Design of an expression system to enhance MBP-mediated crystallization

Tengchuan Jin1,2, Watchalee Chuenchor2, Jiansheng Jiang2, Jinbo Cheng1, Yajuan Li1, Kang Fang1, Mo Huang2, Patrick Smith2 & Tsan Sam Xiao3

Crystallization chaperones have been used to facilitate the crystallization of challenging proteins. Even though the maltose-binding protein (MBP) is one of the most commonly used crystallization chaperones, the design of optimal expression constructs for crystallization of MBP fusion proteins remains a challenge. To increase the success rate of MBP-facilitated crystallization, a series of expression vectors have been designed with either a short flexible linker or a set of rigid helical linkers. Seven death domain superfamily members were tested for crystallization with this set of vectors, six of which had never been crystallized before. All of the seven targets were crystallized, and their structures were determined using at least one of the vectors. Our successful crystallization of all of the targets demonstrates the validity of our approach and expands the arsenal of the crystallization chaperone toolkit, which may be applicable to crystallization of other difficult protein targets, as well as to other crystallization chaperones.

Recombinant expression of fusion proteins containing a target protein and an unrelated expression tags is a common strategy to obtain samples for medical and research applications. Depending on their function and properties, these tags can be divided into several categories including purification tags, stability tags, and function tags. Many proteins have been tested as expression tags including the glutathione S-transferase (GST), small ubiquitin-like modifier (SUMO), thioredoxin (TRX), disulfide oxidoreductase (DsbA), N utilization substance protein A (NusA), B1 domain of protein G (GB1), oligopeptide derived from pancreatic ribonuclease A (S-tag), and maltose-binding protein (MBP).

MBP is one of the most commonly used protein expression tags due to its exceptional performance in enhancing the solubility of the target proteins. It is the periplasmic portion of the ATP-binding cassette (ABC) maltose/maltodextrin transporter found in many bacterial species. MBP has a large protein-protein interaction surface on the same side of the maltose ligand-binding pocket, through which it binds to the transmembrane domain of the ABC transporter. It is commonly used as an affinity tag for purification, which takes advantage of its ability to bind amylose. While the mechanism of the solubility-enhancing property of MBP is still not completely understood, its large size, intrinsic solubility, abundance of flexible loops, and high percentage of exposed hydrophilic residues may be contributing factors. Recent studies indicate that MBP does not have chaperone activity originally thought to facilitate the folding of the fused proteins.

Obtaining diffraction-quality crystals remains the rate-limiting step in solving a novel structure using X-ray crystallography, despite the recent advancements in the areas of molecular cloning, protein production and purification, crystallization screening, data collection using synchrotron radiation, and structure determination programs. Many tools have been developed and employed in the structural biology community to enhance the crystallizability of proteins, including in situ proteolysis, lysine methylation, surface entropy reduction, metal mediated crystallization, polymer-driven crystallization, site-specific mutagenesis, and chaperone-aided crystallization (also see recent reviews). In particular, many high-affinity binding partner molecules have been used to facilitate the crystallization of challenging protein targets such as the G
complex structures have been determined through X-ray crystallography, including the PIDDosome\(^4\), FAS/ 
third of the total 99 death domain superfamily members in the human genome, and seven death-fold protein 
validity of our approach and expands the arsenal of the crystallization chaperone toolkit.

The design of the linker sequences takes into consideration the rigidity of the helical structure, the various spac-
linker sequences, and selected seven death domains as targets, six of which have never been crystallized before.
We hypothesize that optimization of the linker regions between MBP and the target protein will increase the like-
usage of linkers of various lengths and test their effects on crystallization of death domain superfamily members.

**Vector Name** | **Linker Sequence** | **Linker Structure**
--- | --- | ---
V28E | VD | Loop
V28E1 | AAA | Helix
V28E2 | AAAA | Helix
V28E3 | AAAA | Helix
V28E4 | AAAAA | Helix
V28E5 | AARAAAA | Helix
V28E6 | AARAFAAA | Helix

| Vector Name | Linker Sequence | Linker Structure |
|---|---|---|
| V28E1 | AAA | Helix |
| V28E2 | AAAA | Helix |
| V28E3 | AAATA | Helix |
| V28E4 | AAAAA | Helix |
| V28E5 | AARAAAA | Helix |
| V28E6 | AARAFAAA | Helix |

**Methods**
**Data base mining.** A BLAST search using the *E. coli* MBP sequence was performed with the Protein Data 
Bank (www.rcsb.org) to collect information on all deposited structures containing an MBP tag. All of the hits were manually inspected for information of the target proteins such as size and isoelectric point (pI).

