Supporting Information

O-tert-Butyltyrosine, an NMR tag for high-molecular weight systems and measurements of submicromolar ligand binding affinities

Wan-Na Chen, Kekini Vahini Kuppan, Michael David Lee, Kristaps Jaudzems, Thomas Huber, Gottfried Otting

Expression constructs

The DNA nucleotide sequence of the *Bacillus stearothermophilus* (*Bst*) gene for DnaB was codon optimized for expression in *E. coli* using the program OPTIMISER.[1] The construct of the wild-type protein included a protein G B1 domain as a N-terminal solubility enhancement tag,[2] preceded by a MASMTG sequence (coded by 5' nucleotides of the bacteriophage T7 gene 10) and a His_6-tag. A tobacco etch virus (TEV) protease cleavage site was inserted between the GB1 domain and the DnaB (Figure S1).

The *S. aureus* sortase A (SrtA) construct used in this study is the catalytic domain (residues 60-206) of the full-length protein without the signal polypeptide. The catalytic core has been reported to retain the enzymatic activity.[3,4] The sequence was preceded by a MASMTG tag at the N-terminus, and contained a TEV cleavage site and a His_6-tag at the C-terminal end.

The expression construct for the wild-type *E. coli* aspartate/glutamate binding protein (DEBP) was preceded by a MHHHHHHMENLYFQG sequence, resulting in 291 residues.

The genes were cloned into the T7 expression plasmid pETMCSI,[5] which was transformed into the *E. coli* strain TOP10 for plasmid progression. Sites for the incorporation of *O-tert*-butyl-tyrosine (Tby) were selected by inspection of the three-dimensional protein structures to make sure that the *tert*-butyl group would be solvent exposed, using the structures with the PDB accession codes 2R6A for *Bst* DnaB,[6] 1IJA and 2KID for *S. aureus* SrtA,[4,7] and 2VHA for *E. coli* DEBP.[8] Site-directed mutations were generated by two pairs of primers (IDT-DNA Technologies) using Q5 DNA polymerase (New England Biolabs), following a modified overlap extension protocol as described.[9] The PCR amplified products were digested with the restriction enzymes *Nde*I and *Eco*RI for insertion into pETMCSI.

Protein expression and purification

Bst DnaB proteins containing Tby were produced by using a published pUltra vector[10] to express *p*-cyanophenylalanyl-tRNA synthetase (*p*CNF-RS) and the requisite $tRNA_{CUA}$. The pUltra vector was co-transformed into *E. coli* BL21 (DE3) together with the pETMCSI T7 vector harbouring the target *Bst* DnaB mutant. The proteins were produced in a high-cell density protocol to minimize the usage of unnatural amino acid.[11] After the 1L cell culture reached OD_{600} 0.5-1.0, the cells were spun down at 5,000 *g* at 10 ^{o}C and then resuspended in 250 mL fresh Luria-Bertani medium containing 1 mM Tby (Sigma-Aldrich). Following a recovery period of 1-2 h shaking at 37 ^{o}C, isopropyl-*β*-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM for protein overexpression. Perdeuterated DnaB samples were prepared in a similar way, except that the cells from a 1 L culture were washed twice using 25 mL minimal medium made from 70% D_2O and 99% D_2O prior to resuspension in minimal medium prepared in 99% deuterium oxide. Unlabeled ammonium chloride and glucose were used as nitrogen and carbon source, respectively. Following overnight expression at 25 ^{o}C, the cells were lysed by sonication and the lysate was loaded onto a 5 mL Ni-NTA column (GE Healthcare, USA). The target protein was then eluted using a gradient buffer mixture of buffer A (50 mM Tris-HCl, pH 7.5, 300 mM NaCl) and buffer B (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 300 mM imidazole). To separate the truncation product from the full-length DnaB, the protein was dialyzed into buffer C (25 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM DTT), concentrated to 500 µL using a centrifugal filter unit with molecular weight cut-off (MWCO) of 10 kDa (Amicon Ultra, Millipore, Billerica, USA), and loaded onto a Superdex 200 column (GE Healthcare, USA). The protein was eluted with buffer C. Finally, the buffer was exchanged to NMR buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl) using a centrifugal filter device (MWCO 10 kDa).

