Accepted Article

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

Uncovering the mechanism leading to the synthesis of symmetric and asymmetric Tröger's bases

Priscila A. Lanza [a] [§], Diego Dusso [b] [§], Cristina. L. Ramirez [c], Alejandro R. Parise [a], Carlos A. Chesta[d],* E. Laura Moyano [b] ** and D. Mariano A. Vera [a].***

a QUIAMM-INBIOTEC. Department of Chemistry. Facultad de Ciencias Exactas y Naturales. Universidad Nacional de Mar del Plata. Mar del Plata. Argentina.

b INFIQC. Department of Organic Chemistry. Facultad de Ciencias Químicas. Universidad Nacional de Córdoba, Córdoba, Argentina.

c Department of Chemistry. Facultad de Ciencias Exactas y Naturales. Universidad Nacional de Mar del Plata. Mar del Plata. Argentina.

d IITEMA. Department of Chemistry. Facultad de Ciencias Físicoquimicas y Naturales. Universidad Nacional de Rio Cuarto, Las Higueras, Argentina.

§ *Both authors contributed equally.*

*** *corresponding author:* dmavera@yahoo.com

** *corresponding author:* lauramoy@fcq.unc.edu.ar

* *corresponding author:* cchesta@exa.unrc.edu.aring

Abstract

Tröger's bases have been attracting great interest due to their potential applications in nanoelectronics, supramolecular chemistry, molecular recognition, biological activity and auxiliaries for asymmetric synthesis. However, a detailed step by step proposal for the mechanism leading to the production of these compounds is still uncovered. A set of five model syntheses of symmetric and asymmetric Tröger's base derivatives starting from substituted anilines and formaldehyde was done and envisaged as representative for understanding the underlying mechanism. All reasonable pathways were thoroughly scanned by means of DFT calculations. The

1

highest energy TS was associated with the entrance of the first formaldehyde which produces the first out of three key carbocations. The last step, the closure of the methylene-bridged diazocyne heterocycle was also found of kinetic relevance and as a source of stable intermediates or byproducts. The whole mechanistic picture would provide keys for the rational planning of the synthesis of these compounds.

Introduction

Tröger's base was first synthesized by Julius Tröger in 1887 from *p*-toluidine and formaldehyde in acid medium.[1] During the next 48 years its structure remained unrevealed, it begin finally elucidated in 1935.[2] This species has a "V"-shaped form (Scheme 1), with a methylene group bridging the two nitrogens which form the diazocyne:
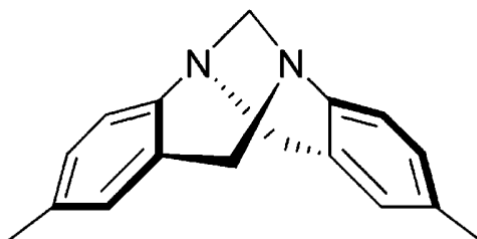
Scheme 1.

Such a folding allows it for being chiral in the absence of asymmetric carbons. Its enantiomers were solved for the first time in 1945 by Vladimir Prelog, who proved that chirality would be promoted by nitrogen atoms as well.[3]

Owning to its symmetry, chirality, rigidity and particular shape, Tröger's base derivatives were subject of remarkable interest in the last decades. These compounds were studied aiming to their applications in molecular recognition,[4] interactions with DNA,[5] biological activity,[6] clathrate formation,[7] chiral auxiliaries for stereoelective synthesis [8] and applications in optics and optoelectronics,[9] among others. More recently, the attention was focused on their photochemical and electrochemical properties. Their potential for the OLEDs fabrication was explored [10] as well as their emissive properties in solution and in solid state.[11] Interestingly, our group has raised attention on the very unusual properties for the charge transport of the Tröger's base as bridge between two electroactive subtituents;[12] in other words in a donor-bridge-acceptor dyad setup. Indeed, we have demonstrated that, sometimes, this aliphatic bridge can act as an efficient

2

connector or "cable" between the redox centers (surprisingly as a conjugated π system would do) and in others as an insulator, that is to say as a bridge with a weak coupling, typical of aliphatic bridges. In particular, since it was possible to efficiently transport holes but not electrons, this feature would allow us to envisage its potential as molecular OR logic gate as well as other applications currently under investigation.[13]
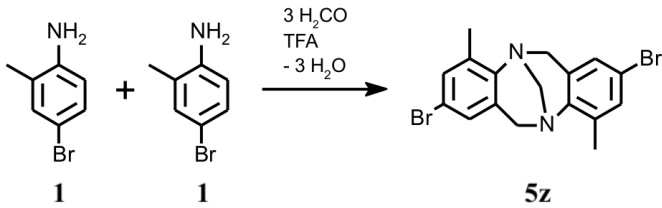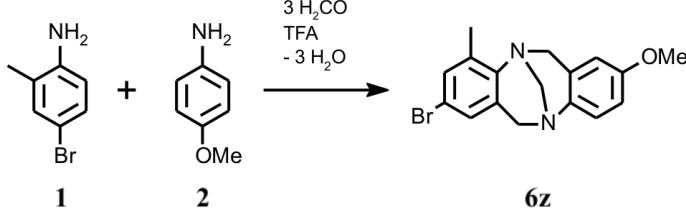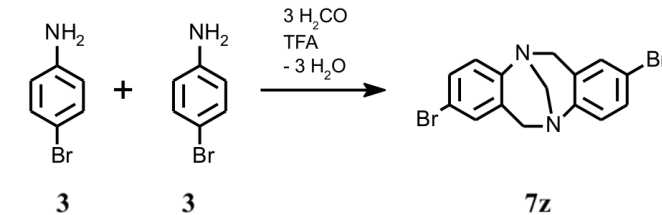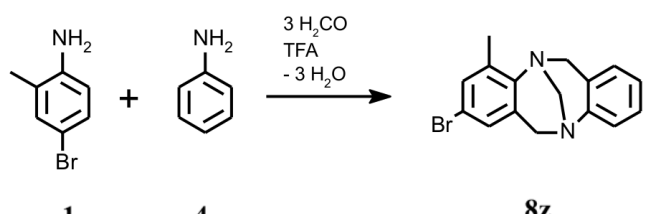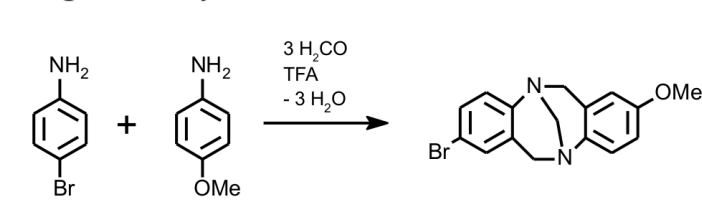
However, the synthesis of arbitrarily complex symmetric and, still more problematic, asymmetric Tröger's base – bridged dyads seems to be a challenging issue. There are descriptions of numerous available synthetic routes for different symmetric species.[14,15] Some intermediates have been isolated and some mechanistic proposal have appeared to explain part of the whole process. [15,24,25,27-31] However, a detailed step by step description of each elementary process leading to the synthesis of these compounds starting from substituted anilines is still uncovered, at least to the best of our knowledge.

We here present the results of five characteristic reactions of synthesis of symmetric and asymmetric Tröger's base derivatives together with a thorough description of each plausible mechanistic pathway leading to their production starting from anilines and formaldehyde. In this aim, each intermediate, transition state and reasonable route connecting them were characterized by means of Density Functional Theory (DFT) quantum calculations and discussed in the light of our experimental results and the evidence of the available intermediates isolated in the last few years.

Results and Discussion

Experimental model syntheses.

Tröger's bases (TB) were prepared from anilines in trifluoroacetic acid (TFA) and characterized as described before.[15] The yields for the five model syntheses are summarized on Scheme 2. The main goal of these syntheses is to understand the underlying mechanism. However, their synthetic potential is interesting *per se* since the bromide atom could be replaced for a great number of functionalities to prepare more complex molecules.[15] In the case of the synthesis of **5z**, **6z** and **8z** starting anilines react quantitatively after 24 h. In contrast, significant yields of TB **7z** and **9z** are only observed after a week.

3

Reactions	Yields (%)	reaction times (/hours)

Scheme 2. Summary of the reactions leading to the Tröger's base derivatives **5z-9z**. Isolated yields. TFA = trifluoroacetic acid. Yields of symmetrical Tröger's base arising from homocoupling: *a*) **1-1:** 23% and **2-2:** 17%; *b*) **1-1:** 20% and **4-4:** 16%; *c*) **2-2:** 15% and **3-3:** 17%.

