GOLDENBRAID 2.0

1. WHY GB2.0?

The new genetic designs used in Biotechnology, and more recently in Synthetic Biology, require the construction of increasingly complex multigenic structures. This is putting pressure in gene synthesis and DNA assembly technologies.

Long DNA strings encoding complex genetic instructions can be custom-built using nucleotides as building blocks. Custom Gene synthesis is becoming increasingly affordable, but the longer the string of DNA, the more difficult and expensive the task becomes. Moreover, direct synthesis gives little room for combinatorial arrangements, something that is highly valued by biotechnologists.

An alternative "building" strategy for complex genetic devices is Modular Construction, that is, the fabrication of new devices by combination of prefabricated standard modules. Modular DNA Construction is an alternative to Custom gene synthesis that brings a number of advantages as speed, versatility, lab autonomy, combinatorial potential and often lower cost. As in any standardized methodology, the more users adopt the standard, the bigger the advantages.

GoldenBraid, version 2.0 (GB2.0) is a Modular DNA Construction method specially designed for Plant Synthetic Biology and Metabolic Engineering. GoldenBraid relies in a very efficient assembly method: 15-20 Kb constructs comprising 4-6 transcriptional units made of dozens of individual pre-fabricated modules (GBparts) can be created routinely in few days. GB2.0 facilitates exchange of genetic modules among laboratories. GB2.0 users can create their own genetic parts and devices following simple and standard construction rules. They also can combine their own parts with those deposited in the **GB2.0 database**, which is an increasingly populated collection of pre-made "GBparts" that conform to the GB standard. In addition, building new GB2.0 constructs will be facilitated by the GB assembler tool, a complete software package that assists users in the design of new GBparts and the construction of more complex genetic devices.

GB2.0 is the latest version of the GoldenBraid system that was developed at the PGB group of the IBMCP-CSIC-UPV in Valencia, Spain (http://www.ibmcp.upv.es/FGB/).

2. WHAT IS GB2.0 ALL ABOUT?

The easiest way to understand what´s GB2.0 is drawing an analogy with a written language. Actually, writing genetic instructions in the form of a DNA string is what GB2.0 is all about.

In written English, letters are joined together to make words following so-called orthographic rules (see also section 8). Then, words are joined together to make sentences following syntactic rules. Finally, sentences can be joined (binarely) one next to the other to make a paragraph. All together makes the English grammar.

In our GB analogy, GB standard parts are equivalent to English words; the GB collection is analogous to an English vocabulary. Sentences (instructions) are equivalent to transcriptional units (TU). And a paragraph is equivalent to a complex genetic instruction made of several transcriptional units.

In the same way that English words are classified in grammatical categories (e.g. nouns, adjectives, verbs, etc) GBparts are also categorized. Each GB category (promoters, coding regions, terminators, etc) occupies a precise position in the genetic sentence (see figure 2.1).

For creating a new English sentence, you take the right words from your vocabulary and assemble them together in a grammatically correct order. Analogously, to create a TU using GB2.0 you just need to take the right standard GBparts from the GB collection, mix them together in a single tube, set up a GB reaction, and the next day you will have clones harboring a grammatically correct sentence. This is called a multipartite GB reaction. You can use your brand new “sentence” to transform your cells. In addition you can keep a copy in the GB collection, so your “sentence” can be re-used to make a paragraph.

Analogously to the composition of and English text, GB TUs can be joined stepwise to make complex instructions: take two of them from the GB collection and mix them in a tube to perform a GB reaction. In less than 24 hours they will be assembled together in a single DNA string. This is called a binary GB reaction. The new assembly can be used directly to transform cells, also kept in the GB collection for further use. You can keep joining new sentences stepwise to your paragraph, or even assemble two small paragraphs together in a single binary reaction.

Words, phrases, sentences and paragraphs in GB collection are all standard and fully exchangeable. Their assembly rules are standardized. Like with natural languages, any speaker can use old words, phrases, sentences and texts kept in the vocabulary and combine them in a single tube reaction to create new and original genetic instructions.

