetically modified enzybiotics created through molecular engineering approaches to deal with the increasing problem of antibiotic resistance prevalence. We present a fully manually curated database, GMEzy, which focuses on GMEs and their design strategies, production and purification methods, and biological activity data. GMEzy collects and integrates all available GMEs and their related information into one web based database. Currently GMEzy holds 186 GMEs from published literature. The GMEzy interface is easy to use, and allows users to rapidly retrieve data according to desired search criteria. GMEzy's construction will increase the efficiency and convenience of improving these bioactive proteins for specific requirements, and will expand the arsenal available for researches to control drug-resistant pathogens. This database will prove valuable for researchers interested in genetically modified enzybiotics studies. GMEzy is freely available on the Web at <http://biotechlab.fudan.edu.cn/database/gmenzy/>.

**Citation:** Wu H, Huang J, Lu H, Li G, Huang Q (2014) GMEzy: A Genetically Modified Enzybiotic Database. PLoS ONE 9(8): e103687. doi:10.1371/journal.pone.0103687

**Editor:** Marc Robinson-Rechavi, University of Lausanne, Switzerland

**Received:** January 16, 2014; **Accepted:** July 4, 2014; **Published:** August 1, 2014

**Copyright:** © 2014 Wu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported in part by the National Major Scientific and Technological Special Project of China for "Significant New Drugs Development" (grant 2013ZX09102057, <http://www.most.gov.cn/>), the China Spark Program (grant 2012GA6900035, <http://www.most.gov.cn/>) and the Shanghai Science and Technology Innovation Action Plan (2012) of China (Grant 12DE1940500, <http://www.stcm.gov.cn/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** Hongyu Wu is affiliated with Shanghai High-Tech United Bio-Technological R&D Co. Ltd (SHUB) as a paid employee and works for projects on State Key Laboratory of Genetic Engineering (SKLGE) at Fudan University as a research assistant. Jinjiang Huang currently majors in Microbiology as a PhD student at SKLGE. Hairong Lu is affiliated with SHUB and currently works on SKLGE for joint postdoctoral research. Guodong Li is the CTO at SHUB. Qingshan Huang is affiliated with Fudan University and as a paid consultant at SHUB. Guodong Li and Qingshan Huang are members of Shanghai Engineering Research Center Of Industrial Microorganisms. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

\* Email: [qshuang@fudan.edu.cn](mailto:qshuang@fudan.edu.cn)

† These authors contributed equally to this work.

## Introduction

The need to develop new classes of antibiotics with novel mechanisms of action against drug-resistant pathogens is urgent, given the increasing prevalence of antibiotic resistance [1]. Thus, the Infectious Diseases Society of America (IDSA) has launched a new collaboration titled the "10 X '20" initiative to develop ten new antibacterial drugs by 2020 [2]. Enzybiotics, in light of their capability to kill susceptible organisms when applied exogenously as recombinant proteins, and because of their low probability for inducing bacterial resistance, have attracted much attention as potential substitutes for conventional antibiotics [3–6].

Enzybiotics primarily consist of endolysins and virion-associated peptidoglycan hydrolases from bacteriophages, bacteriocins and autolysins from bacteria, and lysozymes [7–9]. A number of recent studies have demonstrated the strong potential enzybiotics have for controlling and treating infection caused by varied pathogens, especially gram-positive bacteria [7,10–19]. However, enzybiotics when applied externally, often require modification to improve their properties in specific applications,

particularly in complex environments such as matrices, blood, or on mucous membranes [13]. Several molecular engineering techniques, including domain shuffling, sequence truncation, and random or rational mutagenesis, have been used to optimize enzybiotics by altering their binding or lytic capabilities [13,20]. Among these methods, domain exchange is unique to enzybiotics because of their modular architecture [10,21]. In the early 1990s, Diaz et al. were the earliest to do this, and demonstrated that shuffling functional domains between the pneumococcal autolysin LytA and the phage lysine Cpl-1, switches the catalytic activity and binding properties of the resulting chimeric enzymes [22]. We refer to these types of enzybiotics as genetically modified enzybiotics (GMEs).