Expectedly, during synthesis of asymmetrical TB **6z**, **8z** and **9z** the formation of the symmetrical TBs is also detected. For instance, reaction of anilines **1** and **2** provides **6z** (**1-2**) but also **1-1** and **2-2** according to:[16]

$$1 + 2 \rightarrow x\,(1\text{-}1) + y\,(1\text{-}2) + x\,(2\text{-}2) \qquad (1)$$

4
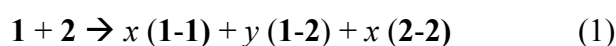
where x and y represent the stoichiometric coefficients of the reaction resulting from combining one mole of each aniline. Relative yields of asymmetrical (a-TB) and symmetrical TBs (s-TB) depend on statistical and kinetical factors. If the specific rates constants for the formation of all the different TBs are identical, statistical factors prevail and the stoichiometric coefficients in eq. (1) become: $x = 0.5$ and $y = 1$. In other words, the theoretical yield of a-TB is 50%. In contrast, if the formation of one (any) of the s-TB is favoured, the yields of both s-TB grows at the expense of that of the a-TB, being able to reach the limit value of $x = 1$ and $y = 0$; *i.e.* no a-TB is obtained. Similarly, when the formation of a-TB is favourable, its relative yield increases, reaching 100% in an extreme favourable case. Theoretical and practical implications of this type of competitive reactions have been extensively discussed in relation to the combination of free radicals in homogeneous and heterogeneous media, see for example Chesta *et al.*.[16] Hence, a-TB reaction yields reported in Scheme 2 (products **6z**, **8z** and **9z**) must be analyzed within the framework of the concepts discussed above. Symmetrical TB **5z** was obtained in ∼96 % yield, which agrees well with previous reports.[15] Compound **6z** was obtained by ∼50% together with the corresponding s-TB, both with ∼25% yields. In this context, the yield of **6z**, close to 50%, results as good as the 96% for **5z** in the sense that both reactions are close to the theoretically expected yield (50 % in a case and 100 % in the other). However, in the case of the asymmetrical synthesis of **8z** and **9z**, a-TB are produced with similar yields: ∼30 % each. Finally, **7z** is obtained in a meagre yield of ∼60%, together with byproducts that could not been identified.

As it can be seen on Scheme 2, there are just a slight difference in the theoretical yields of the symmetric and asymmetric syntheses $1 + 1 \rightarrow$ **5z** and $1 + 2 \rightarrow$ **6z**, the yield of **6z** just being lowered by the statistic probability of $1 + 1$ and $2 + 2$ lateral reactions. Note that the replacement of an electron withdrawing bromide in **1** by a very strong electron donor -OMe group in **2** has practically no effect in the reactivity. This is the first unexpected feature found. Even more surprising is the fact that the simple replacement of and hydrogen by a methyl group in the position 2 of the aniline drastically reduces the reactivity of the aniline. This is clearly noted when it is compared the $3 + 3 \rightarrow$ **7z** against the $1 + 1 \rightarrow$ **5z** reactions (a clean reaction in 24 hours with 96% isolated yield against one with a poor yield, even after 7 days, and the appearance of by products, in the same experimental conditions); a similar decrease in the reactivity upon 2-H- by 2-Me- substitution was observed in the asymmetric case ($1+2 \rightarrow$ **6z** *versus* $3+2 \rightarrow$ **9z**, Scheme 2).

5

Energy surfaces exploration.

The free energy surfaces for the whole mechanistic puzzle were explored using the CAM-B3LYP/6-311+G(d,p) level of theory in a model SCRF solvent (details on Experimental section) . This functional was extensively challenged by us and other authors to correctly describe complex heterocycles chemistries and yielded accurate mechanistic descriptions. For example, it has already been proved that this level of theory could give practically the same energy profiles than the much more computationally expensive MP4(DQ)/6-311+G(d,p) for the syntheses of nitrogenated heterocycles as described recently.[17] More specifically, for a gas phase process (no errors introduced by the solvent model, just the electronic structure description accuracy can be judged) the level of theory used here was very close to the CCSD(T)/6-311+G(d,p)[18,19] (maximum deviation in transition states relative energies below ~3 kcal/mol).[20] Such level of Coupled Cluster theory can be considered as the most accurate pure (not composite or extrapolative) formalism of practical applicability for organic systems of medium size.[21] In view of these antecedents and the goal of this study, no further energy refinement was applied to the CAM-B3LYP/6-311+G(d,p) results.

Besides there are some fragmented mechanistic proposals and some differences between them (*vide infra*), some general agreements can be highlighted: the process is started by the initial electrophilic attack of a formaldehyde to an aniline, which, once activated upon a water molecule release couples with another aniline and then a second a third formaldehydes enter as methylene units for the cyclization to the 1,5-diazocine ring and the generation of the methylene bridge. For sake of simplicity, we shall first describe the simplest case of the symmetric **1 + 1 → 5z** in the deepest detail for further comparison with the other cases. For this process, all the reasonable pathways were calculated and summarized in Scheme 3; all energies are informed as standard free energies relative to the ones of the starting reactants as zero.

Symmetric synthesis and general mechanistic scheme.

Even though the attack of the first formaldehyde molecule has been proposed to occur to the nitrogen of the aniline, this was not found to be the suitable (most nucleophilic) position. Indeed, as it is shown in Figure 1, the Fukui's function , f^+, has the maximum at the position 6 of the aniline aromatic ring (the carbon *ortho* to the amine group), thus this would be the preferred position for the electrophilic attack. Nevertheless, both possibilities were explored and, in fact, the attack to the ring leaded to a transition state (TS) which resulted 11 kcal/mol more stable than the attack to the nitrogen; these are labeled **5TS-b-cc** and **5TS-b-c**, respectively on Scheme 3. This fact suggests that
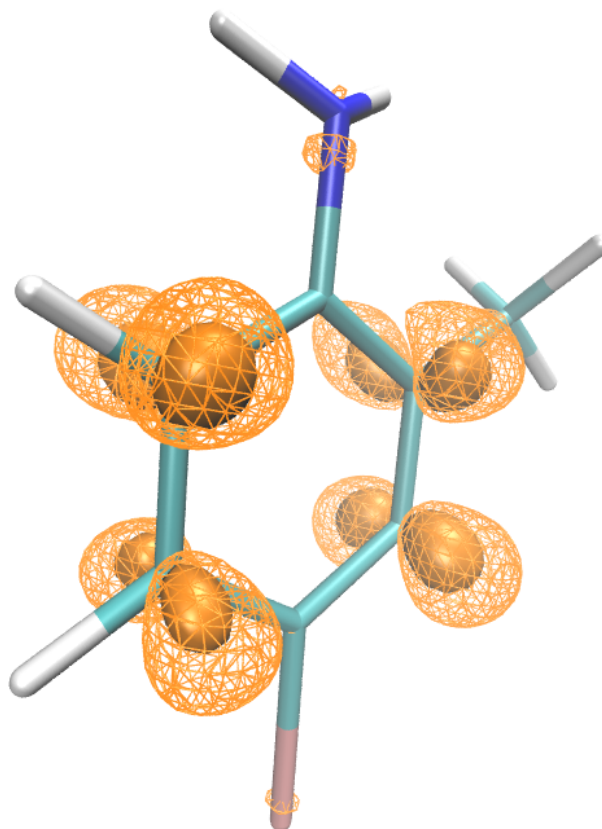
6

the preferential option from the kinetic point of view would be to follow Path I (blue in Scheme 3, **5bb → 5cc → 5dd → 5ee**) instead of Path II (in green , **5c → 5d → 5e**).

7
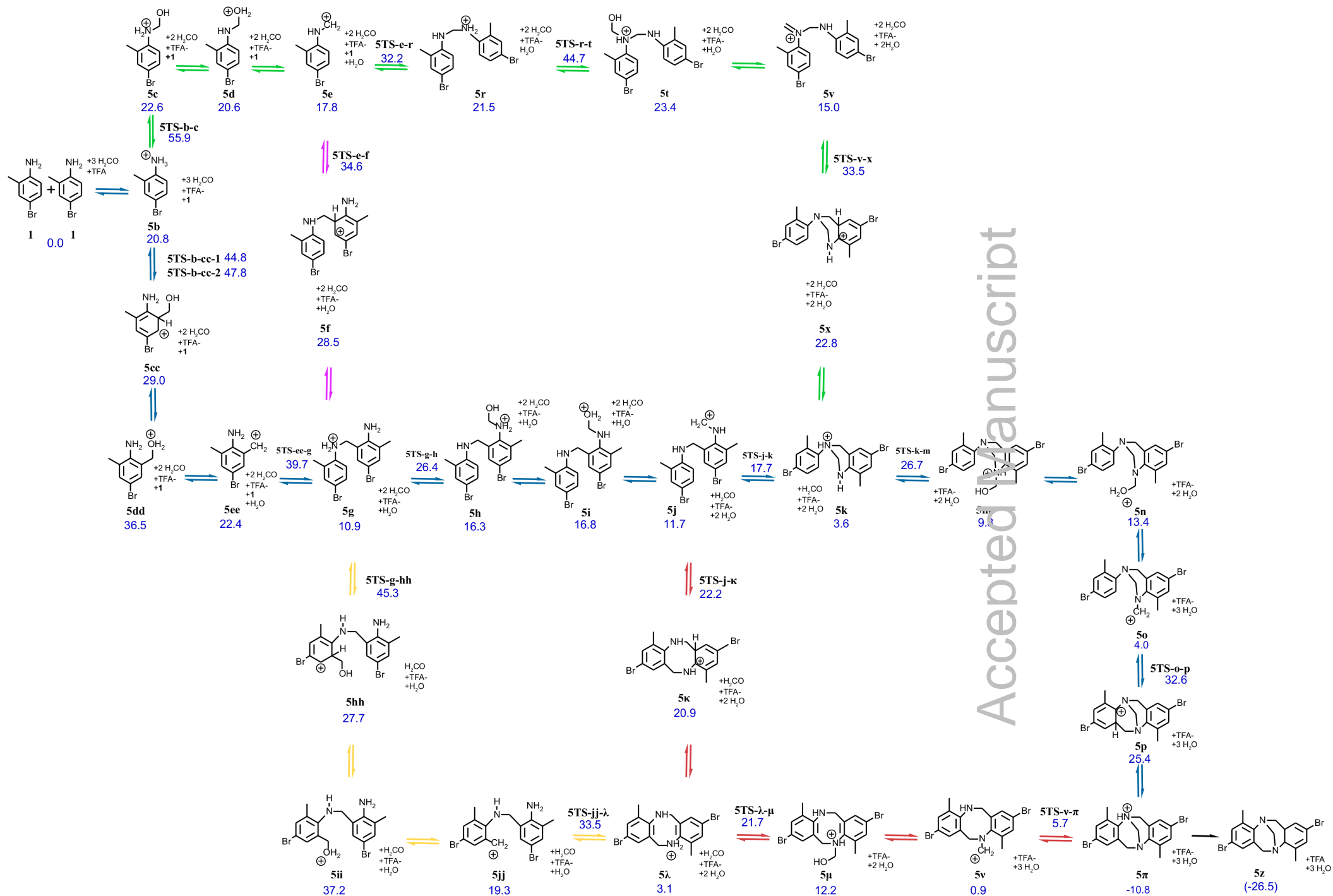
Scheme 3. Calculated mechanistic pathways for the reaction **1** + **1** → **5z**. Relative free energies (ΔG^0) in kcal/mol. The colors of the arrows corresponds to: Path **I (blue)**, Path **II (green)**, Path **III (purple)**, Path **IV (yellow)** and Path **V (red)**.

