Introduction

A variety of glycoconjugates exhibit a range of essential biological functions, such as immune response, intercellular communication, and the bioactive modulation of natural products. Their unstable biochemical properties, such as glycosidic linkages undergoing hydrolysis in the living body, limit their practical development. Notably, the C-aryl glucoside skeleton, which is widely found in nature, has high biological metabolism stability, and C-glycosyl indole derivatives, which are widely used in the field of medicinal chemistry, have unique biological activities. In 1994, the Hofsteenge group first discovered α-C-mannosyltrypophan (α-C-Man-Trp) I from the Trp7 of ribonuclease. Due to their unique structure and biological activities such as cellular signaling and communication, these compounds have attracted considerable attention. Compounds II and III with 3-indolyl-C-glycoside skeleton are sodium-dependent glucose cotransporter 2 (SGLT2) inhibitors and are considered potential drugs for the therapy of type 2 diabetes. Compounds IV and V with 2-indolyl-C-glycoside skeleton were isolated from Isatis indigotica and they show significant cytoxic activities against human liver cancer HepG2 cells, human myeloid leukemia Mata cells and human myeloid leukemia HL-60.

This report describes palladium-catalyzed C–H glycosylation and retro Diels–Alder tandem reaction via structurally modified norbornadienes (smNBDs). smNBDs were proposed to regulate the reactivity of the aryl-norbornadiene-palladacycle (ANP), including its high chemoselectivity and regioselectivity, which were the key to constructing C2 and C3 unsubstituted C4-glycosidic indoles. The scope of this substrate is extensive; the halogenated six-membered and five-membered glycosides were applied to the reaction smoothly, and N-alkyl (primary, secondary and tertiary) C4-glycosidic indoles can also be obtained by this method. In terms of mechanism, the key ANP intermediates characterized by X-ray single-crystal diffraction and further controlled experiments proved that the migration-insertion of smNBDs with phenylpalladium intermediate endows them with high chemo- and regioselectivity. Finally, density functional theory (DFT) calculation further verified the rationality of the mechanism.

Scheme 1 The study of glycosidic indoles was limited to C2 and C3 position substitution.
donors (Scheme 1b).36 However, the construction of 4-indole-C-glycosides has not been reported. Therefore, the development of an efficient C–H glycosylation reaction to construct a 4-glycylanilinodole skeleton has important application value.

The Catellani reaction was discovered in 1997,37 then Lautens expanded the chemical compatibility of Pd/NBE chemistry by utilizing phosphine ligands in 2000.38 Later, it was named the Catellani–Lautens reaction system. The mechanism of the Catellani reaction is rather lengthy,39,40 so there are a lot of competitive reactions to produce a variety of by-products.41 For example, the reductive elimination of aryl-norbornene-palladation,42–44 the ester-substituted NBE reported by Yu et al.45–47 and C6-substituted NBEs that exhibited unique reactivity and higher selectivity compared to the simple unsubstituted NBE.48–50 In recent years, structurally modified norbornene (smNBE) strategy provides a good solution to this problem.

In 2015, Dong et al. first reported the smNBE (C5-amide-substituted),51,52 and since then, Yu,53,54 Zhou,55–57 Gu,58 Ding,59 Cheng,60 Liang61 and others have developed numerous C2-, C5- and C6-substituted NBEs that exhibited unique reactivity and competitive reactions to produce a variety of by-products.53,54 In 2018, Dong et al. made a breakthrough in using the C2-substituted smNBDs as the coupling partner of the Catellani reaction to realize the ortho-C–H glycosylation reaction and used the reverse Diels–Alder reaction of norbornadiene to realize the synthesis of C4-glycoside indole.

Results and discussion

Initially, we used N-(tert-butyl)-2-iodoaniline (1a) as a model substrate and glycosyl chloride donors (2a) as an ortho-glycosylation reagent, using palladium/norbornadiene (NBD) cooperative catalysis to investigate this reaction (Table 1). Unfortunately, the desired product (4a) was isolated in a 23% yield (entries 1) with NBD (N1, 3.5 equiv.), Pd(OAc)2 (10 mol%), tri(2-furyl)phosphane (TFP, 20 mol%), and Cs2CO3 (4.0 equiv.) in 1,4-dioxane under Ar at 90 °C (12 h)–150 °C (24 h), and most of the reaction proceeded toward the intramolecular Buchwald coupling without ortho-glycosylation. We hypothesized that a wide spectrum of NBDs may have significant effects on the

