FEBS Letters

FEBS PRESS science publishing by scientists

Structural and enzymatic analyses of *Anabaena* heterocyst-specific alkaline invertase InvB

Jin Xie, Hai-Xi Hu, Kun Cai, Ling-Yun Xia, Feng Yang, Yong-Liang Jiang, Yuxing Chen and Cong-Zhao Zhou

Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei Anhui, China

Correspondence
C.-Z. Zhou and Y. Chen, Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei Anhui 230027, China
Fax/Tel: +86 551 63600406
E-mails: zcz@ustc.edu.cn (C.Z.Z.); cyxing@ustc.edu.cn (Y.C.)

(Received 5 February 2018, revised 10 March 2018, accepted 15 March 2018, available online 10 April 2018)

doi:10.1002/1873-3468.13041

Edited by Miguel De la Rosa

Anabaena **sp. PCC 7120 encodes two alkaline/neutral invertases, namely InvA and InvB. Following our recently reported InvA structure, here we report the crystal structure of the heterocyst-specific InvB. Despite sharing an overall structure similar to InvA, InvB possesses a much higher catalytic activity. Structural comparisons of the catalytic pockets reveal that Arg430 of InvB adopts a different conformation, which may facilitate the deprotonation of the catalytic residue Glu415. We propose that the higher activity may be responsible for the vital role of InvB in heterocyst development and nitrogen fixation. Furthermore, phylogenetic analysis combined with activity assays also suggests the role of this highly conserved arginine in plants and cyanobacteria, as well as some proteobacteria living in highly extreme environments.**

Keywords: alkaline/neutral invertases; catalytic activity; crystal structure; cyanobacteria; nitrogen fixation; phylogenetic analysis

Sucrose, the major product of photosynthesis in plants and cyanobacteria [1], could be degraded into hexoses or their derivates to provide carbon and energy or to act as signaling molecules, for the growth, development, and defense [2–5]. Besides, the hydrolysis of sucrose was reported to be involved in nitrogen fixation in diazotrophic cyanobacteria [6,7]. More recently, genomic and experimental evidence suggested the existence of sucrose-related genes in proteobacteria [4,8–12], where sucrose is used as a compatible solute, and its osmotic effect can be doubled by sucrose degradation [5,11,13,14]. Sucrose degradation could be catalyzed by two different enzymes. Sucrose synthase (SUS, EC 2.4.1.13) cleaves sucrose into fructose and UDP-glucose or ADP-glucose reversibly, whereas invertase (Inv, EC 3.2.1.26n) catalyzes the irreversible hydrolysis of sucrose into glucose and fructose [15]. Current understanding indicates that these two enzymes have distinct roles. SUS-catalyzed sucrose degradation is mainly involved in the biosynthesis of structural and storage polysaccharide by providing sugar nucleotides as the precursors [2,16]. By contrast, Invs plays a critical role when there is a demand for carbon and energy [2,4].

There are two classes of invertases, which are initially categorized according to their optimum pH: the acid invertases (Ac-Invs) with an optimum pH of 4.0–5.5, and the alkaline/neutral invertases (A/N-Invs) with an optimum pH of 6.5–8.0 [17,18]. Notably, A/N-Invs and Ac-Invs differ from each other in the primary sequence. As the members of glycoside hydrolase (GH) 32 family, Ac-Invs adopt an all β structure with a five-fold

Abbreviations
Ac-Inv, acid invertase; A/N-Inv, alkaline/neutral invertase; ELSD, evaporative light scattering detector; GH, glycoside hydrolase; Inv, invertase; RMSD, root mean square deviation; SeMet, selenomethionine; SUS, sucrose synthase.

β-propeller catalytic domain, and adopt a so-called double displacement mechanism for hydrolysis [19–22]. Ac-Invs can hydrolyze sucrose and other β-fructose-containing oligosaccharides, therefore they are also termed β-fructofuranosidases. Recently, the first structure of an A/N-Inv, *Anabaena* alkaline invertase InvA was solved, which revealed that A/N-Invs share an overall structure of $(\alpha/\alpha)_6$ barrel differing from Ac-Invs [23]. Structural analysis and enzymatic assays identified the catalytic residues, Asp188 and Glu414, and the general acid-base mechanism for hydrolysis [23]. Furthermore, the stringent substrate specificity toward sucrose enabled us to propose that A/N-Invs, which exclusively constitute GH100 family, represent a novel family of glucosidase [23]. Besides, A/N-Invs and Ac-Invs differ a lot in biochemical properties, species distribution, and subcellular location [1,24].

Recent studies revealed that sucrose translocated from vegetative cells to heterocysts as the major carbon carrier molecule plays a central role in connecting carbon metabolism and nitrogen fixation in diazotrophic cyanobacteria [7,15,25]. Curatti *et al.* showed that SUS was specially involved in the cleavage of sucrose in vegetative cells [7,25]. Bioinformatics analysis revealed that nearly all genome-sequenced cyanobacteria encode only A/N-Invs, but not Ac-Invs. Vargas *et al.* [24] identified two A/N-Invs, namely InvA and InvB, in the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120. As the detected optimum pHs were 7.8 and 6.7, they further classified InvA and InvB into alkaline and neutral invertase, respectively. Both InvA and InvB are expressed in vegetative cells; however, only InvB is specifically expressed in heterocysts [6,26]. Moreover, knockout of *invB* (*alr0819*) but not *invA* (*alr1521*) blocked the growth of cyanobacteria at diazotrophic conditions, which suggested that InvB is essential for diazotrophic growth [6,26]. However, why InvB is specifically involved in the diazotrophic growth remains not clear.

It is recognized that A/N-Invs widely spread in plants and cyanobacteria, and plant A/N-Inv might evolve from cyanobacteria after the endosymbiotic origin of chloroplasts. [1,24] Most recently, Wan *et al.* [13] performed detailed phylogenetic analysis of Ac-Invs and A/N-Invs from cyanobacteria and especially green algae, lower and higher plants, and concluded that A/N-Invs are evolutionarily and functionally more stable than Ac-Invs, possibly due to their roles in maintaining cytosolic sugar homeostasis for cellular function. However, there is no report concerning the A/N-Inv activity or detailed phylogenetic analysis in other bacteria except for cyanobacteria so far.

Here, we report the crystal structures of *Anabaena* InvB and its complex with sucrose. Despite sharing an overall structure similar to the subunit of InvA, the $loop_{\beta2-\beta3}$ close to the substrate binding pocket, in addition to the residue Arg430 of InvB differs a lot. Activity assays revealed that the optimum pH of InvB is 7.7, suggesting that InvB is also an alkaline invertase like InvA, but not a neutral invertase as previously reported [24]. Compared to InvA, InvB has a higher affinity toward sucrose and a higher catalytic efficiency (k_{cat}/K_m). Structural analyses combined with activity assays revealed that Arg430 contributes to the high catalytic efficiency of InvB, which is necessary for heterocyst development and nitrogen fixation. Moreover, phylogenetic analysis enabled us to extend the species distribution of A/N-Invs to proteobacteria and further confirmed the role of the conserved arginine.

