A core of three amino acids at the carboxyl-terminal region of glutamine synthetase defines its regulation in cyanobacteria

Lorena Saelices†§, Rocío Robles-Rengel§, Francisco J. Florencio and M. Isabel Muro-Pastor*

Instituto de Bioquímica Vegetal y Fotosíntesis, CSIC-Universidad de Sevilla, Seville 41092, Spain.

Running Title: Mutational and structural analysis of GS/IFs interaction

Key words: Glutamine Synthetase, Inactivating Factors, Posttranscriptional Regulation, Synechocystis 6803, Protein-Protein Interaction

*For correspondence: M. Isabel Muro-Pastor, Instituto de Bioquímica Vegetal y Fotosíntesis, Américo Vespucio 49, E-41092 Sevilla, Spain, Tel. 34-954-489-573, Fax 34-954-460-065, E-mail: imuro@ibvf.csic.es

†Present address: Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH-Hönggerberg, Wolfgang-Pauli-Str. 10, CH-8093 Zürich, Switzerland

§These authors contributed equally to this work.
Summary

Glutamine synthetase type I (GS) is a key enzyme in nitrogen metabolism, and its activity is finely controlled by cellular carbon/nitrogen balance. In cyanobacteria, a reversible process that involves protein-protein interaction with two proteins, the inactivating factors IF7 and IF17, regulates GS. Previously, we showed that three arginine residues of IFs are critical for binding and inhibition of GS. In this work, taking advantage of the specificity of GS/IFs interaction in the model cyanobacteria *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC 7120, we have constructed different chimeric GSs from these two cyanobacteria. Analysis of these proteins, together with a site-directed mutagenesis approach, indicates that a core of three residues (E419, N456 and R459) is essential for the inactivation process. The three residues belong to the last 56 amino acids of the C-terminus of *Synechocystis* GS. A protein-protein docking modeling of *Synechocystis* GS in complex with IF7 supports the role of the identified core for GS/IF interaction.
Introduction

The glutamine synthetases (GSs, GS: E.C. 6.3.1.2) are a family of large, oligomeric enzymes that catalyze the condensation of ammonium and glutamate to form glutamine, the main nitrogen source for protein and nucleic acid synthesis. GS is present in both prokaryotic and eukaryotic organisms because it is critical to nitrogen metabolism (Robertson & Tartar, 2006). The GS superfamily includes three distinct classes, GS type I, II and III, each differing in molecular size and number of subunits in the holoenzyme (Eisenberg et al., 2000). GS type I, encoded by glnA, is a dodecameric enzyme composed of identical subunits (Mr, ca. 50,000) and is found exclusively in bacteria and archaea (Brown et al., 1994, Yamashita et al., 1990).

GS performs the enzymatic mechanism in two steps to yield glutamine and ADP. The first step of the biosynthetic reaction is the formation of the activated intermediate γ-glutamyl phosphate. A metal ion (magnesium or manganese) coordinates the phosphate oxygen atoms of ATP to allow phosphoryl transfer to the carboxylate group of glutamate, yielding the intermediate. In the second step, ammonia attacks the intermediate and releases the products, a free phosphate and glutamine (Eisenberg et al., 2000, Gill et al., 2002). Residues 323-330 of Salmonella typhimurium form a loop, termed “the Glu327 flap”, that closes the glutamate entrance to shield the intermediate from water hydrolysis. After the phosphoryl group is transferred and the ammonia attacks the intermediate, the Glu327 flap opens the entrance to release glutamine.

Countless studies have established that GS occupies a central position in the regulation of nitrogen metabolism (Leigh & Dodsworth, 2007, Reitzer, 2003, Stadtman, 2001). In E. coli, GS is regulated by several mechanisms, including (a) cumulative feedback inhibition by multiple end products of glutamine metabolism, (b) regulated expression of glnA, and (c) reversible covalent modification of each subunit; responding to carbon and nitrogen signals. Thus, in the presence of abundant carbon sources, nitrogen deficiency results in a high level of GS activity. By contrast, when nitrogen is abundant, GS activity is down regulated.
Two types of covalent modifications can regulate the catalytic activity of bacterial GS: adenyllylation/deadenyllylation of a specific tyrosine residue (Ginsburg et al., 1970) and oxidative modification (Levine et al., 1996, Liaw et al., 1993). In E. coli, adenyllylation of Tyr397 leads to alteration of various catalytic properties, including the inactivation of the biosynthetic activity in the presence of Mg²⁺ (Shapiro et al., 1967, Wulff et al., 1967). The degree of adenyllylation depends on the glutamine and 2-OG levels. For instance, higher glutamine level causes more monomers to be adenyllylated, thereby producing lower activity of glutamine synthetase.

In some archaea, GS is regulated by direct interaction with PII signaling proteins, specifically GlnK (Ehlers et al., 2005, Pedro-Roig et al., 2013). In the case of Bacillus subtilis, GS itself has a regulatory role by directly interacting with transcription factors (TnrA and GlnR) (Wray et al., 2001). Although the enzyme is mostly controlled by feedback inhibition, the TnrA protein also inhibits GS activity (Fedorova et al., 2013). In cyanobacteria we have previously shown that under high nitrogen conditions GS activity in Synechocystis sp. PCC 6803 decreases when two small peptides that behave as inactivating factors are present (IF7 and IF17, of 7 kDa and 17 kDa, encoded by gifA and gifB genes, respectively). The analysis of mutant strains lacking one or both IFs revealed that each of these proteins contributes to GS regulation by inactivation in vivo. A maximal level of inactivation of GS was observed when both proteins were present (García-Domínguez et al., 1999). In contrast, the filamentous cyanobacterium Anabaena sp. PCC 7120 possesses a single gifA gene that encode an IF7-like protein, named IF7A (Galmozzi et al., 2010). The C-terminus of IF17 is 37% identical to IF7 and 34% identical to IF7A, whereas IF7 and IF7A are 53% identical.