To explore the effect of the GB1 domain fusion to DnaB on 1D ^{1}H NMR spectra, TEV protease was added after the gel filtration step at 1:100 weight ratio and the solution dialyzed into buffer D (25 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM 2-mercaptoethanol) at 4 ^{o}C for 14 h. The digested mixture was passed onto a 1 mL Ni-NTA gravity column (GE Healthcare, USA) and the flow-through fraction was exchanged into NMR buffer using a centrifugal filter unit (MWCO 10 kDa).

To minimize sample losses, the perdeuterated DnaB samples were applied to a centrifugal filter unit (MWCO 100 kDa) with NMR buffer instead of using gel filtration to separate the truncation product from the hexamer.

S. aureus SrtA and *E. coli* DEBP were produced by continuous exchange cell-free protein synthesis using PCR amplified DNA as template.[9,12-14] RF1-depleted S30 extract was prepared from an engineered *E. coli* BL21(DE3) strain.[13,15] The proteins were purified by diluting the reaction mixture with buffer A, loading the mixture onto a 1 mL Ni-NTA gravity column, washing with buffer A plus 30 mM imidazole, and eluting with buffer B. *S. aureus* SrtA was subsequently dialyzed into NMR buffer (20 mM MES buffer, pH 6.5, 50 mM NaCl). Following the Ni-NTA purification and prior to exchange into NMR buffer, *E. coli* DEBP was unfolded by incubation for 2 h at room temperature in the presence of 6 M guanidinium hydrochloride to release any bound substrate. Protein refolding and removal of denaturant were achieved by dialysis against NMR buffer (20 mM phosphate buffer, pH 7.5, 100 mM NaCl). Finally, the protein samples were concentrated using centrifugal filter units (MWCO 10 kDa).

Ligation with a lanthanide binding tag

The *Bst* DnaB Ala303Cys mutant was ligated with the C1 tag loaded with Ce^{3+} [16] by incubation for 14 h at room temperature with a three-fold excess of C1-Ce tag. Excess tag was removed by washing with NMR buffer using a centrifugal filter unit (MWCO 10 kDa).

NMR spectroscopy and measurements of K_d values

All 1D ^{1}H NMR spectra were recorded of 2-50 µM protein solutions in aqueous buffer containing 10% D_2O at 25 ^oC, using a Bruker 600 MHz or 800 MHz NMR spectrometers equipped with TCI cryoprobes. The spectra of DnaB were recorded using the jump-and-return pulse sequence,[17] while a double spin echo[18] was used to record the spectra of SrtA and DEBP. For accurate determination of peak positions in the titration experiment of DEBP with glutamate, line shape fits were performed for the *tert*-butyl resonance using TopSpin 2. The chemical shifts were fitted by the program Origin 8.0 using equation 1.

Figure S1. *Bst* DnaB expression construct used. The N-terminal fusions increased the molecular mass of the monomeric protein to 58.7 kDa. Due to the flexibility of the linker between GB1 domain and DnaB, this increase in molecular mass is not expected to influence the NMR spectrum of the hexamer very much.