In general, enzybiotics are considered antimicrobial peptides on the basis of their protein nature and antibacterial activities. In fact, enzybiotics are different from antimicrobial peptides in both molecular size and antimicrobial mode of action. Although many databases [23–25] have been developed devoted to antimicrobial peptides, there are presently only two existing databases that focus on enzybiotics: EnzyBase [26] and phiBIOTICS [27]. The former contains 1144 enzybiotics from

216 natural sources, whereas the latter represents the knowledge encompassing 21 enzybiotics and 69 corresponding research studies. Moreover, no existing database provides information on how GMEs can be designed, produced, or obtained. With this observation, we developed the manually curated database, GMEzy.

GMEzy focuses on GMEs and stores not only basic GME information but also design strategy, production and purification method, and activity data. GMEzy will serve as a unique tool for GME studies. The construction of GMEzy will enhance understanding the mechanism of GME action, and increase the efficiency and convenience of developing these bioactive proteins on demand. Eventually, GMEzy could help reduce and/or delay increases in antibiotic resistance worldwide.

## Results and Discussion

### Database Description

GMEzy was created as a useful resource for genetically modified enzybiotics studies. As a web based database, GMEzy provides a user-friendly web interface for users to easily query and retrieve information on GMEs. All the data in GMEzy can be accessed and retrieved directly from a web browser. All data were manually collected from scientific literature and public databases, or were calculated by computer programs. All GMEzy data are categorized and stored in relational tables using MySQL.

GMEzy comprises eleven MySQL relational tables. Figure 1 shows the schema of the database. The table GMEzy contains the protein definition, function, and physiochemical properties information fields. Reference information related to each GME is stored in Reference and Reflink. Information related to creating and producing GMEs is contained in Strategy and Production respectively. The features of each GME and its annotation are included in Feature. The parent enzybiotics which GMEs derived from and their annotations are involved in Source and Gmlink respectively. Available protein sequences are contained in the table Sequence. Antimicrobial activity data *in vitro* or *in vivo* for each GME are stored in Activity. Users can add comments to any of the GMEs, and these comments are added to the Comment table.

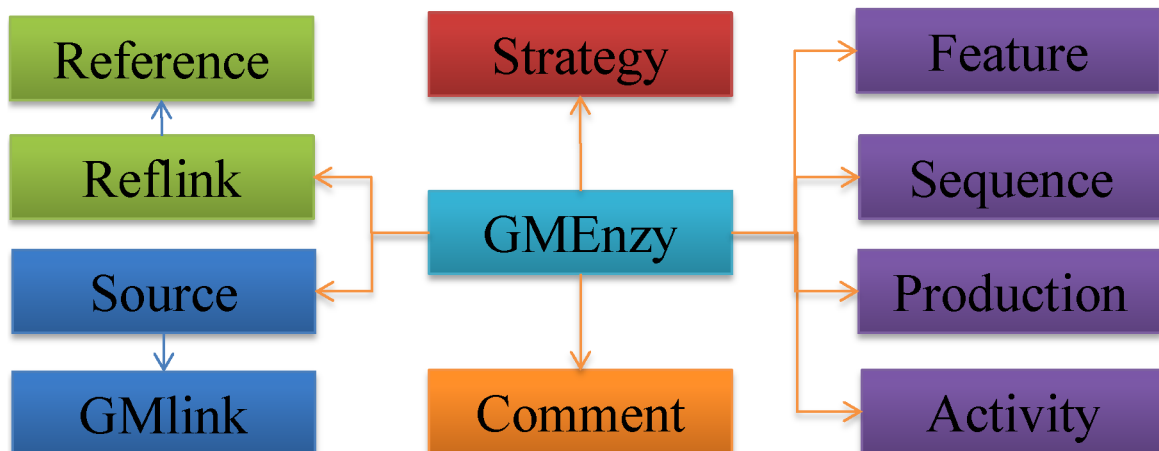
### Database Interfaces

A concise navigational interface that contains the database Browse, Search, Tools, Statistical Info, and Guide options was designed to generate a clearly structured database layout that enables fast and easy navigation (Figure 2).

