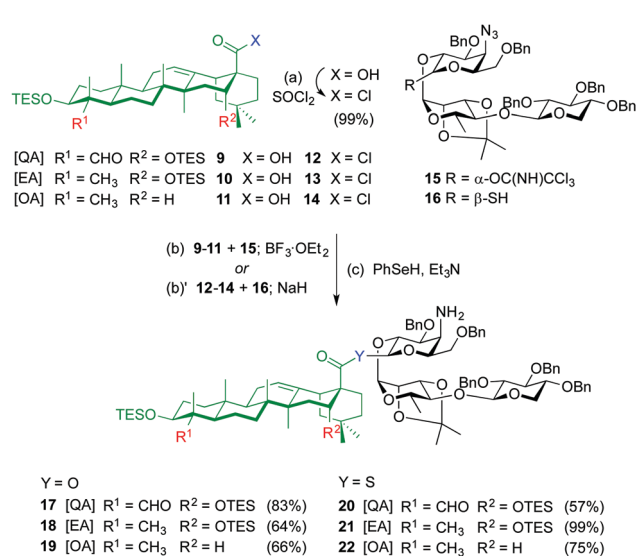


simplified, immunologically silent carboxylic acid acyl chain with a terminal negative charge (carboxylate) to increase synthetic efficiency and polarity, and the right-hand linear trisaccharide as optimal carbohydrate moiety within this domain. Notably, despite the toxicity showed by the central β -glycosyl ester and thioester variants bearing the left-hand trisaccharide,¹⁰ their superior adjuvant activity led us to select these two central linkages to connect the acyl chain-trisaccharide conjugate to the original quillaic acid (QA) core as well as to the echinocystic acid (EA) triterpene, in which the native C4-aldehyde substituent is replaced with a methyl group. To solve the toxicity issue and based on previous SAR,¹² these saponins were designed and synthesised devoid of the entire branched trisaccharide domain, providing new variants through a shortened synthetic route. The corresponding β -glycosyl ester and thioester congeners incorporating oleanolic acid (OA) (5, 8), in which the triterpene C4-aldehyde substituent and the C16-hydroxyl group are lacking, were also synthesised and evaluated for comparative purposes (Fig. 1b). Conformational analysis by NMR and molecular dynamics (MD) simulations on the ester variants (3–5) identified a key role for the C16-hydroxyl group in saponin conformation that correlated with adjuvant activity.

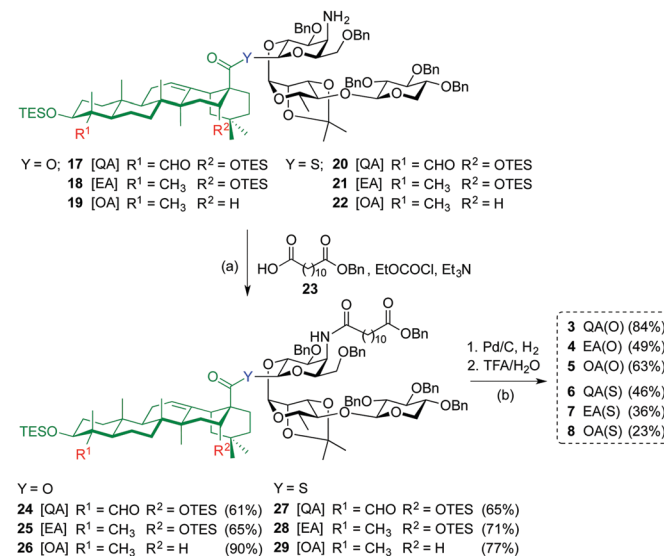
The syntheses of ester variants 3–5 started with Schmidt glycosylation ($\text{BF}_3 \cdot \text{OEt}_2$) of triterpenes 9–11¹¹ with trisaccharide trichloroacetimidate 15⁹ to provide, after benzeneselenol (PhSeH) reduction of the azide to the amine, the QA, EA and OA β -ester intermediates 17–19 in good overall yields (Scheme 1). For the synthesis of the thioester variants 6–8, the triterpene acyl chlorides 12–14 were first prepared by treatment of acids 9–11 with thionyl chloride. Glycosylation of triterpenes 12–14 with trisaccharide thiohemiacetal 16¹⁰ in the presence of NaH, followed by azide reduction gave the corresponding β -thioester derivatives 20–22.