Materials and methods

Cloning, expression, and purification

The gene of InvB (NCBI accession number CAC85155.1, 483 residues) was amplified from the genomic DNA of *Anabaena* sp. PCC 7120. The full-length *invB* and their mutants were individually cloned into a pET29a-derived vector with an N-terminal His_6 tag. Likewise, the truncated InvB covering residues Thr9–Ser457 was constructed for crystallization. Both the wild-type and mutant proteins were overexpressed in *Escherichia coli* strain BL21 (DE3) (Novagen, Madison, WI, USA). Cells were grown in LB culture medium (10 g of tryptone, 5 g of yeast extract and 10 g of NaCl per liter) containing 30 $\mu g \cdot mL^{-1}$ kanamycin at 37 °C until the $A_{600\ nm}$ reached ~ 0.6. Then protein expression was induced with 0.2 mM isopropyl β-D-1-thio-galactopyranoside at 37 °C for another 4 h. Cells were collected and resuspended in 40 mL of lysis buffer (20 mM HEPES, pH 8.0, 100 mM NaCl). After 12 min of sonication and centrifugation at 12 000 × ***g*** for 25 min, the supernatant containing the target protein was loaded onto a nickel-nitrilotriacetic acid column (Qiagen, Hilden, Germany) equilibrated with the binding buffer (20 mM HEPES, pH 8.0, 100 mM NaCl). The target protein was eluted with 300 mM imidazole, and further applied to a HiLoad 16/600 Superdex 200 pg column (GE Healthcare, Uppsala, Sweden) equilibrated with the binding buffer. Samples for enzymatic activity assays were collected at the highest peak fractions without concentration and stored at −80 °C with 50% glycerol. The purity of protein was assessed by gel electrophoresis.

Analytical gel filtration chromatography was used to determine the molecular weight of InvB in solution by a Superdex 75 10/300 GL column (GE Healthcare). The

following standard molecular markers were used for calibration: ribonuclease A (13.7 kDa), ovalbumin (43.0 kDa), conalbumin (75.0 kDa), aldolase (158.0 kDa), ferritin (440.0 kDa), and thyroglobulin (669.0 kDa).

The selenomethionine (SeMet)-substituted full-length and truncated InvB were overexpressed in *E. coli* B834 (DE3) (Novagen). Transformed cells were first cultured in LB medium at 37 °C overnight, then harvested and washed twice with the M9 medium [27]. Then the cells were cultured in SeMet medium (M9 medium with 50 mg·L^{-1} SeMet and other essential amino acids) to an $A_{600\ nm}$ of 0.6–0.8. The following steps in protein expression and purification were the same as those for the native protein in addition to adding 5 mM β-mercaptoethanol during purification.

Crystallization, data collection, and processing

Both native and SeMet-substituted InvB were concentrated to 5–10 mg·mL^{-1} by ultrafiltration (Millipore, Bedford, MA, USA) for crystallization. Crystals were grown at 16 °C using the hanging drop vapor diffusion method, with a drop of 1 μL protein solution mixed with an equal volume of reservoir solution. Microseeding was adopted in crystal optimization. Crystals were obtained against the reservoir solution of 10% polyethylene glycol 6000, 0.1 M sodium citrate, pH 5.6 for the native and SeMet-substituted full-length protein, and 0.7 M lithium chloride, 6% polyethylene glycol 6000, and 0.1 M MES, pH 6.0 for the SeMet-substituted Millipore-truncated InvB. The crystals were transferred to cryoprotectant (reservoir solution supplemented with 30% glycerol) and flash-cooled with liquid nitrogen. For soaking experiments, crystals were transferred to 2 μL of mother liquor containing 30% saturated sucrose and then 3 min later flash-cooled directly, with sucrose as the cryoprotectant. X-ray diffraction data were collected at 100 K using beamline BL17U with an ADSC Q315r CCD detector at the Shanghai Synchrotron Radiation Facility. All of the diffraction data were integrated and scaled with the program HKL2000 [28].

Structure determination and refinement

The crystal structure of SeMet-substituted truncated InvB was determined using the single-wavelength anomalous dispersion method [29]. The AutoSol program of PHENIX [30] was used to search the selenium atoms and to calculate the phase. Then automatic model building was carried out using Autobuild in PHENIX. The initial model was refined using the maximum likelihood method implemented in REFMAC5 [31] of CCP4i program suite [32] and rebuilt interactively using the program COOT [33]. The crystal data of InvB complexed with sucrose were refined against the apo-form InvB structure. The final models were evaluated with the web service MolProbity (http://molprobity.biochem.duke.edu) [34]. Crystallographic parameters were listed in Table 1. The $|F_o|-|F_c|$ omit electron

Table 1. Crystal parameters, data collection, and structure refinement.

	SeInvB	SeInvB-Suc
Data collection		
Space group	*C2*	*C2*
Unit cell		
a, *b*, *c* (Å)	166.6, 82.4, 94.8	166.9, 82.2, 97.4
α, *β*, *γ* (°)	118.9	119.1
Resolution range (Å)	50.00–1.93 (2.00–1.93)[a]	50.00–1.95 (2.02–1.95)
Unique reflections	82 269 (8274)	80 206 (8173)
R_{merge}[b]	0.078 (0.505)	0.096 (0.537)
$I/\sigma I$	15.1 (2.3)	13.1 (2.4)
Completeness (%)	97.6 (98.9)	95.7 (98.1)
Average redundancy	2.7 (2.7)	3.5 (3.5)
Structure refinement		
Resolution range (Å)	31.41–1.93	31.35–1.95
R_{work}[c]/R_{free}[d]	0.206/0.228	0.210/0.239
Number of protein atoms	7,206	7,245
Number of water atoms	287	184
RMSD[e] bond lengths (Å)	0.005	0.005
RMSD bond angles (°)	0.914	1.114
Average B-factors (Å^2)		
Protein	41.3	37.0
Ligand	42.0 (GOL in subunit A only)	32.6/33.9 (Suc in both subunits)
Ramachandran plot[f] (residues, %)		
Most favored	98.58	98.96
Additional allowed	1.42	1.04
Outliers	0	0
MolProbity clashscore[g]/score percentile	0.87/100th percentile (*N* = 742, 1.93 ± 0.25 Å)	2.06/100th percentile (*N* = 821, 1.95 ± 0.25 Å)
MolProbity score[h]/score percentile	0.77/100th percentile (*N* = 11 840, 1.93 ± 0.25 Å)	1.00/100th percentile (*N* = 13 349, 1.95 ± 0.25 Å)
Protein Data Bank entry	5Z73	5Z74

[a]The values in parentheses refer to statistics in the highest bin. [b]$R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value for its unique reflection. Summations are over all reflections. [c]$R_{work} = \sum_h |F_o(h) - F_c(h)| / \sum_h F_o(h)$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. [d]R_{free} was calculated with 5% of the data excluded from the refinement. [e]RMSD from ideal values. [f]The categories were defined by Molprobity. [g]The number of serious steric overlaps (> 0.4 Å) per 1000 atoms. [h]MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.

density map of the ligand contoured at 3.0 σ was calculated by PHENIX. The interface areas were calculated by PDBePISA [35]. All structure figures were prepared with PYMOL [36].

Enzymatic assays

The pH dependence of invertase activity of recombinant InvB was measured in 40 mM MES (pH 5.5–6.5), HEPES (pH 7.0–8.0), Bicine (pH 8.5–9.0), and glycine–NaOH (pH 9.5) buffer containing 100 mM NaCl. Subsequent enzymatic assays were carried out at pH 7.9 [23] and pH 7.7 for InvA and InvB, respectively. A 50-μL reaction mixture containing 100 mM sucrose and an appropriate volume of wild-type enzymes or the mutants was incubated at 30 °C for different times, and then stopped by heating at 95 °C for 10 min. After centrifuged at 12 000 × ***g*** for 10 min, the supernatant was diluted five-fold for HPLC analysis by