We have previously shown that three arginine residues of IF7 (R8, R21 and R28) and their homologous residues of IF17 (R90, R103 and R110) are essential for the interaction with GS (Saelices et al., 2011a). These residues, conserved in all ORFs homologous to IF7 and IF17, are located in the same positions in IF7 and IF7A.
Expression of \textit{gif} genes depends on nitrogen status (Galmozzi et al., 2010, García-Domínguez et al., 2000), which is perceived as changes in the intracellular pool of 2-oxoglutarate (Muro-Pastor et al., 2001). It is worth noting that there is a marked GS/IF specificity between \textit{Synechocystis} and \textit{Anabaena}, despite the high similarity between their GSs. \textit{Synechocystis} and \textit{Anabaena} GS sequences are 77.7% identical. While \textit{Synechocystis} GS (SyGS) can be down regulated by IF7, IF17 and IF7A, \textit{Anabaena} GS (AnGS) is only inactivated by IF7A (Galmozzi et al., 2010). Thus, we hypothesized that only a few residues should be responsible for this specificity, and that those residues would likely be involved in IF recognition and/or interaction.

Considering the different GS/IFs interactions of \textit{Synechocystis} and \textit{Anabaena}, we designed and analyzed different chimeric proteins, as well as mutated variants of SyGS and AnGS. Our results indicate that IFs down regulate GS through the C-terminal residues glutamate 419, asparagine 456 and arginine 459. A computational model of the SyGS structure that predicts that the inactivating factor IF7 binds indeed to this region further supports our studies.

\textbf{Results}

\textit{Cyanobacterial GS inactivation takes place in the C-terminus}. \textit{Anabaena} GS (AnGS) cannot be inactivated by \textit{Synechocystis} IFs; however, IF7A from \textit{Anabaena} fully inactivates \textit{Synechocystis} GS (SyGS) (Galmozzi et al., 2010). First, we constructed chimeric proteins in order to identify the region responsible for this specificity. Three different fragments of \textit{Synechocystis glnA}, encoding the N-terminus of SyGS, were cloned and fused to various fragments of AnGS \textit{glnA}, leading to three chimeric genes (resulting proteins, Chi1, Chi2 and Chi3, are schematized in Fig. 1A). Purified chimeric proteins were enzymatically active and characterized by gel electrophoresis (Fig. S1). They were analyzed for GS inactivation with partners IF7, IF17 or IF7A (Fig. 1B). AnGS and SyGS were analyzed as controls. It is noteworthy that Chi1, Chi2, and Chi3 chimeras behaved as AnGS. They are inactivated by IF7A, but not by IF7 or IF17. Since Chi3 contains only a short sequence from AnGS (Fig. 1A), we inferred that the
region of 56 amino acid residues of the C-terminus corresponds to the region involved in the specificity and hence critical for IF7 and IF17-mediated GS inactivation in *Synechocystis*. As a control, the reverse version of Chi3, containing only the last 56 residues from SyGS (Chi4) was also cloned and analyzed (Fig. 1A and 1B). As expected, Chi4 is inactivated by IF7, IF17 and IF7A, confirming that the C-terminal 56-residue segment from cyanobacterial GS is responsible for the IF specificity *in vitro*.

*In vivo* analysis of chimeric proteins was consistent with the *in vitro* results. The *Synechocystis glnA* gene was replaced by each chimeric version by transformation, generating SChi1-SChi4 strains (Table 1 and Fig. S1 C-D). The correct *glnA* expression in these strains was tested by Northern blot (Fig. S3) and their GS specific activity was similar to that of the wild type *Synechocystis* strain (Fig. S1E). We studied GS inactivation by ammonium addition in the strains containing chimeric versions, in comparison to wild type (Syn6803), and ΔgifAΔgifB strains (García-Domínguez *et al.*, 1999) (Fig. 1C). Consistent with the *in vitro* results, SChi1-SChi3 strains do not show GS inactivation (Fig. 1C). In addition we analyzed *in vivo* accumulation of IF7 and IF17, that requires interaction with GS (Galmozzi *et al.*, 2007). IF7 and IF17 do not accumulate in strains SChi1-SChi3, confirming lack of interaction (Fig 1D). However, the SChi4 strain shows GS regulation kinetics similar to that of the wild-type strain (Fig. 1C). In turn, IF7 and IF17 accumulate in SChi4 strain due to their interaction with GS (Fig. 1D). These results determine that the last 56 residues of the GS drive the specificity of the enzymatic regulation in cyanobacteria.

Two GS residues are involved in the specific GS/IF interaction in *Synechocystis* and *Anabaena*. Taking into account that the GS/IF interaction has an electrostatic nature (Mérida *et al.*, 1991a, Saelices *et al.*, 2011a), the observed specificity between *Synechocystis* and *Anabaena* GS inactivation could be due to differential repulsion and/or attraction pattern. We analyzed the sequence alignment of the last 56 C-terminal residues of SyGS and AnGS (Fig. 2A). We considered every residue that was differentially charged between the two proteins for a mutational analysis (Fig. 2A, shaded residues). In SyGS we made point substitutions of each of the identified residues to that in the corresponding position in AnGS. A fixed amount of each purified GS variant was used in inactivation assays with 2 µM of IF7, IF17 or IF7A.
As shown in Fig. 2B, two single mutations in SyGS that changed completely the inactivation outline, were identified: the SyGS-N456K variant was unable to be inactivated by any inactivation factor (IF7, IF17 or IF7A) and SyGS-R459Q displayed impaired inactivation by IF7 or IF17. We focused on these two residues as the rest of the changes did not alter significantly the pattern of inactivation by the different IFs. Two different experiments tested GS protein function. First, inactivation assays with increasing amounts of IF7, IF17 or IF7A proteins were performed (Fig. 3A). Second, GS/IF interaction was tested by protein-protein gel shift experiments (Fig. 3B). Surprisingly, the introduction of a lysine in the position 456 of SyGS hindered inactivation by IF7A as well, although AnGS presents this lysine in that position. Next, we constructed the double mutant SyGS-N456K/R459Q to check whether the combination of both mutations allows the inactivation by IF7A. Interestingly, the SyGS-N456K/R459Q mutant perfectly mimics the inactivation profile shown by AnGS, inactivation by IF7A but not by IF7 or IF17 (Fig. 2B and Fig. 3A). In addition we constructed a reverse double mutant, AnGS-K457N/Q460R, that contains *Synechocystis* GS residues at the positions homologous to N456 and R459. This mutant is inactivated by IF7, IF17 and IF7A as SyGS (Fig. 3A). These results were consistent with protein-protein gel shift assays, except for IF17, which showed interaction with both SyGS-N456K and SyGS-N456K/R459Q mutants, although this interaction did not cause GS inactivation (Fig. 3B). To study if the residues N456 and R459 of SyGS are also responsible for the specificity *in vivo*, *Synechocystis* glnA gene was replaced by the SyGS-N456K, SyGS-R459Q or SyGS-N456K/R459Q expressing glnA variants by transformation (Table 1 and Fig. S2). The glnA expression level in the resulting strains was tested by Northern blot (Fig. S3). GS inactivation in these strains, compared with the Δgij/ΔgijB and wild type (Syn6803) strains, supported the results obtained *in vitro*. Strains expressing SyGS-N456K, SyGS-R459Q or SyGS-N456K/R459Q variants did not show inactivation by *Synechocystis* IFs (Fig. 4A). We analyzed accumulation of IFs after ammonium addition, IF7 accumulated only in wild type (Syn6803). IF17 accumulation was barely detectable in SN456K and SR459Q strains and undetectable in SN456K/R459Q strain (Fig. 4B). Taking together, these *in vitro* and *in vivo* results strongly indicate that the residues N456
and R459 in SyGS and their analogues K457 and Q460 in AnGS are critical for the specificity of the 