Figure 1. Representation of the Fukui's function f^+ which suggests the position 6 of the aniline **1** as the preferred position for the electrophilic attack. Solid surface corresponds to an isodensity of 0.05 e and the wireframe surface an isodensity of 0.01 e.

Also, to obtain the next **5cc** intermediate, two possible transition states were analyzed. In all cases, the protonation of formaldehyde occurs in a concerted mechanism involving the transfer of a proton from the -NH$_3^+$ group of a protonated aniline. The first possible transition state for the reagents, **5TS-b-cc**-1, consists of the protonation of formaldehyde by an ammonium, which favors the electrophilic character of formaldehyde and its approach to the carbon in *ortho* position of the ring of another non-protonated amine (sparse in acid medium). Another alternative is **5TS-b-cc**-2, which describes the mechanism of proton transfer arranged by the same protonated amine that suffers the electrophilic attack. In this scheme, as the proton of the ammonium group moves away from the N of the protonated aniline and approaches the O of the carbonyl, the carbonylic C begins

9

to gain carbocation character, attacking the ring and breaking its aromaticity. In either case, and also for the other reactions (obtaining **6z-9z**), the formation of this dearomatized intermediate was the highest free energy transition state of the entire process. Its structure is illustrated for **6TS-b-cc-**2 as example in Figure 2.
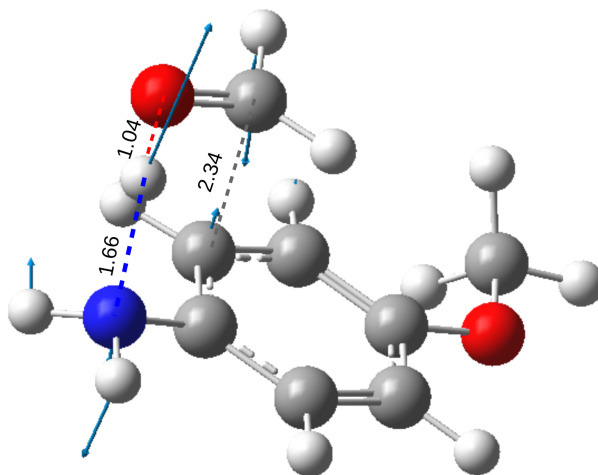

Figure 2. Geometry of the first transition state for the entrance of the first formaldehyde, shown for **6TS-b-cc-**2 as representative. The blue vectors correspond to the normal vibration mode with the imaginary frequency of 308 *i* cm^{-1}, distances in Ångstroms.

This TS leads to the intermediate **5cc**, which, continuing along Path I (blue arrows on Scheme 3), tautomerizes to **5dd**, which release a water molecule to yield **5ee**. Hereafter, all tautomeric isomerizations will be considered as fast equilibria assisted by the protic solvent and will not deserve further discussion nor TS description. Also for the sake of simplicity, the rapid release of water molecules will not be described (however, all intermediates and their energies are informed in the Supporting Information, SI, Table S1). This procedure has recently been fully justified in the study of a similar heterocycle synthesis, even though in that case the solvent was not as good proton donor as in this study.[17]

Continuing down the Path I, the carbocation **5ee** attacks an aniline passing through the **5TS-ee-g** (Scheme 3, blue path) to give **5g**, which is the first intermediate of some stability, of whose existence there is experimental evidence (see below). **5g** could have been formed through either the trajectories in the green and then purple arrows in Scheme 3 (Path II and Path III, respectively), but as it was shown, such processes have a much lower probabilities considering that they come from an extremely high TS (**5TS-b-c**, 55.9 kcal/mol). The free energy profiles of the different paths for **1 + 1 → 5z** are superimposed on Figure 3, with the same colors as in Scheme 3. These considerations

10

become clearer by comparing the profiles blue against green and purple from the reactants until reaching **5g**.
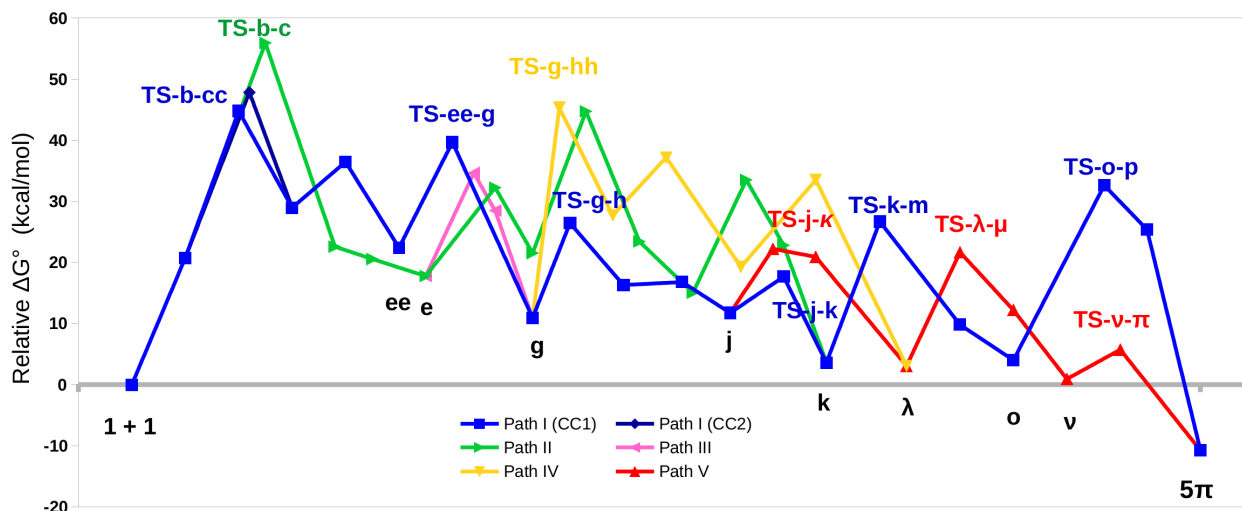
Figure 3. Free energy profiles for the (**1+1→ 5z**). All possible paths are superimposed on a symbolic reaction advance with the same colors as in Scheme 3.

The intermediate **5g** can promote the entrance of the second methylene, that is the second attack of a formaldehyde followed by water release. One possibility is the attack on one of the rings, a possibility that would be ruled out, given that the free energy of this **5TS-g-hh** (Scheme 3) was almost 20 kcal/mol higher than the other alternative for **5g**, which is the attack to an aniline nitrogen passing through **5TS-g-h** to form **5h**, which then tautomerizes and releases water to give **5j**. The notable difference between the **5TS-g-hh** and **5TS-g-h** allows discarding the Path IV (yellow in Scheme 3, note the height of the TSs involved in the green and yellow profiles of Figure 3) and continue on Path I (blue). The fact that the formaldehyde has a preference for the N of amine in this case as well as the remarkably greater ease with which the second formaldehyde enters in comparison with the first one, can be rationalized by virtue of the strong electrostatic stabilization and by the hydrogen bond pattern formed in **5TS-g-h.** As it is shown on Figure 4, one of the nitrogens acts as donor of the proton that activates the H_2CO, then the protonated aldehyde is already perfectly oriented to attack the other nitrogen.