The Browse interface allows users to navigate not only the entire database, but also grouped GMEs classified by specified creation strategies. Moreover, the Browse interface contains an additional link that allows the download of all GMEs in FASTA format.

The Search interface can be used to retrieve specific information using either a Quick or Advanced option. A Quick search only allows keywords, while the Advanced search accepts the specification of up to six separate fields: GM Id, protein name, parent enzybiotic name, modified strategy, domains, and target organism. The user can query the database by either one particular condition or a combination of various conditions. In addition, the search result page provides the function for quick alignment of the results by the ClustalW [28] software when the results more than two. Each GME Search produces a results page with seven sections: summary, construction, annotation, production, activity, references, and comments (Figure 2). Summary information consists of enzybiotic name, construction strategy, and resulting effect. In addition, summary information also provides basic physiochemical properties such as sequence length, calculated protein mass and isoelectric point (pI), and simple functional annotation. Construction information demonstrates how the GME derived from parent proteins and its correlation with functional alterations. The details information describes the parent protein which the GME derived from and the links to databases describing them. The annotation section demonstrates all components within each GME along with brief descriptions and lists function annotations and the deduced sequence. Moreover, a link to BlastP against GMEzy was added below the sequence. Production, activity, and reference information related to each GME is also listed below. Furthermore, a comment form at the bottom of the page provides users the option to submit comments for the displayed GME to facilitate the exchange of information among researchers.

The Tools interface implements the ClustalW and BlastP functions. The ClustalW function provides a simple ClustalW



**Figure 1. Schema of the GMEzy database.**  
doi:10.1371/journal.pone.0103687.g001

**Home** **Browse** **Database Search** Tools Statistical Info Guide  Search

# GMEnzy: A Genetically Modified Enzybiotic Database

## Detail of GM0001

### Summary

The genetically modified enzybiotic, named **ClyS**, constructed by **Domains Assembly** strategy, change or extend lytic spectrum.

**With ClyS**

- Sequence Length: 280 AA.
- Mass: 31998 Da.
- Isoelectric Point: 9.76

Function: ClyS lysed MRSA, vancomycin-intermediate strains of *S. aureus* (VISA), and methicillin-sensitive (MSSA) strains of *S. aureus* in vitro. In a mouse nasal decolonization model, a 2-log reduction in the viability of MRSA cells was seen 1 h following a single treatment with ClyS. One intraperitoneal dose of ClyS also protected against death by MRSA in a mouse septicemia model.

### Construction

GM0001, constructed by **Domains Assembly** strategy.

O56788: N-acetylmuramoyl-L-alanine amidase			A0EX11: Amidase			UniProt
IPR007921.4-141	IPR002502.172-343	IPR003646	IPR007921.4-133			InterPro
GM0001:1-186			GM0001:187-280			GMEnzy

30 60 90 120 150 180 210 240 270 300 330 360 390 420 450 467 30 60 90 120 150 180 210 240 251

Details:

GM0001[1-186]: derive from protein **O56788** sequence 1-186

▶ O56788, 467 AA., the N-acetylmuramoyl-L-alanine amidase from *Staphylococcus* phage Twort

GM0001[187-280]: derive from protein **A0EX11** sequence 158-251

▶ A0EX11, 251 AA., the Amidase from *Staphylococcus* phage phiNM3

### Annotation

1-186	187-280
-------	---------

50 100 150 200 250 300 350 400 450 500 550 600 650 700 718

1-186 : EAD, **CHAP** domain, derived from *S. aureus* Twort phage lysin;

187-280 : CBD, **Cell wall-targeting** domain, derived from *S. aureus* phage lysin (phiNM3);

**Domain and repeats:**

4-141 : IPR007921, CHAP domain;

**GO term prediction:**

```
MKTLKQAESYIKSKVNTGTFDFDGLYGYQCMDLAVDYIYHVTDGKIRMWGNA
KDAINNSFGGTATVYKNYPAPFRPKYGDVVVWTTGNFATYGHIAIVTNPDPY
GDLQYVTVLEQNWNGNGIYKTELATIRTHDYTGITHFIRPNFATESSVKKK
DTKKKPKPSNRDGINKDKIVYDRTNINYNMVKR-GKSASKITVGSKAPYNL
KWSKGAYFNAKIDGLGATSATRYGDNRTNYRFDVVGQAVYAPGTLIYVFEII
DGWCRIYWNHNEWIWHERLIVKEVF
```

