

# Functional assessment of hydrophilic domains of late embryogenesis abundant proteins from distant organisms

Yingying Liu,<sup>1</sup> Heng Zhang,<sup>1</sup> Jiahui Han,<sup>1</sup> Shijie Jiang,<sup>1,2</sup> Xiuxiu Geng,<sup>1,2</sup> Dong Xue,<sup>1</sup> Yun Chen,<sup>1</sup> Chen Zhang,<sup>1</sup> Zhengfu Zhou,<sup>1</sup> Wei Zhang,<sup>1</sup> Ming Chen,<sup>1</sup> Min Lin<sup>1</sup> and Jin Wang<sup>1,\*</sup> 

<sup>1</sup>Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China.

<sup>2</sup>College of Life Science and Engineering, Southwest University of Science and Technology, Mianyang 621000, China.

## Summary

Late embryogenesis abundant (LEA) proteins play a protective role during desiccation and oxidation stresses. LEA3 proteins are a major group characterized by a hydrophilic domain (HD) with a highly conserved repeating 11-amino acid motif. We compared four different HD orthologs from distant organisms: (i) DrHD from the extremophilic bacterium *Deinococcus radiodurans*; (ii) CeHD from the nematode *Caenorhabditis elegans*; (iii) YIHD from the yeast *Yarrowia lipolytica*; and (iv) BnHD from the plant *Brassica napus*. Circular dichroism spectroscopy showed that all four HDs were intrinsically disordered in phosphate buffer and then folded into  $\alpha$ -helical structures with the addition of glycerol or trifluoroethanol. Heterologous HD expression conferred enhanced desiccation and oxidation tolerance to *Escherichia coli*. These four HDs protected the enzymatic activities of lactate dehydrogenase (LDH) by preventing its aggregation under desiccation stress. The HDs also interacted with LDH, which was intensified by the addition of hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>), suggesting a protective role in a chaperone-like manner. Based on these results, the HDs of LEA3 proteins show promise as protectants for desiccation and oxidation stresses, especially DrHD, which is a potential ideal stress-response

element that can be applied in synthetic biology due to its extraordinary protection and stress resistance ability.

## Introduction

Both desiccation and oxidation are widely studied extreme processes that are detrimental to the growth of organisms because they induce protein aggregation and nucleic acid damage (Rajpurohit and Misra, 2013; Atashgahi *et al.*, 2018). Moreover, recent advancements have shown that the organismal responses to both stresses are similar, indicating that the mechanisms for protecting crucial biomolecules from damage are related (Fredrickson *et al.*, 2008; Chakrabortee *et al.*, 2010). Long-term investigations into the molecular mechanism of desiccation resistance emphasized late embryogenesis abundant (LEA) proteins, which account for more than 4% of cellular proteins in cotton orthodox seeds (Hughes and Galau, 1989; Roberts *et al.*, 1993; Shimizu *et al.*, 2010). Furthermore, many reports have revealed that the heterologous expression of LEA proteins can significantly improve recipient oxidation resistance (Liu and Zheng, 2005; Kovacs *et al.*, 2008; Shih *et al.*, 2012; Sasaki *et al.*, 2014; Wang *et al.*, 2017).

Late embryogenesis abundant proteins were first characterized in cotton seeds three decades ago and are hydrophilic (Dure and Galau, 1981); their accumulation simultaneously occurs alongside seed maturation, desiccation and other forms of (a) biotic stresses (Chakrabortee *et al.*, 2010; Hatanaka *et al.*, 2013; Liu *et al.*, 2013; Leeuwen *et al.*, 2016; Bremer *et al.*, 2017). Continuous engagement in research on LEA proteins has shown us that these family proteins are not restricted to plants, as they have also been documented in bacteria, fungi and animals (Hand *et al.*, 2011). Based on the amino acid information and conserved motifs, LEA proteins are classified into at least seven distinct groups (Battaglia *et al.*, 2008). Among them, group 3 LEA (G3LEA) proteins have been extensively studied in vegetative plants, revealing the important roles of these proteins in inducing and maintaining a temporary period of 'anhydrobiosis' or 'cryptobiosis' (Ried and Walker-Simmons, 1993). Molecular shield activity or chaperone-like activity was assumed as the typical function of LEA group 3 as they

Received 23 November, 2018; accepted 6 April, 2019.

\*For correspondence. E-mail wangjin@caas.cn; Tel. +86(10) 82109868; Fax +86(10)82109868.

*Microb Biotechnol* (2019) 12(4), 752–762

doi:10.1111/1751-7915.13416

## Funding Information

This study was supported by the Ministry of Agriculture Transgenic Program (No. 2016ZX08009003-002) and National Natural Science Foundation of China (No. 31800061, 31570080, 31500063).

© 2019 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

probably form physical electrostatic barrier, which thus reduces desiccation-induced aggregation, and then the crucial enzymes, like lactate dehydrogenase (LDH), were protected (Chakrabortee *et al.*, 2010, 2012; Olvera-Carrillo *et al.*, 2011; Kovacs and Tompa, 2012).

Loosely conserved 11-mer amino acid units (motifs) are the main feature of G3LEA proteins and were previously identified as 'TAQAAKEKAGE' (Ried and Walker-Simmons, 1993). The number and size of motifs in G3LEA proteins vary by species (Dure and Galau, 1981; Dure, 1993; Dure, 2001; Swire-Clark and Marcotte, 1999). Structural exploration of G3LEA proteins using circular dichroism (CD) indicated that they are randomly coiled in aqueous solution but adopt a largely  $\alpha$ -helical conformation in the presence of glycerol or after desiccation, which is considered to be the prerequisite for LEA-macromolecule interactions for protection under stresses (Tollete *et al.*, 2007). Although many studies implicated that the hydrophilic regions of LEA proteins were crucial to achieve their structural transformation that contributes to the organisms' stress tolerance acquisition (Haaning *et al.*, 2008; Rajpurohit and Misra, 2013; Cuevas-Velazquez *et al.*, 2016), details of the function of this hydrophilic region remain unclear. Additionally, previous work highlighted the importance of G3LEA protein motifs from soybean that contribute to their enhanced salt resistance (Liu and Zheng, 2005). Unfortunately, their work focused on the role of only 22-mer motifs, which are not the common feature of G3LEA proteins. Therefore, investigating the role of 11-mer motifs became necessary to get a better and whole-scale understanding of G3LEA proteins. Herein, the 11-mer motifs of G3LEA proteins are artificially labelled the hydrophilic domain (HD).

Regarding the exploration of the mechanism of living cells adapting to extreme conditions, *Deinococcus radiodurans* was selected as the research model due to its extraordinary stress resistance, especially under desiccation and oxidation (Makarova *et al.*, 2001; Daly, 2009; Slade and Radman, 2011; Ujaoney *et al.*, 2017). Coincidentally, a gene named *dr\_1172* has been characterized to encode a G3LEA protein bearing distinct 11-mer repeating motifs, and its disruption leads to a 75% growth loss in desiccated cultures (Battista *et al.*, 2001). However, little information provided from previous *dr\_1172* study, including its functional domain identification and physiological activity. Herein, the protein encoded by *dr\_1172* is designated DrG3LEA (*Deinococcus radiodurans* group 3 LEA protein). Henceforth, DrG3LEA, the representative from *D. radiodurans*, together with other three G3LEA proteins from *Caenorhabditis elegans* (animal), *Yarrowia lipolytica* (fungus) and *Brassica napus* (plant) was selected due to their sequence similarity, comparatively conserved HDs and their representativeness. More importantly, comparisons among these four HDs in structure, function and

mechanism would be useful to generalize the ubiquitous properties of G3LEA proteins as well as for screening the anti-stress element for future application in synthetic biology rather than employing varied full-length LEA proteins.