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Optimization of reaction conditionsa</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>2a (2.0 equiv)</td>
</tr>
<tr>
<td>Pd(OAc)2 (10 mol%), TFP (20 mol%), Cs2CO3 (4.0 equiv.) in 1,4-dioxane</td>
<td>1,4-dioxane, 90 °C (12 h)–150 °C (24 h)</td>
</tr>
<tr>
<td>N (3.5 equiv)</td>
<td>N (3.5 equiv)</td>
</tr>
<tr>
<td>3a</td>
<td>1a</td>
</tr>
<tr>
<td>N1 (NBD), 23 %</td>
<td>N2, 19 %</td>
</tr>
<tr>
<td>N3, 7 %</td>
<td>N4, 5 %</td>
</tr>
<tr>
<td>N5, 24 %</td>
<td>N6, 11 %</td>
</tr>
<tr>
<td>N7, 66 %</td>
<td>N8, 79 %</td>
</tr>
<tr>
<td>N9, 62 %</td>
<td>N10, 49 %</td>
</tr>
<tr>
<td>N11, trace</td>
<td>N12, trace</td>
</tr>
<tr>
<td>N13, trace</td>
<td></td>
</tr>
</tbody>
</table>

a Reaction conditions unless otherwise noted: 1a (0.2 mmol), 2a (0.2 mmol), N (3.5 equiv.), Pd(OAc)2 (10 mmol%), TFP (20 mmol%), and Cs2CO3 (4.0 equiv.) in 1,4-dioxane, 90 °C (12 h)–150 °C (24 h) isolated yields. The yields and ratios of the byproduct were observed by GC-MS.
reactivity of this transformation. Then, we tried a series of modified NBDs and found that the intramolecular Buchwald coupling without ortho-glycosylation by-product was inhibited. After intensive investigation, the NBD with substitution para-toluenesulfonyl groups (N8) was the optimal ortho-C–H glycosylation coupling partner for the Catellani reaction, and it gave an isolated yield of 79%.

With optimized reaction conditions in hand, we first examined the substrate scope of o-iodoanilines. The reactions of o-iodoanilines, electron-donating group (–Me), halogen (–F, Cl), and strongly electron-withdrawing group (–NO2, CN, CO2Et, CF3) substituents performed smoothly, affording the desired 4-glycosylindole products (3b–3h) in 66–79% yield. It is noteworthy to mention that 3-iodopyridine-2-amine as a heteroaromatic substrate reacted favourably and 4-glycosyl-7-
Finally, we found that by adding boron trifluoride ether as the Lewis acid at −78 °C, ethylene dithioglycol and DCM as the solvent to react at −10 °C, the yield of the deprotected product (4a–4b) was 79–83%.

Urinary tract infections (UTIs) are mostly caused by Gram-negative uropathogenic Escherichia coli (UPEC) bacteria. C-Mannosides have good inhibitory activity against the type 1 pilus adhesion, FimH.3,7,4 According to the above experimental results, the structure-modified norbornadiene (smNBDs) can effectively regulate the reaction activity of aryl-norbornadiene-palladacycle (ANP), thus successfully realizing the ortho C–H glycosylation. Due to the unique structural characteristics of norbornadiene, the retro-Diels–Alder reaction could further be realized to construct the indole skeleton. However, the modified norbornadiene has two olefins and four migration insertion sites. In theory, four isomers may be formed, among which J1 and J2 have migratory-insertion on the non-p-toluenesulfonyl side olefins, while J3 and J4 are migratory-insertions on the side of the p-toluene sulfonic acid. Therefore, we first carried out experimental verification.

azaindole (3i) was obtained in 61% yield. Subsequently, a range of the groups on the nitrogen atom of o-iodoaniline was also examined under the optimal reaction conditions. Chain alkane (3j), 3-phenylbutane-2-yl (3l) containing aromatic hydrocarbons, benzyl (3m), cycloalkane (3n), tetrahydronaphthalene (3o), and adamantane (3q) with a large steric hindrance were outstandingly tolerated and afforded the desired products in excellent yields. It can be seen that this method synthesizes a variety of N-alkyl substituted indoles, which are difficult to generate by the direct coupling reaction of indole with tertiary or secondary carbon. Inspired by these encouraging results, we then studied a range of glycosyl chlorides. Tetramethyl-, tetra benzyl-, and tetraallyl-protected α-mannosyl chlorides performed excellently and delivered the corresponding product (3r–3t) in 79–85% yield. The corresponding α-C-aryl mannoside (3u) can be obtained using 2,3,4,6-di-O-isopropylidene-α-mannosyl chloride as a substrate in 52% yield. Other mannosyl chlorides with TBDDS are methyl- and benzyl-, such as 3v and 3w, affording the desired products in 65–82% yields with exclusive α-selectivity. Delightfully, α-ribofuranosyl chlorides with different functional groups, such as methyl-, benzyl-, methoxymethyl-, allyl- and TBDPS can also deliver the corresponding products in 63–76% yield (3x–3ab). In addition, α-hamnosyl chloride can also be used as a suitable substrate for this reaction, and the corresponding targeted product can be obtained with an excellent yield of 70–79% (Table 2).

It can be known from the introduction that pharmaceutical compounds and natural products usually have unprotected carbohydrate skeleton structures, so we conducted deprotection experiments on 4-glycosylindole. We tried using ethylene dithioglycol and boron trifluoride ether for deprotection.