[BlastP to GMEnzy](#)

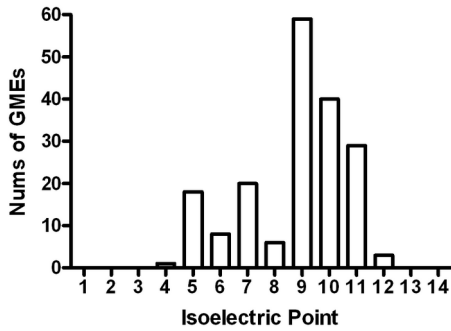
### Production

- Expressed by pJML6 in *E. coli* DH5Alpha, Purified by CM-Sepharose column

### Activity

- In vitro activity of ClyS test on *Staphylococcus aureus* (*S. aureus* 8325-4 cells) by the turbidity-reduction assays showed the OD600 dropped 3- to 4-fold within ~15 min

**Figure 2. Screen shots of a GME result page.** The screen shots show the advanced search and result views. Please note that not all fields are shown. doi:10.1371/journal.pone.0103687.g002



**Figure 3. Distribution of calculated isoelectric points for the GMEs in GMEZy.** Every bar indicates the number of GMEs calculated to have their isoelectric point range from pI-1 to pI. doi:10.1371/journal.pone.0103687.g003

service for users to alignment the GME by ClustalW software and a link to EBI ClustalW which supplied more options for multiple sequence alignment. The BlastP function provides only the BLASTP against GMEZy service and a link to NCBI BLASTP for users who want to blast full datasets in NCBI.

The Statistical info interface supplies statistical data on the physiochemical properties, production, and effect information of the GMEs. (see ‘Statistical description and findings’ section below for further details).

The Guide interface provides simple instructions to potential users on how to use GMEZy functions.

**Statistical Description and Findings**

The current version of GMEZy contains 186 GMEs. Most of GMEs in GMEZy are novel sequences. We used BLASTP [29,30] to align full length GME against UniProt [31], and only 17 sequences showed identical hits from UniProt. The lengths of the GME sequences range from 35 to 879 amino acids. The distribution of pIs for all the GMEs is shown in Figure 3. The majority (98.4%) of the GMEs have calculated pIs that range from 5 to 11, with most of these being basic, rather than acidic.

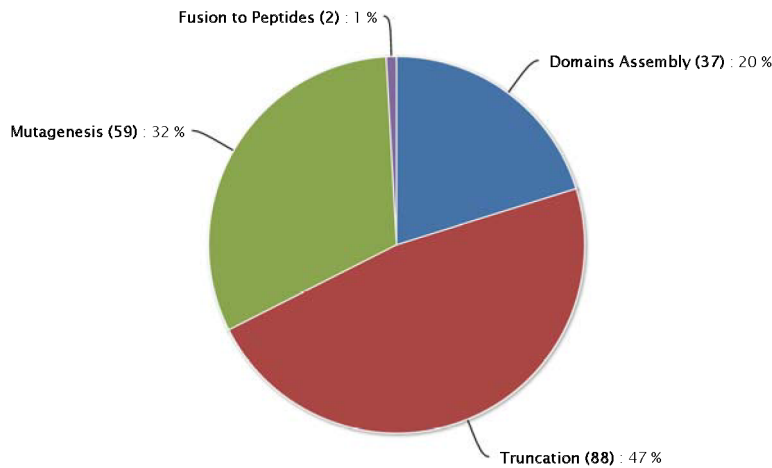
Presently molecular engineering approaches for GME construction mainly include domain exchange, truncation, random