In the present work, we characterized four new HDs from G3LEA proteins using bioinformatic analysis and CD analysis. Heterologous expression further identified that their functionality involves desiccation and oxidation tolerance enhancement of *E. coli*, and their protective functions were well-elucidated according to their anti-aggregation and stress-strengthen interaction ability, in which they were shown to act in the same chaperone-like manner as the G3LEA proteins behaved.

## Results

### *In silico analysis of DrHD, CeHD, YIHD and BnHD*

The hydrophilic domain (HD) of DrG3LEA (NP\_294896) from *D. radiodurans* is named DrHD and contains 8 motifs, and hydrophilic residues account for 62.4% of the HD amino acids, whereas the HD of the group 3 LEA (G3LEA) protein (BAB88877) from *Brassica napus* exhibits the highest hydrophilicity (70.1%), is named BnHD and bears six motifs. The HD of the G3LEA protein (SEI35173.1) from *Y. lipolytica*, designated YIHD, has the highest number of motifs (27), and hydrophilic residues account for 57.9% of the HD amino acids. The HD from the *C. elegans* G3LEA protein (NP\_001256171), named CeHD, has 24 motifs, in which the hydrophilic residues occupy 60.4% of the HD amino acids (Table 1). Obviously, the motif numbers varied by species, and the motif residues were not conserved among HDs as shown in Fig. S2. According to their hydrophilicity analysis shown in Kyte and Doolittle hydrophathy plots (Fig. S1) (Kyte and Doolittle, 1982), these HDs were predominantly hydrophilic with few hydrophobic amino acid residues, which indicates their potential to form  $\alpha$ -helix secondary structures (Dure, 1993).

### *Structural transformation of HDs from intrinsically disordered to ordered with the addition of glycerol or trifluoroethanol (TFE)*

Studies have shown that most late embryogenesis abundant (LEA) proteins are intrinsically disordered proteins (IDPs) under natural conditions that are transformed into ordered structures in dry environments (Tollete *et al.*, 2007; Chakrabortee *et al.*, 2012; Hundertmark *et al.*, 2012; Hatanaka *et al.*, 2013; Cuevas-Velazquez *et al.*, 2016). All four HDs, DrHD, CeHD, BnHD and YIHD, were predicted to be intrinsically disordered using Cspritz and IUPred (data not shown).

Henceforth, to investigate these four HD polypeptides' secondary structure in aqueous solution, their genes

**Table 1.** Comparisons among four LEA proteins based on their amino acids (aa). These four LEA proteins were representative of those in bacteria (*D. radiodurans*), fungi (*Y. lipolytica*), plants (*B. napus*) and animals (*C. elegans*). The HDs were named DrHD, YIHD, BnHD and CeHD respectively.

Accession No. in GenBank	Species	Amino acids of full length(aa)	Name of hydrophilic domain	Amino acids of domain(aa)	Molecular weight of domain(Da)	Numbers of 11-mer motif	Hydrophilic residues percentage of domain (%)
NP_294896	<i>Deinococcus radiodurans</i> R1	298	DrHD	142 (104–245)*	14707	8	62.4
NP_001256171	<i>Caenorhabditis elegans</i>	733	CeHD	477 (226–702)*	50138.6	24	60.4
SEI35173.1	<i>Yarrowia lipolytica</i>	748	YIHD	575 (87–661)*	62262.7	27	57.9
BAB88877	<i>Brassica napus</i>	226	BnHD	121 (42–162)*	12995.5	6	70.1

The asterisk indicates the position from the starting aa to the terminating aa.

were cloned into the expression vector pET28a with a 6xHis tag. The expressing strains (BL21) were induced by IPTG (0.1 mM), and the produced proteins were purified by affinity chromatography. The protein purity reached 90% after gel filtration (Fig. S3), and the molecular mass of the purified DrHD, CeHD, YIHD and BnHD proteins was approximately 14.7 kDa, 50.1 kDa, 62.3 kDa and 13 kDa respectively. However, according to the SDS-PAGE results, the HDs migrated to a higher molecular mass due to their hydrophilic character, which is often used to identify IDPs (Tompa, 2002; Chakrabortee *et al.*, 2010). The proteins were analysed by far-UV circular dichroism (CD) spectroscopy. The spectra of all four HDs had a negative peak at 198 nm in phosphate buffer, indicating that DrHD, CeHD, BnHD and YIHD are natively unstructured in solution (Fig. 1A). As shown by the stacked bar graphs in the right half of Fig. 1, these HDs are folded and formed  $\alpha$ -helices with the addition of glycerol (50%) (generally used to mimic water-loss environment) or TFE (50%) (trifluoroethanol, a well-known  $\alpha$ -helix inducer) (Cuevas-Velazquez *et al.*, 2016), with a high ellipticity at 192 nm and minima at 208 and 220 nm (Fig. 1B and C). In addition, DrHD, YIHD and CeHD are able to gain higher levels of helicity than BnHD (Table S2), implying that, on one hand, these four HDs are the core domain for determining their function, as previous work demonstrated that folding into  $\alpha$ -helix is required for activating their protective function under water-deficient environments (Cuevas-Velazquez *et al.*, 2016); on the other hand, different  $\alpha$ -helix levels of HDs might lead to differences in their physiological function.

