Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Thiolate addition to withaferin A: a GFN2-xTB and DFT study of site-selective glutathione conjugation

Bhargava Reddy

bhargava1506003@gmail.com

Chalapathi Institute of Pharmaceutical Sciences

Hari Prasad

Chalapathi Institute of Pharmaceutical Sciences

P. V. Suresh

Chalapathi Institute of Pharmaceutical Sciences

Research Article

Keywords: withaferin A, glutathione conjugation, Michael addition, GFN2-xTB, density functional theory, electrophilicity index

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11381477/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Thiolate addition to withaferin A: a GFN2-xTB and DFT study of site-selective glutathione conjugation

Bhargava Reddy [1•], Hari Prasad [1], P.V. Suresh [1]

[1] *Department of Pharmaceutical Analysis, Chalapathi Institute of Pharmaceutical Sciences, Chalapathi Nagar, Lam, Guntur, A. P., India.*

• Corresponding author: Bhargava Reddy, bhargava1506003@gmail.com

ORCID – Bhargava Reddy: 0009-0005-0568-6966

Abstract

Withaferin A (WFA), the major steroidal lactone of *Withania somnifera*, alkylates cysteine residues in vimentin, Hsp90 and IKKβ. Glutathione (GSH), the cell's most abundant thiol, serves as the initial step in detoxifying electrophiles. The carbon that interacts with it determines both the rate of clearance and the extent to which a compound remains to alkylate proteins. It is unquantified at any of WFA's four electrophilic carbons: the ring-A enone (C3), the side-chain lactone (C24) and the epoxide carbons (C5, C6). This study computes conformer-ensemble free energies with GFN2-xTB in implicit water with quasi-RRHO corrections, scans the C–S approach coordinate at each site, and recalibrates with M06-2X single point calculations on truncated models. Nucleophiles were modelled as thiolates and adducts as anions; the isolable species is the neutral protonated conjugate. Addition is most exergonic at C3-α ($\Delta G = -11.7$ kcal mol^{-1}), then C6 (−8.9), C5 (−8.3) and C24 (+0.2 to −4.4). Site-specific DFT corrections invert this ordering toward the epoxides. A def2-TZVP check shifts every correction by about +7.8 kcal mol^{-1}, removing the sign difference between reaction classes but, because it is uniform, leaving the ranking intact: the inversion is robust to basis set though the corrections are not. No Hessian-verified saddle was located, so no activation energy is reported; C3 has no interior scan maximum whereas both methods place one at the epoxide, supporting kinetic control without establishing it. The methanethiolate probe reproduced glutathionate's site preference but inverted its facial preference, α by 7.0 against β by 13.3 kcal mol^{-1}, so minimal-thiol screening ranks sites dependably, faces unreliably.

Keywords: withaferin A; glutathione conjugation; Michael addition; GFN2-xTB; density functional theory; electrophilicity index

---

1 Introduction

Withaferin A (WFA; $C_{28}H_{38}O_6$) is a well-studied withanolide, a family of C_{28} steroidal lactones found in *Withania somnifera* (ashwagandha), and among the most thoroughly studied covalent natural products. Chemoproteomic and mutagenesis work has mapped a wide target set that takes in cytoskeletal proteins, molecular chaperones, kinases and redox-signalling nodes, and in every case the common step is covalent modification of a cysteine thiolate [1, 2]. In vimentin the modified residue is Cys328, which sits against the fused A/B-ring face of the bound ligand,

though which carbon actually forms the bond has never been pinned down at atomic resolution [3, 4].

That gap is more than a structural detail, because WFA offers three chemically distinct electrophilic motifs across four carbons: an α,β-unsaturated ketone in ring A (a 2,3-en-1-one, referred to below by its β-carbon C3), an α,β-unsaturated δ-lactone in the side chain (C24), and a 5β,6β-epoxide alongside ring B (C5 and C6). A survey of more than fifty natural withanolides concluded that the ring-A enone is the dominant covalent pharmacophore for protein alkylation, with the epoxide and lactone contributing secondary, target-dependent reactivity [1]. The oldest physical-organic study of WFA, which used ethanethiol, thiophenol and L-cysteine ethyl ester under non-enzymatic conditions, reached the same verdict for C3 on the basis of nuclear magnetic resonance (NMR) evidence [5]. Set against that, fungal biotransformation of a *W. somnifera* extract by *Beauveria bassiana* delivered a real, enzymatically formed WFA–glutathione conjugate, CR-777, which one- and two-dimensional NMR and high-resolution mass spectrometry placed at the C6 epoxide-derived position and not at C3 [6]. Regiochemistry here is therefore a live question, and the enzymatic and non-enzymatic answers need not agree.