Scheme 1 Reagent and conditions: (a) SOCl_2 , Py, DCM, 0 °C to r.t., 2 h, 99%; (b) $\text{BF}_3 \cdot \text{Et}_2\text{O}$, 4 Å MS, DCM, –78 °C, 6 h (b') NaH, THF, 0 °C to r.t., 3 h; (c) PhSeH, Et_3N /THF/toluene 10 : 3 : 3, 40 °C, 3 h, 83% (2 steps) (for 17), 64% (2 steps) (for 18), 66% (2 steps) (for 19); 57% (2 steps) (for 20), 99% (2 steps) (for 21), 75% (2 steps) (for 22).

Acylation of amines 17–22 with dodecanedioic acid mono-benzyl ester 23 to give 24–29, and subsequent global deprotection by hydrogenolysis (H_2 , Pd/C) and acid hydrolysis (TFA/ H_2O) afforded saponin variants 3–8 in good to moderate yields (Scheme 2). To minimize undesired reduction of the thioester functionality to the corresponding C28-aldehyde, the hydrogenolysis reaction of the β -thioester variants was carefully monitored and the time was reduced to 20 minutes.

We next evaluated saponin adjuvant activity (50 μg dose) by vaccinating mice in combination with OVA (20 μg) as a model antigen at days 0, 14, and 28. As controls, mice were immunized with OVA alone (20 μg) (no-adjuvant) and with OVA (20 μg) plus QS-21 (20 μg) (positive control), a QS-21 dose known to induce optimal antibody responses with acceptable toxicity in mice. Antibody titres against OVA in mouse sera were determined by indirect ELISA before the second boost (day 26) and at the time of sacrifice (day 46). On day 26, anti-OVA IgG titres elicited by echinocystic and quillaic acid variants were significantly higher than those of the no-adjuvant control and comparable to (for echinocystic acid variants), or just slightly below (for the quillaic acid derivatives) those of the QS-21 group (Fig. 2a). In all cases, replacement of the central ester linkage with the thioester linker (6–8) did not impair saponin adjuvant activity. Oleanolic acid variants 5 and 8 elicited lower antibody responses, similar to the no-adjuvant control. The observed trend was maintained on day 46 (Fig. 2b), with increased titres for all three ester variants compared to those on day 26. On day 46, however, the anti-OVA IgG titres with quillaic acid thioester 6 were no longer significantly higher than the no-adjuvant group. Antibody subtyping of the anti-OVA IgG isotypes (day 46) showed that both echinocystic acid variants 4 and 7 elicited the highest IgG1, IgG2b and IgG2c titres among the synthetic



Scheme 2 Reagent and conditions: (a) EtOCOC(OMe)_2 , Et_3N , THF, 0 °C to r.t., 4 h, 61% (for 24), 65% (for 25), 90% (for 26); 65% (for 27), 71% (for 28), and 77% (for 29); (b) 1: Pd/C, H_2 (1 atm), EtOH/THF 1 : 1, r.t., 16 h (for 24–26); 20 min (for 27–29); 2: TFA/ H_2O 3 : 1, 0 °C, 1 h, 84% (for 3), 49% (for 4), 63% (for 5); 46% (for 6), 36% (for 7), and 23% (for 8).



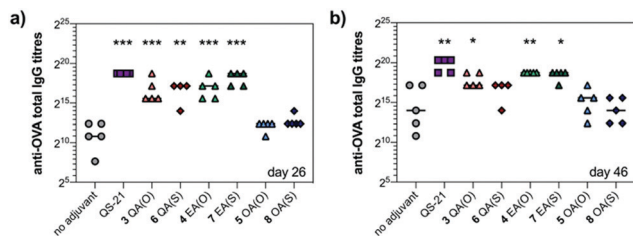


Fig. 2 Antibody responses in mice elicited by saponins **3–8** with OVA antigen (a) on day 26 and (b) on day 46 after first immunization. Horizontal bars indicate median titres. Statistical significance versus no-adjuvant group using unpaired Student's *t*-test: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

variants (Fig. S1, ESI[†]). In general, the mouse IgG1 subclass is associated with Th2 cell responses (humoral immunity), whereas the IgG2 is related to Th1 responses (cellular immunity).¹³ For the quillaic acid variants, the ester analogue **3** showed IgG1 and IgG2b titres significantly higher than the no-adjuvant group, albeit lower than its echinocystic acid counterpart **4**. In contrast, the thioester analogue **6** elicited IgG1 and IgG2b titres similar to the inactive oleanolic acid variants **5** and **8** and the negative control (Fig. S1, ESI[†]).