11
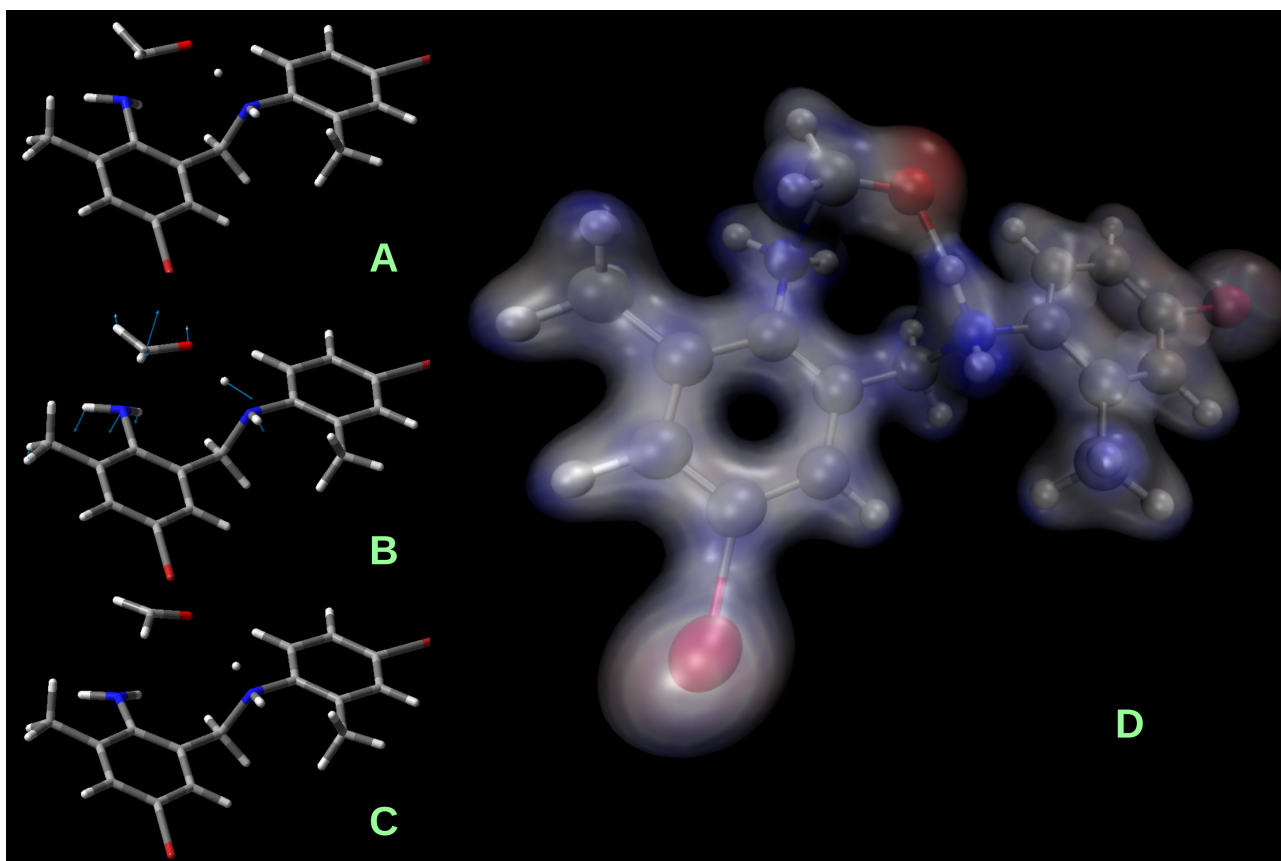
Figure 4. Transition state **5TS-g-h** for the attachment of the second formaldehyde through the Path I. A) Structure on the intrinsic reaction coordinate (IRC) 0.5 a. u. before the TS. B) Structure of the TS, the blue vectors correspond to the normal vibrational mode with the imaginary frequency of 268 i cm^{-1}. C) structure 0.8 a. u. after the TS. D) Isosurface of electron density of 0.05 e, colored according to the molecular electrostatic potential, from red (- 0.01) to blue (+ 0.05 a.u.).

The intermediate **5j** (Scheme 3) has an activated electrophilic center which can now attack a nitrogen for condensing either a six-membered ring, the intermediate **5k** on Path I, or a phenyl carbon for closing an eight-membered ring of **5κ** and continuing through Path V. The preference for one of these branches would be determined by the difference in energy between the corresponding TSs, which is less than 5 kcal/mol between **5TS-j-k** (blue) and **5TS-j-κ** (red on Scheme 3 and Figure 3). The difference which favors continuing Path I is much smaller than the differences previously discussed and the two pathways are connected by reversible steps and fast tautomerizations; on the other hand, the difference is even smaller for other model syntheses, thus the possibility of continuing through Path V will be further considered.

Continuing along Path I, the intermediate **5k** promotes the third and last methylene addition after crossing the **5TS-k-m** (Scheme 3 and Figure 3). Again, the formaldehyde molecule enters in a very similar way to that shown in Figure 4, that is with the formaldehyde H-bonded, polarized and

12

oriented between the two nitrogens. As a general trend, also observed for the synthesis of **6z-9z**, the entrances of the second and third formaldehyde molecules are much easier than the first one, with relative energies ranging from 10 to 25 kcal/mol against 45 kcal/mol and more, respectively.

Once the last formaldehyde was added, tautomerization and water loss steps follow, generating a relatively stable intermediate, **5o**. The final step to complete the aliphatic bicycle of the Tröger bridge is the intramolecular attack of the **5o** cation on the ring thus forming the last C-C bond (Figure 5). This step involves the highly energetic transition state **5TS-o-p** (relative energy 32.6 kcal/mol, activation of 28.6 kcal/mol, blue profile in Figure 3), which forms a Meisenheimer-like intermediate **5p**. This intermediate rapidly tautomerizes to give **5π,** thus recovering its aromaticity.
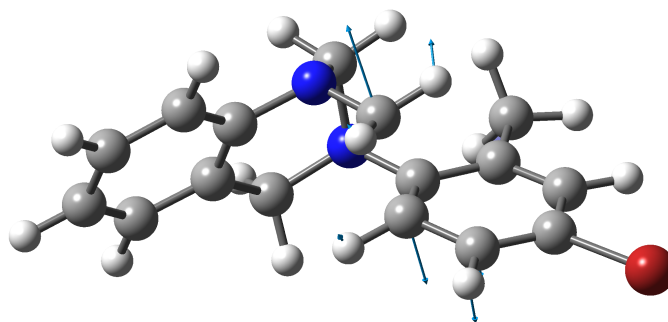
Figure 5. Final transition state for the closing of the aliphatic Tröger bridge; for clarity it is shown for the case of the smallest molecule, **9TS-o-p**.The blue vectors correspond to the normal vibration mode with the imaginary frequency of 175 i cm^{-1}.

The cation **5π** (-10.8 kcal/mol) is finally the conjugate acid of the desired product **5z** (ΔG^0_{rel} = -26.5 kcal/mol), which is obtained by neutralizing the reaction (**5z** is precipitated with NH_4OH). The transition state **5TS-o-p** may have a certain participation in the control of kinetics in this way. Although the highest TS is **5TS-b-cc** (for the first entry of formaldehyde), after the formation of a stable intermediate like **5o** (which is even more stable in other syntheses, such as those of **6z**, **7z** and **9z**, even more than the starting material, $\Delta G^0_{rel} < 0$), a high transition state immediately follows. This scenario would be rationalized in terms of two speed determining regions [one involving (**1+1** → → ... **5o**) and another of (**5o** → **5TS-o-p** → → ... product)] instead of a single determinant step. This analysis of regions or determining states [22] instead of a single determining step would be more general and has been discussed in detail in the references.[17, 20, 22] Besides the case of **1+1**→ **5z**, and taking into account that our group is interested in obtaining a wide variety of asymmetric Tröger's

13

base compounds with different substituents, the possibility of going down the Path V should still be considered in this framework. Even though Path I was probably more likely than V at the level of the crossroads corresponding to **TS-j-k** versus **TS-j-κ** for the case of the synthesis of **5z**, Path V and I are similar in some aspects. Once one of the aliphatic branches of the Tröger bicycle was condensed at the point of the fork to Path V (**5j** → **TS-j-κ** → **κ** → **λ**, Scheme 3), the last formaldehyde enters, it tautomerizes and eliminates water (**5λ** → **TS-λ-μ** → **μ** → **ν**). Intermediate **5ν** should finally close the methylene bridge of the bicycle (**5ν** → **TS-ν-π** → **π**). At difference to the Path I, this last step requires less activation (4.8 kcal/mol) than in the case of Path I (**5TS-o-p**), since it does not involve the de-aromatization of an aromatic ring.

Comparison of symmetric and asymmetric pathways

The asymmetric syntheses were explored at the same level of detail, however just **1 + 2** → **6z** will be discussed as representative and the production of **8z** and **9z** will be compared when relevant and all the details are deferred as Supporting Information (SI). Besides the trivial reduction of yields due to the parallel reactions **1 + 1** and **2 + 2**, the main (*i. e.* the lower energies pathways) routes are described on Scheme 4 (the complete scheme for most reasonable paths calculated are summarized in SI Schemes S1 and S2).