**Design of the MBP expression vectors.** The coding sequence for an *E. coli* MBP protein mutant “E” 
(D82A/K83A/E172A/N173A/K239A)\(^1\) was amplified by PCR and ligated into a T7 promoter-based bacterial 
protein expression vector pET30a (Novagen) using restriction enzyme sites Nde I and Xho I (sequence is shown 
in Table S2). A Sal I site was introduced following the MBP residue A369. This vector was designated “V28E” that 
features a short linker of residues VD encoded by the Sal I site. Building on vector V28E, vectors V28E1-E6 were 
designed to harbor helical links with one residue increment (Table 1). Gene of interest can be ligated into this set of vectors using the same restriction enzymes Not I and Xho I, with the Not I site encoding residues AAA that adopts
helical structures. A hexa-histidine tag sequence following the Xho I site was retained from the pET30a backbone to facilitate the purification of the fusion protein through immobilized metal affinity chromatography (IMAC).

**Protein expression, purification and crystallization.** Expression and purification of the MBP-tagged proteins have been described previously\(^ {47,48} \). Briefly, transformed BL21 (DE3) Codon Plus RILPI cells (Stratagene, Santa Clara, CA) were grown at 37 °C and protein expression was induced at 18 °C for at least two hours with 0.3 mM IPTG. Cells were lysed by sonication and the recombinant protein was purified by IMAC columns followed by size-exclusion chromatography. The purified protein was concentrated to 20–50 mg/ml in the presence of 10 mM maltose before setting up crystallization screening.

Coding sequences for seven death domain superfamily members (hNLRP1-CARD, hAIM2-PYD, hCARD8-CARD, zIGBP1-CARD, hNLRP1-PYD, hNLRP12-PYD, and hMNDA-PYD) were ligated into each of the V28E and V28E1-E6 vectors. Several domain boundaries were tested for each of the targets, which resulted in more than 50 expression constructs (Table S2). Crystallization of these fusion proteins was screened using the Mosquito crystallization robot (TTP Labtech, United Kingdom). For a particular target protein, once the structure was determined using one of the constructs, further exploration of crystallization for the same target was terminated. It is therefore possible that some additional crystallization conditions for the target might be missed.

**X-ray diffraction data collection and structure determination.** X-ray diffraction data were collected at GM/CA beamline at the Advanced Photon Source, Argonne National Laboratory (ANL) (Table S3). Data were processed with the HKL2000 program suite\(^ {49} \) and XDS\(^ {50} \). Structures were solved using molecular replacement program Phaser\(^ {51} \) from the CCP4 program suite\(^ {52} \) and an *E. coli* MBP structure (3DM0) as the search model. The structures were determined through alternative manual model building with Coot\(^ {53} \) and refinement with Phenix\(^ {54} \). The crystal structures were validated by the Molprobity server\(^ {55} \). Figures were prepared using PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.).

**Results**

**A survey of using MBP as a crystallization chaperone.** A survey of the Protein Data Bank (www.rcsb.org) showed that many groups have successfully adopted the MBP fusion strategy in the crystallization of their proteins of interest (Table S1). To date, there are over 50 primary research articles reporting the structures of MBP-tagged fusion proteins and there is still a clear upward trend (Fig. 1a). There are over 100 deposited MBP fusion protein structures in the PDB by the middle of 2016, making MBP one of the most utilized crystallization chaperone (also see Waugh’s recent review ref. \(^ {37} \)). The MBP-tagged targets vary in their overall folds, function, and origins (Table S1). The size of the target proteins falls in the range of 12 residues to 431 residues, with the
median of 105 residues (Fig. 1b). The isoelectric points of target proteins are distributed between 4.0 and 11.0, with a medium value of 6.6 (Fig. 1b).