and/or rational mutagenesis, and fusion to peptides, although specific modification methods are varied [13]. Thus, we categorized all GMEs into four groups by construction strategy. Figure 4 demonstrates these categories and the proportion of GMEs in each. In GMEZy the truncation and mutagenesis method accounts for most (79%, 147/186) of the cases. The domain assembly method includes 20% (37/186) of the cases, while the fusion to peptides modification only contains 1% (2/186) of the cases (Figure 4). This distribution is in accord with the current state of GME research. In the past GMEs were mainly created for the control of gram-positive bacterial infections. More recently the fusion to peptide approach was designed as a new orientation for killing gram-negative pathogens. Furthermore, Díez-Martínez et al. have introduced an efficient method for improving the lethal effect of Cpl-7 by inverting the net charge of its cell wall-binding module [20]. This technique could be adopted to help enzybiotics fight gram-negative pathogens. Finally, about 27% (50/186) of the entries in the database are GMEs that exhibit ideal or improved properties, such as increased lytic activity or an extended lytic spectrum, after modification (Table 1). Our data suggest that among these approaches, domain assembly modification exhibits 91% (34/38) positive results, and is the best method for genetic modification (other than the fusion to peptides technique, which has only two cases).

Bacteria are the most commonly used expression host, according the data in our database. Bacteria are hosts in more than 99% of all cases, while only one case used *S. cerevisiae* (Figure 5). The bacterium *E. coli* is the most popular expression host because of attractive features of the organism itself (e.g. fast growth, low cost, and convenient operation). Additionally, to facilitate rapid purification, most (73%, 136/184) of the GMEs are normally expressed as fusion proteins with either histidine (His) or glutathione-S-transferase (GST) purification tags (Figure 6), as well as data not disclosed. Ion-exchange chromatography and size exclusion chromatography purification techniques represent 19% (35/184) and 1% (2/184) of all GMEs respectively. The traditional ammonium sulfate precipitation method has almost been abandoned, as only one case in our database uses it to purify a GME.

**Limitations and Future Prospects**

Currently GMEZy only holds 186 GMEs. Moreover, because of rapidly emerging drug-resistance in bacteria, GMEs have



**Figure 4. Methods for GME construction.** doi:10.1371/journal.pone.0103687.g004

**Table 1.** Strategies for GME construction and their resulting effect.

Strategy	Effect	Nums of GMEs	Positive rate (%)*
Domains assembly	<b>Change or extend lytic spectrum</b>	<b>18</b>	91
	<b>Improve the lytic activity</b>	<b>10</b>	
	<b>Enhance the solubility, change or extend lytic spectrum</b>	<b>2</b>	
	Lose the activities of cell targeting	1	
	<b>Enhance the solubility</b>	<b>1</b>	
	Decrease the lytic activity	1	
	Keep the lytic activity	1	
Truncation	Decrease the lytic activity	38	16
	<b>Improve the lytic activity</b>	<b>14</b>	
	Lose the activities of cell targeting	14	
	Lose the lytic activity	8	
	Keep the lytic activity	7	
	Keep the activities of cell targeting	6	
	Undisclosed	2	
Mutagenesis	Decrease the lytic activity	26	5
	Keep the lytic activity	25	
	Decrease the binding activity	4	
	Change or extend lytic spectrum	2	
	Lose the activities of cell targeting	2	
	Keep the activities of cell targeting	1	
	<b>Improve the lytic activity</b>	<b>1</b>	
Fusion to peptides	<b>Change or extend lytic spectrum</b>	<b>2</b>	100

\*, positive rate indicates the proportion of GMEs with improving lytic activity or extending lytic spectrum (with bold font) to all calculated GMEs.  
doi:10.1371/journal.pone.0103687.t001

attracted increasing interest as potential antimicrobial agents. Therefore, available data will rapidly increase. We will collect each new GME's activity information and related data and incorporate them into GMEnzy by updating the database twice a year in the future.

## Conclusions

GMEnzy is a web accessible database that focuses on the genetically modified enzybiotics. The current version of GMEnzy has 186 entries (as of January, 2014). The database can be queried either by simply using keywords or by combinatorial conditions searches. GMEnzy will not only aid in expanding our current understanding of GMEs and their mechanisms of action, but may have implications in the development of new GMEs to fight against increasingly drug-resistant pathogens. GMEnzy is now available at <http://biotechlab.fudan.edu.cn/database/gmenzy/>.