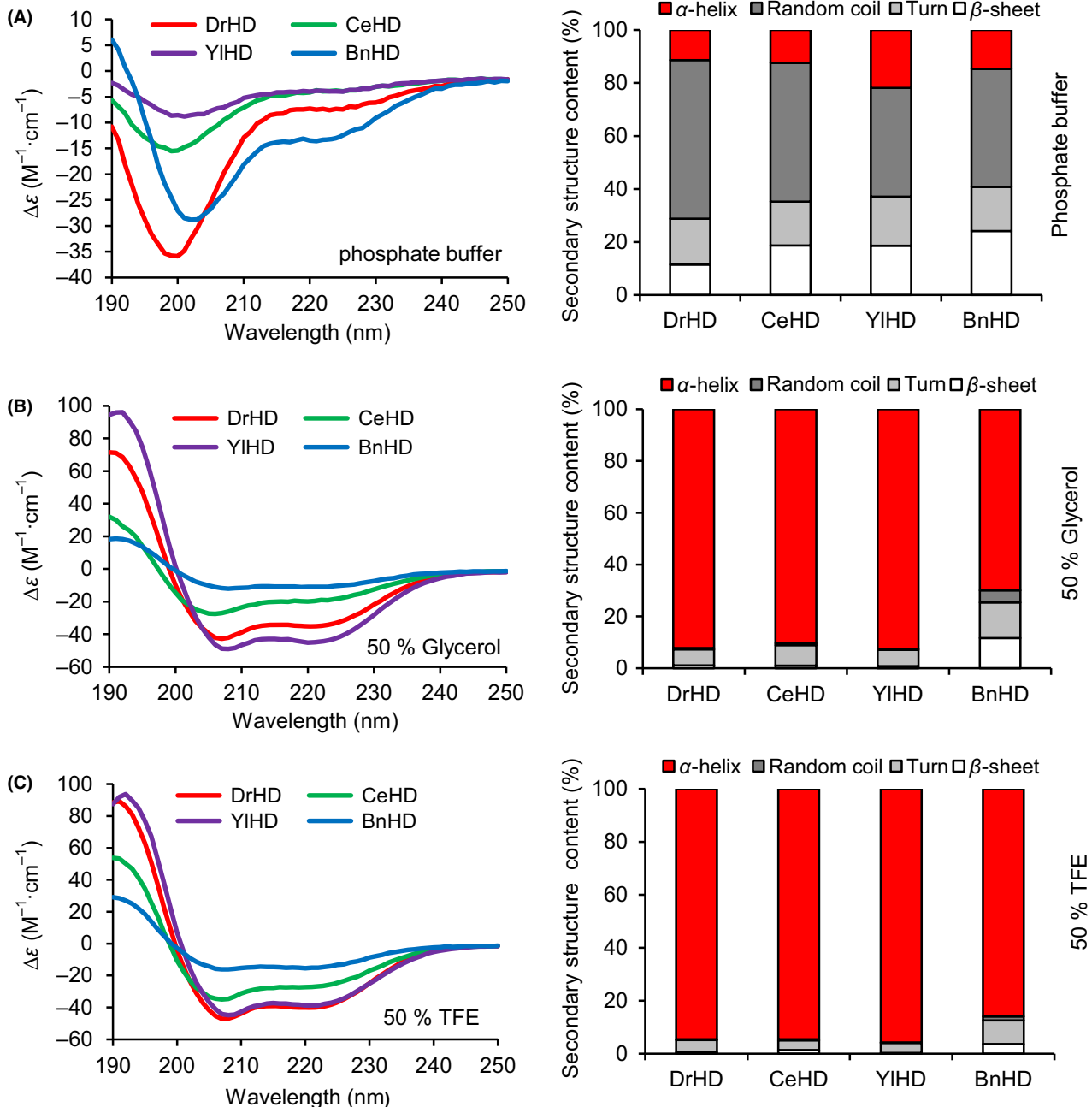
#### Hydrophilic domains (HDs) confer improved desiccation and oxidation resistance to *E. coli*

To explore the function of HDs against desiccation and oxidation, four *E. coli* transformants harbouring HD fragments were employed, and transformants harbouring pET28a served as the control. Clearly, compared to the control, the four transformants exhibited enhanced desiccation resistance after 10 days of treatment. In

particular, BL/DrHD displayed the strongest viability, approximately one order of magnitude higher than that of BL/CeHD (Fig. 2). Similarly, BL/DrHD or BL/CeHD exhibited improved oxidation tolerance that was approximately one order of magnitude stronger than those of BL/YIHD and BL/BnHD, which seemingly conform to the same trend as desiccation. With regard to oxidation resistance, further analysis revealed that both BL/DrHD and BL/CeHD transformants, compared to BL/YIHD and BL/BnHD, displayed minimum loss of antioxidant activity, as shown in Fig. S4. Taken together, these results not only indicated the capability of HDs to enhance desiccation and oxidation tolerance of *E. coli* but also proved that the level of  $\alpha$ -helix formation in G3LEA proteins tends to impact their physiological function, increased the  $\alpha$ -helix level, and improve stress resistance, combining with the secondary structure measurements (Fig. 1 and Table S2).

#### HDs reduce desiccation-induced enzyme aggregation and inactivation

These HDs from LEA proteins were highly hydrophilic and disordered, which were known to reduce desiccation-induced enzyme aggregation, such as lactate dehydrogenase (LDH), which catalyses the conversion of lactate to pyruvic acid (Goyal *et al.*, 2005; Chakrabortee *et al.*, 2007; Liu *et al.*, 2013). Therefore, we speculated that these HDs might perform stabilizing activity *in vitro* on LDH that was commonly used for aggregation and enzymatic activity protection assay (Reyes *et al.*, 2005), as well as bovine serum albumin (BSA) serving as the positive control. It was clear to see that LDH marked aggregation subjected to cycles of desiccation (Fig. 3A), addition of HDs can significantly decrease the aggregation of LDH compared to BSA mixture and negative control (only LDH), respectively, even after four cycles of desiccation, suggesting their aggregation-preventing ability. Noticeably, DrHD displayed the better job inhibiting the aggregation of LDH. The enzymatic activity of LDH was impaired under the dry conditions, especially after



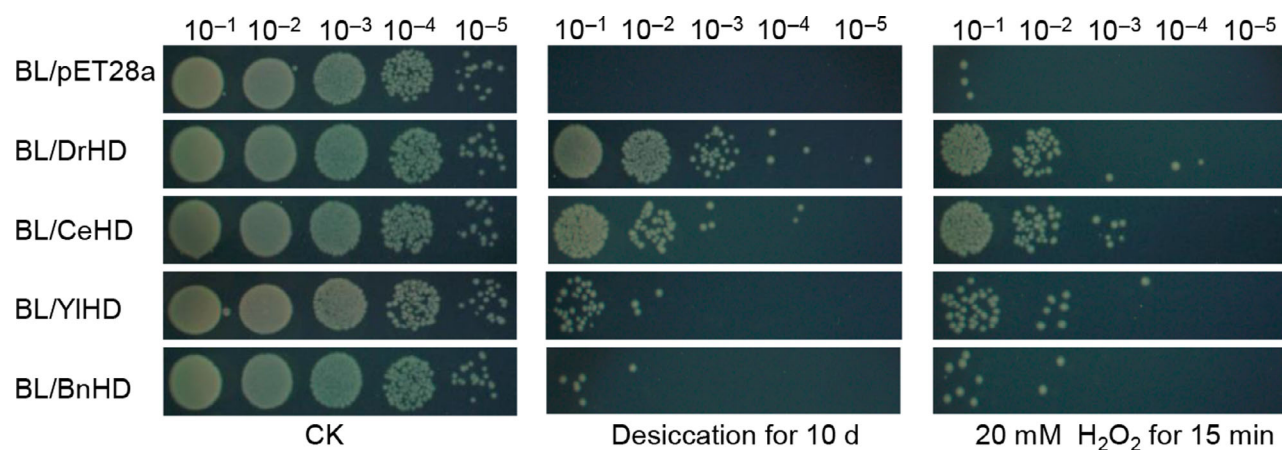
**Fig. 1.** CD analysis of HDs. The structural transformation of HDs was promoted by the water-deficient reagents 50% glycerol (B) and 50% TFE (C) compared to phosphate buffer (A). The secondary structure composition was estimated from CD spectra. Percentages of random coil (dark grey), turn (light grey), beta-sheet (white) and  $\alpha$ -helix (red) are deduced from CD spectra using the CDPPro program.

four cycles, and < 20% of the undried LDH activity was retained. Conspicuously, the loss of LDH enzymatic activity was significantly neutralized with the additional HDs and BSA. DrHD and CeHD performed excellently, preserving up to 60% of the LDH activity, and even BnHD protected up to 40% of the activity, which was slightly higher than the 38% protected by BSA (Fig. 3B). The HD protection of LDH enzymatic activity under oxidation stress followed a trend similar to that in