During our survey of the linker sequences between MBP and the target proteins, we noticed that most linkers are short. This suggests that the relative orientations of MBP and the target proteins may be restricted despite the largely random coil structures of the linker. In agreement with a recent review ref. 37, we have not been able to find consensus linker sequence in these structures, although a linker with the sequence AAA appears to be common.

**Crystallization of fusion proteins with a short VD linker.** In our study of the death domain superfamily members, obtaining highly purified samples for crystallization has been challenging, as documented for many of these protein-protein interaction domains56. Our initial design of the V28E vector (Fig. 2a) was based on an E. coli MBP protein mutant harboring five mutations that reportedly enhance its crystallizability 13,46. The short linker of VD residues followed by the first helical residues of the target proteins may help enhance crystallization.

Figure 2. Crystallization of MBP fusion proteins with a short VD linker. (a) The cloning site of the V28E expression vector. Coding sequences for genes of interest can be ligated into the V28E vector using restriction enzyme sites of Sal I and Not I. The other six vectors V28E1-E6 share the overall feature as V28E but use the Not I and Xho I sites for insert ligation. (c) Superposition of the MBP tags in the three fusion protein structures. The three target proteins show different conformation relative to the MBP tag due to the flexibility of VD linker. (b) Different conformations of the linker loops from the three structures in (c).
Even though this strategy has resulted in success for some of our targets such as the NLRP1 CARD and the AIM2 PYD domain (Fig. 2b,c), it has failed for other targets such as the zebrafish IGBP1-CARD and human NLRP1-PYD (sequences are listed in Table S2). After extensive testing of the types and concentrations of different commonly used cryo-protectant solutions, we failed to collect useful X-ray diffraction data suitable for structural determination despite the fact that sizable crystals were readily obtained (Fig. 3a,b). The poor diffraction quality of these MBP fusion proteins may be related to the VD linker that introduces internal flexibility of the fusion proteins and micro-scale heterogeneity of the crystals. Indeed, low resolution diffraction is often observed for crystals of large proteins or protein complexes due to structural complexity and internal flexibility.

As expected, the VD linker region forms a loop bridging the MBP and target proteins in the three solved structures above (Fig. 2b,c). However, conformations of the linker loops among the three structures are drastically different, suggesting that the flexible VD linker may not be optimal for crystallization of some targets such as the IGBP1-CARD and NLRP1-PYD. In addition, target proteins, even those of similar size and fold, may significantly influence the overall structure of the fusion proteins.

Crystallization of fusion proteins containing a series of helical linkers. To design linkers that adopt rigid structures, which can potentially reduce the conformational heterogeneity of the MBP fusion protein, we sought to model the MBP-IPS-1 CARD structure. Here the linker sequence TNSA adopts a continuous helix from the last helix of MBP to the first helix of the IPS-1 CARD which may pose minimal perturbation of the helical structures for either MBP or IPS-1 CARD (Fig. 4a). The potential drawback of such a “rigid” linker is that the relative position and orientation of the MBP and the target may be “fixed” such that the chances of packing into ordered crystals may be limited. To explore broader crystallization space, we designed a series of protein expression vectors V28E1-V28E6 containing helical linkers of different lengths (Table 1). The one residue increment of the linker length may change the orientation and location of the target proteins relative to MBP, similar to those of successive residues in an α-helix. Therefore, an increment of one residue in the linker region may allow the target protein to rotate ~100° and translate ~1.5 Å relative to MBP (Fig. 4b). As a result, in fusion proteins expressed in vectors V28E1 to V28E6, the orientation of the target protein may change by one and a half turns along the helical axis of the last helix from MBP combined with a translation of ~9 Å. This varied spacing and orientation between MBP and targets may facilitate the presentation of different surface features of the fusion proteins and the sampling of more crystallization space (Table 1).