## Materials and Methods

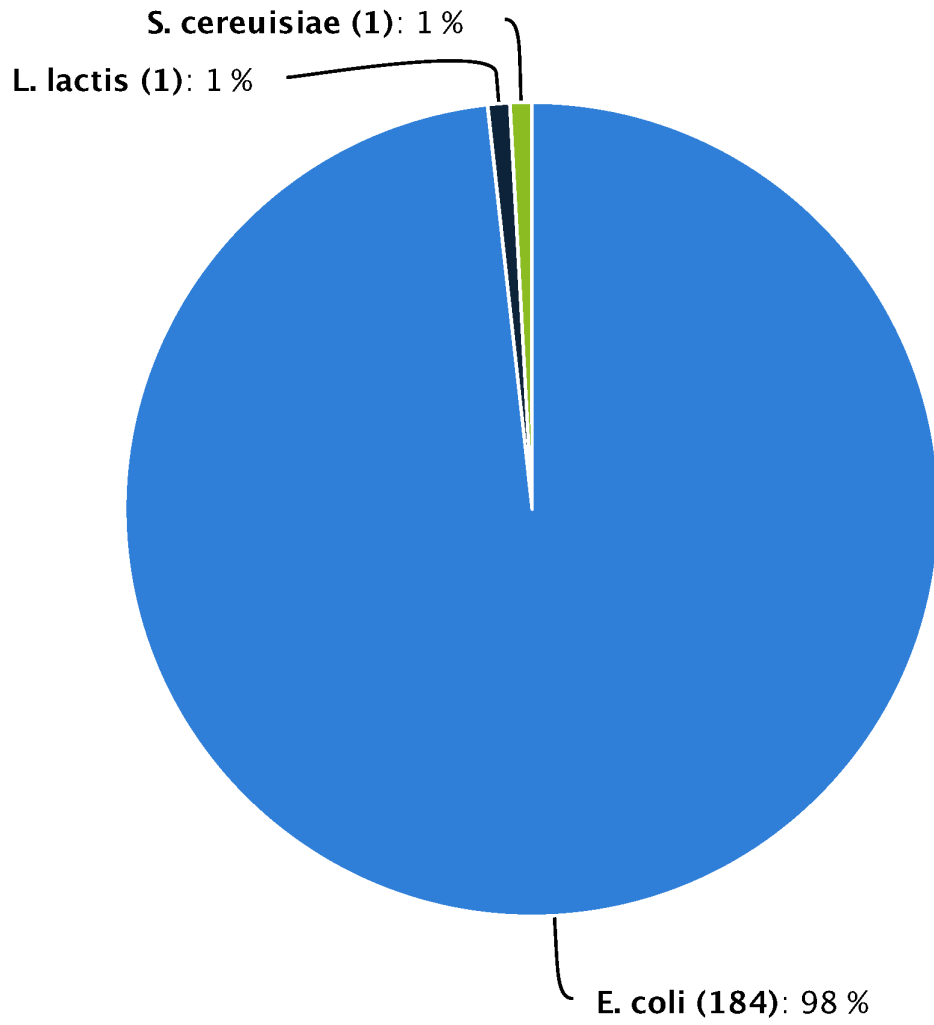
### Data Collection and Organization

We collected all available GME information from scientific literature by searching the Web of Science and Google scholar with the following query: (keyword: "hydrolyse" OR "endolysin" OR "lysine" OR "lysozyme" OR "enzybiotics" AND "fusion" OR "chimeric"), and also by exploring the references listed in the