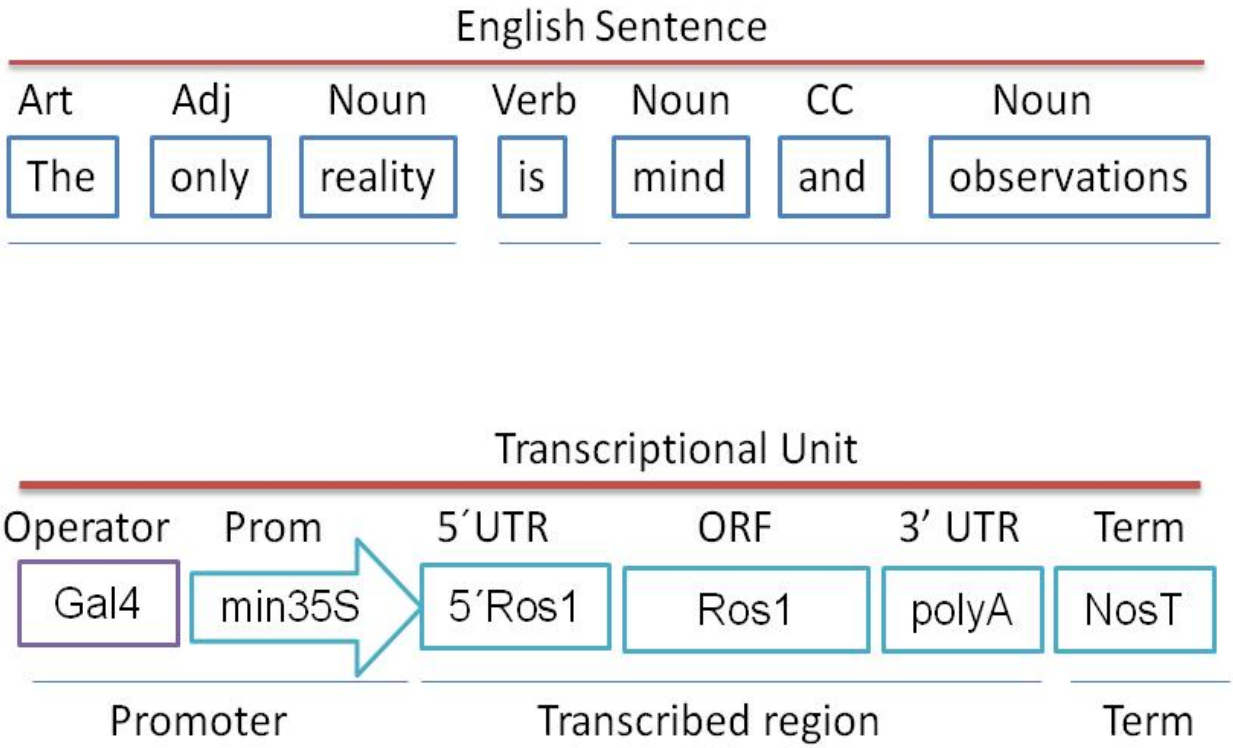


Fig 2.1. Analogy between English sentences and transcriptional units

3. THE CHEMISTRY UNDERNEATH

GB builds increasingly complex DNA strings by means of restriction-ligation reactions between pre-fabricated standard DNA parts. GB uses Type IIS restriction enzymes to generate four-nucleotide sticky ends flanking each DNA piece, so they can be easily and efficiently joined together using an enzyme called T4 ligase.

All GB reactions are performed is the same way: by mixing in a single tube a number of DNA parts (basic or composite parts) and a destination vector in the presence of T4 ligase and appropriate type IIS restriction enzymes. The mix is incubated in a thermocycler that alternates cycles of 16ºC and 37ºC. In these conditions, the restriction enzyme releases the

DNA parts leaving 4 nucleotides overhangs. If the combinations of parts and destination vectors have been chosen correctly, the parts will be orderly ligated by T4 ligases into the destination plasmid. The result is a well-constructed DNA string where each DNA part has occupied its position according to the GB grammar.

The way Type IIS enzymes can be used to efficiently assembly DNA pieces was earlier described by Engler and co-workers (Engler et al., 2009) and the basic process is known as Golden Gate assembly. Golden Gate is an extremely efficient assembly method. Its high efficiency is one of the strong points of GoldenBraid system. GoldenBraid brings Golden Gate a step forward, into a double iterative loop (braid) that ensures the indefinite growth of the multigenic structures. Moreover, GoldenBraid provides standardization to the assembly method.

4. THE COMPONENTS OF GB2.0 SYSTEM

If you want to start using GB2.0, let us introduce you first the components of the system. We have created a GB database from where all components can be consulted.

4.1 GB-destination vectors

The destination vectors (pDGBs) are the recipients of the new assemblies. GB destination vectors are binary plasmids containing a specially designed cloning structure name the “GB cassette”. This cassette contains the blue/white selection lacZ gene flanked by recognition sites of the Type IIS restriction enzymes BsmBI and BsaI. During a GB reaction, the GB cassette of the GB destination vector is replaced by the new DNA assembly. All destination plasmids are labeled as “pDGBs”. The numbers and the Greek letters serve to identify each destination plasmid according to the flanking overhangs left by BsaI and BsmBI digestion respectively.

There are eight types of pDGB:

- Four of them are known as level α plasmids. They are used as destination plasmids in BsaI GB-reactions. These are pDGB_α1, pDGB_α2, pDGB_α1R and pDGB_α2R. Regularly, you

will use only pDGB_α1 and pDGB_α2. However, if you are interested in assembling TUs in reverse orientation, you will be using the other two plasmids (see section 7.1).

- The remaining four plasmids are known as level Ω plasmids (Fig 4.1.1). They are used as destination plasmids in BsmBI-GB reactions. These are pDGB_Ω1, pDGB_Ω2, pDGB_Ω1R and pDGB_Ω2R. Regularly, you will use only pDGB_Ω1, pDGB_Ω2. However, if you are interested in assembling TUs in reverse orientation, you will be using the other two plasmids.

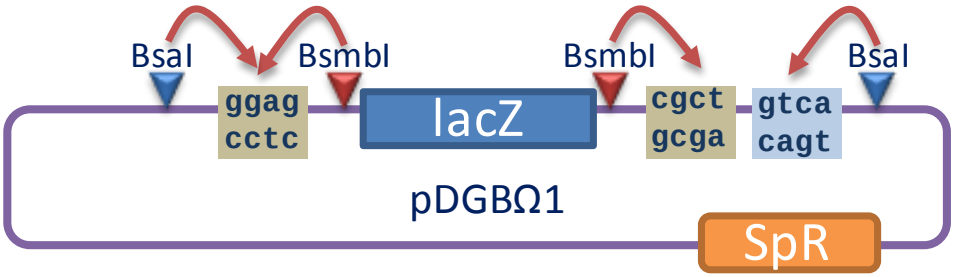


Figure 4.1.1. Schematic view of a destination vector

4.2 GB-basic parts

GB basic DNA pieces named "GBparts" (e.g. promoters, terminators, CDSs, etc), are fragments of DNA flanked by fixed 4 pb overhangs. The flanking overhangs in each part are ordinarily produced after cleavage with a Type IIS restriction enzyme, either BsaI or BtgZI, and therefore GB parts ordinarily come in the form of an insert within a specially designed circular vector named "Universal GB Part Domesticator" or pUPD. Upon BsaI or BtgZI digestion, the GBpart is released from the plasmid leaving 4 nucleotides overhangs. The design of pUPD plasmid is such that both BsaI and BtgZI digestions release exactly the same piece of DNA containing the same 4bp overhangs regardless of the enzyme used. As a consequence, GBparts can be assembled using either BsaI GB-reactions or BtgZI GB-reactions indistinctly. See section 9 to learn how to design a new GBpart (Note: this feature is a major improvement in version GB2.0).

As you probably have anticipated, the 4bp flanking sites define each part´s relative position in the multipartite assembly. In order to ensure that all GB users can exchange their parts, it is necessary that we all use the same flanking sites for the same category. In other words, we

need to use the same ”syntax” within the GB “grammar”, so we can communicate. Therefore we defined a number of functional categories and arbitrarily assigned flanking 4bp to each of these basic categories. See section 6 and Fig 6.1 for a simplified version of GB grammar. See section 8 and Fig 8.2 for a complete view of GB syntax. See section 7.1 to learn how to assemble a new TU using the GB multipartite assembly.