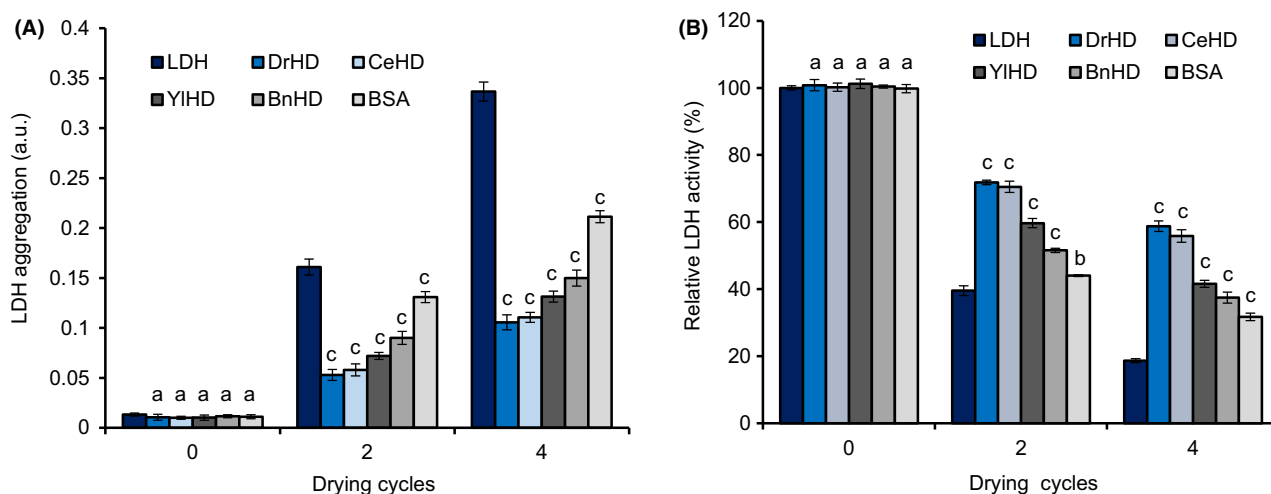
desiccation (Fig. S5). DrHD and CeHD clearly exhibited the strongest protection ability (over 65% of LDH activity retained) of the HDs.

#### *HDs can interact with lactate dehydrogenase (LDH)*

Many research articles have reported that the enzymatic activity protection of G3LEA proteins is achieved due to their binding of the target molecules and then enveloping



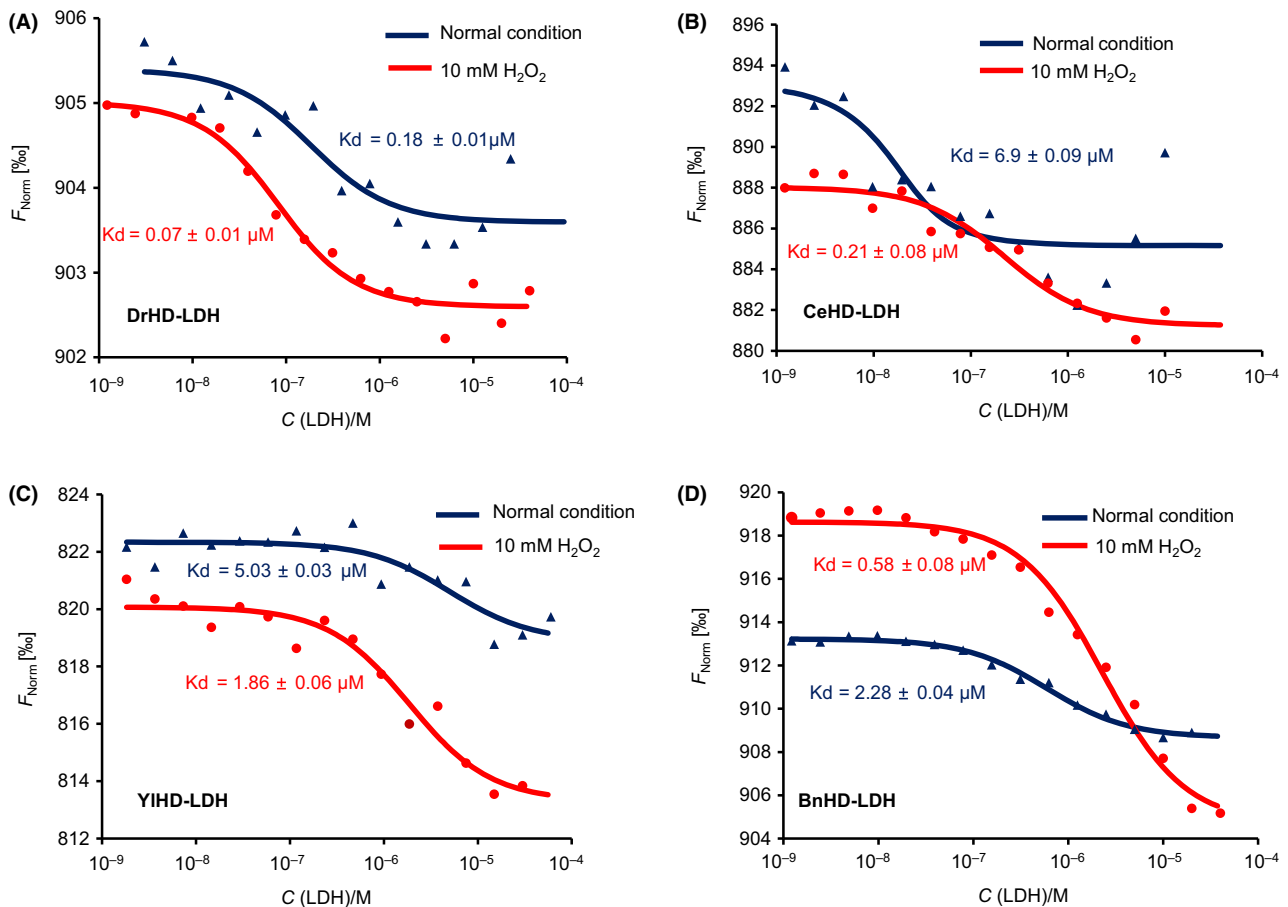
**Fig. 2.** Survival phenotype plate assay of *E. coli* recombinant strains under desiccation and oxidation. BL/DrHD, BL/CeHD, BL/YIHD and BL/BnHD are the expressing strains, and BL/pET28a recombinant carried an empty pET28a vector. Serial 10-fold dilutions of OD-standardized recombinant strains ( $OD \approx 0.6$ ) were spotted onto LB plates ( $6 \mu\text{l}$ ) after desiccation for 10 days and 20 mM  $\text{H}_2\text{O}_2$  treatment for 15 min; CK stands for untreated culture control.



**Fig. 3.** Effect of desiccation on LDH aggregation and activity. LDH aggregation (A) and (B) activity on repeated desiccation. The molar ratio of LDH and tested proteins was 1:3 on aggregation assay, while at 1:2.5 molar ratio on enzyme activity measurement. These measurements were performed three times for each case, and their  $P$ -values were calculated based on Tukey multiple comparisons using by R statistics. The letters 'a', 'b' and 'c' represent 'not significantly different ( $P > 0.05$ )', 'significantly different ( $0.001 < P < 0.01$ )' and 'extremely significantly different ( $P < 0.001$ )', respectively, compared to LDH alone.

them and allowing them to avoid direct exposure to harsh conditions (Chakrabortee *et al.*, 2010; Hundertmark *et al.*, 2011; Sasaki *et al.*, 2014). Thus, a protein–protein interaction assay was introduced to explore the binding ability of HDs. With regard to the dissociation constant ( $K_d$ ), which represents the affinity of the binding participants, the interaction was strengthened with decreasing  $K_d$  value. Accordingly, the protection levels of HDs were deliberately demonstrated. As the micro-scale thermophoresis results are shown in Fig. 4, all four HDs can bind LDH under normal conditions, and the

interaction became strengthened with the addition of 10 mM hydrogen peroxide, which not only revealed their chaperone-like function but also dynamically disclosed the complex configurational change. More importantly, binding curves illustrated that the DrHD/LDH pair displayed the strongest interaction, with a lower  $K_d$  value ( $K_d = 0.18 \pm 0.01 \mu\text{M}$ ) than CeHD ( $6.9 \pm 0.09 \mu\text{M}$ ), YIHD ( $5.03 \pm 0.03 \mu\text{M}$ ) and BnHD ( $2.28 \pm 0.04 \mu\text{M}$ ) under normal conditions, suggesting that DrHD had the strongest protein binding ability and that their binding was enhanced in the presence of 10 mM  $\text{H}_2\text{O}_2$  (Fig. 4A).