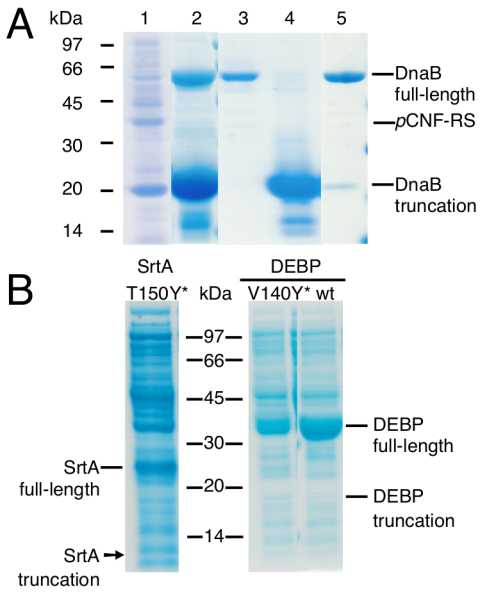

Figure S2. Protein production and purity analyzed by 15% SDS-polyacrylamide gel electrophoresis. The gels were stained with Coomassie blue. The positions of the expected full-length and truncated samples are indicated. (A) *Bst* DnaB produced *in vivo* with Tby incorporated at position 104. Lane 1: Crude protein fraction after IPTG induction. Lane 2: Elution from Ni-NTA affinity purification. Lane 3: Elution at 8 mL from Superdex 200 gel filtration. Lane 4: Elution at 12 mL from Superdex 200 gel filtration. Lane 5: Sample after Ni-NTA column and filtration with a centrifugal filter unit (MWCO 100 kDa), without using gel filtration. (B) Samples of SrtA and DEBP produced by cell-free synthesis. The expected positions of the truncation products are indicated. Left panel: cell-free reaction mixture containing SrtA with Tby (denoted Y*) at position 150. There is no evidence for truncation product. Right panels: DEBP with Tby at position 140 and wild-type protein. Notably, although the amber stop codon at site 140 was followed by a thymidine in the DNA sequence (Figure S7), the expression level of the Val140Tby mutant was not much reduced compared with the wild-type protein, with only little evidence of truncation in contrast to expectations based on previous literature.[19]

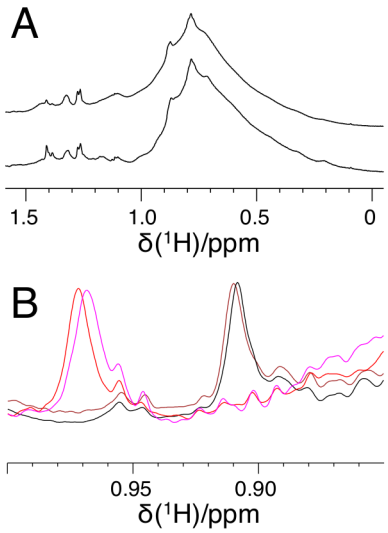

Figure S3. Control experiments to assess the effect of the fusion with GB1 on the structural integrity of wild-type *Bst* DnaB and the effect of DMSO on the chemical shift of the *tert*-butyl resonance in SrtA. (A) 40 µM solution of *Bst* DnaB in 25 mM Tris-HCl, pH 7.5, 150 mM NaCl. Top: after cleavage of the GB1 domain with TEV protease; bottom: before TEV protease cleavage. A few more signals can be observed in the uncleaved product, suggesting that the GB1 moiety is mobile relative to the DnaB. The conservation of the overall appearance of the spectrum suggests that the hexamer remains unaffected by the fusion with GB1. (B) *tert*-butyl resonance measured of a 20 µM solution of SrtA with Tby incorporated at position 150. Black: SrtA in 20 mM MES buffer, pH 6.5, 50 mM NaCl, 5 mM $CaCl_2$; brown: same as the black spectrum, but with 1% DMSO; red: SrtA in 20 mM MES buffer, pH 6.5, 50 mM NaCl, 1 mM EDTA; magenta: same as red spectrum, but with 1% DMSO. The chemical shift changes induced by DMSO are much smaller than the chemical shift changes generated by the inhibitor.

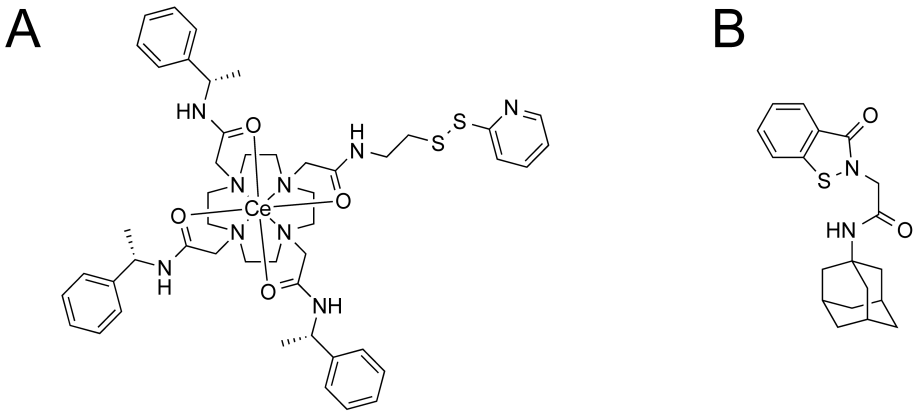

Figure S4. Chemical structures of compounds used. (A) C1-Ce tag.[16] (B) SrtA inhibitor N-(adamantan-1-yl)-2-(3-oxo- 2,3-dihydro-1,2-benzothiazol-2-yl)-acetamide.[20]