*Synechocystis* and *Anabaena* GS/IF interaction and GS inactivation processes.

**Impaired GS-regulation leads to altered Gln/Glu balance.** We have previously shown that the addition of ammonium to nitrate-grown *Synechocystis* cells provokes a quick and dramatic change in the intracellular pools of Glu and Gln, that are completely restored to the original levels ~30 min after ammonium shift (Mérida *et al.*, 1991b). This restoration to the steady-state levels is impaired in the ΔgifAΔgifB strain and therefore it is the consequence of the GSI inactivation (Muro-Pastor *et al.*, 2001). We hypothesized that mutant strains harboring GS variants not susceptible to inactivation by ammonium must also be impaired in restoring amino acid homeostasis upon this strong change in nitrogen availability. As expected, similar to what happens in a ΔgifAΔgifB strain (Muro-Pastor *et al.*, 2001) and in contrast to what happens in the wild type strain (Syn6803), in the SN456K strain the Gln pool increased continuously after ammonium was added (Fig. 4C).

**Inactivation of SyGS is driven by a core of three residues in vitro.** In order to address the question of which residues of SyGS are involved in enzyme inactivation, we decided to examine the biochemical environment of N456 and R459, using the crystal structure of SyGS (PDB ID 3NG0). It is noteworthy the remarkable number of negatively-charged solvent-exposed residues close to N456 and R459. Since the GS/IFs interaction is electrostatic (Mérida *et al.*, 1991a), N456 and R459, together with other charged, solvent-exposed residues adjacent to them (Fig. 2A, squared residues) were substituted by the nonpolar amino acid alanine or by an oppositely-charged residue; residue N456 was mutated to both charges. These two types of substitutions would allow us to check both the contribution of a particular charged residue (alanine substitutions) and, given the electrostatic nature of the GS/IF interaction, the putatively more drastic effect of the introduction of an opposite charge (change of charge substitutions). The *in vitro* GS inactivation analysis of single mutants showed that N456, R459, and a third amino acid residue E419 are key sites for interaction and inactivation of GS mediated by both IF7 and IF17 (Fig. 5B and Table 2). Substitution of E419, N456 or R459 by alanine entailed partial or total loss of IF7 and/or IF17-mediated
inactivation. In order to determine whether an accumulative effect controls the GS-IF complex formation, we designed and analyzed the triple alanine mutant SyGS-E419A/N456A/R459A that shows total loss of IF7 and partial loss of IF17-mediated GS inactivation (Table 2). It is worth noting that some amino acid substitutions differentially affect IF7 and IF17 function. SyGS-E423K and SyGS-E448K mutants were mainly impaired in IF7-mediated inactivation while SyGS-D452K mutant is much more affected in IF17-mediated inactivation (Table 2). However, the substitutions E419K, N456K, and R459E caused the total loss of GS inactivation mediated either by IF7 or IF17 (Fig. 5B). For these three key residues (E419, N456 and R459) we analyzed additional mutants harboring conservative changes. The replacements E419D, N456Q or R459K (Table 2) caused a partial decay in GS inactivation. These data suggest that the specific side chain identity is important for the regulation mechanism. These results were consistent with protein-protein gel shift assays, except for IF17, which showed interaction with all mutants, although this interaction did not always cause GS inactivation (Fig. 5C).

GS inactivation is also coordinated by the three-residue core in vivo. To corroborate whether the three-residue core identified in vitro is also critical for GS regulation in vivo, the Synechocystis glnA gene was replaced by mutated variants of these residues (Table 1 and Fig. S2). glnA expression was verified by Northern blot (Fig. S3). Consistent with the in vitro data, GS inactivation after ammonium addition was impaired in strains expressing SyGS variants with E419, N456 or R459 substitutions (Fig. 6A). Strains with change-of-charge substitutions exhibited no GS inactivation while those with single alanine substitutions showed partial GS inactivation. The strain with the triple alanine mutant displays little GS inactivation (Fig. 6A). Accordingly, strains expressing GS variants able to bind IF7 or IF17 in gel shift assays also showed IF7 or IF17 accumulation in vivo, respectively (Fig. 6B). Thus, SE419K, SN456K and SR459E strains did not accumulate IF7 and slightly accumulated IF17. SE419A and SR459A strains accumulate a significant amount of IF7 and IF17. The SN456A and SE419A/N456A/R459A strains accumulated high amount of IF17 but no IF7. The results so far strongly indicate that IF/GS interaction
and inactivation require the coordination of the inactivating factors by the triangle formed by E419, N456 and R459 residues (Fig. 5A).