**Fig. 4.** Interaction of HDs with LDH. A protein–protein interaction method, microscale electrophoresis, was employed to characterize the relationship between LDH and DrHD (A), CeHD (B), YIHD (C) and BnHD (D). The  $K_d$  value represents the dissociation constant indicating their affinity. The binding curve under normal conditions (phosphate buffer) is shown in dark blue, and the binding curve under 10 mM  $\text{H}_2\text{O}_2$  treatment is shown in red (phosphate buffer supplemented with 10 mM  $\text{H}_2\text{O}_2$ ).

To some extent, the close complex formation between HDs and LDH under oxidation conditions reflected their structural flexibility.

## Discussion

Extensive studies focusing on an organism's adaptation against harsh environments have highlighted the importance of late embryogenesis abundant (LEA) proteins found in many species ranging from plants and animals to microbes as they were highly induced in dry environments (Tompa, 2002; Gal *et al.*, 2004; Tolleter *et al.*, 2007; Chakrabortee *et al.*, 2010; Hand *et al.*, 2011). The accumulated LEA proteins act in a chaperone-like manner to protect the organisms from various stresses, for example, salt, desiccation and oxidation (Hoekstra *et al.*, 2001; Shimizu *et al.*, 2010; Cuevas-Velazquez *et al.*, 2016; Wang *et al.*, 2017). Further bioinformatic and structural analyses revealed that 11-mer motifs specific to group 3 LEA (G3LEA) proteins (HD) might be the crucial regions

determining their whole-length functions (Dure, 1993; Xue *et al.*, 2011). Unfortunately, detailed research focusing on this hydrophilic region, such as the study of its physiological function, structural transformation and protective mechanism, has rarely been reported. The annotation of the genome of *D. radiodurans*, the world's most tenacious extremophile model against radiation, desiccation and oxidation (Chen *et al.*, 2017; Jeong *et al.*, 2017; Schmier *et al.*, 2017; Ujaoney *et al.*, 2017; Zhou *et al.*, 2017), indicated that there were three putative LEA homologous proteins (Makarova *et al.*, 2001), in which a gene named *dr\_1172* bearing classic 11-mer repeating motifs (hydrophilic domain, HD) was functionally characterized and responsible for the strain's desiccation resistance (Battista *et al.*, 2001). Therefore, HD from *D. radiodurans* together with other three sequence-similar HDs, CeHD, BnHD and YIHD, representing animal, plant and fungus, respectively, was investigated in this study.

Protein sequence analysis revealed that the motifs were conserved among these four HDs (Table 1), which

to some extent indicated their structural and functional similarity. *In silico* analysis revealed that these four HDs exhibited disordered structures that were abundant with random coils, which was verified by using circular dichroism (CD) spectrometry in phosphate buffer (Fig. 1). Due to their high proportion of hydrophilic amino acid residues, these LEA proteins were mainly converted from disordered status into regular  $\alpha$ -helices under desiccation condition (Hand *et al.*, 2011) such as the presence of glycerol or TFE, which are commonly used reagents used to characterize  $\alpha$ -helix formation potential (Cuevas-Velazquez *et al.*, 2016). Structural identification showed the potential of these HDs to form  $\alpha$ -helices, which was in agreement with the LEA structural features (Hand *et al.*, 2011), implicating that the decisive role of HD holding the whole-length structure conversion under water-loss environments. More interestingly, the  $\alpha$ -helix formation level varied and might be correlated with their physiological function, as compared to its counterparts shown in Fig. 1 and Fig. 2, and BnHD performed slight desiccation and oxidation enhancement with the least  $\alpha$ -helix formation. The high configurational flexibility of intrinsically disordered proteins (IDPs), including LEA, was identified to function in a chaperone-like manner by preventing the protein's aggregation and protecting the enzymes' activity (Chakrabortee *et al.*, 2007, 2010; Toller *et al.*, 2007). Apparently, all four of these HDs reduced desiccation- and oxidation-induced enzyme aggregation and inactivation (Fig. 3 and Fig. S5), from which their important function similar to intensively studied chaperones, such as HSPs, was indicated (Dure, 1993; Kovacs and Tompa, 2012; Salleh *et al.*, 2012; Hatanaka *et al.*, 2013). But Intrinsically, the protective role of HD, same with LEA proteins that had been previously termed as a molecular shield or chaperone-like activity (Chakrabortee *et al.*, 2010, 2012; Kovacs and Tompa, 2012), was distinct from identified canonical chaperones. In structure, these four HDs same to are naturally in an unstructured state, while the classic HSPs are well-structured with just a small disordered segment (Kovacs and Tompa, 2012). As for mechanism, these HSPs were assembled together to form a cooperative machinery undergoing protein recognition, disaggregation and ATP-driving refold processes (Chakrabortee *et al.*, 2007; Hand *et al.*, 2011), whereas our HD studies revealed that they can weakly bind to enzymes such as LDH as indicated in Fig. 4, which is similar to a previous study explained as steric interference (Chakrabortee *et al.*, 2012; Kovacs and Tompa, 2012), and intriguingly, their interaction was strengthened with addition of H<sub>2</sub>O<sub>2</sub>, which might be induced by the partially binding of HD with misfold to prevent further collisions (Kovacs and Tompa, 2012). Besides, HDs can also bind to MDH (data not shown) implicating that there are no specific

binding sites, which is totally different with molecular chaperone, like HSP70, co-immunoprecipitating with a polyQ-containing protein (Chakrabortee *et al.*, 2012). In addition, these HDs can inhibit LDH aggregation and protect its enzymatic activity without additional energy (Olvera-Carrillo *et al.*, 2011; Schopf *et al.*, 2017) as well as without carbohydrates, such as trehalose (Kovacs and Tompa, 2012). Combining together, HD served as the molecular basis of G3LEA protection mechanism as shown in Fig. S6.

Interestingly, these HDs had *in vivo* antioxidant activity, which was seen as a moonlighting function supplementing their anti-desiccation ability (Haaning *et al.*, 2008). Certainly, the relationship between oxidation and desiccation was clearly demonstrated before the impact of desiccation was determined by the levels of biomolecule oxidation (Fredrickson *et al.*, 2008). In other words, the fewer the oxidized metabolites produced, the higher the desiccation resistance. Apart from that, the similar effect trend under both conditions through heterologous expression and LDH enzymatic activity protection was noticeable. Oxidation can cause direct damage of proteins by destroying their three-dimensional structure, leading to misfolding and aggregation, which is intrinsically different from the generated effect from glycerol or TFE that were used to identify LEA or IDPs potential forming  $\alpha$ -helix as well as used to stabilize the target proteins when exposed to harsh conditions (Corrêa and Farah, 2007; Cuevas-Velazquez *et al.*, 2016), and there was rarely reported that both of them were seen as damage-caused reagents on biomolecules. Therefore, it was not difficult to understand that the addition of glycerol or TFE cannot cause the binding affinity change in LDH with HDs.