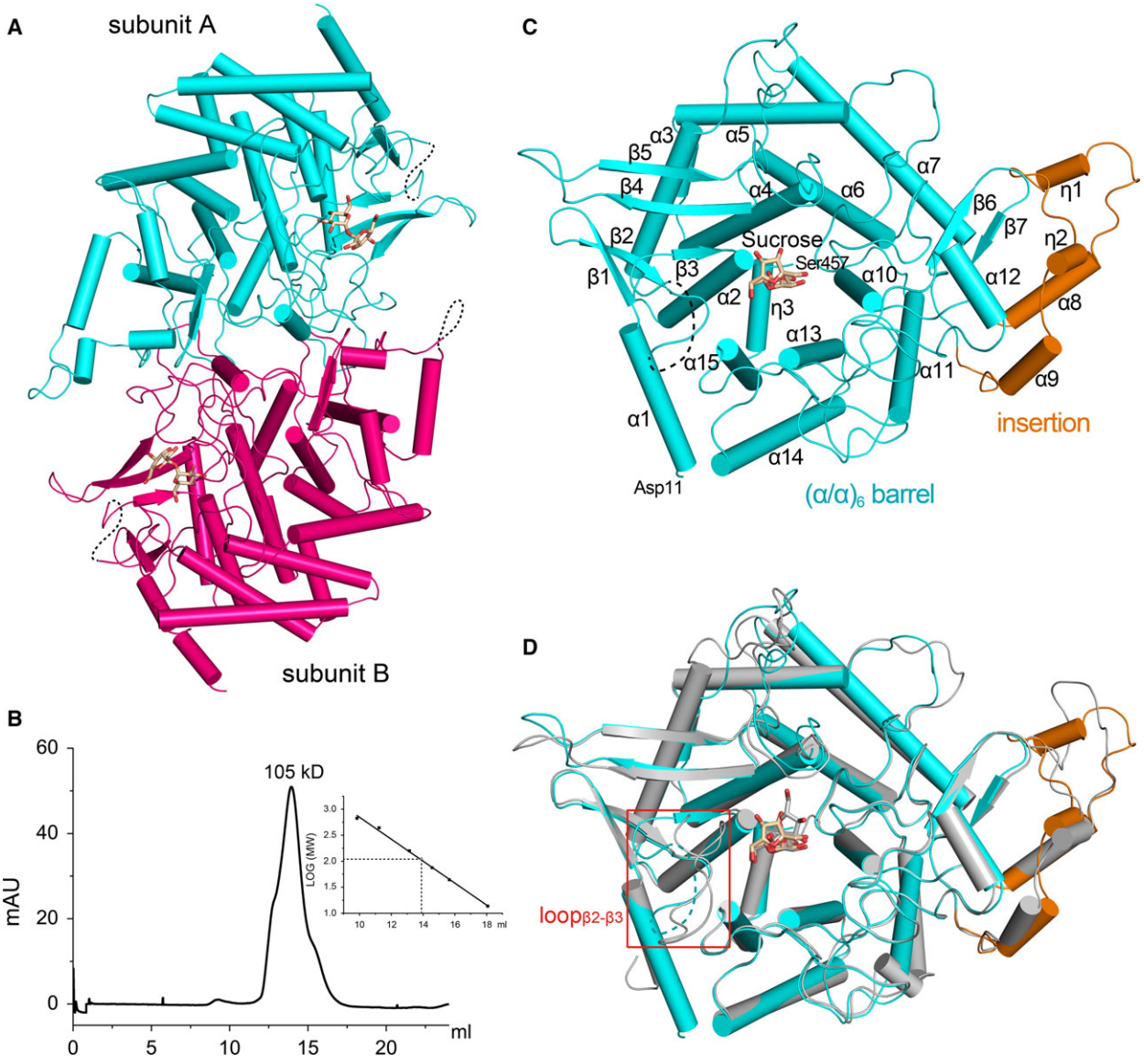


Fig. 1. Overall structure of InvB. (A) Dimeric structure. Helices are shown as cylinders, and sucrose molecules in the active site are shown as wheat sticks. The subunits A and B are colored in cyan and hotpink, respectively. The disordered regions are shown as dotted lines. (B) Molecular mass determination of InvB by analytical gel filtration chromatography. (C) A close view of InvB monomer. Subunit A in InvB-Suc is used as an example. The core $(\alpha/\alpha)_6$-barrel and the insertion part are colored in cyan and orange, respectively. The sucrose molecule is shown in sticks. (D) Superposition of InvB monomer against InvA (PDB entry 5GOP) monomer, which is shown in gray. The different loop close to the active site is highlighted in red rectangle.

Agilent 1200 Series coupled with an evaporative light scattering detector (ELSD, Alltech 2000ES). The samples were injected in volumes of 10 μL onto a Prevail™ Carbohydrate ES column (4.6 × 250 mm, 5 μm, GRACE) and the column temperature was kept at 25 °C. An acetonitrile/water (70 : 30, v/v) solution was used as the mobile phase at 1 mL·min^{-1}. For ELSD, the temperature of the detector nebulizer was set to 85 °C, and the gas flow was 2.5 L·min^{-1}. The kinetic determinations of wild-type enzymes and the mutants were performed at different sucrose concentrations, and the K_m and k_{cat} values were calculated by nonlinear fitting to the Michaelis–Menten equation using the program ORIGIN 8 (OriginLab, Northampton, MA, USA). Three independent determinations were made to calculate the means and standard deviations.

Phylogenetic analysis

The sequence of InvB was used in the BLAST search against the nonredundant protein sequence database (http://blast.ncbi.nlm.nih.gov/Blast.cgi). In total 60 A/N-Inv homologs from plants including green algae, cyanobacteria, and proteobacteria were selected to do multiple sequence alignment by the program MULTALIN [37] and the result was visualized by ESPript 3.0 [38]. Then the rooted phylogenetic tree was constructed using the neighbor-joining method or maximum likelihood method by MEGA 6 [39] with the α-1,6 glucosidase domain of glycogen debranching enzymye of *Giardia lamblia* ATCC 50803 as the outgroup. Confidence limits were assessed by bootstrap analysis using 1000 replicates.

Results

Overall structure of InvB

We obtained crystals of the full-length native and SeMet-substituted InvB, but failed in optimizing because of the twinned crystals. Partial proteolysis combined with multiple sequence alignment indicated that both N- and C-termini are not conserved and most likely unstructured. Thus, a truncated version of InvB (residues Thr9–Ser457) was constructed in a similar way to that of InvA [23], and eventually the SeMet-substituted crystal structure of the apo-form was successfully determined at 1.93 Å resolution in the space group *C2*. Afterward, we obtained the sucrose-complexed crystals by soaking crystals of the truncated InvB with sucrose. The 1.95 Å complex structure of InvB with sucrose was solved by using the apo-form SeMet-substituted InvB structure as the initial model. The parameters for data collection and structure determination are listed in Table 1.

In the sucrose-complexed structure (InvB-Suc), each asymmetric unit contains two molecules of InvB, which form a stable dimer (Fig. 1A) with a buried interface area of ~ 5500 Å^2 as calculated by PDBePISA. The dimerization in crystal structure is consistent with the result obtained from the size exclusion chromatography (Fig. 1B). In fact, InvA exists as a hexamer composed of three dimers [23]. Besides, the plant A/N-Invs also assemble into oligomers, either tetramers [40–43] or octamers [44]. Therefore, we propose that the dimer may be the basic structural unit of these A/N-Invs. The InvB monomer has an overall structure similar to InvA, both of which possess an $(\alpha/\alpha)_6$-barrel core structure composed of 12 α-helices and an insertion structure containing helices α8, α9, η1, and η2 (Fig. 1C) [23].

The apo-form structure of InvB shares an overall structure almost identical to that of the sucrose complex, with a root mean square deviation (RMSD) of 0.52 Å over 856 Cα atoms of the dimer. Notably, a glycerol molecule, which was used as the cryoprotectant, occupies the substrate binding pocket.