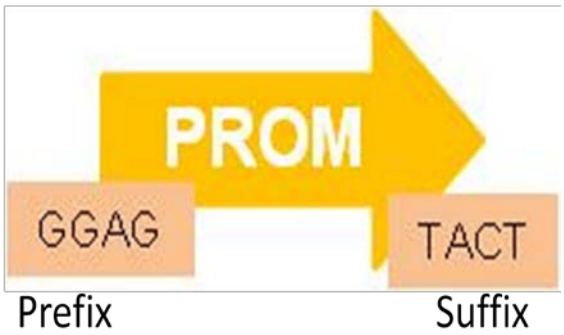


Fig 4.2.1. An example of a GB part, flanked by 4 nucleotides overhangs

4.3 GB-Composite parts (individual TUs and grouped TUs)

As we build new genetic designs, we also incorporate them to our database. Therefore, beside basic GBparts, the GB collection also contains individual TUs resulting from previous multipartite assemblies as well as groups of TUs resulting from binary assemblies. All of them conform to the GB standards and therefore are fully exchangeable. In general, all these complex structures are known as GB composite parts. The magic of GoldenBraid is that every new design (composite part), once tested and characterized, can be directly used as a template to add new genetic elements following the same “grammar”. Therefore the collection of composite parts is an important section of the GB database, and the expectations are that the importance of this collection will grow exponentially as more and more genetic devices are built.

The composite parts can be hosted in any of the eight destination vectors, depending on the route chosen for its assembly (assembly history). In order to combine two composite parts a set of very basic rules need to be followed. That is what we call “binary assembly rules” or “binary GB syntax” and will be explained in section 7.2.

4.4 Additional components: GB-domesticators and twister plasmids

There are a number of secondary components that work as tools to maximize the number of operations that can be performed with GB assembly. They are not strictly required for basic operations, yet can be very useful for more complex achievements.

GB-domesticators: Obviously, the GB collection cannot contain all useful DNA parts. Most probably you will need to generate also your own DNA parts to complete your specific designs. In order to generate a new standardized GBpart, you will need to PCR-amplify your DNA fragment from a suitable template and adapt it to the GB grammar. This is what we call to "domesticate" a DNA fragment, and basically consists in adding flanking restriction sites and eventually removing internal restriction sites. This is usually a very easy step, yet it can be made easier with some help. For this reason we have created the Universal GBpart Domesticator plasmid (pUPD), a unique plasmid that will assist you in the generation of new GBparts. Notice that the same single plasmid is used to domesticate any GBpart regardless of its category. The universality of the domesticator is a new feature introduced in GB2.0. You will find a detailed description on how to domesticate your own DNA parts in section 9.

The universal domesticator can make more for you than simply assisting you in domesticating your own parts. It can serve as an extra level for combinatorial arrangements. Following our usual analogy, the Universal GB-Part Domesticator can make new words by combination of smaller units, let´s say "morphemes". The universal domesticator is in some way the "orthography" of the GB grammar. If you want to learn more on how to use pUPD as an additional combinatorial level please jump to section 11.

The GB collection also contains destination vectors. You may want to create your own set of destination vectors. If this is so, you will find our Universal GB-Vector Domesticator (pUVD) a very useful tool. Just chop your plasmid of choice into pieces, clone each piece into pUVD and then reassemble it as a new pDGB. You will find a detailed description on how to create your own set of GB destination vectors in section 10 (currently unsupported).

Twister plasmids are a particular type of composite parts. They are rarely used, but may be very helpful if you want to combine two pre-made composite parts which were not originally designed for being assembled together. Twister plasmids consist on a small intergenic

region, named stuffer fragment (SF) inserted within a destination plasmid. Syntactically, they have the same structure than a composite part, but they are actually "jokers" that facilitate cloning compatibility. There are a total of 4 twister plasmids. There are 2 plasmids for α level (pDGB_α1_SF, pDGB_α2_SF,) and a similar set for Ω level. You can find how useful twister plasmids can be by visiting section 7.3.

5. RAPID INSTRUCTIONS

If you want to start using GB2.0 right now, we provide you in this section a simplified overview of the GB2.0 mechanism. GB comprises three basic actions, depicted in Fig 5.1; all three functions are software-assisted. The first one is the domestication of new GBparts (panel 1 in Fig 5.1), a prerequisite to introduce your own parts in the assembly. The second function is the multipartite assembly (panel 2) used to create new TU from basic GBparts. The third function (panel 3) is the binary assembly used to assemble TUs and/or other composite parts. The binary part of the GB scheme works as double loop of iterative assemblies (as depicted by the GB logo). As a result, new composite parts can be endlessly added to your genetic devices.

The GB software will guide you step by step through the assembly process: the GB software generates two types of outputs:

1. **A detailed lab protocol** to achieve the assembly of your choice. You only need to carefully follow the indications of the protocol in order to get your assembly.

2. **A sequence file** (gb file) containing the final assembled sequence. You can open your file with your favorite sequence viewer and keep detailed track of your assembly at the nucleotide level.

Just let the program guide you through the assemble process and enjoy!

The GB Assembler software can be accessed from http://gbcloning.org/.

In addition, if you want to know more about the guts of the GB system, you can continue reading the next sections.