We wanted to investigate the reason way IF17 is able to interact, to some extent, with almost all the GS variants, although in some cases such interaction does not cause inactivation. Previous results of our group suggested that binding of IF17 to the GS is modulated not only by its C-terminal portion (homologous to IF7 and essential for enzyme inactivation), but also by its N-terminus (82-residue-long amino-terminal part not present in IF7) (Saelices et al., 2011b). To analyse if IF17 N-terminal region, not involved in GS inactivation, mediates its interaction with SyGS-N456K we have constructed a SN456K strain expressing a previously characterized chimeric IF, containing the amino-terminal part of IF17 fused to IF7 (IF17N/IF7) (Saelices et al., 2011a), in a genetic background devoid of IF17. IF7 does not accumulate in cells harbouring SyGS-N456 substitutions (Fig. 6B) because it does not interact with these GS versions. Therefore we hypothesized that if the IF17N/IF7 protein accumulates in a SyGS-N456K expressing mutant, it must be interacting with GS by its 82-residue-long IF17N region. The results indicate, as expected, that the SN456K,IF17N/IF7 strain is not susceptible of GS inactivation but the chimeric IF accumulates in the cells after ammonium addition and therefore, must interact with the enzyme (Fig. 6C). This strongly suggests that the amino terminal part of IF17 is responsible for the observed interaction of this factor even with GS versions not susceptible to inactivation.

*Modeled structure of the GS-IF7 complex supports mutational conclusions.* We achieved protein-protein docking modeling using the SyGS structure (PDB ID 3NG0). Both IF7 and IF17 are intrinsically disordered proteins (Muro-Pastor et al., 2003, Saelices et al., 2011b) and share three arginine residues critical for their function (Saelices et al., 2011a). We used the computer algorithm Phyre (Kelley & Sternberg, 2009) to generate a hypothetical structural model of the segment IF7(1-38), which contains the three critical arginine residues. We decided to use only a segment in order to increase the flexibility of the molecule during docking. SyGS and modeled IF7 structures were used to generate a protein-protein docking model of the complex, using the program ClusPro (Kozakov et al., 2010). Among all the
outcomes, we selected the most energetically favorable model (Fig. 7). In this model, IF7 is allocated in a polar pocket between each two GS rings, interacting with the three-residue (E419, N456 and R459) triangle that we have identified in this work. It is worth noting that although the program found hundreds of different possible complex structures, all of them presented the inactivating factor attached to the same outer and polar pocket of the enzyme, with changes within IF7 orientation or packing. The stoichiometry exhibited in this model is two monomers of GS per each inactivating factor, consistent with the calculation made by CD (Saelices et al., 2011b). Furthermore, the three critical arginines of IF7(1-38), R8, R21 and R28, appear to interact with GS and stabilize the IF helix (Fig. 7B). R8, together with Q4, of IF7(1-38) shows interactions with E419, N456 and R459 residues in GS, mainly by hydrogen bonds and ionic interactions (Fig. 7C). R21 of IF7(1-38) seems to maintain the folding by an intra-molecular hydrogen bonding interaction with the E32 residue. More interestingly, R28 of IF7(1-38) interacts with E330 of GS forming a hydrogen bond (2.6 Å) (Fig. 7D). The computational model of the complex GS-IF7(1-38) strongly supports the conclusions of the mutational analysis, that demonstrates that the pocket formed by E419, N456 and R459 is essential for GS inactivation.

Discussion

We demonstrate here that GS inactivation by IFs occurs through the C-terminal region of the enzyme. In addition, we propose an interaction SyGS/IFs mediated by the charged outer surface of the enzyme, involving a three-residue core (E419, N456 and R459) and "the Glu327 flap". According to our in vitro and in vivo mutational analysis, only two amino acid residues are involved in the specificity between Synechocystis and Anabaena GS to the corresponding IFs, N456 and R459 from SyGS and K457 and Q460 from AnGS. The replacement of both residues in SyGS by those present in AnGS causes the loss of regulation by IF7 and IF17, and therefore mimics AnGS inactivation pattern. Conversely, the replacement of K457 and Q460 in AnGS by the residues present in SyGS makes the enzyme susceptible of inactivation by IF7, IF17 and IF7A like SyGS. Several filamentous cyanobacteria genomes, including those of the Anabaena genus, show a proximal localization of the GS/IF coding genes.
(Galmozzi et al., 2010). This fact may be related to genome reorganization phenomena or co-evolutionary processes that are responsible for the observed GS/IF interaction specificity. A comparative analysis of the GS C-terminal sequence from different cyanobacteria clearly shows that both residues N456 and R459 are not conserved, regardless of whether these cyanobacteria contain two, one or none IF gene. This could also support possible coevolution of the genes encoding the GS/IFs system in different cyanobacterial species.

The oceanic genus *Prochlorococcus* is interesting in this context because its GS C-terminus contains similar or identical residues to SyGS in both positions (N456 and R459), but it lacks IF encoding sequences. Probably the GS regulation mechanism would not be necessary in a relatively unchanged environment as is the ocean (Garcia-Fernandez et al., 2004) and therefore IF genes were lost in this genus. In this regard, it would be interesting to check if the *Prochlorococcus* GS is susceptible to inactivation by IF7 and/or IF17.

The total lack of GS inactivation in the strain harboring the SyGS-N456K variant is also supported by the amount of Gln and Glu intracellular pools after ammonium addition (Fig. 4C). The SN456K strain has similar behavior to that reported for the ΔgifAΔgifB strain, in which GS regulation is impaired because of the absence of IFs (Muro-Pastor et al., 2001). These observations demonstrate that the single substitution N456K completely abolishes GS inactivation *in vivo* and confirm that this mechanism is responsible for the maintenance of Gln and Glu cellular homeostasis during high nitrogen levels.