InvB has a high structural similarity to InvA with a RMSD of 0.98 Å over 414 Cα atoms. The loop between β2 and β3 strands (designated as $loop_{β2–β3}$), which is close to the substrate binding pocket, in InvB is different from that in InvA (Fig. 1D). Superposition

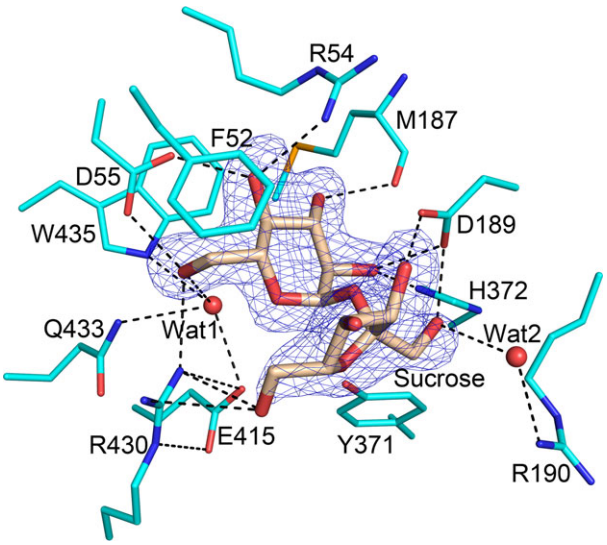


Fig. 2. The substrate binding pocket of InvB. The involved residues are shown as cyan sticks, the sucrose molecule is shown as wheat sticks, and the water molecules are shown as red spheres. The polar interactions are indicated by dashed lines. The $|F_o|-|F_c|$ omit electron density map of the ligand contoured at 3.0 σ is shown as blue mesh.

of the dimeric InvB against the corresponding dimer of the hexameric InvA revealed a difference of about 10° for the intersubunit angles (Fig. S1a). In consequence, the insertion moieties that contribute to the majority of the interdimer interactions of InvA hexamer [23] shift away, resulting in a dimeric form of InvB. Besides, sequence analysis revealed that the residues at the interdimer interface of InvA show a relatively lower sequence conservation among different A/N-Invs (Fig. S1b). Notably, the charged Arg251 and Glu286 that form a couple of interactions with Arg422 from the neighboring dimer in InvA (Fig. S1c) are missing in InvB.

The active-site pocket

In the sucrose-complexed structure InvB-Suc, each subunit binds one sucrose molecule (Fig. 2) which adopts a same conformation as that of subunit C of InvA-Suc structure [23]. InvB shares a nearly identical active-site pocket as InvA except for Arg430, Asn47, and Tyr48 (Figs 3A and S2). Site-directed mutagenesis combined with activity assays showed that Asp189 and Glu415, corresponding to Asp188 and Glu414 of InvA, are catalytic residues. As stated above, superposition revealed that InvA and InvB mainly differ in the $loop_{\beta2-\beta3}$ close to the substrate binding pocket

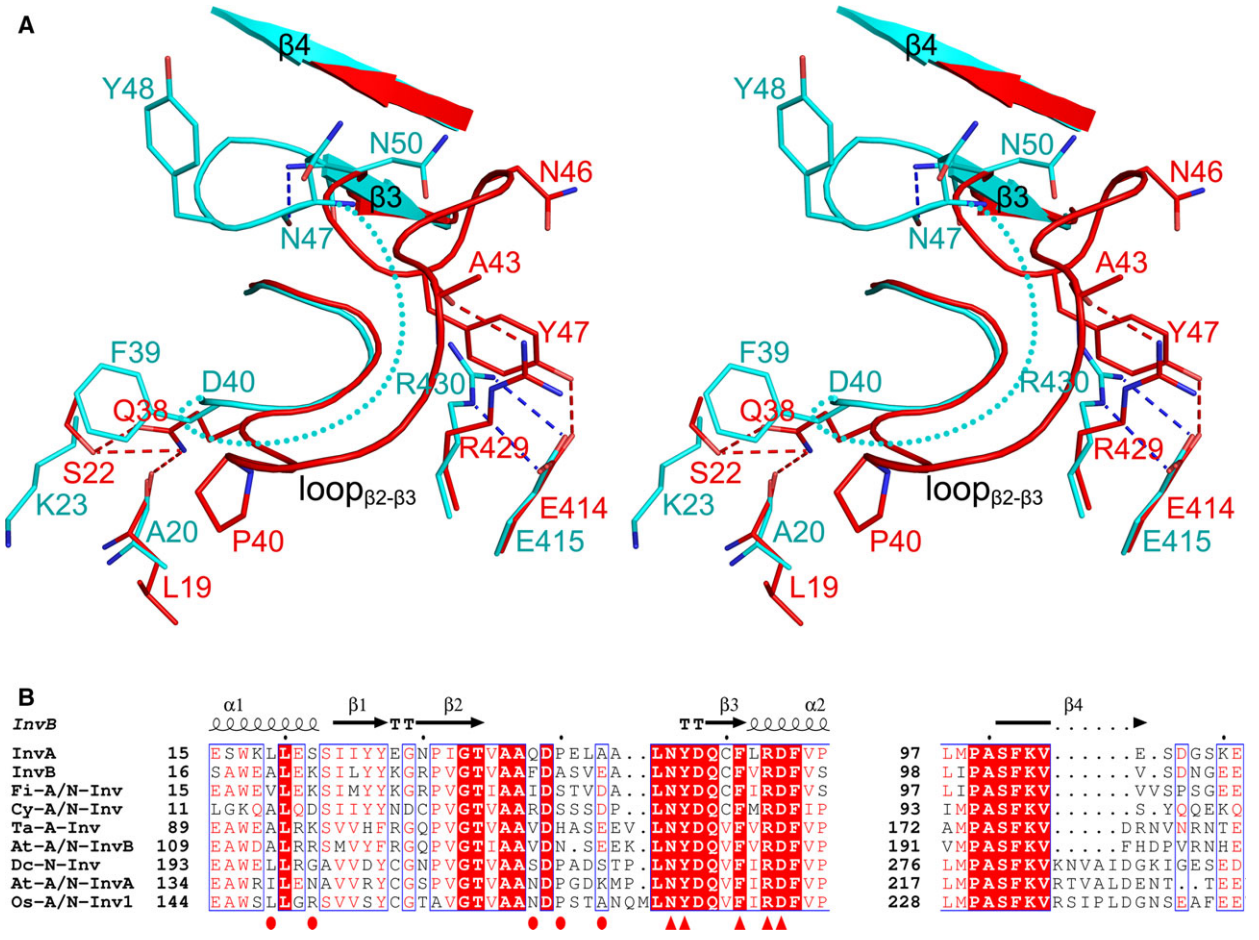


Fig. 3. The difference of the $loop_{\beta2-\beta3}$ close to the substrate binding pocket between InvA and InvB. (A) Stereo representation of $loop_{\beta2-\beta3}$. Subunit A in InvB-Suc structure and subunit A in InvA-Suc structure (PDB entry 5GOP) are shown as cartoon in cyan and red, respectively. The disordered region is shown as dotted lines, with the flanking residues labeled. The key residues are shown as sticks. The polar interactions are indicated by dashed lines. (B) Sequence alignment of the $loop_{\beta2-\beta3}$ in A/N-Invs. Secondary structure elements of InvB are shown at the top of the alignment. The substrate binding residues and the residues probably involved in determining the $loop_{\beta2-\beta3}$ structure are depicted by red triangles and red circles, respectively. The sequences of A/N-Invs are from the following species: *Anabaena* sp. PCC 7120 (InvA, WP_010995690.1; InvB, CAC85155.1), *Arabidopsis thaliana* (At-A/N-InvA, NP_176049.1; At-A/N-InvB, NP_195212.1), *Cyanothece* sp. PCC 7822 (Cy-A/N-Inv, WP_013325329.1), *Daucus carota* (Dc-N-Inv, CAA76145.1), *Fischerella* sp. JSC-11 (Fi-A/N-Inv, ZP_08987807.1), *Oryza sativa* subsp. Japonica (Os-A/N-Inv1, NP_001049936), and *Triticum aestivum* (Ta-A-Inv, CAL26914.1).