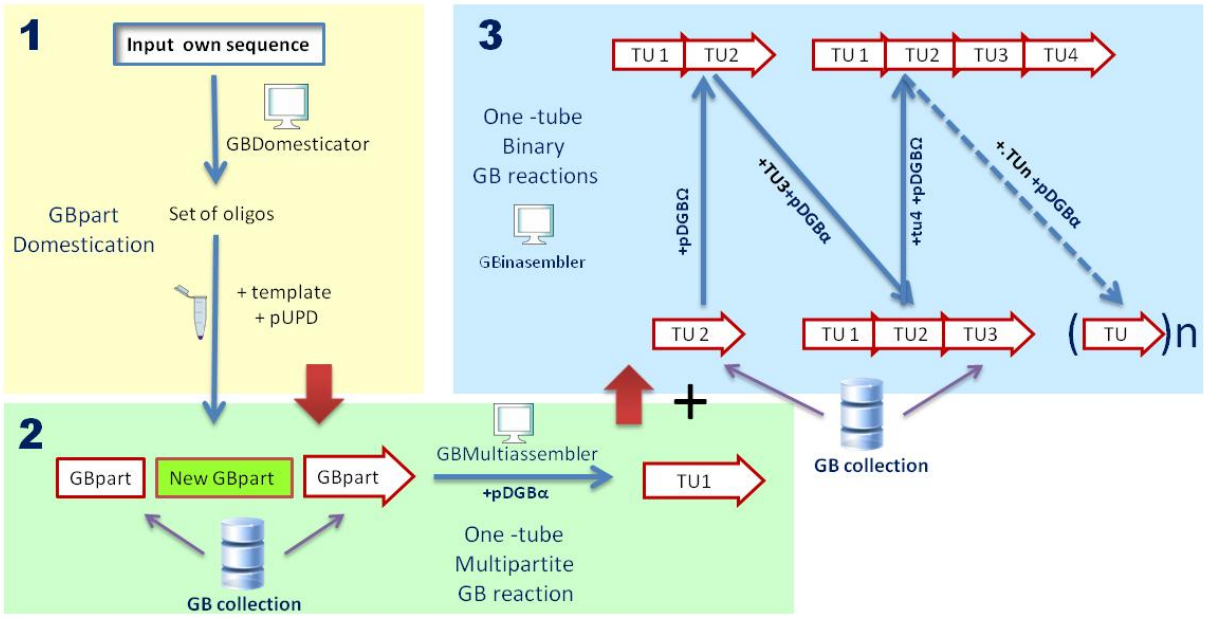


Figure 5.1. Basic scheme of the complete GB2.0 procedure. Computer screens indicate software-assisted steps. GB parts and TUs can be built by the user or collected from the database.

6. THE BASICS: GB SIMPLIFIED GRAMMAR

The following sections will provide you additional information about the mechanisms of GB assembly. To facilitate the understanding of the mechanisms, we will start using a simplified version of the GB grammar. In the most basic setup, GBparts are classified in three categories: promoters (which include the transcription origin and 5´UTR), coding regions (CDS), and terminators (which includes 3´-UTR and polyA site and transcriptional stop signal) (Figure 6.1). With only three standard parts it is easier to visit the guts of the system. However keep in mind that the standard parts PRO, CDS and TER are actually "supercategories", i.e. groups of smaller basic categories. For a detailed view of the full grammar please refer to section 8 (Figure 8.2)

Each category corresponds to a relative position in the transcriptional unit. Parts belonging to the same category share the same 4bp overhang. The exact 4nt overhangs that flank each of the three supercategories can be seen in Figure 6.1.

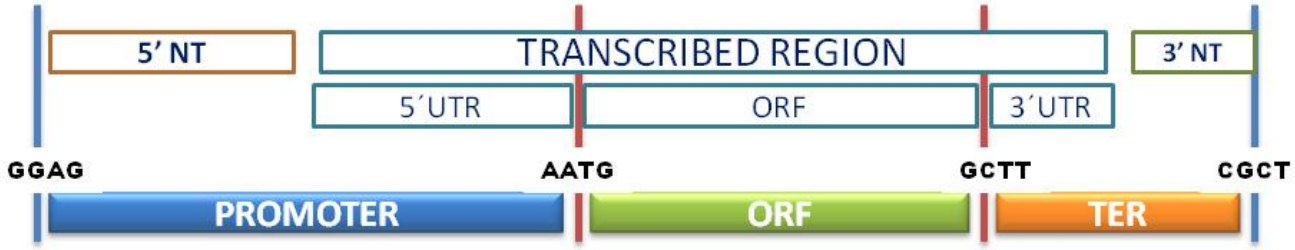


Figure 6.1. Simplified GB grammar

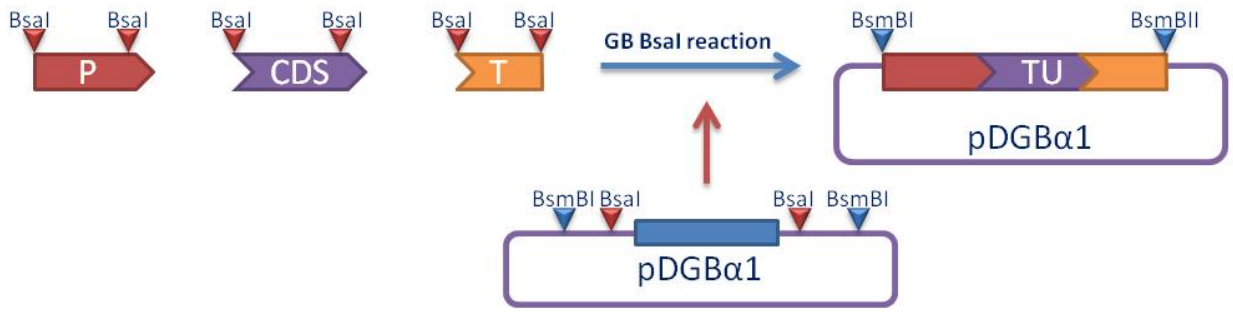
7. THE ASSEMBLY PROCESS

Once GBparts are ready you can start assembling: here we show you how. First you can make TUs by multipartite assembly of basic GBparts. Next, you make composite parts by binary assemblies of TUs and/or other composite parts.

7.1 Multipartite assemblies

The minimal instruction that you can introduce in an organism is a TU. In order to create a new TU you need to join GBparts using a Multipartite GB assembly reaction. Multipartite assemblies of basic GBparts can be performed in two ways: (i) by using the level α destination plasmids (α-PDGBs) in a so-called "BsaI GB-reaction" (Figure 7.1.1A), or (i) by using level Ω destination plasmids (Ω-pDGBs) in a so-called "BsmBI/BtgZI GB-reaction" (Figure 7.1.1B). Upon ligation, the recognition sites of the enzymes used in the assembly will disappear, therefore the composite part cannot be broken apart anymore. pDGBs are themselves binary plasmids, and therefore the resulting composite parts are ready for plant transformation.

A



B

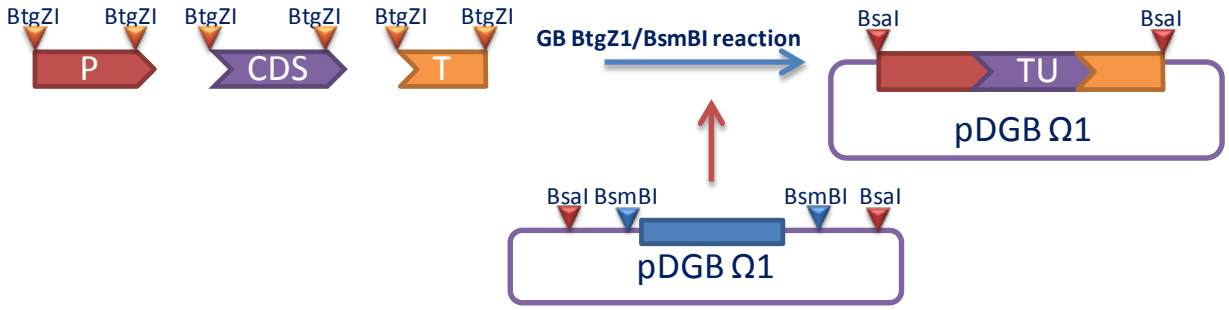


Figure 7.1.1. Multipartite GB reactions. (A) Multipartite BsaI reaction of three basic GBparts using a pDGBα plasmid as recipient. (B) Alternative Multipartite BtgZI/BsmBI reaction using a pDGBΩ destination plasmid. P, promoter; CSD, coding region; T, terminator.