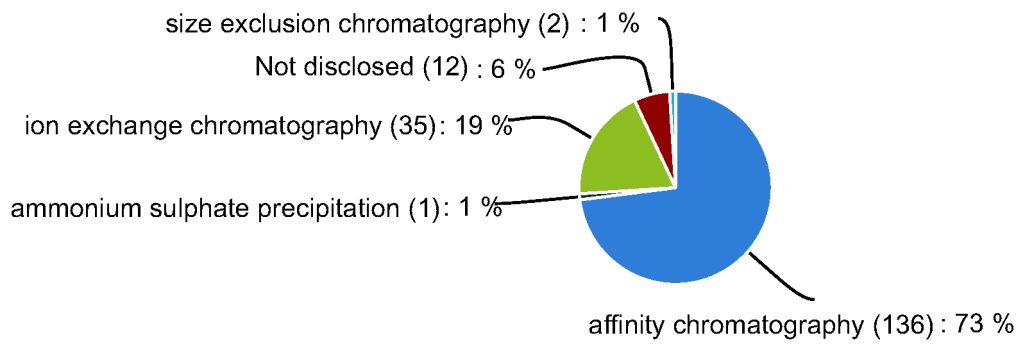
literature discovered with this first pass. Each citation returned from the search was further checked manually to assure it contained the desired information. All GMEs abstracted from the literature were classified into four groups based on the protein engineering strategies: domain assembly, truncation, mutagenesis, and fusion to peptides. All GME sequences were constructed from the UniProt database according to descriptions from the literature. Additional physicochemical data for each GME was either calculated via Bioperl programs or obtained from the literature. All of the collected information was classified and filled into eleven relational tables in MySQL. GMEnzy currently contains 186 GMEs acquired from more than 100 published papers.

### Web Interface and Application

GMEnzy was built on a 64 bit Windows (2008 R2) server running WAMPSEVER (V2.2d), which integrates the Apache HTTP Server (V2.2.21) with PHP (V5.3.10) and the MySQL Server (V5.5.20). All entries are stored in a MySQL database. The application was coded in PHP, using the jQuery JavaScript Library (V1.6.2), the Highchart jQuery Plugin, and Cascading Style Sheets (CSS) for the web design. Apache, MySQL, PHP, and jQuery were preferred as they are open-source software and platform independent, making them suitable for academic use. The web server and all parts of the database are hosted at the Information Office of Fudan University, Shanghai, China.



**Figure 5. Hosts for producing GMEs.**  
doi:10.1371/journal.pone.0103687.g005



**Figure 6. Purification methods for GMEs product.**  
doi:10.1371/journal.pone.0103687.g006

## Acknowledgments

We would like to thank all of our colleagues at the State Key Laboratory of Genetic Engineering at Fudan University and at Shanghai High-Tech United Bio-Technological R&D Co., Ltd., of China for their contributions in the literature search and discussions regarding this manuscript.