Using the set of V28E1-V28E6 vectors, we were able to express, purify and crystallize both the zIGBP1-CARD and the hNLRP1-PYD as MBP-fusion proteins harboring a helical linker. In contrast to crystals of poor diffraction quality using MBP-fusion protein expressed with the V28E vector, X-ray diffraction data of 1.47 Å resolutions were collected for the zIGBP1-CARD crystals using protein expressed with the V28E5 vector (Fig. 4c). This led to structural determination of the first CARD domain from the zebrafish genome (Fig. 4d). As shown in Fig. 4d, the V28E5 linker is nearly a perfect α-helix, which joins the last helix of MBP and the first helix of the CARD domain as designed. This is also one of the highest resolution structures for an MBP fusion protein, which may be attributed to the rigidity of the helical linker.

Of note, none of the previously deposited over 100 MBP fusion protein structures were crystallized with more than one linker sequence, probably due to the challenge of the linker design and the lack of systematic approach for linker sequence selection. The set of expression vectors harboring helical linkers (V28E1-V28E6) provide researchers new opportunities to obtain crystals. Indeed, both hNLRP12-PYD and hMNDa-PYD, two targets...
that failed to crystallize using fusion proteins expressed with the V28E vector, crystallized using fusion proteins expressed with these new vectors (V28E1-V28E6). Remarkably, we were able to crystallize and solve their structures using MBP-fusion proteins expressed with more than one vector. In particular, the hNLRP12-PYD was crystallized as fusion proteins using both V28E4 and V28E6 vectors (Fig. 5a). In the structures, the V28E4 linker adopts a straight \( \alpha \)-helix connecting the last helix of MBP to the first helix of PYD, whereas the V28E6 linker is slightly bent. Structural comparison reveals that the PYD rotates 220° from the V28E4 structure to the V28E6 structure (Fig. 5b). Even though the hNLRP12-PYD domains in the two crystal forms adopt different orientations relative to MBP, their structures are essentially identical with a root-mean-square deviation (rmsd) of 0.35 Å (Fig. 5b,c). This suggests that the different linker sequences and crystal lattice packing do not significantly perturb the structures of the target proteins, further validating the linker design strategy.

Figure 4. Design of helical linkers for MBP fusion proteins. (a) Cartoon representation of MBP-hIPS-1-CARD fusion protein structure (pdb 2VGQ). The last helix of MBP extends and forms a continuous helix with the first helix of CARD. (b) Design of helical linkers. Based on the MBP-hIPS-1-CARD, one residue is inserted in the linker region incrementally from V28E1 to V28E5. The view is shown from the projection of the last helix of MBP. For clarity, only the H3 helix of the CARD is shown in cartoon and in different colors (red for V28E1, orange for V28E2, yellow for V28E3, green for V28E4 and blue for V28E5). The rotation of target protein is \( \sim 100° \) through the insertion of one residue in the linker. (c) X-ray diffraction image of zebrafish IGBP1-CARD as an MBP-fusion protein expressed with the V28E5 vector. By manipulating the linker sequences, X-ray diffraction by the crystals for the IGBP1-CARD fusion protein improved from worse than 6 Å to 1.47 Å. (d) Surface rendering of one MBP tagged IGBP1-CARD structure is shown. The MBP is shown in grey ribbons, the linker residues coded in the V28E5 vector in magenta and the IGBP1-CARD in green. A bound maltotetraose molecule at the MBP is shown as sticks.
Similar to the hNLRP12-PYD, the human MNDA PYD domain was crystallized as MBP fusion proteins using three different protein expression vectors V28E3, V28E4, and V28E6. All of these fusion protein crystals diffract very well which led to their structure determination (Fig. 6a). Structural analysis shows that the designed “helical” linkers are slightly distorted probably attributed to crystal packing (Fig. 6b). Nonetheless, the target protein structures from all three different crystals are essentially identical with an rmsd of 0.30–0.40 Å (Fig. 6c), demonstrating again that the MBP fusion strategy enhances the crystallizability without alteration of the structural features of the target proteins.