An important feature of GB assembly is that TUs can be combined indefinitely to grow multigene structures. This is due to the special pDGBs design, which ensures that any newly assembled composite part (a TUs or a group of TUs) ends up being flanked by target sites of the alternative restriction enzyme. For instance, when you perform a BsaI GB-reaction, the resulting TU will be integrated in a α-level plasmid flanked by BsmBI sites; analogously, if a BsmBI GB-reaction is performed to assemble a composite part, the resulting composite part will be integrated in a Ω-level flanked by BsaI sites (as exemplified in Fig 7.2.1). To allow selection, α and Ω-level destination plasmids have different resistance markers. The iteration between α and Ω-level destination plasmids enables the indefinite growth of GB structures.

The election of destination plasmid depends on what do you plan to do next with your new assembled TU. If you plan to assemble it next with a TU from the database (TU2), which is

cloned in pDGBΩ1, pDGBΩ2 is your best choice, as it will allow you to assemble TU1 and TU2 in a single step (see section 7.2). If you plant to assemble your TU in reverse orientation with respect to TU2, then your best choice is pDGBΩ2R.

7.2 Binary assemblies

The assembly of composite parts (TUs or groups of TUs) involves always only two elements and therefore is known as **binary assembly**.

In order to be assembled together, the two composite parts need to be cloned in complementary pDGBs, (e.g. pDGBΩ1 and pDGBΩ2) so complementary sticky ends can be generated upon digestion. For this reason the GB system comprises a minimum of four plasmids: you need two plasmids per level to ensure binary assemblies (see Fig 7.2.1).

Often there are specific requirements on the relative orientation between the two units to be assembled. To enable assemblies in all possible orientations (direct, reverse, tail-to-tail or head-to-head), we included sets of reverse destination plasmids available in the collection (see section 4.1).

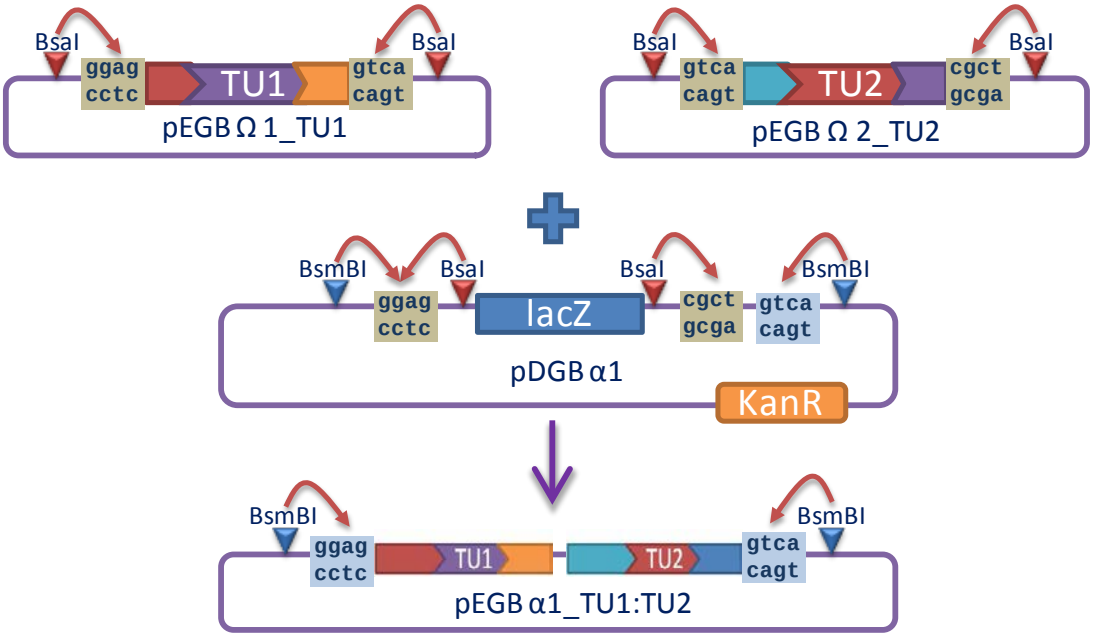


Figure 7.2.1 . Example of a grammatically correct binary assembly: a transcriptional unit cloned in a type 1 plasmid from the level Ω (TU1) can be combined with a complementary TU (TU2) that is cloned in a complementary type 2 plasmid from the same level Ω. To do so,

it has to be incubated with a destination plasmid (type 1 or type 2 indistinctively) of the level α. The result of the assembly will be a composite part inserted in a level α plasmid backbone.

7.3 The iterative double loop of GoldenBraid

As a result of the GB assembly design, the new composite part (pDGBα1_CP1) ends up being flanked by additional recognition sites (BsmBI in the example of Fig 7.2.1). Therefore it can be combined binarely with another GB-assembled multigene construct cloned in a α2 plasmid (pDGBα2_CP1) using a level Ω destination plasmids (pDGBΩ) in a BsmBI GB-reaction. As can be noticed, the cloning strategy turns back to level Ω, closing an endless iterative assembling loop (braid) (Fig 7.3.1).