As a standard initial overall toxicity assessment, no mouse died due to immunization with any of the saponin variants (**3–8**) and no weight loss was observed among the analysed groups. This is notable for quillaic acid variants **3** and **6**, since their parent compounds incorporating the branched trisaccharide were toxic at 20 μg (1/5 mouse dead for the ester variant⁹ and 2/5 mice dead for the thioester variant¹⁰). To assess potential toxic side effects beyond mortality and weight loss, we analysed hepatic alanine transaminase (ALT) levels in blood serum on day 46 (Fig. S2, ESI[†]). The levels of this liver enzyme remained low for all groups,

with no significant differences compared to the no adjuvant control, confirming the non-toxicity of these saponin variants.

Taken together, these results have identified echinocystic acid variants **4** and **7** as the most potent saponin variants, eliciting higher antibody responses than their quillaic acid counterparts, at levels closer to those of QS-21. These data highlight the dispensability of the C4-aldehyde substituent and the importance of the C16-hydroxyl group of the triterpene for elicitation of potent humoral responses by these synthetic saponins.

Next, to seek a potential role of the triterpene C16-hydroxyl in saponin conformational preferences that could explain its importance in adjuvant activity, we performed a conformational analysis of the synthetic saponins by combining experimental NMR data with molecular dynamics (MD) simulations. Given the similar adjuvant activity showed by the ester and thioester congeners, we carried out 0.5 μs MD simulations on the three ester variants **3–5** using AMBER 18 implemented with GAFF and GLYCAM 06 force-fields.¹⁴ In parallel, to validate the MD simulations, we performed 2D-NOESY experiments on the three variants in methanol, as their solubility in water was not sufficient for these NMR studies. Whereas NOE cross-peaks between the carbohydrate and the triterpene were negative, positive NOEs for the distal acyl chain methylene protons were observed (Fig. S3, ESI[†]), implying that the acyl chain displays different local motion and higher flexibility than the rest of the molecule, in accordance with the MD simulations (Fig. S6, ESI[†]). Moreover, the observed NOE cross-peaks were compatible with the main conformations derived from the MD in methanol (Fig. S10, ESI[†]). These conformations differ in the orientation of the ester linkage and thus in the orientation of the Gal residue (and the trisaccharide) relative to the triterpene. Fittingly, the folded conformation that is predicted by MD in