It is worth noting that substitution N456A leads to a partial loss of GS inactivation *in vivo* but has a drastic effect on IF7/GS interaction both *in vitro* and *in vivo*. In fact no IF7 could be detected in cells of the SN456A or the triple alanine substitution strains (Fig. 5B). Taking into account our previous data demonstrating that target protection is required for IF7 accumulation *in vivo* (Galmozzi et al., 2007, Saelices et al., 2011a), these results indicate that IF7 does not interact with the SyGS-N456A variant. On the other hand, IF17 accumulation in SN456A or the triple alanine substitution strains is higher than in the wild type strain. This observation suggests that both IF7 and IF17 somehow compete in their binding
to the GS. In addition, the results obtained with the IF17N/IF7 chimeric protein indicate that IF17 has an additional anchorage site, most probably mediated by its N-terminus, which wouldn't take part in the inactivation process.

Taking into account the results of our mutational analysis and structural model of the complex, there are three different processes that might explain the mechanism underlying GS regulation. First, a change in the GS quaternary structure could lead to a dramatic alteration of the active sites to block enzymatic activity. Second, key residues for substrates and/or products coordination are directly or indirectly involved in the formation of interactions within the GS-IF complex. Finally, the positioning of IFs along the outer surface of the enzyme could hinder the transit of molecules in/out of the active sites. The GS/IF7 docking model suggests a possible mode of inactivation. The strong interaction between R28 of IF7(1-38) and E330 of GS would give the enzyme a permanent closed or 'taut' state. This interaction could hinder the entrance of the substrates or prevent the release of the products. This mechanism of GS regulation is in some way similar to the one mediated by adenylylation/deadenylylation in enterobacteria, which regulates intermediate formation by preventing closure of the E327 flap (Gill et al., 2002). Interestingly the residue adenylylated in enterobacteria, Y397 of E. coli GS, is also located in the C-terminal region of the enzyme (Stadtman, 1991). Additionally, amino acid substitutions of Y397 that alter the environment around it are sufficient to induce changes in GS activity (Luo et al., 2005).

In Bacillus subtilis, mutational analysis has demonstrated that the interaction of TnrA with GS involves a surface-exposed α-helix, next to the Tyr residue homologous to the adenylylated Y397 of enterobacteria (Fisher et al., 2002). This is the same region identified here to be involved in GS/IFs interaction. Interestingly the TnrA/GS interaction also has an inhibitory effect on GS activity (Fedorova et al., 2013).

In summary, in our model the GS/IFs interaction is electrostatic and the three critical arginines of IFs decisively participate in the interactions with GS or IF stabilization. Thereby, IF7 interaction with the GS outer surface seems to create a hydrogen bond between R28 of IF7(1-38) and E330 of GS, closing the gate Glu330 to entrance and/or release of substrates and/or product. Finally, the residues Glu419, Asn456
and Arg459 of SyGS were found to be critical for enzyme regulation and play an important role in the interaction with IF7 through R8 and the helix wherein it is allocated.

Experimental Procedures

Strains and growth conditions. Synechocystis and Anabaena derivative strains were grown photoautotrophically at 30ºC on BG11 medium (Rippka, 1988), using nitrate as nitrogen source, supplemented with 1 g l\(^{-1}\) NaHCO\(_3\) (BG11C) and bubbled with a continuous stream of 1% (v/v) CO\(_2\) in air, under continuous illumination (50 µmol of photons m\(^{-2}\) s\(^{-1}\); white light). Ammonium treatment of cultures was performed by addition of 10 mM NH\(_4\)Cl, and the medium was buffered with 20 mM TES (pH 7.5). For plate cultures, BG11C liquid medium was solidified using 1% (w/v) agar. Strains used in this work are listed in Table 1.

GS assay. GS activity was determined in situ by using the Mn\(^{2+}\)-dependent γ-glutamyl-transferase assay in cells permeabilized with mixed alkyltrimethylammonium bromide (MTA) (Mérida et al., 1991b). For the analysis of the in vitro GS/IFs interaction, binding reactions were performed in a final volume of 20 µl containing purified GS and increasing amounts of IF7, IF17 or IF7A, in Hepes-NaOH buffer, pH 7.0/50 mM KCl. Each sample was subjected to the same GS assay mentioned above for in situ samples, but without MTA addition. One unit of GS activity corresponds to the amount of enzyme that catalyzes the synthesis of 1 µmol min\(^{-1}\) of γ-glutamylhydroxamate.

Plasmid and mutant constructions for purification. Synechocystis glnA was cloned into pBS-SK(+) plasmid (Stratagene) together with five histidine codons inserted after ATG start codon (Galmozzi et al., 2007), generating pSyGS. By standard PCR with oligonucleotides An7F and An1R, Anabaena glnA was amplified bearing a His\(_5\)-tag after the ATG start codon. A SalI-SalI 1,441-bp fragment was cloned into the SalI restriction site of pBS-SK(+) plasmid, giving pAnGS. Using these two plasmids, pSyGS and pAnGS as templates, four chimeric genes were constructed using various fragments of Synechocystis and Anabaena glnA. For this purpose, we used a two-step PCR method that entails the synthesis of overlapping fragments (Higuchi et al., 1988, Saelices et al., 2011a). In addition, primers were designed to
produce site-specific mutations of *Synechocystis* glnA or *Anabaena* glnA. Mutagenesis was performed using the same overlapping PCR method but incorporating the mutations into central overlapping primers as previously described (Saelices et al., 2011a). *Bcl*-XbaI 796-bp fragment from *Synechocystis* glnA of pSyGS was replaced by the different PCR fragments. *Nhe*-PacI 162-bp fragment from *Anabaena* glnA of pAnGS was replaced by the PCR fragment. All DNA constructs were confirmed by DNA sequencing.

**Protein purification.** *Synechocystis* and *Anabaena* GS and IFs expression and purification was carried out as previously described (Galmozzi et al., 2010, Saelices et al., 2011a). Expression vectors used for purification of the different GS or IF variants are listed in the supplemental Table S1. The proteins purified in this work are listed in the supplemental Table S3.

**Protein-Protein Band Shift assay.** The binding reactions were carried out in a final volume of 20 µl containing 1.5 µg (0.12 µM) of purified GS and increasing quantities of IF7, IF17 or IF7A, in Heps-NaOH buffer (pH 7.0), 50 mM KCl. GS-IF complexes were allowed to form during 5 min at room temperature. After the GS-IF complex formation, samples were subjected to 6% nondenaturing polyacrylamide gels run at 25 °C in 25mM Tris-192 mM Glycine (pH 8.3), at 150 V for 2 h. Complexes were visualized by gel staining with Coomassie blue.