Path I (blue in Scheme 4) is similar to that obtained in the symmetric case, it was verified that the lowest free energy involves starting with the attack to the most nucleophilic aniline, **2**, with a first transition state (of greater kinetic relevance) similar to case of **1 + 1**. Then a series of steps of coupling of anilines and entry of the second formaldehyde is followed until reaching **6j**-2 (another possible **6j** is with the formal positive charge on the skeleton that came from **1**, details in the SI). Then comes the same dilemma as in the symmetrical case, with 6.5 kcal/mol in favor of the TS that closes the aliphatic cycle of six-membered instead of the eight-membered ring (**6TS-j-k**-2 vs. **6TS-j-κ**-2. Scheme 4). Continuing on Path I (blue), the second and third formaldehydes enter then reaching the stable intermediate **6o** (ΔG^0_{rel} = -3.4 kcal/mol).

The main difference with respect to the case of **5o**, is that the relative energy of the **6TS-o-p**-2 for the final closure of the Tröger bridge is now lower than in the case of the symmetric Tröger (25.2 *vs*. 32.6 kcal/mol, although the activation barriers are similar).
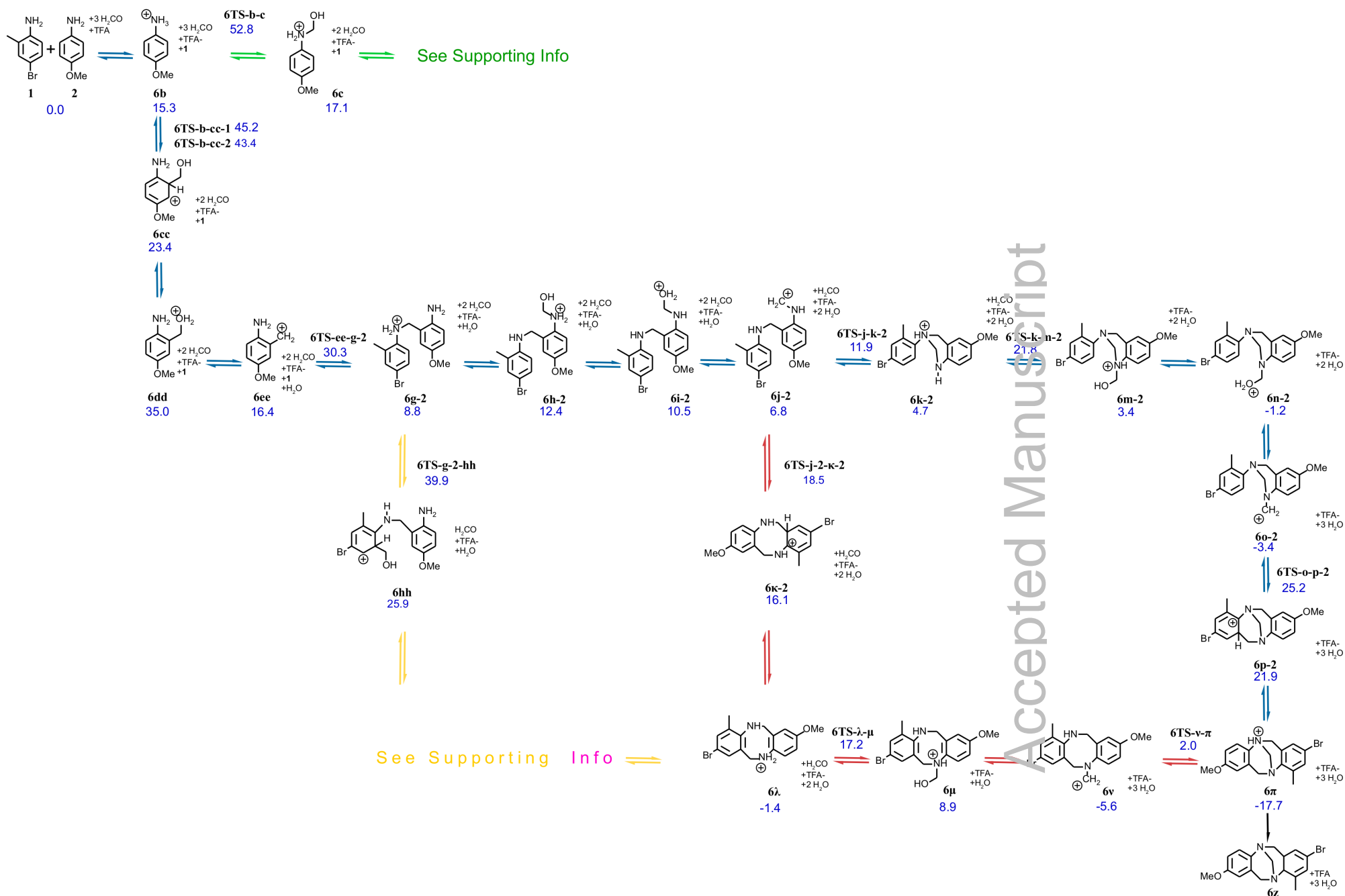
Scheme 4. Simplified summary of pathways for the process **1 + 2 → 6z**. Colors are as in Scheme 3 (Paths **I (blue)**, **II (green)**, **III (purple)**, **IV (yellow)** and **V (red)**. The complete description of all the paths are available as SI (Schemes S1 and S2), as well as the corresponding free energy profiles superimposed for all pathways (Figure S1).

For similar TSs, in the case of **8TS-o-p** and **9TS-o-p**, lower relative energies are also obtained (24.7 and 24.5 kcal/mol respectively) compared with the symmetric ones. The comparison of the relative free energy profiles in this way for **1 + 1 → 5z** and **1 +2 → 6z** is shown in Figure 6 (the comparison of the different paths leading to **6z** are compared in Figure S1, SI; also the whole scheme for obtaining the other asymmetric dyads **8z** and **9z** available as Schemes S4 and S5). No significant differences were observed in the kinetically most critical regions of the profiles, which is in good agreement with the reactivity and distribution of products observed in the experiments.
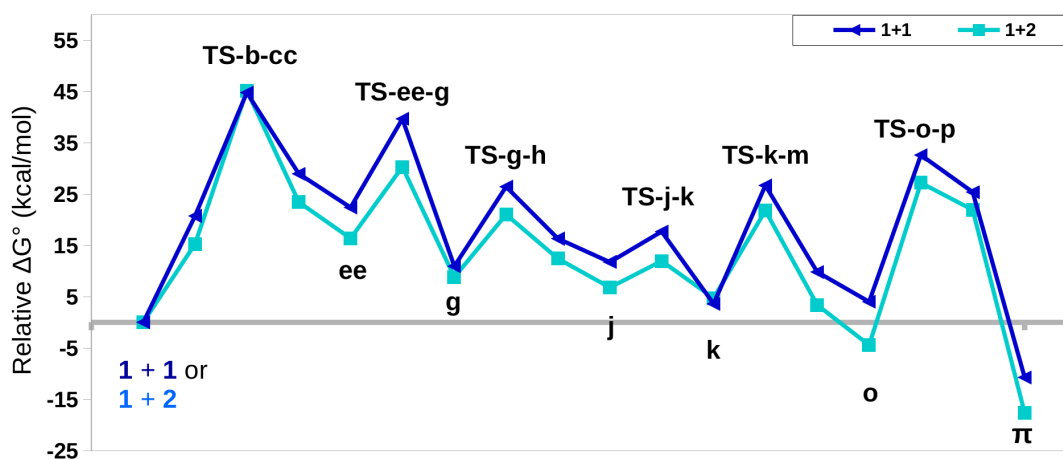

Figure 6. Comparison of the ΔG^{0}_{rel}'s profiles of the main pathways for the processes **1+1→ 5z** *vs*. **1+2 → 6z**.

Impact of the *ortho-* substitution in the aniline.

In relation to the experimental observations (Scheme 2), the remarkable decrease in the reactivity between the aniline **1** and its counterpart with the free *ortho-* position, **3**, still remains intriguing. Besides the relatively low 60 % yield for **3 + 3 →7z** in this work, other authors could not obtain significantly better yields, those ranging within 50-65 %).[23] A deeper insight to the **3 + 3 → 7z** process would shed light on this fact. It would also explain the appearance of the byproducts observed by several authors.