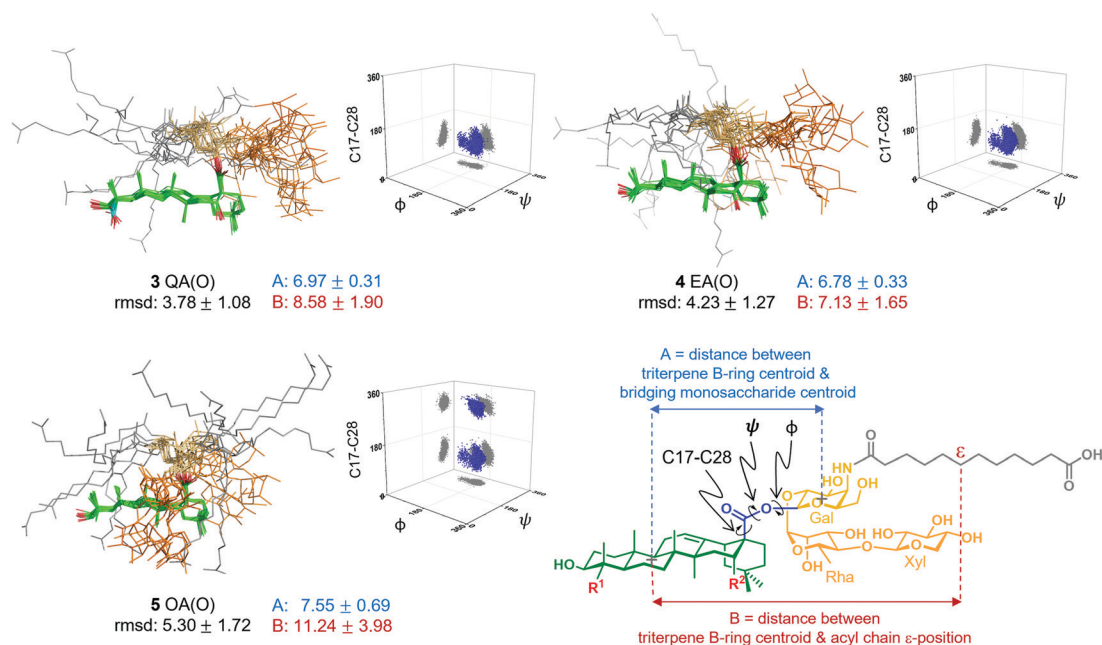


Fig. 3 Structural ensembles derived from unrestrained 0.5 μs MD simulations of ester variants **3–5** in explicit water, showing rmsd (\AA) for heavy atoms relative to the average structure, key distances (\AA), and 3D plots of torsion angle distributions around the central glycosidic linkage.



methanol (Fig. S6, ESI[†]) was experimentally confirmed by NMR. Indeed, NOE cross-peaks between protons at the Gal and Rha residues and at the triterpene were observed (Fig. S4 and S5, ESI[†]). Additionally, the existence of weak cross-peaks between the Gal protons and a methyl group of the triterpene for **3** and **4** (Fig. S4, ESI[†]) strongly suggests that the extended conformation is also present, although in a minor proportion. These experimental results validate our computational approach, which was also applied to analyse the conformational behaviour of these derivatives in water. Thus, MD simulations of **3** QA(O), **4** EA(O) and **5** OA(O) in explicit water were carried out. According to our simulations, both adjuvant-active variants, **3** and **4**, exhibited a main conformation in water in which the acyl chain (grey) was mainly folded back over the triterpene (green), while the linear trisaccharide (orange) was oriented away from the centre of the triterpene, in a rather extended orientation (Fig. 3). In contrast, the inactive variant **5** displayed a distinct conformation. The folded presentation of the acyl chain over the triterpene was barely present, instead pointing away from the triterpene. Moreover, its linear trisaccharide showed a more disorganised conformation, folded over the triterpene. The different geometries were quantitatively characterized by measuring the torsion angles around the central glycosidic linkage (*i.e.* triterpene C17 to Gal C1) and the distances A and B, as defined in Fig. 3.

Interestingly, the distributions obtained for the central linkage torsion angle for the active variants **3** QA(O) and **4** EA(O) were basically identical. In these clusters, the Gal residue displays its β -face approximately perpendicular to the rear edge of the triterpene, favouring the folding of the acyl chain over the triterpene. Strikingly, the lack of the C16-hydroxyl group in the inactive OA(O) variant (**5**) modifies the three-dimensional shape of this saponin in comparison to those deduced for **3** and **4**. Thus, **5** exhibited an additional population for the central linkage torsion angle distribution (Fig. 3), and its linear trisaccharide displayed a different conformational behaviour (Fig. S7, ESI[†]), with a more flexible glycosidic linkage between the xylose and rhamnose units. Moreover, **5** exhibited longer distances between the triterpene B-ring and the Gal residue (A, blue) and between the triterpene B-ring and the acyl chain (B, red), compared to active variants **3** and **4** (Fig. 3). Therefore, the presentation of the polar and non-polar moieties of these molecules is markedly different for the active and inactive saponins, and the different orientation of the carbohydrate and acyl chain in **3** QA(O) and **4** EA(O) might facilitate the hydration of the C16-hydroxyl group. The above conformational features for these truncated analogues are in line with those of the saponin variants incorporating the branched trisaccharide domain,¹⁰ and provide a molecular-level rationale for the attenuated adjuvant activity in **5** OA(O) associated to the conformational changes induced by the lack of the C16-hydroxyl group.