Obviously, CeHD and DrHD played an outstanding role in greatly increasing *E. coli* viability, preventing protein aggregation, minimizing the loss of LDH enzymatic activity and increasing binding ability. Compared to CeHD, DrHD is very advantageous for synthetic biology due to its small size and its equal or better efficacy. Further sequence analysis revealed that both DrHD and CeHD motifs were more conserved than those of the other two HDs (Fig. S2), especially key residues (Dure, 1993, 2001). Therefore, explorations of the crucial key residues would be meaningful to unravel the structural and functional basis of HDs. Our results implicated that these 11-mer repeating motifs were responsible for the full-length physiological function, providing guidance for future investigations as to whether their function is related to the motif numbers since the pioneering work indicative of the increased salt tolerance of duplicated motifs (Liu and Zheng, 2005).

In conclusion, we first characterized four putative HDs that were indispensable to G3LEA protein function

undergoing structural transformation and providing chaperone-like protection. These selected HDs enhanced *E. coli* desiccation and oxidation resistance, and furthermore, these HDs can inhibit desiccation-induced aggregation and protect enzyme activity. Most importantly, the great performance of DrHD encouraged us to continue exploring its specificity and the hidden robustness of the G3LEA proteins, and inspired us to look for brief functional elements in future work for biosynthetic applications.

## Experimental procedures

### Sample acquisition and growth conditions

The bacterial strains and plasmids used in this study are described in Table 2. *E. coli* strains were cultured at 37°C in LB broth (1% tryptone, 0.5% yeast extract and 0.5% NaCl) or on LB plates with 50 µg ml<sup>-1</sup> kanamycin or the appropriate antibiotic.

### Construction of heterologous expression system for DrHD, CeHD, YIHD and BnHD in *E. coli*

DrHD was amplified from the genomic DNA of *D. radiodurans* using specific primers. Similarly, the original nucleotide sequences of the CeHD, YIHD and BnHD genes were engineered and synthesized by BGI (Beijing, China). All forward primers contained a BamHI site

(underlined), and the reverse primers contained a XhoI site (underlined). Four PCR products were cloned into a pET28a vector using the Seamless Assembly Cloning Kit (Cat# C5891-25; Clone Smarter, Houston, Texas, USA). All recombinant plasmids (pET-DrHD, pET-CeHD, pET-YIHD, pET-BnHD) and an empty pET28a vector were introduced into *E. coli* host strain BL21 (DE3) to generate the transformed strains BL/DrHD, BL/CeHD, BL/YIHD, BL/BnHD and BL/pET28a (control). Primers were listed in Table S1.

### Circular dichroism (CD) spectroscopic analysis

Far-UV circular dichroism (CD) spectra were generated using a Jasco-720 spectropolarimeter (Jasco Instrument, Tokyo, Japan) with a 0.1 mm path length cuvette at wavelengths of 190–250 nm as previously described in detail (Chakrabortee *et al.*, 2010). Recombinant proteins were diluted to 0.2 mg ml<sup>-1</sup> in 10 mM phosphate buffer (pH 7.0). Samples were placed into the spectrometer on a single window and scanned immediately. Relative secondary structure content of all proteins was analysed with CDPPro software.

### Desiccation and oxidation stress tolerance in *E. coli*

*Escherichia coli* recombinant strain cultures were induced with IPTG at a final optimal concentration of 0.1 mM for 4 h, and the cells were desiccated and treated with H<sub>2</sub>O<sub>2</sub> as previously described (Rajpurohit and Misra, 2013; Jiang *et al.*, 2017; Wang *et al.*, 2017). Different serial dilutions of these cells were plated onto LB agar plates and incubated at 37°C for 1 day before colonies were observed and counted. All experiments were performed three times. See Supporting Experimental procedures for more details.

### Lactate dehydrogenase (LDH) aggregation and activity assay in vitro

Lactate dehydrogenase (LDH) aggregation and activity was assayed as described previously with some modifications (Goyal *et al.*, 2005; Hatanaka *et al.*, 2013; Liu *et al.*, 2013). For the aggregation measurement, 100 µl of 2 mg ml<sup>-1</sup> LDH was mixed with or without purified recombinant proteins at molar ratio of 3:1 (test protein: LDH in an Eppendorf tube. And then samples were set in a speed-vac VC-960 for desiccation by centrifugal concentration under reduced pressure for 4 h. Once completely desiccated, these samples were rehydrated in the same volume of 10 mM phosphate buffer. Samples were then dried, and the rehydration cycle was repeated. Aggregation was assessed by measuring apparent absorbance at 340 nm. All assays were conducted in a total starting volume of

**Table 2.** Bacterial strains and plasmids used in this study.

Strains/plasmids	Description	Source
<b>Strains</b>		
Trans10	Host for cloning vectors	TransGen Biotech
BL21	F <sup>-</sup> ompT hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) gal dcm (DE3)	TransGen Biotech
BL/pET28a	The control strain harbouring the empty vector pET28a, Kan <sup>r</sup>	This study
BL/DrHD	The recombinant BL21 strain containing pET-DrHD, Kan <sup>r</sup>	This study
BL/CeHD	The recombinant BL21 strain containing pET-CeHD, Kan <sup>r</sup>	This study
BL/YIHD	The recombinant BL21 strain containing pET-YIHD, Kan <sup>r</sup>	This study
BL/BnHD	The recombinant BL21 strain containing pET-BnHD, Kan <sup>r</sup>	This study
<b>Plasmid</b>		
pET28a (+)	pBR322 and f1 ori, KmR, commercial vector for protein overexpression	Novagen
pET-DrHD	pET28a-derived plasmid carrying the DrHD gene with HD	This study
pET-CeHD	pET28a-derived plasmid carrying the CeHD gene with HD	This study
pET-YIHD	pET28a-derived plasmid carrying the YIHD gene with HD	This study
pET-BnHD	pET28a-derived plasmid carrying the BnHD gene with HD	This study



200  $\mu\text{l}$  and were assayed in triplicate. For enzymatic activity measurement, LDH and additives were diluted in 25 mM Tris–HCl (pH 7.5) to a final concentration of 8.3  $\mu\text{g ml}^{-1}$ . For desiccation treatment, the enzyme mixture was left in a desiccator box for 24 h and then rehydrated in the same volume of 25 mM Tris–HCl (pH 7.5). LDH activity was monitored as the rate of absorbance change at 340 nm for 1 min due to the conversion of NADH to NAD<sup>+</sup> at 25°C. The rate determined for the untreated samples was considered 100% in all graphs. See more details in Supporting Experimental procedures.