Conclusion and Discussion

Despite the abundance of structures for MBP fusion proteins, the design of the fusion protein expression constructs has largely been a trial-and-error process without a clear focus on the sequence or length of the linker region. There has been no report on the rational design of the linker sequences or systematic crystallization screening of MBP fusion proteins, which may have hindered the wider application of this approach in structural biology. In this study, we designed a set of seven vectors for expression of MBP-fusion proteins containing either a short flexible linker or six helical linkers. The six helical linkers were designed based on the following three considerations: first, the α helix is a relatively rigid structure which contrast with the flexible VD linker, therefore may reduce conformational heterogeneity of the fusion proteins; second, helical linkers of various lengths allows the target proteins to be positioned not only at different distances relative to the MBP protein, but also at different orientations because of the 100-degree turn and 1.5 Å translation for each successive residue at a helix. This facilitates the presentation of different surface features of the fusion proteins and the sampling of more crystallization space; third, a helical linker may facilitate the connection between the last helix of MBP and the first helix of a death domain superfamily members with minimal disruption of the helices from MBP or death domain family members. This has proven successful in determining the MBP-IPS-1-CARD structure57.

We tested and successfully crystallized seven targets from the challenging death domain superfamily. Two of them were crystallized in more than one crystal forms containing different linker sequences, demonstrating the power of this linker optimization strategy. We were able to determine the first crystal structures for five of them, including the hAIM2-PYD, hCARD8-CARD, zGBP1-CARD, hNLRP12-PYD, and hMNDA-PYD. Structures of the hNLRP12-PYD and hMNDA-PYD determined with different linker sequences in different crystal lattices show no significant perturbation of the PYD structures. The structure of the hNLRP1-CARD was determined previously using a selenomethionine-labeled form at 3.1 Å resolution (3KAT). Because essentially identical
structures were determined using our MBP-mediated crystallization strategy, and we observed minimal interactions between the MBP tag and the targets in our current studies, our data suggest that fusion protein tags such as MBP are unlikely to interfere with the structure or folding of death domains, and maybe other protein targets as reported previously. Even though we have only tested crystallization of the death domain superfamily members, our MBP fusion strategy may promote crystallization of diverse protein targets, as was demonstrated by previously determined structures of MBP fused with targets of various sizes or folds. Nonetheless, such fusion protein strategy may not be universally applicable for certain difficult protein targets, such as intrinsically disordered proteins that may adopt different conformations depending on different fusion tags they are connected with. Therefore, the fusion protein strategy is one of many different approaches to be tested to promote crystallization.

The linker helix is relatively rigid but can still be influenced by crystal packing forces including the interactions between target protein-target protein, MBP-MBP and MBP-target protein. This may bend/distort the helix to some extent. Indeed, the bending and distortion of the linker helix is observed in some of structures in this study. For example, between the two structures of hNLRP12-PYD fusion proteins, the helix in V28E4 is a straight helix, while that in V28E6 was slightly bent. Structural comparison reveals that the PYD rotates ~220° from the V28E4 structure to the V28E6 structure (Fig. 5a).

Highly flexible connection between the MBP tag and the target protein may be detrimental to crystallization. To best take advantage of our designed sets of vectors, we recommend removal of all flexible residues at the N-termini of the targets. Our tests demonstrated that the flexible and helical linkers may be complementary to each other in promoting crystallization. Some of our target proteins could only be crystallized with the VD linker while others only with one or more of the rigid helical linkers. Therefore, it is important to test the different linker sequences to enhance the likelihood of crystallization.

The different linker sequences may also facilitate the crystallization of protein complexes. The short linker sequence may allow the MBP tag to stabilize the target protein, but the closely positioned MBP may prevent the target protein from interacting with its partner proteins due to steric hindrance. In such cases, longer and more flexible linker sequences may be tested. The design of the helical linkers with different lengths is based on the α-helical structure that may facilitate the exposure of different surfaces of the target proteins in the context of different fusion proteins. It is therefore possible that the specific protein-protein interaction surfaces may be exposed in one or more of the designed fusion proteins.

The usage of MBP as a crystallization chaperone is perhaps partially due to its high solubility and expansive hydrophilic surface that provides versatile crystal packing interface. There are other proteins such as NusA that may possess similar properties and may be systemically tested as crystallization chaperones. It would be beneficial to test a diverse set of crystallization chaperones for a difficult protein target, because different fusion tags
can provide distinct packing interfaces. One caveat of using a helical linker sequence is that the C-terminus of a different protein tag may not end with an extruded helix as in the case of MBP. In such cases, both helical and non-helical linker sequences may need to be explored.