The complete scheme for the production of **7z** starting from **3 + 3** up to reaching the intermediate **7o** resembles the case of **1 + 1→ 5z** (Scheme 3) and it is available as SI Scheme S3. The energies and properties of the intermediates and TSs are very similar to those for its *ortho*-Me-substituted parent (**1 + 1 → 5z**). The main difference with respect to the other symmetric synthesis (and with respect to **1 + 2** as well) is the greater activation energy of the last closing of the Tröger

16

bicycle, the **7TS-o-p**, with a relative energy of 33.1 kcal/mol. The stable intermediate **7o** would live long enough to receive a fourth formaldehyde and then be coupled with another aniline as shown in alternative Path VII (brown) in Scheme 5.
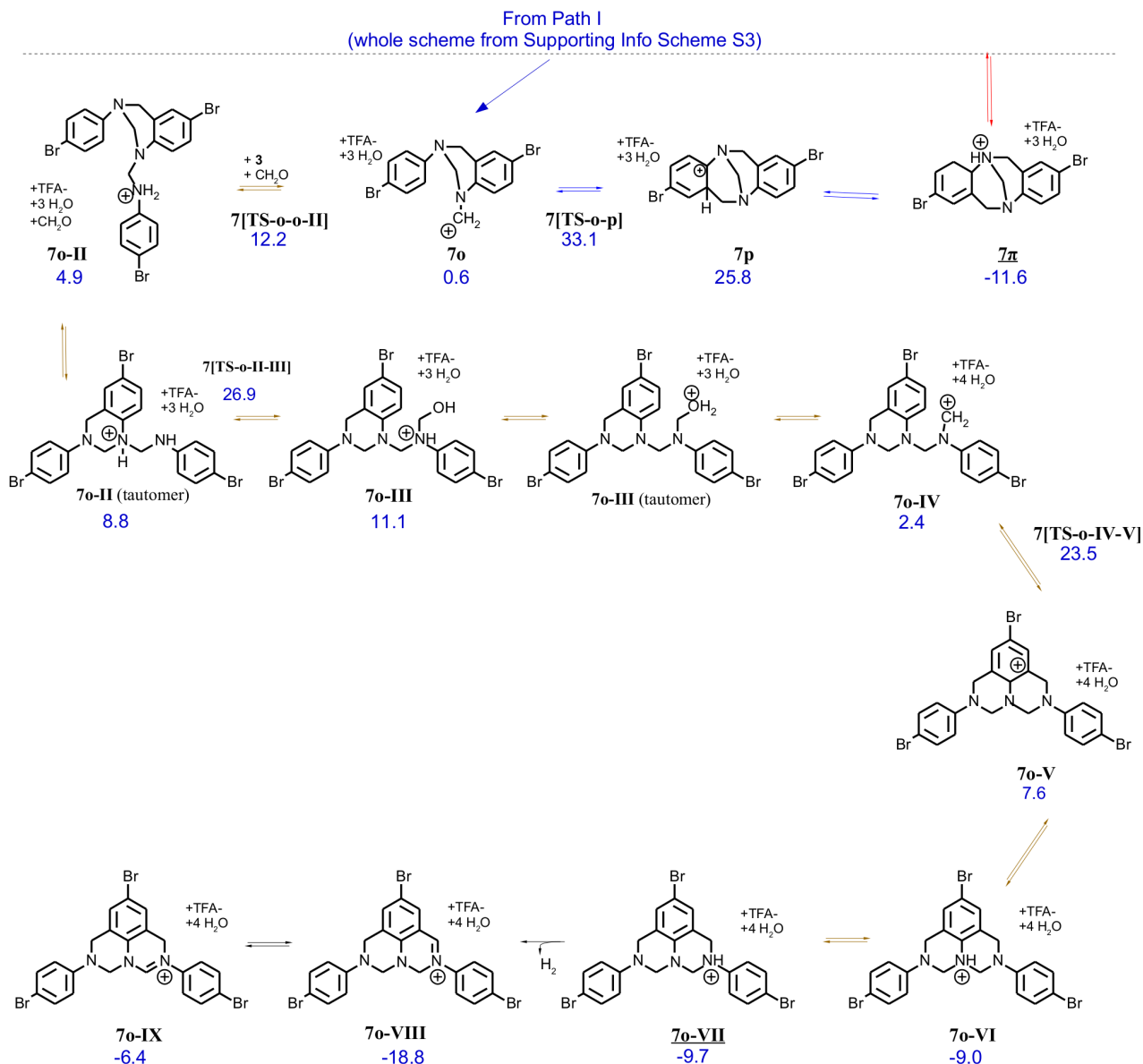

Scheme 5. Lateral Path (Path VII, brown) available when the starting aniline is not substituted in the position 2 of the ring (see also SI Scheme S3).

In fact, its conjugate base (deprotonated **7o-VII**) is characterized; it even has a trivial name, diazajulolidine. Also its derivative **7o-VIII** was identified by mass / electrospray.[24,25] (It is noted

that **7o-VIII** and **7o-IX** are included in the scheme even though they could not be formed under the acid catalysis conditions employed, but are included to compare relative stabilities. The release of H_2 is not observed under our experimental conditions, but it was observed by other authors.) For these adducts to form, it is necessary to go through steps with TSs whose relative energies are lower than that of **7TS-o-p.** It could be seen on Figure 7 that the highest TS along the brown path is smaller than **7TS-o-p** through the blue path, which is the one that leads to the desired product **7π**. The actual extent to which this path competes with the formation of **7π** could not be addressed in a straightforward way, since the bimolecular encounter of **7o** with a third amine (sparse at this point) is required to follow through the brown path, whilst the **7TS-o-p**, although higher, involves an intramolecular process. In any case, its *o*-methyl-substituted counterpart, **1**, lacks such a disadvantage to achieve high yields of the product sought, **5π**. Indeed, for the **1 + 1** reaction, the pathway itself is simply not available, since the position that should be electrophilically attacked is substituted with a methyl group. For the case of reaction **1 + 2**, the higher TS on the lateral path is at higher energy than its **6TS-o-p**.

This lateral pathway finally leads to the adduct of higher molecular weight **7o-VII**, which is about 2 kcal/mol thermodynamically less stable than the desired product **7π** (Scheme 5, profiles in Figure 7).
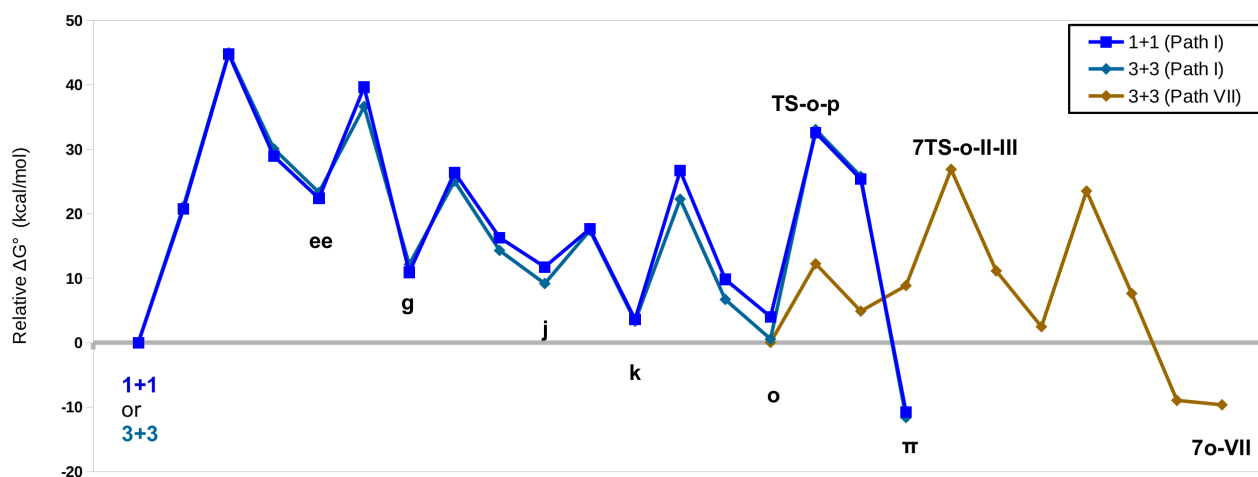
Figure 7. Comparison of the relative free energy profiles for the (**1 + 1 → 5z**) and (**3 + 3 → 7z**) reactions. Path VII (in brown) is only available if the aniline is not substituted at the *ortho* position (**3 + 3**). The lateral path (brown) would be competitive by avoiding the **7TS-o-p**.

18

Byproducts and intermediates.

Wilcox et al. have proposed the intermediacy of species similar to **g** (Schemes 3 or 4) in various symmetric and asymmetric cases since these authors have obtained the TBs using the neutral amine parent of **g** as starting material.[26] Sergeyev *et al*.[14] have obtained the X-ray structure of the intermediate **k,** more precisely of its conjugate base, when neutralized using an aniline similar to **1** (position 2 of the ring subtituted), 2,4-dibromoaniline (Scheme 6). In our case, the conjugated base of **k**, is the compound that we named **l**, with ΔG^0_{rel} of -10.8, -12.9, -11.9, -10.4 kcal/mol for **5l**, **6l**, **7l** and **8l**, respectively (details on the SI, Part1, Scheme S8, and tables on Part 2). There is also evidence of the neutral forms of species like λ and **g**, even if the final product was obtained starting from λ in a symmetrical case (Scheme 3).[27-29]
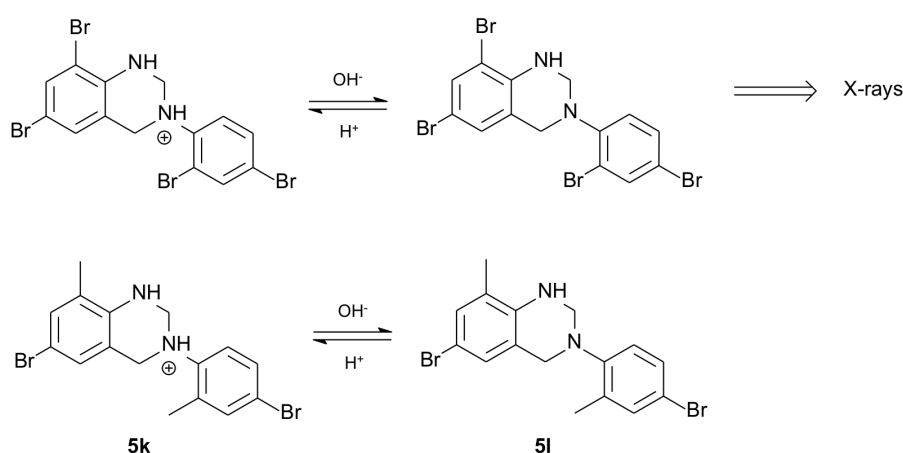
Scheme 6.