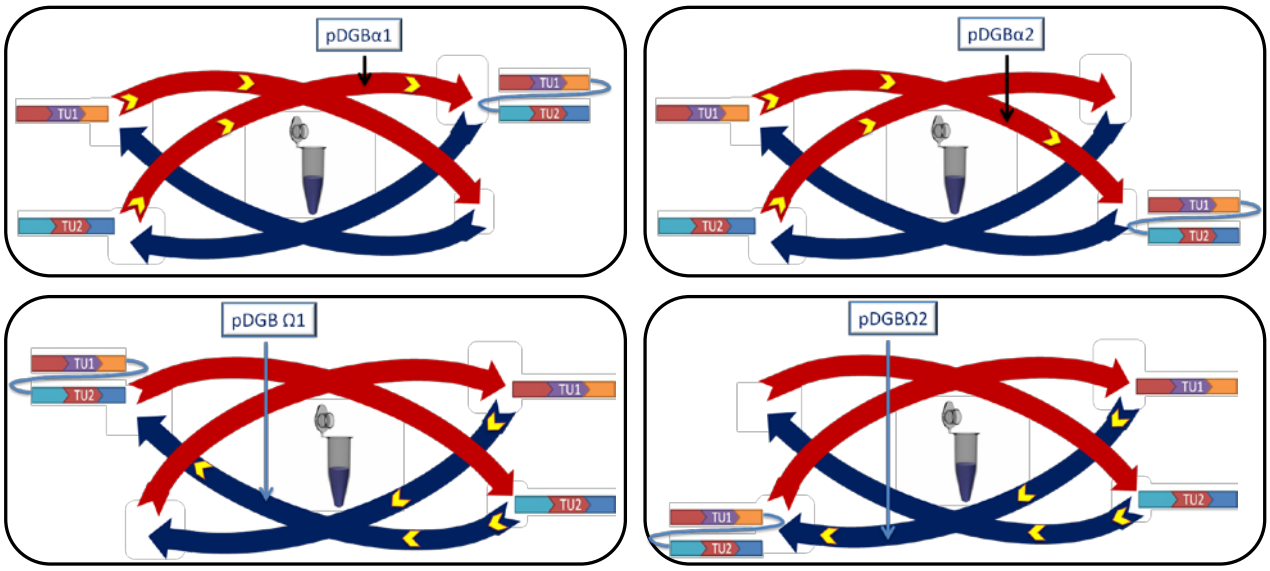


Figure 7.3.1. The double GB loop is dissected in its 4 possible assembly reactions. Each reaction represents a different path in the GB double loop (highlighted with yellow arrows), and it is mediated by a different destination vector. Only 4 destination vectors are sufficient to allow the system to grow indefinitely (four additional plasmids for reverse orientation).

What happens if you want to assemble two composite parts which are cloned in non-compatible destination vectors? Then you simply need to move your CPs along the iterative cloning loop until have them both located in compatible plasmids. A group of plasmids, named twister plasmids, will assist you in this task. Here is an example:

Imagine that you want to clone together two composite parts, CP1 and CP2. However, CP1 and CP2 are inserted in non-compatible plasmids: CP1 is inserted in pDGBα1 plasmid

(pDGBα1-CP1) and CP2 is inserted in pDGB_Ω2 plasmid (pDGB_Ω2-CP2). In this case the steps to follow are:

1. Perform a GB2.0 binary reaction with pDGBα1-CP1, pDGBα2-SF and pDGBΩ1 as destination plasmid. pDGBα2-SF is a twister plasmid containing an small intergenic region (Stuffer Fragment) instead of a composite part. The result will be pDGBΩ1-CP1-SF
2. Perform a second GB2.0 binary reaction with pDGBΩ1-CP1-SF, pDGB_Ω2-CP2 and a compatible destination plasmid of your choice (e.g. pDGBα1). The result will be pDGBα1-CP1-SF-CP2, that is CP1 and CP2 assembled together separated by a small intergenic region.

By wisely using twister plasmids with CPs, all CPs can be assembled together with a maximum of two GB-steps.

8. THE ADVANCED GB2.0 GRAMMAR

Yes, you are right! The basic GB grammar is a bit disappointing. If each TU is divided in more functional parts instead of only three, then we could take full advantage of the extraordinary modularity of GB-assembly!!

Once you got used to the basic 3-units grammar, we will propose you a more complex yet powerful stuff: the advanced GB2.0 grammar.

Let´s keep on with the analogy with natural languages. In the English grammar, contiguous words can be grouped in super-categories named phrases, like a subject, a predicate, a direct complement, etc. In order to construct a sentence (an instruction), the writer can combine individual words, but also has the possibility of using pre-arranged phrases in his/her vocabulary pool. An example of this flexible combination of words and phrases is shown in Fig 8.1.

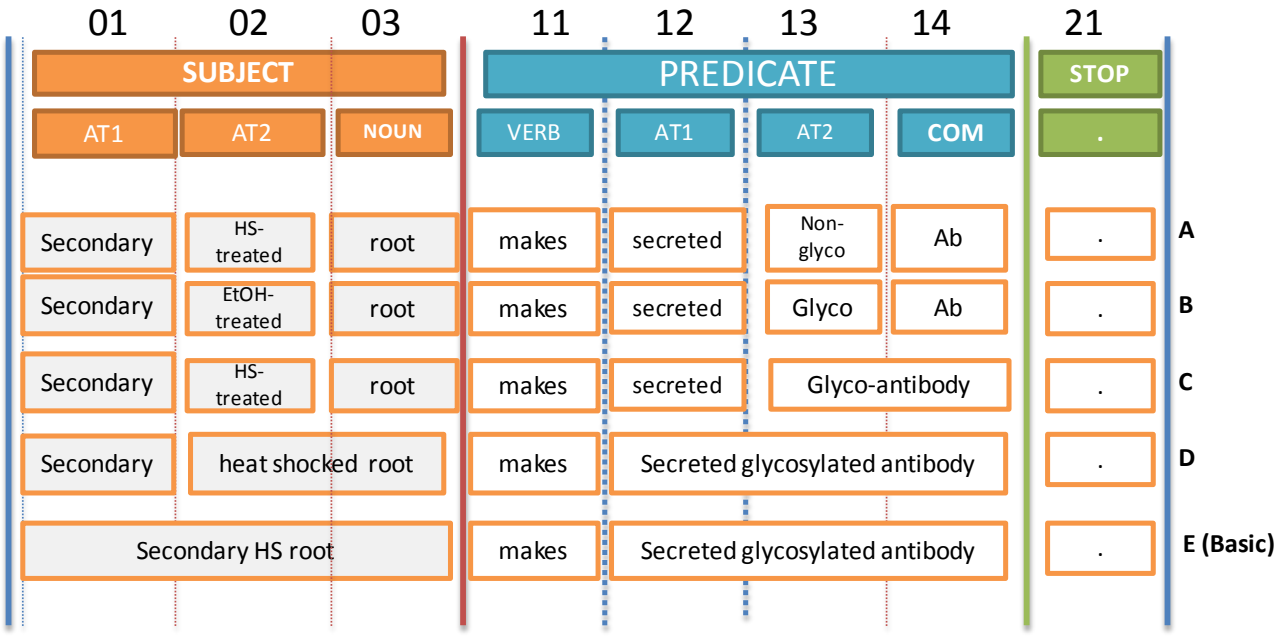


Figure 8.1 An example of multipartite assemblies using a simplified English grammar comprising eight word categories. The Instructions to produce an antibody in roots can be written by assembling all eight basic words (A, B). This gives maximum combinatorial power in the use of standard attributes like specific expression in secondary roots, heat shock (HS) or ethanol (EtOH) regulation. Alternatively, the use of pre-arranged phrases (C, D, E) facilitates the assembly process. AT: attributes; COM: complement.