## References

- Moellering RC Jr. (2011) Discovering new antimicrobial agents. *Int J Antimicrob Agents* 37: 2–9.
- IDSA (2010) The 10 x '20 Initiative: pursuing a global commitment to develop 10 new antibacterial drugs by 2020. *Clin Infect Dis* 50: 1081–1083.
- Borysowski J, Weber-Dabrowska B, Gorski A (2006) Bacteriophage endolysins as a novel class of antibacterial agents. *Exp Biol Med (Maywood)* 231: 366–377.
- Loessner MJ (2005) Bacteriophage endolysins—current state of research and applications. *Curr Opin Microbiol* 8: 480–487.
- Fischetti VA (2005) Bacteriophage lytic enzymes: novel anti-infectives. *Trends Microbiol* 13: 491–496.
- Hermoso JA, Garcia JL, Garcia P (2007) Taking aim on bacterial pathogens: from phage therapy to enzybiotics. *Curr Opin Microbiol* 10: 461–472.
- Borysowski J, Gorski A (2010) Enzybiotics and their potential applications in medicine. In: Villa TG, Veiga-crespo P, editors. *Enzybiotics: antibiotic enzymes as drugs and therapeutics*. 1th ed. New Jersey: John Wiley & Sons, Inc., Publication. 199–218.
- Rodriguez-Rubio L, Martinez B, Donovan DM, Rodriguez A, Garcia P (2013) Bacteriophage virion-associated peptidoglycan hydrolases: potential new enzybiotics. *Crit Rev Microbiol* 39: 427–434.
- Nelson D, Loomis L, Fischetti VA (2001) Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proc Natl Acad Sci U S A* 98: 4107–4112.
- Thallinger B, Prasetyo EN, Nyanhongo GS, Guebitz GM (2013) Antimicrobial enzymes: an emerging strategy to fight microbes and microbial biofilms. *Biotechnol J* 8: 97–109.
- Pastagia M, Schuch R, Fischetti VA, Huang DB (2013) Lysins: the arrival of pathogen-directed anti-infectives. *J Med Microbiol* 62: 1506–1516.
- Szweda P, Schielmann M, Kotowski R, Gorczyca G, Zalewska M, et al. (2012) Peptidoglycan hydrolases-potential weapons against *Staphylococcus aureus*. *Appl Microbiol Biotechnol* 96: 1157–1174.
- Schmelcher M, Donovan DM, Loessner MJ (2012) Bacteriophage endolysins as novel antimicrobials. *Future Microbiol* 7: 1147–1171.
- Fischetti VA (2010) Bacteriophage endolysins: a novel anti-infective to control Gram-positive pathogens. *Int J Med Microbiol* 300: 357–362.
- Fischetti VA (2008) Bacteriophage lysins as effective antibacterials. *Curr Opin Microbiol* 11: 393–400.
- Veiga-Crespo P, Ageitos JM, Poza M, Villa TG (2007) Enzybiotics: a look to the future, recalling the past. *J Pharm Sci* 96: 1917–1924.
- Yang H, Zhang Y, Yu J, Huang Y, Zhang XE, et al. (2013) A novel chimeric lysin with high antimicrobial activity against methicillin-resistant *Staphylococcus aureus* in vitro and in vivo. *Antimicrob Agents Chemother*.
- Saravanan SR, Paul VD, George S, Sundarajan S, Kumar N, et al. (2013) Properties and mutation studies of a bacteriophage-derived chimeric recombinant staphylolytic protein P128: Comparison to recombinant lysostaphin. *Bacteriophage* 3: e26564.
- Rodriguez-Rubio L, Martinez B, Rodriguez A, Donovan DM, Gotz F, et al. (2013) The phage lytic proteins from the *Staphylococcus aureus* bacteriophage vB\_SauS-phiPLA88 display multiple active catalytic domains and do not trigger staphylococcal resistance. *PLoS One* 8: e64671.
- Diez-Martinez R, de Paz H, Bustamante N, Garcia E, Menendez M, et al. (2013) Improved Lethal Effect of Cpl-7, a Pneumococcal Phage Lysozyme of Broad Bactericidal Activity by Inverting Net Charge of its Cell Wall-Binding Module. *Antimicrob Agents Chemother*.
- Schmelcher M, Shabarova T, Eugster MR, Eichenseher F, Tchang VS, et al. (2010) Rapid multiplex detection and differentiation of *Listeria* cells by use of fluorescent phage endolysin cell wall binding domains. *Appl Environ Microbiol* 76: 5745–5756.
- Diaz E, Lopez R, Garcia JL (1990) Chimeric phage-bacterial enzymes: a clue to the modular evolution of genes. *Proc Natl Acad Sci USA* 87: 8125–8129.
- Zhao X, Wu H, Lu H, Li G, Huang Q (2013) LAMP: A Database Linking Antimicrobial Peptides. *PLoS One* 8: e66557.
- Hammami R, Fliss I (2010) Current trends in antimicrobial agent research: chemo- and bioinformatics approaches. *Drug Discov Today* 15: 540–546.
- Waghu FH, Gopi L, Barai RS, Ramteke P, Nizami B, et al. (2014) CAMP: Collection of sequences and structures of antimicrobial peptides. *Nucleic Acids Res* 42: D1154–1158.
- Wu H, Lu H, Huang J, Li G, Huang Q (2012) EnzyBase: a novel database for enzybiotic studies. *BMC Microbiol* 12: 54.
- Hojkova K, Stano M, Klucar L (2013) phiBIOTICS: catalogue of therapeutic enzybiotics, relevant research studies and practical applications. *BMC Microbiol* 13: 53.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.
- Schaffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, et al. (2001) Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res* 29: 2994–3005.
- Wu CH, Apweiler R, Bairoch A, Natale DA, Barker WC, et al. (2006) The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res* 34: D187–191.

## Author Contributions

Conceived and designed the experiments: HW GL QH. Performed the experiments: HW JH HL. Analyzed the data: HW JH HL. Contributed reagents/materials/analysis tools: HW JH HL. Wrote the paper: HW JH.