Recently, two by-products or intermediaries have been isolated and identified by Yu Wan *et al.* [30] that correspond to a stable intermediate, **o**, which is followed by a high TS and then it has critical relevance in the presented discussion (**5TS-o-p**, Scheme 3 and **7TS-o-p,** Fig, 12). These authors also have isolated an intermediate that would be a stable by-product in the reaction **3 + 3 → 7Z** (Figure 12, brown Path VII). The compounds in question, which were isolated as tetrafluoroborate salts to determine their X-ray structures are shown in Scheme 7:[30]

19
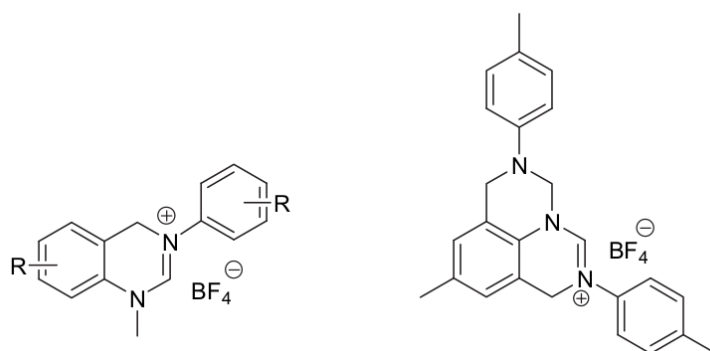
Scheme 7.

The one with the lower molecular weight has been explained by the authors as coming from **o** isomerized in the conditions of ionic liquids in which they performed this reaction. The heavier compound is precisely the salt of **7o-VIII** (Scheme 5), indicated as derivative of **7o-VII**. Abella *et al.* [31] on the other hand, directly isolated and characterized **7o-VII** deprotonated (diazajulolidine). Both Wan and Abella have made symmetrical Tröger's derivatives and found **7o-VII**, which in our mechanistic proposal generates a problem that will decrease the performance for the anilines with the two free *ortho* positions. Indeed, all the reactions carried out by these two groups involve anilines with free *ortho* positions (Wan, for example, has substituents 4-Me-, 4-F-, 4-Cl-, 4-Br-, 3-Cl-4-F- and 3-Cl-4-Me, none of them being 2-substituted). By contrast, the authors discussed above never saw **7o-VII** or its derivatives related to this pathway, working with substituted anilines in the *ortho* position.

All these observations, as well as the yields and distribution of products obtained would be in excellent agreement with the mechanisms proposed for the symmetric and asymmetrical syntheses of Tröger dyads. As far as we know, the only apparent contradiction that we could expect is the proposal of an intermediary **5e** [31] (Path II, green on Scheme 3) discarded in our calculations. But in sharp contrast to the identified intermediaries discussed so far, which were isolated, purified and even characterized by X-rays in some cases, this proposal would be based on previous literature and the masses obtained in ESI-MS. But clearly **7e** and **7ee** have the same mass because they are isomers. It is worth to note a difference of about 10 kcal/mol in the highest transition state of the system (Scheme 3 and Figure 3 and SI figures for **5z-9z**, blue vs. green profiles); such difference would be well above the expected errors for this level of theory, if we consider their small discrepancies obtained between this setup and the higher level CCSD(T)/6-311+G (d,p) in a similar chemical puzzle.[20]

20

Conclusions

The mechanisms that lead to the formation of the Tröger's base derivatives through various ways were proposed and reliably supported by highly correlated DFT calculations. The effects of the substituents on the starting aniline were also analyzed and discriminated according to the role played in each step of the complete mechanism.

The calculations revealed that the studied routes show reaction coordinates with particularities in common. The first TS involving the entry of the first formaldehyde corresponds to the kinetically most expensive step. This would be a TS that could control the reaction rate, at least partially. The possibility of a rate control zone up to **o** and another from the **TS-o-p** was also discussed. In some cases, a stable intermediate such as **o** can reversibly make it possible to have some involvement with Path V together with Path I, which would not affect in general the predominance of the desired product. However, a nonproductive path can be followed starting from **o** and then continuing through a region with TSs of lower energy than **TS-o-p,** that can enable the production of the heavier compound **o-VII**. This complication would be affecting in part the synthesis **3 + 2** and markedly the **3 + 3**, as observed experimentally. This problem would be easily solved by blocking one of the *ortho* positions of the aniline. Indeed, the appearance of **o-VII** and / or its derivatives or tautomers were just observed in experiments for anilines with the two *ortho* positions free.

In regards to other intermediaries, most of them having relative stabilities and preceding high energies TSs were in fact supported by experimental evidence, in some cases even isolated and solved by crystallography (for them or their derivatives).

The mapping of these potential energy surfaces will allow the rational choice of anilines that facilitate higher yields. The limits of the versatility of formaldehyde as a source of methylenes were explored and the collection of stationary points characterized would allow simple calculations in the most critical surface regions in order to consider different synthetic possibilities (different source of methylene, catalyst, substituents in the anilines between other conditions). Beyond the practical interest in the framework of the Tröger-based dyads in our group, which will allow us to more easily generate a variety of derivatives with interesting applications, it is noted that this is, as far as we know, the first mechanistic study that describes in detail the whole mechanism for the synthesis of Tröger's bases, both symmetric and asymmetric from anilines.

21

Experimental Section

General setup for the model syntheses.

TBs were prepared from anilines in trifluoroacetic acid (TFA); the procedure and the characterization of the products were described in detail previously in ref. 15.In brief, synthesis of symmetrical TB (**5z** and **7z**) started 5 mmol of the corresponding aniline and with 2.5 mmol each aniline in the case of the asymmetrical compounds (**6z**, **8z** and **9z**). Anilines were dissolved in 8 ml of trifluoroacetic acid (TFA), 10 mmol of formaldehyde added and the mixture allow to react at room temperature for 24 h (or 168 h, *vide infra*). Finally, the reaction was quenched by adding the mixture dropwise to 30 ml 20% ammonium hydroxide and the precipitate recovered by vacuum filtration. TB were purified by column chromatography as previously reported.

Computational procedure.

The exploration of the potential energy surface was performed using the Gaussian09 package for the DFT calculations.[32] All the structures of reagents, transition states (TS), intermediaries and products were optimized with the solvent model SCRF-IEFPCM[33] (acetic acid) at the level of theory CAM-B3LYP/6-311+G(d,p).

The stationary points were characterized by their Hessian matrix diagonalized to obtain the harmonic frequencies and then the zero point corrections to energy, enthalpy and free energy. The procedure was previously described and tested elsewhere.[17,20] The relative free energy at normal condition used in the discussions is defined as the standard free energy of the species in question minus the standard free energy of the starting reagents, taken as reference ($\Delta G^0_{rel}=0$). In relevant cases, an IRC was done using a mass weighted step of 0.018 atomic units and recalculating the Hessian each 20 or 30 steps.[34]

Occasionally, to help us find encounter complexes, Fukui functions defined from the approximation of discontinuity [35] were used as indicators of reactivity / regioselectivity.[36] The Fukui's functions were defined as usual within this approach, as :

$$\left(\frac{\partial \rho(r)}{\partial N}\right)^+_{v_{ext}} \approx \left(\frac{\Delta \rho(r)}{\Delta N}\right)^+_{v_{ext}} \approx \frac{\rho_{N+1}(r)-\rho_N(r)}{(N+1)-N} = \rho_{N+1}(r)-\rho_N(r) \equiv f^+(r)$$

$$\left(\frac{\partial \rho(r)}{\partial N}\right)^-_{v_{ext}} \approx \left(\frac{\Delta \rho(r)}{\Delta N}\right)^-_{v_{ext}} \approx \frac{\rho_N(r)-\rho_{N-1}(r)}{N-(N-1)} = \rho_N(r)-\rho_{N-1}(r) \equiv f^-(r)$$

22

where $\rho_N(r)$ is the density of the species in question, $\rho_{N+1}(r)$ is the electron density of its radical anion on the equilibrium coordinates of the neutral molecule (because v_{ext}, the potential of the nuclei, should be constant), $\rho_{N-1}(r)$ that of the radical cation, and the value of f^+ and f^- will give an idea of which are the preferential space zones for the attack of an electrophile or a nucleophile, respectively.[35] When these graphs were plotted, the total densities were used instead of their approximate shapes based on the Kohn-Sham orbitals of HOMO and LUMO (one-electron densities),[36] since this way full advantage is taken of the accuracy of the functional to describe open shell species without the problems of the approach of Koopmans.[37]

Fukui's functions numerical data was calculated using the final densities manipulated with the cubman utility of Gaussian 09; also its cubegen utility was used for preparing the cube files for rendering the molecular electrostatic potentials. Visualizations and graphics were prepared with Gabedit 2.4.8 [38] and VMD 1.8.9.[39]

Acknowledgments.