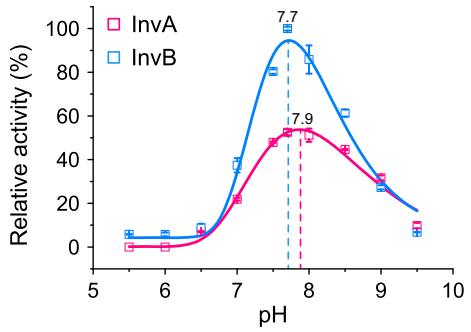


Fig. 4. The optimum pH of InvA (red) and InvB (blue). The invertase activity was determined with 100 mM sucrose in a pH range from 5.5 to 9.5. The relative activity of InvB at pH 7.7 was set to 100%. The error bars denote standard deviations of the mean calculated from three independent experiments.

(Fig. 1D). In subunit A of InvA-Suc structure, the $loop_{\beta2-\beta3}$ which can be fully traced in the electron density map inserts into the pocket through Asn46 and Tyr47 to participate in the binding of sucrose, and Tyr47 is also stabilized by π-stacking with Arg429 (Fig. 3A). But in the corresponding loop of InvB-Suc structure, residues from Ala41 to Leu46 are missing in the electron density map. Besides, residues Asn47 and Tyr48 of InvB, which are corresponding to Asn46 and Tyr47 of InvA, shift outwards from the catalytic pocket. Structural comparison combined with sequence alignment revealed several other factors that determine the different conformations of $loop_{\beta2-\beta3}$ in InvA and InvB (Fig. 3A, B). InvB possesses longer β3 and β4 strands and a preceding β-turn between Asn47 and Asn50, whereas polar interactions of Ala43 with Arg429, in addition to Gln38 with Ser22 and Leu19 could be found in InvA. Moreover, the residue Pro40 makes the loop of InvA adopt a relatively rigid conformation. In contrast, InvB has a more flexible $loop_{\beta2-\beta3}$, a part of which is invisible in the electron density map.

Optimum pH of InvB

To compare the catalytic properties of InvB with InvA, we first performed a series of enzymatic assays. We detected that InvB has an optimum pH at 7.7 (Fig. 4), which is close to the optimum pH of InvA [23], but different from the result reported by Vargas *et al.* [24]. In fact, an earlier report of Schilling *et al.* identified an alkaline invertase with an optimum pH of 7.5–7.8 in *Anabaena* sp. ATCC 29413, which is equivalent to InvB for their involvement with heterocysts [45]. Therefore, InvB should be also an alkaline invertase like InvA, but not the neutral invertase reported before [24].

Key residue contributing to the enzymatic properties of InvB

Although InvA and InvB have a similar optimum pH, InvB shows an obviously higher catalytic activity than InvA between pH 6.5 and pH 9.0 (Fig. 4). Furthermore, kinetic determinations revealed that InvB has a higher affinity toward sucrose and a higher catalytic efficiency (k_{cat}/K_m) of about five times to that of InvA (Table 2). As stated above, Arg430 of InvB close to the catalytic glutamate residues has a different orientation, compared to its counterpart residue Arg429 of InvA. From the InvB structure complexed with sucrose, Arg430 forms a salt bridge with the catalytic residue Glu415, and also forms five hydrogen bonds with Glu415, Gln433, and the sucrose molecule (Fig. 5A). In the apo-form InvB structure, Arg430 possesses the same orientation (Fig. 5B). Similar to the previous proposals [46,47], we presume that Arg430 could provide the positive charge to decrease the pK_a of the catalytic residue Glu415, which is more susceptible for the deprotonation to trigger the nucleophilic attack. But in the two InvA structures complexed with sucrose molecules adopting different conformation [23], the corresponding residue Arg429 deviates away from Glu414 and is fixed by polar interactions with

Table 2. Kinetic constants of wild-type and mutant InvA and InvB toward sucrose

InvA				InvB			
Enzyme	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (10^2 $M^{-1}\cdot s^{-1}$)	Enzyme	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (10^2 $M^{-1}\cdot s^{-1}$)
WT	20.8 ± 1.7	73.7 ± 1.9	35.4	WT	8.9 ± 0.5	163.8 ± 2.1	184.0
R429A	89.3 ± 5.4	29.1 ± 0.7	3.3	R430A	ND[a]	ND	ND
R429N	27.1 ± 2.3	18.9 ± 0.4	7.0	R430N	18.1 ± 1.6	5.3 + 0.1	2.9

[a]ND, no detectable activity.

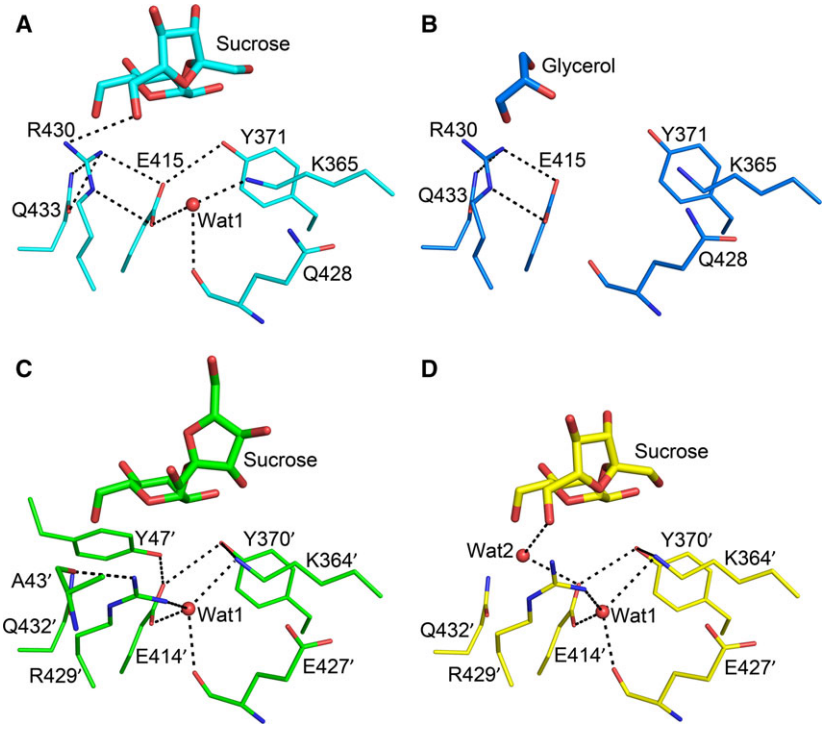
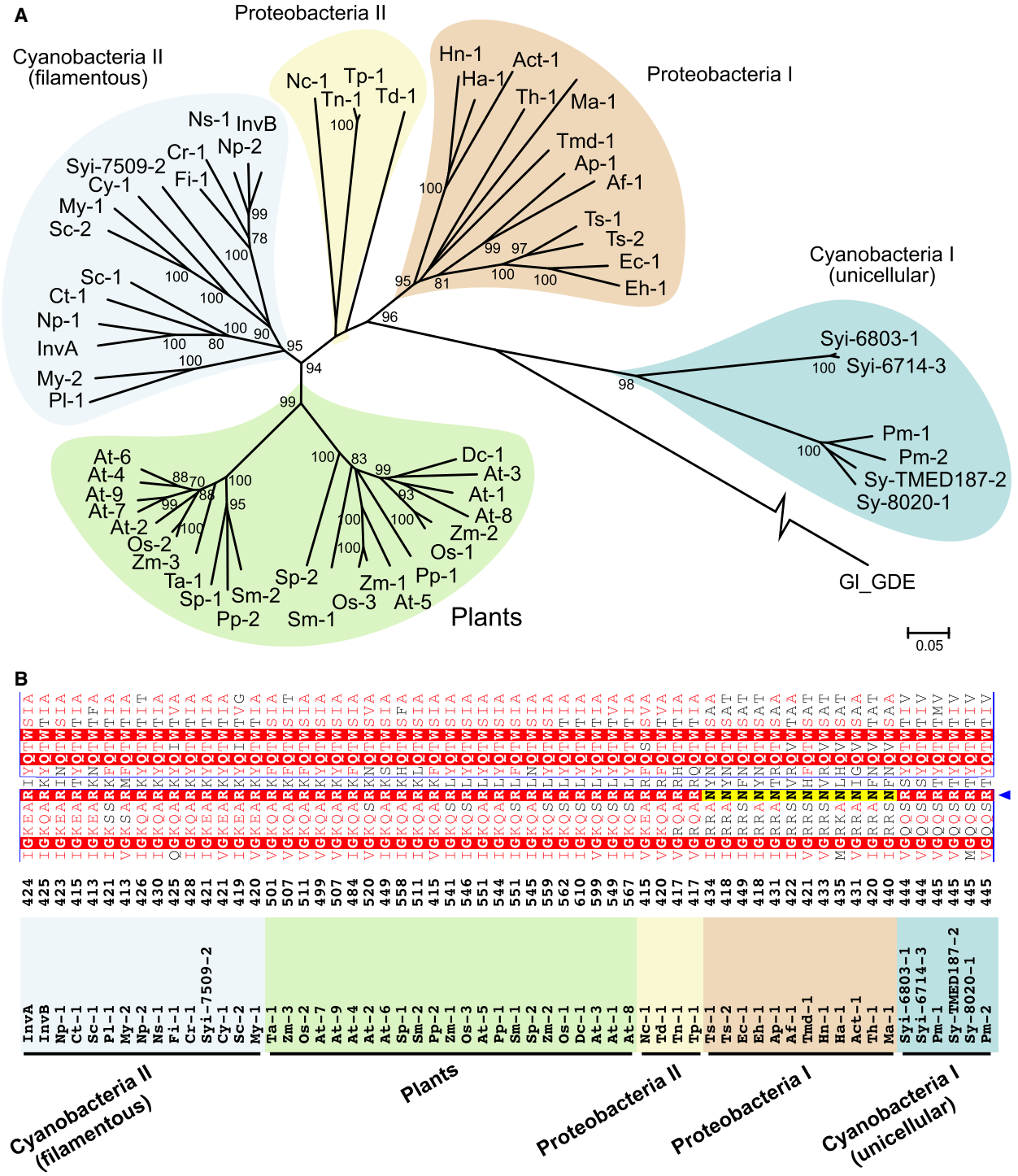