In conclusion, our study provides a strategy for testing MBP and other crystallization chaperones through optimization of the linker sequences. We believe that the use of versatile crystallization chaperones will significantly enhance crystallizability of many challenging protein targets.

References
1. Terpe, K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Applied microbiology and biotechnology 60, 523–533, doi: 10.1007/s00253-002-1158-6 (2003).
2. Malhotra, A. Tagging for protein expression. Methods in enzymology 463, 239–258, doi: 10.1016/S0076-6879(09)63016-0 (2009).
3. Kapust, R. B. & Waugh, D. S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci 8, 1668–1674, doi: 10.1101/ps.8.8.1668 (1999).
4. Oldham, M. L. & Chen, J. Crystal structure of the maltose transporter in a pretranslocation intermediate state. Science 332, 1202–1205, doi: 10.1126/science.1200767 (2011).
5. Kankanurussi, S. & Waugh, D. S. The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. PLoS One 7, e49589, doi: 10.1371/journal.pone.0049589 (2012).
6. Dong, A. et al. In situ proteolysis for protein crystallization and structure determination. Nature methods 4, 1019–1021, doi: 10.1038/nmeth1118 (2007).
7. Abkharov, R. N. et al. Combining in-situ proteolysis and microseed matrix screening to promote crystallization of P3c-nanobody complexes. Protein engineering, design & selection: PEDS 24, 737–741, doi: 10.1093/protein/gxq017 (2011).
8. Gheyi, E., Sauder, J. M. & Burley, S. K. Mass spectrometry guided in situ proteolysis to obtain crystals for X-ray structure determination. Journal of the American Society for Mass Spectrometry 21, 1795–1801, doi: 10.1016/j.jasms.2010.06.015 (2010).
9. Slend, P. et al. New surface contacts formed upon reducive lysine methylation: improving the probability of protein crystallization. Protein Sci 19, 1395–1404, doi: 10.1002/pro.420 (2010).
10. Walter, T. S. et al. Lysine methylation as a routine rescue strategy for protein crystallization. Structure 14, 1617–1622, doi: 10.1016/j.str.2006.09.005 (2006).
11. Kobayashi, M., Kubota, M. & Matsuura, Y. Crystallization and improvement of crystal quality for x-ray diffraction of maltoligosyl trehalose synthase by reductive methylation of lysine residues. Acta crystallographica. Section D, Biological crystallography 55, 931–933 (1999).
12. Goldschmidt, L., Cooper, D. R., Derewenda, Z. S. & Eisenberg, D. Toward rational protein crystallization: A Web server for the design of crystalizable protein variants. Protein Sci 16, 1569–1576, doi: 10.1101/ps.72914007 (2007).
13. Moon, A. F., Mueller, G. A., Zhong, X. & Pedersen, L. C. A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. Protein Sci 19, 901–913, doi: 10.1002/pro.368 (2010).
14. Derewenda, Z. S. It’s all in the crystals. Acta crystallographica. Section D, Biological crystallography 67, 243–248, doi: 10.1107/S0907444910077971 (2011).
15. Laganowski, A. et al. An approach to crystallizing proteins by metal-mediated synthetic symmetrization. Protein Sci 20, 1876–1890, doi: 10.1002/pro.727 (2011).
16. Nauli, S. et al. Polymer-driven crystallization. Protein Sci 16, 2524–2551, doi: 10.1100/ps.70374207 (2007).
17. Ruggiero, A., Smaldone, G., Squeglia, F. & Berisio, R. Enhanced crystallizability by protein engineering approaches: a general overview. Protein and peptide letters 19, 732–742 (2012).
18. Derewenda, Z. S. Application of protein engineering to enhance crystallizability and improve crystal properties. Acta crystallographica. Section D, Biological crystallography 66, 604–615, doi: 10.1107/S09074449100644X (2010).
19. Bukowska, M. A. & Grutter, M. G. New concepts and aids to facilitate crystallization. Current opinion in structural biology 23, 409–416, doi: 10.1016/j.sbi.2013.03.003 (2013).
20. Czabotar, P. E., Hothorn, M., Scheffzek, K., Sattler, M. & Stier, G. Thioredoxin as a fusion tag for carrier-driven crystallization. Structure of the human smoothened receptor bound to an antitumour agent. Structure 20, 1266–1273, doi: 10.1012/scon.1150609 (2007).
21. Chun, E. et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. Structure 20, 967–976, doi: 10.1012/pro.727 (2012).
22. Wang, C. et al. Structure of the human smoothened receptor bound to an antimitour agent. Nature 497, 338–343, doi: 10.1038/nature12167 (2013).
23. Liu, W. et al. Structural basis for allosteric regulation of GPCRs by sodium ions. Science 337, 232–236, doi: 10.1126/science.1219218 (2012).
24. Kuge, M. et al. Use of a fusion protein to obtain crystals suitable for X-ray analysis: crystallization of a GST-fused protein containing the DNA-binding domain of DNA replication-related element-binding factor, DREF. Protein Sci 6, 1783–1786, doi: 10.1002/pro.3560086822 (1997).
25. Corsini, L., Hothorn, M., Scheffzek, K., Sattler, M. & Stier, G. Thioredoxin as a fusion tag for carrier-driven crystallization. Protein Sci 17, 2070–2079, doi: 10.1101/ps.37564.108 (2008).
26. Razbata, P. E. et al. Bax crystal structures reveal how BH3 domains activate Bax and nucleate its oligomerization to induce apoptosis. Cell 152, 519–531, doi: 10.1016/j.cell.2012.12.031 (2013).
27. Menetrey, J. et al. Processive steps in the reverse direction require uncoupling of the lead head lever arm of myosin VI. Molecular cell 48, 75–86, doi: 10.1016/j.molcel.2012.07.034 (2012).
28. Sault, A. & Vulliez-le Normand, B., Lema, F. & Bentley, G. A. Crystal structure of a dominant B-cell epitope from the preS2 region of hepatitis B virus in the form of an inserted peptide segment in maltodextrin-binding protein. Journal of molecular biology 280, 185–192 (1998).
29. Center, R. J. et al. Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. Protein Sci 7, 1612–1619, doi: 10.1002/pro.3560070715 (1998).
30. Kuge, M. et al. Center, R. J., Kemp, B. E. & Poutousbouris, P. C. Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. Proceedings of the National Academy of Sciences of the United States of America 96, 4319–4324 (1999).
31. Pioszak, A. A. & Xu, H. E. Molecular recognition of parathyroid hormone by its G protein-coupled receptor. Proceedings of the National Academy of Sciences of the United States of America 105, 5034–5039, doi: 10.1073/pnas.0801027105 (2008).
Competing financial interests: The authors declare no competing financial interests.