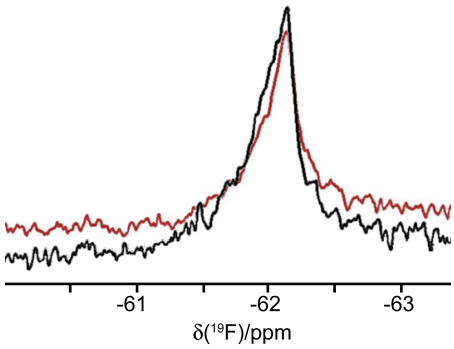

Figure S5. ^{19}F NMR spectra of 70 μM solutions of *Bst* DnaB with trifluoromethyl-phenylalanine (tfmF) incorporated at site 104. The samples were prepared using the same protocol as for the preparation of the Tby samples described above, except that Tby was replaced by 1 mM tfmF. The buffer used contained 50 mM Tris-HCl, pH 7.4, 1 mM EDTA, 1 mM DTT, and 200 mM NaCl. The spectra were recorded at room temperature on a Varian Inova 500 MHz NMR spectrometer in 4 h each. Spectrum calibration was relative to external tfmF (-62.1 ppm). Black: after TEV protease cleavage of the GB1 solubilization domain. Brown: before TEV protease cleavage.

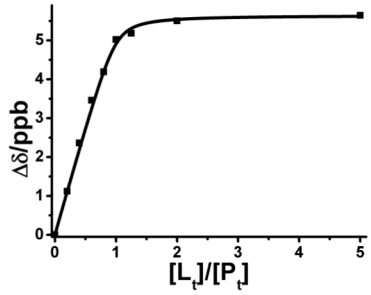

Figure S6. Binding affinity measurement of 10 μM *E. coli* DEBP for glutamate. The spectra were recorded in 20 mM phosphate buffer, pH 7.5, 100 mM NaCl at 25 ^{o}C, using a Bruker 600 MHz NMR spectrometer. The fit yielded a K_d value of 0.16 μM with a standard error of 0.04 μM. This result is unreliable because the K_d value is more than 10 times smaller than the protein concentration.

ATGCACCATCACCATCACCATATGGAAAACCTGTATTTTCAGGGCATGGCCGCAGGCAGCACGCTGGACAAAATCGCCAA
AAACGGTGTGATTGTCGTCGGTCACCGTGAATCTTCAGTGCCTTTCTCTTATTACGACAATCAGCAAAAAGTGGTGGGTT
ACTCGCAGGATTACTCCAACGCCATTGTTGAAGCAGTGAAAAAGAAACTCAACAAACCGGACTTGCAGGTAAAACTGATT
CCGATTACCTCACAAAACCGTATTCCACTGCTGCAAAACGGCACTTTCGATTTTGAATGTGGTTCTACCACCAACAACGT
CGAACGCCAAAAACAGGCGGCTTTCTCTGACACTATTTTCGTGGTCGGTACGCGCCTGTTGACCAAAAAGGGTGGCGATA
TCAAAGATTTTGCCAACCTGAAAGACAAAGCCGTAGTCGTCACTTCCGGCACTACCTCTGAA**GTT**TTGCTCAACAAACTG
AATGAAGAGCAAAAAATGAATATGCGCATCATCAGCGCCAAAGATCACGGTGACTCTTTCCGCACCCTGGAAAGCGGTCG
TGCCGTTGCCTTTATGATGGATGACGCTCTGCTGGCCGGTGAACGTGCGAAAGCGAAGAAACCAGACAACTGGGAAATCG
TCGGCAAGCCGCAGTCTCAGGAGGCCTACGGTTGTATGTTGCGTAAAGATGATCCGCAGTTCAAAAAGCTGATGGATGAC
ACCATCGCTCAGGTGCAGACCTCCGGTGAAGCGGAAAAATGGTTTGATAAGTGGTTCAAAAATCCAATTCCGCCGAAAAA
CCTGAACATGAATTTCGAACTGTCAGACGAAATGAAAGCACTGTTCAAAGAACCGAATGACAAGGCACTGAACTAA