Supporting Information (SI) Available

Part 1: Complete mechanistic schemes (as Scheme 3) for all syntheses are available as SI schemes S1-S7. Other tautomers and conjugated bases of main intermediates are summarized on scheme S8. *Part 2:* Summary of total free energies for all stationary points in main text and SI schemes in atomic units. *Part 3*: Extra Figures. *Part 4:* Compounds characterization. *Part 5: XYZ* coordinates of main stationary points.

23

References

[1] J. Tröger, *J. Prakt. Chem.* **1887**, *36*, 225–245.

[2] M. A. Spielman, *J. Am. Chem. Soc.* **1938***, 57*, 583–585.

[3] V. Prelog, P. Wieland, *Helv. Chim. Acta.* **1944**, *27*, 1127-1134.

[4] a) M. D. Coward, I. Sucholeiki, R. R. Bukownik, C. S. Wilcox, *J. Am. Chem. Soc.*, **1988***, 110*, 6204–6210; b) M. J. Crossley, L. G. Mackay, A. C. Try, *Chem. Commun*, **1995***,* 1925–1927; c) Y Kubo, T. Ohno, J. Yamanaka, S. Tkita, T. Iida, Y. Ishimaru, *J. Am. Chem. Soc.*, **2001***, 123,* 12700–12701; d) S. Satishkuma, M. Periasamy, *Tetrahedron*: *Asymmetry*, **2009***, 20*, 2257–2262.

[5] a) A. Tatibouët , M. Demeunynck, C. Andraud, A. Collet *et al*., *J. Chem Commun,* **1999**, *23,* 161–162; b) M. ValÍk, J. Malina, L. Palivec, J. Foltýnová, M. Tkadlecová, M. Urbanová, V. Brabec, V. Král, *Tetrahedron* **2006**, *62,* 8591–8600.

[6] N. Classens, F. Perard, C. Bresson, C. Moucheron, A. K. D. Mesmaeker, *J. Inorg. Biochem*. **2007**, *101*, 987-996.

[7] a) C. S. Wilcox, L. M. Gree, V. J. Lynch, *J. Am. Chem. Soc.* **1987**, *109*, 1865–1867; b) J. Artacho , P. Nilsson, K.-E. Bergquist, O. F. Wendt, K. Wärnmark, *Chem. Eur. J.,* **2006**, 1*2,* 2692–2701; c) M. Valík, R. M. Strongin, V. Král, *Supramol Chem*, **2005**, *17*, 347–367.

[8] M. S. Sigman, D. R. Jensen, *Acc. Chem. Res.* **2006**, *39*, 221-229.

[9] R. Yuan, M. Li, J. Xu, S. Huang, S. Zhou, P, Zhang, J. Liu, H. Wu, *Tetrahedron* **2016**, *72*, 4081-4084.

[10] a) C. X. Yuan, X. T. Tao, Y. Ren, Y. Li, J. X.Yang, W.-T. Yu, L. Wang, M. H. Jiang*, J. Phys. Chem*. *C* **2007**, *111*, 12811-12816; b) Q. Xin, X. T. Tao, F.Z. Wang, J.-L. Sun, D.-C. Zou, F.-J. Wanga, H.-J. Liu, Z. Liu, M. H. Jiang, *Organic Electronics*, **2008**, *9*, 1076–1086.

[11] T. Braukyla, N. Sakai, M. Daskeviciene, V. Jankauskas, E. Kamarauskas, R. Komskis, T. Malinauskas, S. Jursenas, H. J. Snaith, V. Getautis, *J. Phys. Chem. C* **2017**. *121,* 10267-10274.

[12] C. L. Ramirez, L. Trupp, A. Bruttomesso, V. T. Amorebieta, D. M. A. Vera, A. R. Parise, *Phys. Chem. Chem. Phys,* **2011**, *13*, 2076-2080.

[13] C. L. Ramirez, R. Procaccini, C. A. Chesta, A. R. Parise, D. M. A. Vera, *Organic Electronics*, **2013**, *14*, 2564-2572.

[14] a) S. Sergeyev, *Helvetica Chimica Acta* **2009**, *92*, 415-444; b) B. E. Koene, D. E. Loy, M. E. Thompson, *Chem. Mater.* **1998**, *10*, 2235-2250; c) J. L. Díaz, A. Dobarro, B. Villacampa, D. Velasco, *Chem. Mater*. **2001**, *13*, 2528-2536; d) O. V. Rúnarsson, J. Artacho, K. Wärnmark, *Eur. J. Org. Chem.* **2012**, 7015–704.

24

[15] D. Dusso, C. Ramirez, A. R. Parise, P. A. Lanza, D. M. A. Vera, C. A. Chesta, E. L. Moyano, N. G. Akhmedov, *Magnetic Resonance in Chemistry* **2019**, *57*, 423-454.

[16] C. A. Chesta, J. Mohanty, W. M. Nau, U. Bhattacharjee, R. G. Weiss, *J. Am. Chem. Soc.* **2007**, *129*, 5012-5022.

[17] J. P. Colomer, M.-L. Sciú, C. L. Ramirez, D. M. A. Vera, E. L. Moyano, *Eur. J. Org. Chem.* **2018** (13), 1514–1524.

[18] R. Krishnan, M. J. Frisch, J. A. Pople, *J. Chem. Phys.* **1980**, *72*, 4244–4245.

[19] J. Pople, M. Head-Gordon, K. Raghavachari, *J. Chem. Phys.* **1987**, *87*, 5968-5975.

[20] I. Malvacio, E. L. Moyano, D. M. A. Vera, *RSC Adv.* **2016**, *6*, 83973–88981.

[21] G. Scuseria, C. Janssen, H. Schaefer, *J. Chem. Phys.* **1988**, *89*, 7382-7387.

[22] S. Kozuch, J. Martin, *ChemPhysChem*, **2011**, *12*, 1413-1418.

[23] D. Didier, B. Tylleman, N. Lambert, C. M. L. Vande Velde, F. Blockhuys, A. Collas, S. Sergeyev, *Tetrahedron* **2008**, *64*, 6252–6262.

[24] W. V. Farrar, *J. Appl. Chem.* **1964**, *14*, 389–399.

[25] R. A. Johnson, R. R. Gorman, R. J. Wnuk, N. J. Crittenden, J. W. Aiken, *J. Med. Chem.* **1993**, *36*, 3202–3206.

[26] T. H. Webb, C. S. Wilcox, *J. Org. Chem.* **1990**, *55*, 363-365.

[27] F. C. Cooper, M. W. J. Partridge, *J. Chem. Soc.* **1955**, Part 3, 991–994.

[28] E. C. Wagner, *J. Am. Chem. Soc.* **1935**, *57*, 1296–1298.

[29] T. R. Miller, E. C. Wagner. *J. Am. Chem. Soc.* **1941**, *63*, 832–836.

[30] Y. Wan, R. Yuan, W. Zhang, Y. Shi, W. Lin, W. Yin, R. Bo, J. Shi, H. Wu, *Tetrahedron*, **2010**, *66*, 3405–3409.

[31] C. A. M. Abella, M. Benassi, L. S. Santos, M. N. Eberlin, F. Coelho, *J. Org. Chem.*, **2007**, *72*, 4048–4054.

[32] Gaussian 09, Revision A.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski,

25

R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, **2009.**

[33] G. Scalmani, M. J. Frisch, *J. Chem. Phys.* **2010**, *132*, 114110–114115.

[34] H. P. Hratchian, H. B. Schlegel, in Theory and Applications of Computational Chemistry: The First 40 Years, Ed. C. E. Dykstra, G. Frenking, K. S. Kim, G. Scuseria. Elsevier*, Amsterdam. 2*005, 195-249.

[35] P. Geerlings, F De Proft, W. Langenaeker, *Chem. Rev*. **2003**, *103*, 1793-1874

[36] N. Sablom, F. De Proft, *J. Chem. Phys.* **2007**, *126*, 224108-1 – 224108-6.
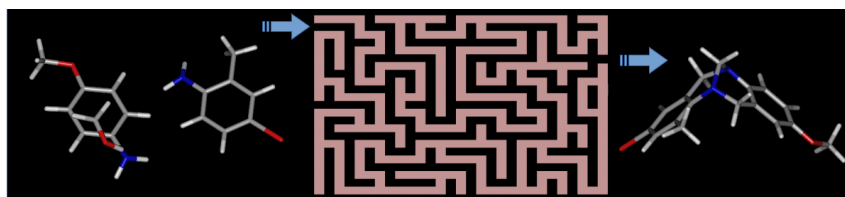
[37] J. L. Borioni, M. Puiatti, D. M. A. Vera, A. B. Pierini, *Phys. Chem. Chem. Phys.* **2017**. *19*, 9189-9198.

[38] A. R. Allouche, *J. Comput. Chem.* **2011***, 32*, 174-182.

[39] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph*. **1996**, *14*, 33-38.

26

TOC Entry

Solving the mechanistic maze through DFT calculations: how do the syntheses of symmetric and asymmetric Tröger's base derivatives work

27