**Generation of mutant strains of Synechocystis.** Targeting vector to obtain GS mutant strains, are listed in supplemental Table S1. Previously generated pBS-SK(+) (Stratagene) containing *Synechocystis* glnA locus (pMA1) (Mérida et al., 1992) was used for mutant constructions. The *Kpn*-DraI fragment of the glnA locus was cloned into pBS-SK(+) digested by *Kpn*-SmaI. After removing the XbaI site of the multiple cloning site of the original plasmid pBS-SK(+), targeting vectors were generated by replacing the 796-bp *Bcl*-XbaI fragment, by the mutant variants obtained by site-directed mutagenesis as described above. An Sm′ Spl′ C.S3 cassette (Prentki & Krisch, 1984) from pRL463 (pUC18/19 containing L.HEH1 and C.S3, nomenclature of Elhai & Wolk, (Elhai & Wolk, 1988) was cloned in the *Xba*I site of glnA locus. In the case of Chi4 the strategy was different. The Chi4 chimeric gene used in the *in vitro* study was joined to the upstream *Synechocystis* glnA region by PCR, using oligonucleotides SyChi4 5′ HindIII,
SyChi4 5’ R, SyChi4 and SyChi4R. This PCR-synthesized fragment was cloned into pBS-SK(+) digested by HindIII-KpnI. A downstream Synechocystis glnA region was PCR-synthesized using oligonucleotides SyChi4 3’ KpnI and SyChi4 3’Xhol-KpnI. This fragment was cloned in the KpnI site of the above described plasmid. Finally, an SmI’ Sp’ C.S3 cassette (Prentki & Krisch, 1984) from pRL463 (pUC18/19 containing L.HEH1 and C.S3, nomenclature of Elhai & Wolk, (Elhai & Wolk, 1988) was cloned in the Xhol site incorporated previously in the SyChi4 3’Xhol-KpnI oligonucleotide. The resulting targeting plasmids containing the mutant variants of glnA gene were used to transform the wild type Synechocystis strain. For the generation of a Synechocystis strain expressing the SyGS-N456K and IF17N/IF7 proteins (SN456K,IF17N/IF7), the targeting vectors pS-SyGS-N456K(2) and pCHV (Saelices et al., 2011a) were used to transform the ΔgifB Synechocystis strain (García-Domínguez et al., 1999). All DNA constructs were confirmed by DNA sequencing. Correct recombination was verified by PCR analysis (Fig. S1 and S2). Oligonucleotides used for strains construction and verification are summarized in supplemental Table S2.

RNA isolation and Northern-blot analysis. For Northern-blot analysis, total RNA was isolated and extracted as previously described (García-Domínguez & Florencio, 1997). The concentration of total RNA in each sample was quantified spectrophotometrically at 260 nm. RNA integrity was confirmed by visualization of intact rRNA under UV light. Northern-bLOTS were performed as previously described (Saelices et al., 2011a). PCR-synthesized fragments, encompassing the glnA, gifA or gifB genes were used as probes. As a control the filters were reprobed with a 580-bp DNA fragment containing the constitutively expressed RNase P RNA gene (rnpB) from Synechocystis (Vioque, 1992). Hybridization signals were quantified with a Cyclone Phosphor System (Packard).

Western blot analysis. Anti-IF7, anti-IF17 and Anti-TrxA antisera were obtained previously according to standard immunization protocols (Galmozzi et al., 2007, Marqués et al., 1992, Navarro et al., 2000). For Western blot analysis, proteins were fractionated on 12-15% SDS-PAGE according to the method of Laemmli (Laemmli, 1970) and immunoblotted with anti-IF7 (1:2,000), anti-IF17 (1:2,000) or anti-TrxA.
The ECL Plus immunoblotting system (GE Healthcare) was used to detect the different antigens with anti-rabbit secondary antibodies.

Preparation of crude extracts from Synechocystis cells. For analysis of IF abundance in Synechocystis cells grown under different conditions, crude extracts were prepared using glass beads as previously described (Reyes et al., 1995) in 50 mM Hepes-NaOH buffer (pH 7.0), 50 mM KCl. Equal volumes (typically 10 µl) of the processed samples were loaded on SDS-PAGE. Protein concentration in cell-free extracts or purified protein preparations was determined by the method of Bradford, using ovalbumin as a standard (Bradford, 1976).

Amino acid determination. Cells from 2 ml of culture were recovered by centrifugation, and cell lysates were obtained by adding 0.45 ml of 0.2 N HCl, followed by vigorous shaking and incubation for 15 min on ice. After centrifugation, supernatant was filtered through an Amicon Ultra-0.5ml, Ultracel-10K centrifugal filter (Millipore) for deproteinization. The method used for the analysis of glutamate and glutamine concentration in the deproteinized lysate involves a derivatization of amino acids with phenylisothiocyanate (PITC) (Heinrikson & Meredith, 1984), which binds to primary or secondary amines producing a derivative, phenylthiocarbamyl, that is detected by measuring the absorbance at 254 nm. Sixty microliters of sample were mixed with 60 µL of the derivatizing solution (ethanol:H₂O:triethanolamine:PITC, 7:1:1:1), incubated at room temperature for 30 min, and dried under flowing N₂. The pellet was resuspended in 60 µL of 4 mM sodium phosphate (pH 7.4) and 2% acetonitrile and injected in a HPLC Elite LaChrom (Hitachi) system. The separation was performed using a LichroCART 125-4 column. Amino acids were separated using a linear gradient from 70 mM sodium acetate, 5% acetonitrile buffer (pH 6.55) to acetonitrile/water (50:50). Retention times for glutamate and glutamine was 1.69 and 3.92 min, respectively.

Secondary structure prediction and protein-protein docking analysis. The 38 residue long primary structure of the amino-terminus of the unfolded protein IF7 was used to generate a secondary structure
model by the application Phyre (Kelley & Sternberg, 2009). We selected a short segment in order to increase flexibility of IF7 in the docking. The (1-38) segment of IF7 includes the three critical arginines (Saelices et al., 2011a). Protein-protein docking analysis was performed using the structure of Synechocystis GS (PDB ID 3NG0) and the computational model of IF7. The two structures were subjected to docking experiments using ClusPro (Kozakov et al., 2010). The docking outputs were analyzed on energy provided by the application. Among all docking results, we selected the first and most energetically favorable model.