In conclusion, we have designed and synthesised streamlined saponin variants exploiting optimal structural variations at the acyl chain, triterpene and central linker regions. The synthetic strategy enabled efficient access to optimized minimal saponins

in reduced steps [*i.e.* 22 total steps for **3** QA(O) and **4** EA(O)]. Overall, replacement of the ester central linker by a thioester modification did not significantly impair adjuvant activity. Echinocystic acid variants **4** and **7** were found to be superior to their quillaic acid congeners, which makes them attractive leads for further development and provides a more affordable and sustainable triterpene source for potent saponin adjuvants. These multidisciplinary studies emphasise the dispensability of the C4-aldehyde and the importance of the C16-hydroxyl in the adjuvant activity of these truncated variants. Notably, conformational analysis of the ester variants by NMR and molecular dynamics simulations highlighted a key role of the triterpene C16-hydroxyl group in saponin conformation that correlated with adjuvant activity, suggesting a mechanism of action involving interaction with discrete molecular targets. In all, the superiority of the echinocystic acid variants presented in this work makes them leading scaffolds for future mechanistic studies and synthetic vaccines based on saponin adjuvants.

Funding from the ERC (ERC-2016-STG-716878 to A. F.-T.) and Spanish MCIU (CTQ2017-87530-R, RYC-2015-17888 to A. F.-T.; SAF2015-65327-R and RTI2018-096494-B-100 to J. A.; RTI-2018-099592-B-C21 to F. C.; Severo Ochoa Accreditation SEV-2016-0644 to CIC bioGUNE) is acknowledged. A. F.-T. thanks Raquel Fernández for inspiration.

Conflicts of interest

A. F.-T. is co-inventor on patents and patent applications that include saponin molecules presented in this work.

Notes and references

- 1 G. Del Giudice, R. Rappuoli and A. M. Didierlaurent, *Semin. Immunol.*, 2018, **39**, 14–21.
- 2 C. R. Kensil, U. Patel, M. Lennick and D. Marciani, *J. Immunol.*, 1991, **146**, 431–437.
- 3 M. A. Lacaille-Dubois, *Phytomedicine*, 2019, **60**, 152905.
- 4 G. Ragupathi, J. R. Gardner, P. O. Livingston and D. Y. Gin, *Expert Rev. Vaccines*, 2011, **10**, 463–470.
- 5 A. Fernández-Tejada, *Pure Appl. Chem.*, 2017, **89**, 1359–1378.
- 6 P. Wang, Y.-J. Kim, M. Navarro-Villalobos, B. D. Rohde and D. Y. Gin, *J. Am. Chem. Soc.*, 2005, **127**, 3256–3257.
- 7 A. Fernández-Tejada, D. S. Tan and D. Y. Gin, *Acc. Chem. Res.*, 2016, **49**, 1741–1756.
- 8 M. M. Adams, P. Damani, N. R. Perl, A. Won, F. Hong, P. O. Livingston, G. Ragupathi and D. Y. Gin, *J. Am. Chem. Soc.*, 2010, **132**, 1939–1945.
- 9 E. K. Chea, A. Fernández-Tejada, P. Damani, M. M. Adams, J. R. Gardner, P. O. Livingston, G. Ragupathi and D. Y. Gin, *J. Am. Chem. Soc.*, 2012, **134**, 13448–13457.
- 10 W. E. Walkowicz, A. Fernández-Tejada, C. George, F. Corzana, J. Jiménez-Barbero, G. Ragupathi, D. S. Tan and D. Y. Gin, *Chem. Sci.*, 2016, **7**, 2371–2380.
- 11 A. Fernández-Tejada, E. K. Chea, C. George, N. Pillarsetty, J. R. Gardner, P. O. Livingston, G. Ragupathi, J. S. Lewis, D. S. Tan and D. Y. Gin, *Nat. Chem.*, 2014, **6**, 635–643.
- 12 A. Fernández-Tejada, E. K. Chea, C. George, J. R. Gardner, P. O. Livingston, G. Ragupathi, D. S. Tan and D. Y. Gin, *Bioorg. Med. Chem.*, 2014, **22**, 5917–5923.
- 13 F. Nimmerjahn and J. V. Ravetch, *Science*, 2005, **310**, 1510–1512.
- 14 K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley and R. J. Woods, *J. Comput. Chem.*, 2008, **29**, 622–655.