According to the above experimental results, the structure-modified norbornadienes (smNBDs) can effectively regulate the reaction activity of aryl-norbornadiene-palladacycle (ANP), thus successfully realizing the ortho C–H glycosylation. Due to the unique structural characteristics of norbornadiene, the retro-Diels–Alder reaction could further be realized to construct the indole skeleton. However, the modified norbornadiene has two olefins and four migration insertion sites. In theory, four isomers may be formed, among which J1 and J2 have migratory-insertion on the non-p-toluenesulfonyl side olefins, while J3 and J4 are migratory-insertions on the side of the p-toluene sulfonic acid. Therefore, we first carried out experimental verification.

Fig. 1  Computed free energy surface of the oxidation addition reaction. The relative free energies are presented in kcal mol⁻¹.

Scheme 3  Synthetic applications and biochemical experiment.
The barrier of the reverse process of the alkene insertion was 24.1, 27.3, 31.3, and 31.6 kcal mol\(^{-1}\) generating complex benzene ring coordinates with the Pd for forming four isomers of 14.8 kcal mol\(^{-1}\) respectively, while releasing 12.6, 14.2, 15.0, and 15.6 kcal mol\(^{-1}\) as suggested in Liu’s work,\(^\text{55}\) releasing around 56.8 kcal mol\(^{-1}\). The C=C of N8 coordinates with Pd forming four isomers F1, F2, F3 and F4, as shown in Fig. 2.

The alkene insertion occurred with a barrier of 27.2, 27.1, 29.6, and 20.4 kcal mol\(^{-1}\) through G1-ts, G2-ts, G3-ts and G4-ts, respectively, while releasing 12.6, 14.2, 15.0, and 24.7 kcal mol\(^{-1}\) forming H1, H2, H3 and H4, respectively. The benzene ring coordinates with the Pd for H1, H2, H3 and H4. The barrier of the reverse process of the alkene insertion was 24.1, 27.3, 31.3, and 31.6 kcal mol\(^{-1}\) from the complex H series through the G-ts series. We found that the alkene insertion is a reversible process. C–H activation occurred with barriers of 18.3, 28.4, 23.9, and 37.2 kcal mol\(^{-1}\) for H1, H2, H3, and H4, respectively, generating five-membered ring complex J1, J2, J3, and J4. The C–H activation barrier for H1 was 18.3 kcal mol\(^{-1}\), which is the smallest barrier in the C–H activation barriers. The C–H activation occurred from H1 to I2-ts most favorably generating K1, agreeing with the experiment. The reverse C–H activation barrier is –36.0, –49.9, –41.4, and 56.2 kcal mol\(^{-1}\) from complex J series to complex H series.

In conclusion, we directly proved that the migratory insertion of p-toluenesulfonyl substituted norbornadiene with the phenylpalladium intermediate was highly chemo- and regioselective manner by X-ray single-crystal diffraction and density functional theory (DFT) calculation. Achieving high chemo- and regioselectivity was crucial for the subsequent C–H glycosylation, and it can be inferred that the regulation of the norbornadiene p-toluenesulfonyl group on the reaction was the induction effect. In particular, the chemo- and regioselectivity of p-toluenesulfonyl substituted norbornadiene was also very important for the formation of target products, avoiding the formation of p-toluenesulfonyl substituted indoles. Of course, the key step of C–H glycosylation, which is the transition state between ANP intermediates and halogenated glycosides, was complex and still under further exploration.

**Conclusion**

This report describes the palladium-catalyzed C–H glycosylation and retro Diels–Alder tandem reaction via structurally modified norbornadiene (smNBDs). A highly functionalized 4-
glycosylindole was synthesized via the three-component cross-coupling of o-iodoaniline, glycosyl chloride and structurally modified norbornadiene. And the skeleton structure of unprotected glycosyl indole was obtained by the deprotection experiment. In terms of mechanism, the key ANP intermediates characterized by X-ray single-crystal diffraction and further control experiment proved that the migration-insertion reaction of smNBDs with phenyl palladium intermediate has high chemoselectivity. Finally, density functional theory (DFT) calculations were used to study the main five-membered aryl-norbornadiene-palladacycle (ANP) intermediate formation process.

Data availability
All Data associated with this article are available in the ESI.

Author contributions
Y. A. and B.-S. Z. performed the methodology, synthesis, characterization, analysis, and wrote the manuscript, and these authors contributed equally. Y.-N. D., Z. Z., X.-Y. G. and X.-S. L. analyzed the data, discussed the results. X. W. performed the biological activity experiment. Y. L. provided carried out DFT calculations. Y.-M. L. designed the project and supervised the whole experiment. All authors read and approved the final manuscript.

Conflicts of interest
The authors declare no conflicts of interest.

Acknowledgements
This work was supported by a grant from the National Natural Science Foundation of China (NSF 21532001 and 21772075), and the Innovation Fund Project of the Gansu Education Department (2021A-004). Computations reported in this paper were performed on the computer clusters at the Centre for Scientific Modeling and Computation, CUHK.

References