#### Microscale thermophoresis (MST) measurements

Microscale thermophoresis (MST) experiments were performed as described previously (Zhan *et al.*, 2016). The purified proteins (DrHD, CehD, YIHD and BnHD) were labelled with the Monolith NTTM Protein Labeling Kit RED-NHS according to the supplied labelling protocol (Cat# L011; NanoTemper Technologies Co., Ltd., Chaoyang, Beijing, China). Labelled proteins were used at a concentration of 2  $\mu\text{M}$ . Label-free LDH was used as the ligand. LDH was diluted from 71  $\mu\text{M}$  in binding buffer (20 mM HEPES, 150 mM KCl, pH 7.4). For the oxidation treatment, H<sub>2</sub>O<sub>2</sub> solution was added to the enzyme–protein mixture to obtain a final concentration of 10 mM. Ten microlitres of the LDH solution was mixed with 10  $\mu\text{l}$  of labelled protein. After mixing, the samples were loaded on standard treated capillaries. The measurements were taken using a Monolith NT.115 (NanoTemper Technologies) that was operated at 20% light-emitting diode (LED) and 60% MST power. Data analyses were performed using NanoTemper Analysis software v.1.2.101 (NanoTemper Technologies).

#### Acknowledgements

This study was supported by the Ministry of Agriculture Transgenic Program (No. 2016ZX08009003-002) and National Natural Science Foundation of China (No. 31800061, 31570080, 31500063).

#### Conflict of interest

None declared.

#### References

- Atashgahi, S., Sánchez-Andrea, I., Heipieper, H.J., van der Meer, J.R., Stams, A.J.M., and Smidt, H. (2018) Prospects for harnessing biocide resistance for bioremediation and detoxification. *Science* **360**: 743–746.
- Battaglia, M., Olvera-Carrillo, Y., Garcarrubio, A., Campos, F., and Covarrubias, A.A. (2008) The enigmatic LEA proteins and other hydrophilins. *Plant Physiol* **148**: 6–24.
- Battista, J.R., Park, M.-J., and McLemore, A.E. (2001) Inactivation of two homologues of proteins presumed to be involved in the desiccation tolerance of plants sensitizes *Deinococcus radiodurans* R1 to desiccation. *Cryobiology* **43**: 133–139.
- Bremer, A., Wolff, M., Thalhammer, A., and Hinch, D.K. (2017) Folding of intrinsically disordered plant LEA proteins is driven by glycerol-induced crowding and the presence of membranes. *FEBS J* **284**: 919–936.
- Chakrabortee, S., Boschetti, C., Walton, L.J., Sarkar, S., Rubinsztein, D.C., and Tunnacliffe, A. (2007) Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. *Proc Natl Acad Sci USA* **104**: 18073–18078.
- Chakrabortee, S., Meersman, F., Kaminski Schierle, G.S., Bertocini, C.W., McGee, B., Kaminski, C.F., and Tunnacliffe, A. (2010) Catalytic and chaperone-like functions in an intrinsically disordered protein associated with desiccation tolerance. *Proc Natl Acad Sci USA* **107**: 16084–16089.
- Chakrabortee, S., Tripathi, R., Watson, M., Kaminski Schierle, G.S., Kurniawan, D.P., Kaminski, C.F., *et al.* (2012) Intrinsically disordered proteins as molecular shields. *Mol Biosyst* **8**: 210–219.
- Chen, A., Contreras, L.M., and Keitz, B.K. (2017) Imposed environmental stresses facilitate cell-free nanoparticle formation by *Deinococcus radiodurans*. *Appl Environ Microbiol* **83**: e00798–17.
- Corrêa, F., and Farah, C.S. (2007) Different effects of trifluoroethanol and glycerol on the stability of tropomyosin helices and the head-to-tail complex. *Biophys J* **92**: 2463–2475.
- Cuevas-Velazquez, C.L., Saab-Rincón, G., Reyes, J.L., and Covarrubias, A.A. (2016) The unstructured N-terminal region of *Arabidopsis* group 4 late embryogenesis abundant (LEA) proteins is required for folding and for chaperone-like activity under water deficit. *J Biol Chem* **291**: 10893–10903.
- Daly, M.J. (2009) A new perspective on radiation resistance based on *Deinococcus radiodurans*. *Nat Rev Microbiol* **7**: 237–245.
- Dure, L. III (1993) A repeating 11-mer amino acid motif and plant desiccation. *Plant J* **3**: 363–369.
- Dure, L. III (2001) Occurrence of a repeating 11-mer amino acid sequence motif in diverse organisms. *Protein Pept Lett* **8**: 115–122.
- Dure, L. III, and Galau, G.A. (1981) Developmental biochemistry of cottonseed embryogenesis and germination. XIII. Purification and properties of principal storage proteins. *Plant Physiol* **68**: 187–194.
- Fredrickson, J.K., Li, S.M.W., Gaidamakova, E.K., Matrosova, V.Y., Zhai, M., Sulloway, H.M., *et al.* (2008) Protein oxidation: key to bacterial desiccation resistance? *ISME J* **2**: 393–403.
- Gal, T.Z., Glazer, I., and Koltai, H. (2004) An LEA group 3 family member is involved in survival of *C. elegans* during exposure to stress. *FEBS Lett* **577**: 21–26.
- Goyal, K., Walton, L.J., and Tunnacliffe, A. (2005) LEA proteins prevent protein aggregation due to water stress. *Biochem J* **388**: 151–157.