Fig. 5. Conformational comparison of Arg430 of InvB and Arg429 of InvA. (A) In the structure of InvB-Suc, Arg430 directly interacts with the catalytic residue Glu415. The involved residues are shown as lines, the sucrose molecule is shown as sticks, and the water molecule is shown as sphere. The polar interactions are indicated by dashed lines. (B) In the apo-from InvB structure, Arg430 adopts an orientation similar to that in InvB-Suc, despite the substrate binding pocket is occupied by a glycerol molecule. In (C) subunit A and (D) subunit C of InvA-Suc structure, Arg429 interacts with Glu414 *via* a water molecule. The residues in InvA-Suc structure are labeled with a prime.

Fig. 6. Phylogenetic analysis of A/N-Invs and conservation analysis of Arg430. (A) The neighbor-joining method was adopted to construct the rooted phylogenetic tree. The maximum likelihood method produced identical topology. The α-1,6 glucosidase domain of glycogen debranching enzymye of *Giardia lamblia* ATCC 50803 (Gl-GDE, XP_001705733.1) was used as the outgroup. The five major groups are shaded in different colors. Bootstrap values higher than 70% are shown at the nodes. The scale bar represents 0.05 changes per amino acid. (B) The site corresponding to Arg430 of InvB is depicted by a blue triangle, with conserved arginine and asparagine residues shaded in red and yellow, respectively. The grouping of A/N-Invs sequences is consistent with the phylogenetic tree in (A). The sequences of A/N-Invs are from the following species: *Anabaena* sp. PCC 7120 (InvA, WP_010995690.1; InvB, CAC85155.1), *Acidithiobacillus thiooxidans* ATCC 19377 (Act-1, WP_010639088.1), *Acidihalobacter ferrooxidans* (Af-1, WP_076837606.1), *Acidihalobacter prosperus* (Ap-1, WP_065089328.1), *Arabidopsis thaliana* (At-1, NP_176049.1; At-2, NP_195212.1; At-3, NP_187302.2; At-4, NP_564177.1; At-5, NP_197643.1; At-6, NP_177345.1; At-7, NP_174791.2; At-8, NP_001326532.1; At-9, NP_567347.1), *Cylindrospermopsis raciborskii* CS-505 (Cr-1, ZP_06306902.1), *Chroococcidiopsis thermalis* (Ct-1, WP_015156189.1), *Cyanothece* sp. PCC 7822 (Cy-1, WP_013325329.1), *Daucus carota* (Dc-1, CAA76145.1), *Ectothiorhodospira* sp. PHS-1 (Ec-1, WP_083838783.1), *Ectothiorhodospira haloalkaliphila* (Eh-1, WP_026623645.1), *Fischerella* sp. JSC-11 (Fi-1, ZP_08987807.1), *Halothiobacillus* sp. LS2 (Ha-1, WP_066101158.1), *Halothiobacillus neapolitanus* c2 (Hn-1, WP_012823125.1), *Magnetofaba australis* IT-1 (Ma-1, WP_085442439.1), *Myxosarcina* sp. GI1 (My-1, WP_036482288.1; My-2, WP_036476871.1), *Nitrosomonas cryotolerans* (Nc-1, SFQ17342.1), *Nostoc punctiforme* ATCC 29133 (Np-1, CAD37134.1; Np-2, CAD37133.1), *Nodularia spumigena* CCY9414 (Ns-1, ZP_01631199.1), *Oryza sativa* subsp. Japonica (Os-1, NP_001049936; Os-2, XP_015625957.1; Os-3, XP_015627031.1), *Pleurocapsa* sp. PCC 7319 (Pl-1, WP_019507642.1), *Prochlorococcus marinus* (Pm-1, WP_011131014.1; Pm-2, WP_036892687.1), *Physcomitrella patens* (Pp-1, XP_001754878.1; Pp-2, XP_001758344.1) *Stanieria cyanosphaera* (Sc-1, WP_041619725.1; Sc-2, WP_015195523.1), *Selaginella moellendorffii* (Sm-1, XP_002978791.1; Sm-2, XP_002968256.1), *Spirogyra pratensis* (Sp-1, GBSM01000698.1; Sp-2, GBSM01023190.1; obtained from the Transcriptome Shotgun Assembly database of NCBI and referred to reference [13]), *Synechococcus* sp. WH 8020 (Sy-8020-1, WP_048348470.1), *Synechococcus* sp. TMED187 (Sy-TMED187-2, OUW48797.1), *Synechocystis* sp. (PCC 6803, Syi-6803-1, CAD33848.1; PCC 7509, Syi-7509-2, WP_009632367.1; PCC 6714, Syi-6714-3, WP_028948477.1), *Triticum aestivum* (Ta-1, CAL26914.1), *Thiohalorhabdus denitrificans* (Td-1, WP_054965954.1), *Thiohalospira halophila* DSM 15071 (Th-1, WP_093427069.1), *Thiohalomonas denitrificans* (Tmd-1, WP_092994335.1), *Thioalkalivibrio nitratireducens* (Tn-1, WP_043739311.1), *Thioalkalivibrio paradoxus* (Tp-1, WP_006747994), *Thioalkalivibrio sulfidophilus* (Ts-1, WP_012637287.1; Ts-2, WP_026289890.1), and *Zea mays* (Zm-1, XP_008661659.1; Zm-2, XP_008655058.1; Zm-3, NP_001130493.1).