Acknowledgments

This work was supported by grants BFU 2010-15708 and BFU 2013-41712, cofinanced by FEDER, from the Spanish Ministerio de Economía y Competitividad and by Junta de Andalucía (Bio-284). L.S and R.R.-R were recipients of fellowships from Ministerio de Ciencia e Innovación. We thank Alicia M. Muro-Pastor and Lisa M. Johnson for a critical reading of the manuscript.

References


### Table 1: Cyanobacterial strains used in this study

<table>
<thead>
<tr>
<th>Synechocystis strains</th>
<th>Description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synechocystis sp. PCC 6803</td>
<td>wild type</td>
<td>Garcia-Dominguez et al., 1999</td>
</tr>
<tr>
<td>ΔgifAΔgifB</td>
<td>ΔgifA::npt, ΔgifB::cat, Km(^R),Cm(^R)</td>
<td></td>
</tr>
<tr>
<td>SChi1</td>
<td>glnASy(726)::glnAAn(693)::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SChi2</td>
<td>glnASy(1104)::glnAAn(315)::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SChi3</td>
<td>glnASy(1251)::glnAAn(168)::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SChi4</td>
<td>glnAAn(1254)::glnASy(168)::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SE419A</td>
<td>glnAE419A::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SE419K</td>
<td>glnAE419K::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SE419A/N456A/R459A</td>
<td>glnAE419A/N456A/R459A::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SN456A</td>
<td>glnAN456A::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SN456D</td>
<td>glnAN456D::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SN456K</td>
<td>glnAN456K::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SR459A</td>
<td>glnAR459A::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SR459E</td>
<td>glnAR459E::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SR459Q</td>
<td>glnAR459Q::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SN456K/R459Q</td>
<td>glnAN456K/R459Q::aadA(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>SN456K,IF17N/IF7</td>
<td>glnAN456K::npt, gifB(1-246)::gifA::aadA(^R), Km(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
</tbody>
</table>

Abbreviations: \(^R\) denotes resistance to the indicated antibiotic: Cm, chloramphenicol; Km, kanamycin; Sm, streptomycin; and Sp, spectinomycin.
Table 2: *In vitro* GS inactivation assay for the wild-type enzyme and different mutant versions.

<table>
<thead>
<tr>
<th>Protein name</th>
<th>Proportion (%)</th>
<th>+ IF7 (%)</th>
<th>+ IF17 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SyGS</td>
<td>100.0</td>
<td>6.9</td>
<td>5.8</td>
</tr>
<tr>
<td>SyGS-E419A</td>
<td>100.0</td>
<td>58.2</td>
<td>19.2</td>
</tr>
<tr>
<td>SyGS-E419K</td>
<td>100.0</td>
<td>86.2</td>
<td>90.9</td>
</tr>
<tr>
<td>SyGS-E419D</td>
<td>100.0</td>
<td>45.4</td>
<td>10.8</td>
</tr>
<tr>
<td>SyGS-E423A</td>
<td>100.0</td>
<td>7.0</td>
<td>4.8</td>
</tr>
<tr>
<td>SyGS-E423K</td>
<td>100.0</td>
<td>83.3</td>
<td>26.1</td>
</tr>
<tr>
<td>SyGS-E426A</td>
<td>100.0</td>
<td>8.8</td>
<td>6.1</td>
</tr>
<tr>
<td>SyGS-E426K</td>
<td>100.0</td>
<td>34.9</td>
<td>5.9</td>
</tr>
<tr>
<td>SyGS-E430A</td>
<td>100.0</td>
<td>14.3</td>
<td>9.7</td>
</tr>
<tr>
<td>SyGS-E430K</td>
<td>100.0</td>
<td>22.8</td>
<td>8.8</td>
</tr>
<tr>
<td>SyGS-D441A</td>
<td>100.0</td>
<td>9.2</td>
<td>4.9</td>
</tr>
<tr>
<td>SyGS-D441K</td>
<td>100.0</td>
<td>9.4</td>
<td>6.9</td>
</tr>
<tr>
<td>SyGS-E444A</td>
<td>100.0</td>
<td>8.2</td>
<td>8.3</td>
</tr>
<tr>
<td>SyGS-E444K</td>
<td>100.0</td>
<td>24.0</td>
<td>25.2</td>
</tr>
<tr>
<td>SyGS-E448A</td>
<td>100.0</td>
<td>14.6</td>
<td>10.0</td>
</tr>
<tr>
<td>SyGS-E448K</td>
<td>100.0</td>
<td>42.5</td>
<td>23.4</td>
</tr>
<tr>
<td>SyGS-D452A</td>
<td>100.0</td>
<td>8.0</td>
<td>27.0</td>
</tr>
<tr>
<td>SyGS-D452K</td>
<td>100.0</td>
<td>17.0</td>
<td>84.8</td>
</tr>
<tr>
<td>SyGS-N456A</td>
<td>100.0</td>
<td>98.8</td>
<td>14.6</td>
</tr>
<tr>
<td>SyGS-N456K</td>
<td>100.0</td>
<td>98.2</td>
<td>99.9</td>
</tr>
<tr>
<td>SyGS-N456D</td>
<td>100.0</td>
<td>95.7</td>
<td>93.0</td>
</tr>
<tr>
<td>SyGS-N456Q</td>
<td>100.0</td>
<td>85.6</td>
<td>35.0</td>
</tr>
<tr>
<td>SyGS-R459A</td>
<td>100.0</td>
<td>83.3</td>
<td>50.3</td>
</tr>
<tr>
<td>SyGS-R459E</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>SyGS-R459K</td>
<td>100.0</td>
<td>76.0</td>
<td>40.5</td>
</tr>
<tr>
<td>SyGS-N456A/R459A</td>
<td>100.0</td>
<td>93.4</td>
<td>36.8</td>
</tr>
<tr>
<td>SyGS-E419A/N456A/R459A</td>
<td>100.0</td>
<td>98.8</td>
<td>45.2</td>
</tr>
</tbody>
</table>