- Haaning, S., Radutoiu, S., Hoffmann, S.V., Dittmer, J., Giehm, L., Otzen, D.E., and Stougaard, J. (2008) An unusual intrinsically disordered protein from the model legume *Lotus japonicus* stabilizes proteins in vitro. *J Biol Chem* **283**: 31142–31152.
- Hand, S.C., Menze, M.A., Toner, M., Boswell, L., and Moore, D. (2011) LEA proteins during water stress: not just for plants anymore. *Annu Rev Physiol* **73**: 115–134.
- Hatanaka, R., Hagiwara-Komoda, Y., Furuki, T., Kanamori, Y., Fujita, M., Cornette, R., et al. (2013) An abundant LEA protein in the anhydrobiotic midge, PvLEA4, acts as a molecular shield by limiting growth of aggregating protein particles. *Insect Biochem Mol Biol* **43**: 1055–1067.
- Hoekstra, F.A., Golovina, E.A., and Buitink, J. (2001) Mechanisms of plant desiccation tolerance. *Trends Plant Sci* **6**: 431–438.
- Hughes, D.W., and Galau, G.A. (1989) Temporally modular gene expression during cotyledon development. *Genes Dev* **3**: 358–369.
- Hundertmark, M., Dimova, R., Lengefeld, J., Seckler, R., and Hinch, D.K. (2011) The intrinsically disordered late embryogenesis abundant protein LEA18 from *Arabidopsis thaliana* modulates membrane stability through binding and folding. *Biochim Biophys Acta* **1808**: 446–453.
- Hundertmark, M., Popova, A.V., Rausch, S., Seckler, R., and Hinch, D.K. (2012) Influence of drying on the secondary structure of intrinsically disordered and globular proteins. *Biochem Biophys Res Commun* **417**: 122–128.
- Jeong, S.W., Yang, J.E., Im, S., and Choi, Y.J. (2017) Development of Cre-lox based multiple knockout system in *Deinococcus radiodurans* R1. *Korean J Chem Eng* **34**: 1728–1733.
- Jiang, S., Wang, J., Liu, X., Liu, Y., Guo, C., Zhang, L., et al. (2017) DrwH, a novel WHy domain-containing hydrophobic LEA5C protein from *Deinococcus radiodurans*, protects enzymatic activity under oxidative stress. *Sci Rep* **7**: 1–10.
- Kovacs, D., and Tompa, P. (2012) Diverse functional manifestations of intrinsic structural disorder in molecular chaperones. *Biochem Soc Trans* **40**: 963–968.
- Kovacs, D., Agoston, B., and Tompa, P. (2008) Disordered plant LEA proteins as molecular chaperones. *Plant Signal Behav* **3**: 710–713.
- Kyte, J., and Doolittle, R.F. (1982) A simple method for displaying the hydrophobic character of a protein. *J Mol Biol* **157**: 105–132.
- Leeuwen, M.R., Wyatt, T.T., van Doorn, T.M., Lugones, L.G., Wösten, H.A., and Dijksterhuis, J. (2016) Hydrophilins in the filamentous fungus *Neosartorya fischeri* (*Aspergillus fischeri*) have protective activity against several types of microbial water stress. *Environ Microbiol Rep* **8**: 45–52.
- Liu, Y., and Zheng, Y. (2005) PM2, a group 3 LEA protein from soybean, and its 22-mer repeating region confer salt tolerance in *Escherichia coli*. *Biochem Biophys Res Commun* **331**: 325–332.
- Liu, Y., Wang, L., Xing, X., Sun, L., Pan, J., Kong, X., et al. (2013) ZmLEA3, a multifunctional group 3 LEA protein from maize (*Zea mays* L.), is involved in biotic and abiotic stresses. *Plant Cell Physiol* **54**: 944–959.
- Makarova, K.S., Aravind, L., Wolf, Y.I., Tatusov, R.L., Minton, K.W., Koonin, E.V., and Daly, M.J. (2001) Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics. *Microbiol Mol Biol Rev* **65**: 44–79.
- Olvera-Carrillo, Y., Reyes, J.L., and Covarrubias, A.A. (2011) Late embryogenesis abundant proteins: versatile players in the plant adaptation to water limiting environments. *Plant Signal Behav* **6**: 586–589.
- Rajpurohit, Y.S., and Misra, H.S. (2013) DR1769, a protein with N-terminal beta propeller repeats and a low-complexity hydrophilic tail, plays a role in desiccation tolerance of *Deinococcus radiodurans*. *J Bacteriol* **195**: 3888–3896.
- Reyes, J.L., Rodrigo, M.-J., Colmenero-Flores, J.M., Gil, J.-V., Garay-Arroyo, A., Campos, F., et al. (2005) Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects in vitro. *Plant, Cell Environ* **28**: 709–718.
- Ried, J.L., and Walker-Simmons, M.K. (1993) Group 3 late embryogenesis abundant proteins in desiccation-tolerant seedlings of wheat (*Triticum aestivum* L.). *Plant Physiol* **102**: 125–131.
- Roberts, J.K., DeSimone, N.A., Lingle, W.L., and Dure, L. III. (1993) Cellular concentrations and uniformity of cell-type accumulation of two lea proteins in cotton embryos. *Plant Cell* **5**: 769–780.
- Salleh, F.M., Evans, K., Goodall, B., Machin, H., Mowla, S.B., Mur, L.A.J., et al. (2012) A novel function for a redox-related LEA protein (SAG21/AtLEA5) in root development and biotic. *Plant, Cell Environ* **35**: 418–429.
- Sasaki, K., Christov, N.K., Tsuda, S., and Imai, R. (2014) Identification of a novel LEA protein involved in freezing tolerance in wheat. *Plant Cell Physiol* **55**: 136–147.
- Schmier, B.J., Chen, X., Wolin, S., and Shuman, S. (2017) Deletion of the *ml* gene encoding a nick-sealing RNA ligase sensitizes *Deinococcus radiodurans* to ionizing radiation. *Nucleic Acids Res* **45**: 3812–3821.
- Schopf, F.H., Biebl, M.M., and Buchner, J. (2017) The HSP90 chaperone machinery. *Nat Rev Mol Cell Biol* **18**: 345–360.
- Shih, M.D., Hsieh, T.Y., Jian, W.T., Wu, M.T., Yang, S.J., Hoekstra, F.A., and Hsing, Y.I.C. (2012) Functional studies of soybean (*Glycine max* L.) seed LEA proteins GmPM6, GmPM11, and GmPM30 by CD and FTIR spectroscopy. *Plant Sci* **196**: 152–159.
- Shimizu, T., Kanamori, Y., Furuki, T., Kikawada, T., Okuda, T., Takahashi, T., et al. (2010) Desiccation-induced structuralization and glass formation of group 3 late embryogenesis abundant protein model peptides. *Biochemistry* **49**: 1093–1104.
- Slade, D., and Radman, M. (2011) Oxidative stress resistance in *Deinococcus radiodurans*. *Microbiol Mol Biol Rev* **75**: 133–191.
- Swire-Clark, G.A., and Marcotte, W.R. Jr (1999) The wheat LEA protein Em functions as an osmoprotective molecule in *Saccharomyces cerevisiae*. *Plant Mol Biol* **39**: 117–128.
- Tolletter, D., Jaquinod, M., Mangavel, C., Passirani, C., Saulnier, P., Manon, S., et al. (2007) Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell* **19**: 1580–1589.

- Tompa, P. (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* **27**: 527–533.
- Ujaoney, A.K., Padwal, M.K., and Basu, B. (2017) Proteome dynamics during post-desiccation recovery reveal convergence of desiccation and gamma radiation stress response pathways in *Deinococcus radiodurans*. *Biochim Biophys Acta* **1865**: 1215–1226.
- Wang, H., Wu, Y., Yang, X., Guo, X., and Cao, X. (2017) *SmLEA2*, a gene for late embryogenesis abundant protein isolated from *Salvia miltiorrhiza*, confers tolerance to drought and salt stress in *Escherichia coli* and *S. miltiorrhiza*. *Protoplasma* **254**: 685–696.
- Xue, R., Liu, Y., Zheng, Y., Wu, Y., Li, X., Pei, F., and Ni, J. (2011) Three-dimensional structure and mimetic-membrane association of consensus 11-amino-acid motif from soybean LEA3 protein. *Biopolymers* **98**: 59–66.
- Zhan, Y., Yan, Y., Deng, Z., Chen, M., Lu, W., Lu, C., *et al.* (2016) The novel regulatory ncRNA, NfiS, optimizes nitrogen fixation via base pairing with the nitrogenase gene *nifK* mRNA in *Pseudomonas stutzeri* A1501. *Proc Natl Acad Sci USA* **113**: E4348–E4356.
- Zhou, Y., Shen, P., Lan, Q., Deng, C., Zhang, Y., Li, Y., *et al.* (2017) High-coverage proteomics reveals methionine auxotrophy in *Deinococcus radiodurans*. *Proteomics* **17**: 13–14.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** Hydropathic index plot of the HD amino acid sequences analyzed by using the Kyte-Doolittle algorithm. Regions with a hydropathy score below zero are hydrophilic.

**Fig. S2.** Comparison of the repeating 11-mer motif in HDs.

**Fig. S3.** SDS-PAGE analysis of four purified HD proteins. Lane 1, purified BnHD; lane 2, purified DrHD; lane 3, purified CeHD; lane 4, purified YIHD; lane M, molecular weight marker (kDa).

**Fig. S4.** The total antioxidant activity of the transformants was detected by a spectrometer.

**Fig. S5.** The role of HDs in protecting LDH against H<sub>2</sub>O<sub>2</sub> stress.

**Fig. S6.** Chaperone-like function model of HD. HD can bind with native proteins, especially crucial enzymes in metabolic pathways, in a weak manner under normal conditions.

**Table S1.** List of primers used in this study.

**Table S2.** Secondary structure content in four HD proteins was obtained by far-UV CD spectrometry and calculated with the CDPro program.