Ala43 and a water molecule, in addition to π-stacking against Tyr47 (Fig. 5C, D). Consequently, the activation of Glu414 in InvA is driven by the water molecule stabilized by Arg429 and Lys364. As expected, the mutant R430A of InvB completely loses the hydrolysis activity, whereas the mutant R429A of InvA maintains a residual activity of about 9.3% to the wild-type protein (Table 2). Notably, upon mutation of the arginine to asparagine, the enzymatic activities of InvB and InvA decreased to about 1.6% and 19.8%, respectively (Table 2). All together, the residue Arg430 is important for the activity of InvA and InvB; and moreover,

its conformation may also make InvB differ from InvA in catalytic efficiency.

Discussion

The higher activity of InvB might be necessary for the heterocyst development and nitrogen fixation

Recent studies reported that *Anabaena* sp. PCC 7120 *invA* is only expressesed in vegetative cells, whereas *invB* is expressesed in both vegetative cells and heterocysts [6]. The spatiotemporal expression assay showed that the amount of InvB increased higher than that of InvA and particularly accumulated in heterocysts upon deprivation of combined nitrogen [26]. Interestingly, Vargas *et al*. found a possible DNA binding site of NtcA, a global nitrogen responsive regulator, in the promoter region of *invB* [6]. Besides, the increase of *invB* expression upon combined nitrogen deprivation was impaired in the *ntcA* mutant strain [26]. These findings implied that the expression of InvB may be regulated by nitrogen starvation. Moreover, knock-out of *invB* impaired diazotrophic growth, but no affect was detected in *invA* mutant strain [6,26].

Nitrogen fixation requires lots of ATP, reductant, and carbon skeletons, which come from the carbohydrate metabolism, most likely starting with sucrose imported from vegetative cells [7,48]. As there is no SUS or Ac-Inv activity in heterocysts, A/N-Inv, namely InvB becomes vital for nitrogen fixation. What is more, rather than InvA, InvB of higher activity could fulfill the necessity for large amount of carbon and energy *via* sucrose degradation. Moreover, the hexoses produced from sucrose could also be processed through the oxidative pentose pathway to yield α-ketoglutarate and NADPH. Notably, the reductant NADPH is necessary for the nitrogenase, and also needed for the synthesis of cell-envelope components, such as polysaccharide and glycolipids of heterocysts [26,48].

Extend the species distribution of A/N-Invs to proteobacteria

To date, the prokaryotic A/N-Invs are only found in cyanobacteria. To further decipher the evolutionary hints of A/N-Invs, we performed the phylogenetic analysis, which classified A/N-Invs into five major groups, including two groups in cyanobacteria, two groups in proteobacteria, and one group in plants (Fig. 6A). Notably, in addition to the previously identified clades in cyanobacteria and plants [24,49], we identified two new clades of A/N-Invs in proteobacteria. Sequence analysis suggested that proteobacteria A/N-Invs share 41–54% sequence identities with InvB and harbor conserved catalytic and substrate binding residues (Fig. S3). Two distant cyanobacteria groups representing unicellular and filamentous cyanobacteria, respectively, are separated by proteobacteria A/N-Invs from chemolithoautotrophic bacteria or anoxygenic photosynthetic bacteria. In addition, these proteobacteria A/N-Invs can be divided into two sister clades. The inconsistence of the topology between A/N-Invs phylogeny and the organismal phylogeny suggested that proteobacteria are likely to have acquired A/N-Inv genes from cyanobacteria through horizontal gene transfer, similar to other sucrose-metabolizing enzymes [4,12,50,51].

In fact, cyanobacteria and proteobacteria are almost the only prokaryotes known to synthesize and metabolize sucrose [4,8–12]. In agreement with it, here we identified putative A/N-Invs in proteobacteria. Specifically, multiple sequence alignment revealed that all A/N-Invs share a conserved arginine residue corresponding to Arg430 of InvB except for the first clade of proteobacteria A/N-Invs substituted by asparagine (Fig. 6B). Our results showed that this arginine is important for the invertase activity, and the asparagine mutant only retains weak activity. Base on the phylogenetic analysis, we propose that the acquisition of the A/N-Inv genes by proteobacteria probably occurred in a single event and in the early period during the evolution of A/N-Invs among cyanobacteria. Thus, most proteobacteria A/N-Invs (clade I) kept a proposed ancient asparagine at Arg430 site with a relatively lower activity. However, the proteobacteria from the second clade (clade II) that live in highly extreme environments, such as hypersalinity (up to 4 M NaCl) [52,53], extremely alkaline pH (optimum at 10.0) [53], or cold temperature (as low as −5 °C) [54,55], changed the asparagine into arginine. This substitution might enhance sucrose hydrolysis to provide osmoprotective substances and energy in order to overcome the harsh environments, and it was also adopted in cyanobacteria and plants with high demand of sucrose metabolism.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grants No. 31630001, 31370757, and 31500598). We appreciate the staff at the Shanghai Synchrotron Radiation Facility (SSRF) and the Core Facility Center for Life Sciences in University of Science and Technology of China for technical assistance. We are grateful to all of the

developers of the CCP4 SUITE, PYMOL, PHENIX, ESPript, and MEGA.

Author contributions

JX, HXH, YC, and CZZ designed the study; JX, KC, LYX, and FY performed the experiments; JX, HXH, YLJ, YC, and CZZ analyzed the data; JX, YC and CZZ, wrote the paper.

References

1 Vargas WA and Salerno GL (2010) The Cinderella story of sucrose hydrolysis: alkaline/neutral invertases, from cyanobacteria to unforeseen roles in plant cytosol and organelles. *Plant Sci* **178**, 1–8.

2 Ruan Y-L, Jin Y, Yang Y-J, Li G-J and Boyer JS (2010) Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. *Mol Plant* **3**, 942–955.

3 Koch K (2004) Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Curr Opin Plant Biol* **7**, 235–246.

4 Salerno GL and Curatti L (2003) Origin of sucrose metabolism in higher plants: when, how and why? *Trends Plant Sci* **8**, 63–69.

5 Ruan Y-L (2014) Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annu Rev Plant Biol* **65**, 33–67.

6 Vargas WA, Nishi CN, Giarrocco LE and Salerno GL (2011) Differential roles of alkaline/neutral invertases in *Nostoc* sp. PCC 7120: Inv-B isoform is essential for diazotrophic growth. *Planta* **233**, 153–162.

7 Curatti L, Flores E and Salerno G (2002) Sucrose is involved in the diazotrophic metabolism of the heterocyst-forming cyanobacterium *Anabaena* sp. *FEBS Lett* **513**, 175–178.

8 Lunn JE (2002) Evolution of sucrose synthesis. *Plant Physiol* **128**, 1490–1500.

9 Diricks M, De Bruyn F, Van Daele P, Walmagh M and Desmet T (2015) Identification of sucrose synthase in nonphotosynthetic bacteria and characterization of the recombinant enzymes. *Appl Microbiol Biotechnol* **99**, 8465–8474.

10 Syamaladevi DP, Jayaraman N and Subramonian N (2013) Evolutionary relationship between available homologous sequences of sucrose phosphate phosphatase (SPP) enzyme. *Sugar Tech* **15**, 136–144.

11 Empadinhas N and da Costa MS (2008) Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. *Int Microbiol* **11**, 151–161.

12 Wu R, Asención Diez MD, Figueroa CM, Machtey M, Iglesias AA, Ballicora MA and Liu D (2015) The crystal structure of *Nitrosomonas europaea* sucrose synthase reveals critical conformational changes and insights into sucrose metabolism in prokaryotes. *J Bacteriol* **197**, 2734–2746.

13 Wan H, Wu L, Yang Y, Zhou G and Ruan YL (2017) Evolution of sucrose metabolism: the dichotomy of invertases and beyond. *Trends Plant Sci* **23**, 163–177.

14 Khmelenina VN, Kalyuzhnaya MG, Sakharovsky VG, Suzina NE, Trotsenko YA and Gottschalk G (1999) Osmoadaptation in halophilic and alkaliphilic methanotrophs. *Arch Microbiol* **172**, 321–329.