The same amount of each GS version (1.5 μg) was assayed alone (Ø) or after incubation (5 min) with 2 μM IF7 or IF17. 100% represent GS activity of each enzyme variant. The percentage of remaining activity after incubation is showed in each case. The values represent arithmetic means from three independent experiments.
**Figure Legends**

**Figure 1. Analysis of chimeric proteins between Anabaena GS (AnGS) and Synechocystis GS (SyGS).**

A. Scheme of the chimeric proteins constructed. Numbers into the boxes indicate the residues corresponding to the GS of each organism. **B. In vitro** inactivation assay of the different chimeric proteins. A fixed amount of each GS was incubated without (Ø) or with 2 µM of IF7, IF17 or IF7A. GS-IF complexes were allowed to form during five minutes and GS transferase activity was determined. **C.** Kinetics of the GS inactivation in Synechocystis strains expressing glnA chimeric genes. 10 mM NH₄Cl was added to Synechocystis cells cultivated with nitrate as nitrogen source. Aliquots were withdrawn and GS transferase activity was measured. The curves represent arithmetic means from three independent experiments and their standard deviation values. **D.** Representative Western blot of IF7 and IF17 along the GS inactivation for each strain. As a protein loading control, membranes were incubated also with anti-TrxA. Thioredoxin A (TrxA) is constitutively expressed in Synechocystis cells.

**Figure 2. In vitro analysis of residues differentially charged between Synechocystis and Anabaena GS C-terminal region.**

A. Sequence alignment of the last C-terminal 56 residues of SyGS and AnGS. ‘*’ indicates positions which have identical residues in the two sequences. Grey shadow represents difference of charge. Charged and exposed residues spatially adjacent to N456 and R459 in the crystal structure of SyGS are framed with a line. **B. In vitro** inactivation assays of the different GSs. Each GS protein (1.5 µg) was incubated without (Ø) or with 2 µM IF7, IF17 or IF7A. GS-IF complexes were allowed to form during five minutes and GS transferase activity was determined.

**Figure 3. In vitro inactivation assays of SyGS, AnGS, and those mutants that are relevant for specificity.**

A. Each GS protein (1.5 µg) was incubated with increasing amounts of IF7 (black squares), IF17 (white squares) or IF7A (black circles). GS-IF complexes were allowed to form during five minutes and GS transferase activity was determined. **B.** GS variants (1.5 µg) were incubated with an excess of
IF7, IF17 or IF7A (11, 12 and 14 μM, respectively). Then GS-IF complexes, together with GS alone (0), were separated in a 6% non-denaturing polyacrylamide gel and stained with Coomassie blue.

Figure 4. Analysis of the GS inactivation in Synechocystis strains expressing SyGS-N456K, SyGS-R459Q and SyGS-N456K/R459Q variants. A. Kinetics of the GS inactivation in Synechocystis wild-type (Syn6803), ΔgifAΔgifB and strains expressing SyGS-N456K, SyGS-R459Q and SyGS-N456K/R459Q variants. 10 mM NH₄Cl was added to Synechocystis cells cultivated with nitrate as nitrogen source. Aliquots were withdrawn and GS transferase activity was measured. The curves represent arithmetic means from three independent experiments and their standard deviation values. B. Representative Western blot of IF7 and IF17 along the GS inactivation for each strain. C. Change in the intracellular Gln and Glu pools upon ammonium up-shift in wild-type (Syn6803) and SyGS-N456K expressing strains. Intracellular concentrations of Gln and Glu pools, relative to total protein, were determined before (t = 0) and after ammonium addition.

Figure 5. In vitro analysis of SyGS-E419, SyGS-N456 and SyGS-R459 variants. A. Structural overview of the C-terminal domain of SyGS (colored in pink). Charged and exposed residues together with N456 are labeled in the close view on the right. B. In vitro inactivation assays of GS variants. Each GS protein (1.5 μg) was incubated with increasing amounts of IF7 (black squares) or IF17 (white squares). GS-IF complexes were allowed to form during five minutes and GS transferase activity was determined. C. GS variants (1.5 μg) were incubated with IF7 or IF17 (11 and 12 μM, respectively). Then GS-IF complexes, together with GS alone (0), were separated in a 6% non-denaturing polyacrylamide gel and stained with Coomassie blue.

Figure 6. Analysis of the GS inactivation in Synechocystis strains expressing SyGS-E419, SyGS-N456 and SyGS-R459 variants. A. Kinetics of the GS inactivation in Synechocystis strains expressing different GS mutant variants. 10 mM NH₄Cl was added to Synechocystis cells cultivated with nitrate as nitrogen source. Aliquots were withdrawn and GS transferase activity was measured. The curves...
represent arithmetic means from three independent experiments and their standard deviation values. B. Representative Western blots of IF7 and IF17 along the GS inactivation for each strain. C. Analysis of the GS inactivation in the SN456K,IF17N/IF7 strain. Kinetics of the GS inactivation after ammonium addition in the SN456K,IF17N/IF7 and wild type strains. Representative Western blots of IF7, IF17 and IF17N/IF7 along the GS inactivation for each strain.

Figure 7. Protein-protein docking modeling of the complex GS-IF7. A. Lateral view of the electrostatic surface of GS, together with IF7, represented in light grey. The 38-residue-long IF7 structural model and the dodecamer from SyGS structure (PDB ID 3NG0) were used to generate a protein-protein docking model of interaction. IF7 appears bound to the belt of the dodecamer, attached to a charged pocket. A square marks the segment zoomed in B. Close-up view of the GS/IF7 interaction region. Orange residues correspond to the three arginines critical for IF function (Saelices et al., 2011a). C. Close view of the binding between IF7-R8 and the three-residue core from GS. Interactions with less than 3.2 Å of distance are marked with dotted lines. IF7-R8 and IF7-Q4 are coordinating the interaction with E419, N456 and R459. D. Close view of the hydrogen bonding between IF7-R28 and SyGS-E330.