Acknowledgements
The authors would like to thank Xinhua Ji at the National Cancer Institute for suggestions and discussions, Lars C. Pedersen at the National Institute of Environmental Health Sciences (NIEMS) for providing us the pMAL-MBP wild type and mutant plasmids. X-ray diffraction data were collected at the Advanced Photon Source (APS) GM/CA beamline. Some data was collected at the Advanced Photon Source (APS) GM/CA beamline, which is funded in whole or in part by funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Sciences (Y1-GM-1104). Some data was collected at the National Synchrotron Light Source (NSLS) beamline X-25 at the Brookhaven National Laboratory (BNL). Use of the NSLS is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-98CH10886. We’d like to thank the beamline scientists at both synchrotrons for their assistance. T.S.X. is supported by the Case Research Institute and the University Hospitals Case Medical Center. T.J. is supported by the Fundamental Research Funds for the Central Universities and the 100 Talents Program of CAS. Y.L. is supported by the China Postdoctoral Science Foundation (Grant No.: 2015MS580207).

Author Contributions
T.J. contributed conception, design, data collection and analysis, and manuscript writing; W.C., J.J., J.C., Y.L. K.F., M.H., and P.S. contributed protein purification, data collection, data analysis, and manuscript writing; T.S.X. contributed conception, design, supervision, data interpretation and manuscript writing.

Additional Information
Accession codes: Coordinates and structure factors have been deposited in the Protein Data Bank, and the accession codes are presented in Table S3.

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.
How to cite this article: Jin, T. et al. Design of an expression system to enhance MBP-mediated crystallization. Sci. Rep. 7, 40991; doi: 10.1038/srep40991 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2017