Let´s go back now to GB grammar. We defined a total of 11 basic standard categories (words) for GB multipartite assemblies. The prefix and suffix defining each category are chosen arbitrarily. However, the GB vocabulary contains, in addition to individual words, also pre-arranged “phrases”. When creating a new sentence, you can also include phrases instead of individual words in the multipartite GB reaction, and also combine phrases and words. Indeed, the PROM part used in the basic grammar is a “phrase” made of the combination of five basic word categories (01, 02, 03, 11 and 12, see fig 8.2).

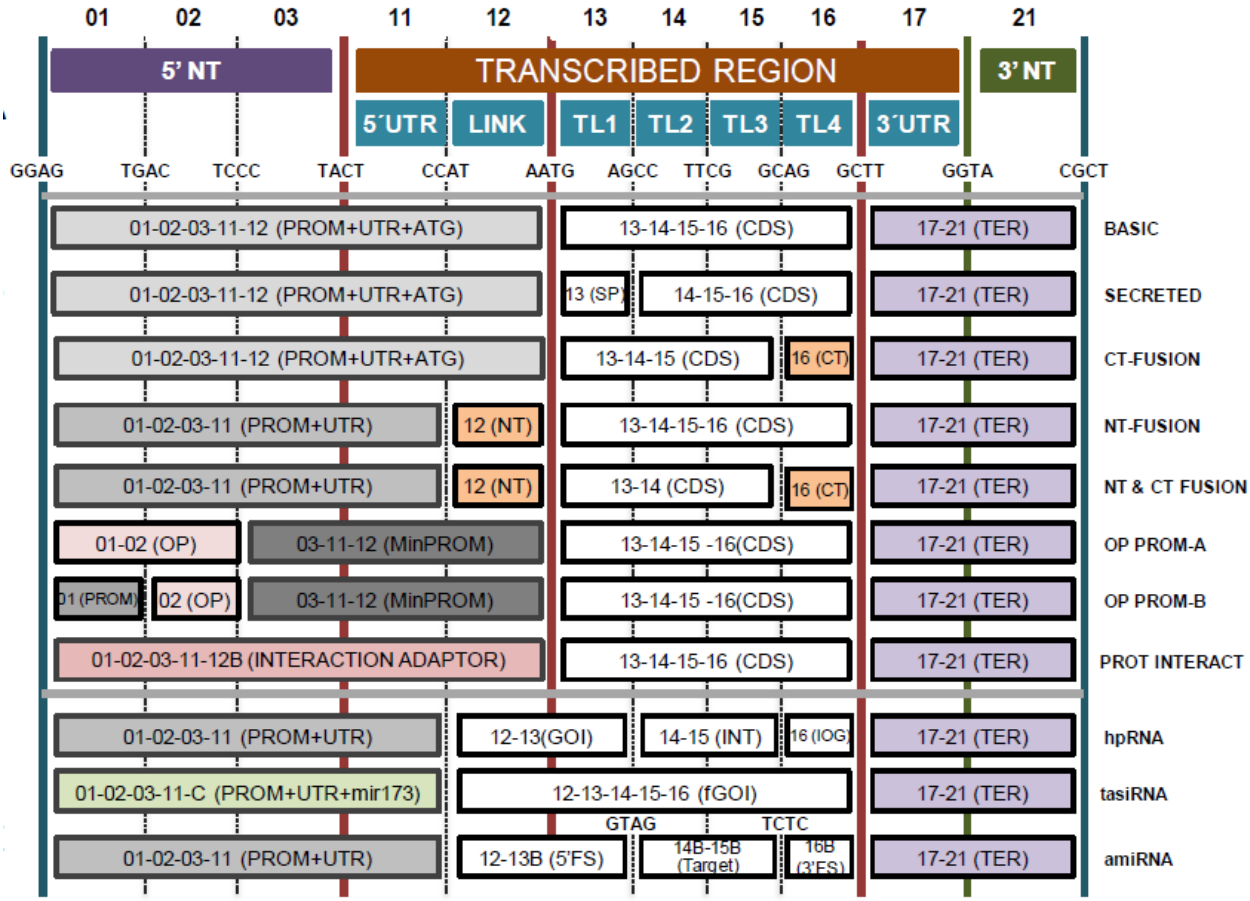


Figure 8.2. An schematic view of the advanced GB2.0 grammar.

All together, Fig 8.2 shows the Syntax of GoldenBraid 2.0. Although GB syntax is very systematic, there are a few special rules for creating special transcriptional units as amiRNAs, Protein Interaction adaptors etc. These cases will be dicussed in detail in a later section (unfinished).

Using GB advanced grammar for multipartite assemblies is as simple as using its basic grammar. You just need to be careful in putting together in a single tube all necessary GBparts (words or phrases) so they can assemble a grammatically correct TU.

We realized that, as our collection grows, finding the right set of GBparts to mount our genetic design is becoming more complicated. Therefore we decided to produce a software tool that assists you in the database search and in the design of new genetic devices (see http://gbcloning.org/).

9. DESIGNING NEW PARTS

This section helps you in the design of new GBparts. As mentioned above, parts are simply DNA fragments flanked by BsaI and BtgZI recognition sites. Normally, we keep them in the form of a circular plasmid, where the plasmid backbone is usually the pUPD vector and the insert contains the part itself. The design of pUPD plasmid is such that both BsaI and BtgZI digestions release exactly the same piece of DNA containing the same 4bp overhangs regardless of the enzyme used. As a consequence, GBparts can be assembled using either BsaI GB-reactions or BtgZI GB-reactions indistinctly.