Figure S7. Nucleotide sequence of the construct of *E. coli* DEBP used in the present work. The codon replaced by the amber stop codon is in bold and underlined.

References

(1) Puigbo, P.; Guzman, E.; Romeu, A.; Garcia-Vallve, S. *Nucleic Acids Res*. **2007**, *35* (Web Server issue): W126-131.

(2) Zhou, P.; Wagner G. *J. Biomol. NMR* **2010**, *46*, 23-31.

(3) Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O. *Science* **1999**, *285*, 760-763.

(4) Ilangovan, U.; Ton-That, H.; Iwahara, J.; Schneewind, O.; Clubb, R. T. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6056-6061.

(5) Neylon, C.; Brown, S. E.; Kralicek, A. V.; Miles, C. S.; Love, C. A.; Dixon, N. E. *Biochemistry* **2000**, *39*, 11989-11999.

(6) Bailey, S.; Eliason, W. K.; Steitz, T. A. *Science* **2007**, *318*, 459-463.

(7) Suree, N.; Liew, C. K.; Villareal, V. A.; Thieu, W.; Fadeev, E. A.; Clemens, J. J.; Jung, M. E.; Clubb, R. T. *J. Biol. Chem*. **2009**, *284*, 24465-24477.

(8) Hu, Y. L.; Fan, C.-P.; Fu, G. S.; Zhu, D. Y.; Jin, Q.; Wang, D.-C. *J. Mol. Biol*. **2008**, *382*, 99-111.

(9) Wu, P. S. C.; Ozawa, K.; Lim, S. P.; Vasudevan, S. G.; Dixon, N. E.; Otting, G. *Angew. Chemie Int. Ed.* **2007**, *46*, 3356-3358.

(10) Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G. *Biochemistry* **2013**, *52*, 1828-1837.

(11) Sivashanmugam, A.; Murray, V.; Cui, C. X.; Zhang, Y. H.; Wang, J. J.; Li, Q. Q. *Protein Sci*. **2009**, *18*, 936-948.

(12) Ozawa, K.; Headlam, M. J.; Schaeffer, P. M.; Henderson, B. R.; Dixon, N. E.; Otting, G. *Eur. J. Biochem.* **2004**, *271*, 4084-4093.

(13) Apponyi, M. A.; Ozawa, K.; Dixon, N. E.; Otting, G. *Methods Mol. Biol.* **2008**, *426*, 257-268.

(14) Jia, X.; Ozawa, K.; Loscha, K. V.; Otting, G. *J. Biomol. NMR* **2009**, *44*, 59-67.

(15) Loscha, K. V.; Herlt, A. J.; Qi, R.; Huber, T.; Ozawa, K.; Otting, G. *Angew. Chem. Int. Ed*. **2012**, *51*, 2243-2246.

(16) Graham, B.; Loh, C. T.; Swarbrick, J. D.; Ung, P.; Shin, J.; Yagi, H.; Jia, X., Chhabra, S.; Pintacuda, G.; Huber, T.; Otting, G. *Bioconjugate Chem*. **2011**, *22*, 2118-2125.

(17) Plateau, P.; Guéron, M. *J. Am. Chem. Soc*. **1982**, *104*, 7310-7311.

(18) Hwang, T. L.; Shaka, A. J. *J. Magn. Reson. A* **1995**, *112*, 275-279.

(19) Smolskaya, S.; Zhang, Z. J.; Alfonta, L. *PLoS One* **2013**, *8*, e68363.

(20) Zhulenkovs, D.; Rudevica, Z.; Jaudzems, K.; Turks, M.; Leonchiks, A. *Bioorg. Med. Chem*. **2014**, *22*, 5988-6003.