15 Cumino AC, Marcozzi C, Barreiro R and Salerno GL (2007) Carbon cycling in *Anabaena* sp. PCC 7120. Sucrose synthesis in the heterocysts and possible role in nitrogen fixation. *Plant Physiol* **143**, 1385–1397.

16 Coleman HD, Yan J and Mansfield SD (2009) Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proc Natl Acad Sci U S A* **106**, 13118–13123.

17 Sturm A (1999) Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. *Plant Physiol* **121**, 1–8.

18 Liu S, Lan J, Zhou B, Qin Y, Zhou Y, Xiao X, Yang J, Gou J, Qi J, Huang Y *et al.* (2015) HbNIN2, a cytosolic alkaline/neutral-invertase, is responsible for sucrose catabolism in rubber-producing laticifers of *Hevea brasiliensis* (para rubber tree). *New Phytol* **206**, 709–725.

19 Lammens W, Le Roy K, Van Laere A, Rabijns A and Van den Ende W (2008) Crystal structures of *Arabidopsis thaliana* cell-wall invertase mutants in complex with sucrose. *J Mol Biol* **377**, 378–385.

20 Alberto F, Jordi E, Henrissat B and Czjzek M (2006) Crystal structure of inactivated *Thermotoga maritima* invertase in complex with the trisaccharide substrate raffinose. *Biochem J* **395**, 457–462.

21 Alberto F, Bignon C, Sulzenbacher G, Henrissat B and Czjzek M (2004) The three-dimensional structure of invertase (beta-fructosidase) from *Thermotoga maritima* reveals a bimodular arrangement and an evolutionary relationship between retaining and inverting glycosidases. *J Biol Chem* **279**, 18903–18910.

22 Reddy A and Maley F (1996) Studies on identifying the catalytic role of Glu-204 in the active site of yeast invertase. *J Biol Chem* **271**, 13953–13958.

23 Xie J, Cai K, Hu H-X, Jiang Y-L, Yang F, Hu P-F, Cao D-D, Li W-F, Chen Y and Zhou C-Z (2016) Structural analysis of the catalytic mechanism and substrate specificity of *Anabaena* alkaline invertase InvA reveals a novel glucosidase. *J Biol Chem* **291**, 25667–25677.

24 Vargas W, Cumino A and Salerno GL (2003) Cyanobacterial alkaline/neutral invertases. Origin of

sucrose hydrolysis in the plant cytosol? *Planta* **216**, 951–960.
25 Curatti L, Giarrocco L and Salerno GL (2006) Sucrose synthase and RuBisCo expression is similarly regulated by the nitrogen source in the nitrogen-fixing cyanobacterium *Anabaena* sp. *Planta* **223**, 891–900.
26 López-Igual R, Flores E and Herrero A (2010) Inactivation of a heterocyst-specific invertase indicates a principal role of sucrose catabolism in heterocysts of *Anabaena* sp. *J Bacteriol* **192**, 5526–5533.
27 Walden H (2010) Selenium incorporation using recombinant techniques. *Acta Crystallogr D Biol Crystallogr* **66**, 352–357.
28 Otwinowski Z and Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.
29 Brodersen DE, de La Fortelle E, Vonrhein C, Bricogne G, Nyborg J and Kjeldgaard M (2000) Applications of single-wavelength anomalous dispersion at high and atomic resolution. *Acta Crystallogr D Biol Crystallogr* **56**, 431–441.
30 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221.
31 Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F and Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.
32 Collaborative Computational Project, N (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760–763.
33 Emsley P and Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
34 Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS and Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21.
35 Krissinel E and Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
36 Delano WL (2002) The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA.
37 Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**, 10881–10890.
38 Robert X and Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–W324.
39 Tamura K, Stecher G, Peterson D, Filipski A and Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
40 Lin C-L, Lin H-C, Wang A-Y and Sung H-Y (1999) Purification and characterization of an alkaline invertase from shoots of etiolated rice seedlings. *New Phytol* **142**, 427–434.
41 Chen JQ and Black CC (1992) Biochemical and immunological properties of alkaline invertase isolated from sprouting soybean hypocotyls. *Arch Biochem Biophys* **295**, 61–69.
42 Liu C-C, Huang L-C, Chang C-T and Sung H-Y (2006) Purification and characterization of soluble invertases from suspension-cultured bamboo (*Bambusa edulis*) cells. *Food Chem* **96**, 621–631.
43 Van den Ende W and Van Laere A (1995) Purification and properties of a neutral invertase from the roots of *Cichorium intybus*. *Physiol Plant* **93**, 241–248.
44 Lee HS and Sturm A (1996) Purification and characterization of neutral and alkaline invertase from carrot. *Plant Physiol* **112**, 1513–1522.
45 Schilling N and Ehrnsperger K (1985) Cellular differentiation of sucrose metabolism in *Anabaena variabilis*. *Z Naturforsch C* **40**, 776–779.
46 Guillén Schlippe YV and Hedstrom L (2005) A twisted base? The role of arginine in enzyme-catalyzed proton abstractions. *Arch Biochem Biophys* **433**, 266–278.
47 Harris TK and Turner GJ (2002) Structural basis of perturbed p*K*a values of catalytic groups in enzyme active sites. *IUBMB Life* **53**, 85–98.
48 Haselkorn R (2007) Heterocyst differentiation and nitrogen fixation in cyanobacteria. In Associative and Endophytic Nitrogen-Fixing Bacteria and Cyanobacterial Associations (Elmerich C and Newton WE, eds), pp. 233–255. Springer, Dordrecht, The Netherlands.
49 Vargas WA, Pontis HG and Salerno GL (2008) New insights on sucrose metabolism: evidence for an active A/N-Inv in chloroplasts uncovers a novel component of the intracellular carbon trafficking. *Planta* **227**, 795–807.
50 Kolman MA, Torres LL, Martin ML and Salerno GL (2012) Sucrose synthase in unicellular cyanobacteria and its relationship with salt and hypoxic stress. *Planta* **235**, 955–964.
51 Chua TK, Bujnicki JM, Tan TC, Huynh F, Patel BK and Sivaraman J (2008) The structure of sucrose phosphate synthase from *Halothermothrix orenii* reveals its mechanism of action and binding mode. *Plant Cell* **20**, 1059–1072.

52 Sorokin DY, Tourova TP, Galinski EA, Muyzer G and Kuenen JG (2008) *Thiohalorhabdus denitrificans* gen. nov., sp. nov., an extremely halophilic, sulfur-oxidizing, deep-lineage gammaproteobacterium from hypersaline habitats. *Int J Syst Evol Microbiol* **58**, 2890–2897.
53 Sorokin DY, Tourova TP, Lysenko AM, Mityushina LL and Kuenen JG (2002) *Thioalkalivibrio thiocyanoxidans* sp. nov. and *Thioalkalivibrio paradoxus* sp. nov., novel alkaliphilic, obligately autotrophic, sulfur-oxidizing bacteria capable of growth on thiocyanate, from soda lakes. *Int J Syst Evol Microbiol* **52**, 657–664.
54 Jones RD, Morita RY, Koops H-P and Watson SW (1988) A new marine ammonium-oxidizing bacterium, *Nitrosomonas cryotolerans* sp. nov. *Can J Microbiol* **34**, 1122–1128.
55 Rice MC, Norton JM, Stein LY, Kozlowski J, Bollmann A, Klotz MG, Sayavedra-Soto L, Shapiro N, Goodwin LA, Huntemann M *et al.* (2017) Complete genome sequence of *Nitrosomonas cryotolerans* ATCC 49181, a phylogenetically distinct ammonia-oxidizing bacterium isolated from arctic waters. *Genome Announc* **5**, e00011–e00017.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1. Different oligomerization patterns of InvA and InvB.

Fig. S2. The active site comparison of InvA (yellow) and InvB (cyan).

Fig. S3. Multiple sequence alignment of A/N-Invs.