As it was mentioned before, the 4bp flanking sites define each part´s relative position in the multipartite assembly. In order to ensure that all GB users can exchange their parts, it is necessary that we all use the same flanking sites for the same category. In other words, we need to use the same "grammar", so we can communicate.

We call "domestication" to the adaptation of basic DNA parts to the GB grammar. Domestication comprises the addition of flanking BsaI and BtgZI sites as indicated earlier, but also the REMOVAL of INTERNAL BsaI, BsmBI and BtgZI sites. We strongly recommend building BsaI/BsmBI/BtgZI-free parts to facilitate GB-reactions, especially for larger assemblies. If you don't want to eliminate the sites, you still can succeed in your assembly, because GB reactions are extremely efficient. However, you will get much less correct colonies, and for large assemblies involving many pieces your efficiency may be too low. Furthermore, if you are planning to reuse your TUs in the future to build more complex devices, the presence of undomesticated pieces will only cause you problems.

There are several possible procedures to eliminate internal BsaI and BsmBI sites. We establish a very simple and efficient one using the Universal Domesticator plasmid pUPD. In this simple protocol, the sequence of interest is amplified in the form of overlapping PCR segments (named DNA patches), each segment mutagenized in its internal sites and flanked by built-in BsmBI restriction sites (BsmBI is a type IIS restriction enzyme). DNA patches are sewn together in a BsmBI multipartite reaction using pUPD as destination plasmid. The whole process takes only a few hours and the result is a domesticated GBpart clone within pUPD backbone (Fig 9.1).

The special design of pUPD makes possible to use this single recipient plasmid to domesticate parts belonging to all the standard categories defined in GB2.0. The clue to get this feature is that the 4 nt overhangs are not incorporated in the plasmid. Instead they are included in the oligonucleotides used to amplify the DNA fragments that will integrate the new DNA part. For software–assisted GB part design please visit http://gbcloning.org/. There you will be assisted in the design of the oligonucleotide pairs needed for the amplification of each DNA patch.

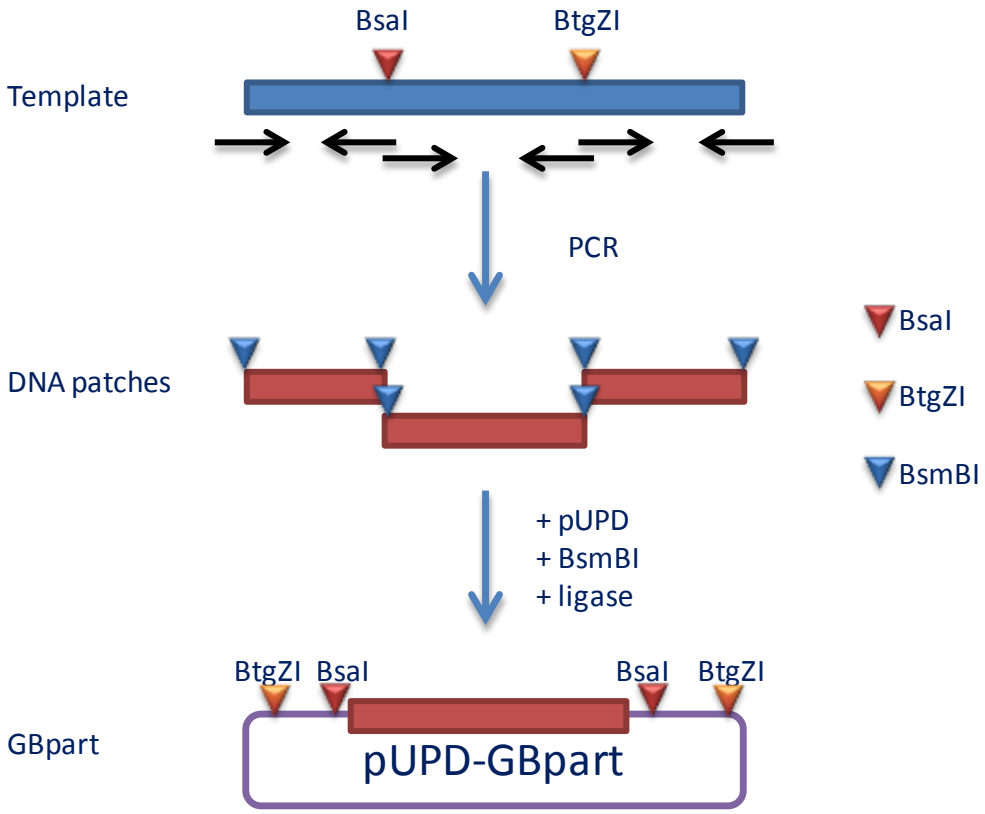


Fig 9.1 Process of GBpart domestication using the Universal Part domesticator vector (pUPD)

10. DESIGNING NEW VECTORS

At the present stage, the GB2.0 vector design tool is not available for public use.

11. THE UNIVERSAL DOMESTICATOR AS A NEW COMBINATORIAL LEVEL: THE GB2.0 ORTHOGRAPHY.

What if eleven standard words are not enough for me? What happens if I want to mount my own modular proteins and design my own modules? Can I still use GB2.0 as my assembly system? The answer is yes, you can. The 11 words syntax we made suits well to most of our general purposes. However, to facilitate most exigent combinatorial needs we created a new level of assembly that operates underneath the syntax rules described before. In some way, the basic units of this new assembly level are analogous to morphemes in a natural language. Morphemes are minimal meaningful units in a language. Words are made of morphemes. Because orthography is the part of grammar that rules the proper formation of words, we call this assembly level the “orthography” of GB.

As the matter of fact, if you read before section 9 of this manual, you are already familiar with this assembly level. Yes, part domestication is a particular case of GB orthography, where you chose to assemble a word or a phrase starting from smaller units (DNA patches, or morphemes) in order to remove internal restriction sites. But you can also use the universal domesticator to assemble your own protein modules PCR-amplified as DNA patches. There is no need for standardization at this level. You can choose overhangs at your will. In this way you can design seamless assemblies for your modules, or perhaps define your own standard orthographic rules. The free option of the software tool “GB Domesticator” will assist you in creating new words and phrases by “orthographic” seamless assembly of user-defined morphemes. You can find all the GB assembly tools at http://gbcloning.org/.