It also matters that the three motifs are not interchangeable. The ring-A 2,3-en-1-one behaves as a textbook soft Michael acceptor: the nucleophile attacks the β-carbon, the developing charge spreads onto the C1 carbonyl oxygen, and retro-Michael elimination makes the step reversible in principle. The side-chain δ-lactone belongs to the same class on paper, but its β-carbon is more heavily substituted and sits on a mobile side chain, so it differs both in steric access and in how much stabilisation the incipient enolate can draw on. Epoxide opening is another reaction altogether. Attack comes in from the back of a strained three-membered ring, strain release rather than conjugation supplies the driving force, the product is an alkoxide instead of an enolate, and the step is effectively irreversible. Comparing sites across these motifs therefore means comparing reactions that differ in charge distribution, steric demand and reversibility, which is exactly why no single reactivity descriptor will settle the question and why explicit energetics are needed.

Glutathione (γ-L-glutamyl-L-cysteinylglycine, GSH) reaches 1–10 mM in hepatocytes, making it the most abundant low-molecular-weight thiol in the cell and its main defence against electrophilic xenobiotics. Conjugation to an electrophile, whether spontaneous or catalysed by glutathione *S*-transferase (GST), usually opens Phase II detoxification and ends in a urinary mercapturic acid. For comparable Michael acceptors the non-enzymatic kinetics are well mapped: pseudo-first-order ultraviolet and chromatographic measurements on cyclic and open-chain chalcones yield second-order rate constants that track the electron-withdrawing power of the acceptor and the steric shielding of its β-carbon [7]. Nothing equivalent exists for WFA, or indeed for any withanolide. The omission matters, since GSH conjugation is chemically the same process that drives WFA's on-target protein alkylation, and the balance between detoxification and target engagement turns on which site reacts and how quickly.

Treating the problem from first principles runs into two difficulties. WFA is conformationally and stereochemically awkward, with six fused or pendant rings and nine defined stereocentres, so reaction energies have to be built on conformer ensembles rather than single structures, and every new C–S bond creates a stereocentre whose diastereomers need enumerating. Meanwhile a 34-heavy-atom steroid paired with a 20-heavy-atom tripeptide, repeated over four sites, two faces and two nucleophiles, lies out of reach of hybrid DFT or correlated-wavefunction methods

except for a handful of single points. Semi-empirical tight binding bridges the gap. GFN2-xTB is parameterised against a large reference set and recovers geometries and reaction energies for medium-sized organic systems at a small fraction of DFT cost [8], while its analytical linearised Poisson–Boltzmann (ALPB) solvation model carries that into the condensed phase [9]. Descriptor calculations of this type have already been used to rank withanolide electrophilicity against a kinase target [10]. GFN2-xTB does, however, carry documented errors of 2–6 kcal mol^{-1} on reaction energies and larger, less predictable errors on barriers, which makes DFT calibration a requirement rather than a refinement. Quasi-rigid-rotor-harmonic-oscillator (qRRHO) thermochemistry is applied throughout, since harmonic entropy diverges for the low-frequency modes that both partners have in abundance [11].

This work maps the addition of thiolate to WFA, categorised by site, face, and stereoisomer. A minimal methanethiolate probe covers all four electrophilic carbons, authentic glutathionate confirms the outcome at the dominant site, M06-2X/def2-SVP recalibrates the semi-empirical energetics everywhere rather than at one site, and the approach coordinate is examined for a kinetic barrier. Four questions shape the analysis. Which carbon gives the most stable conjugate, and by a margin large enough to survive the method's error bar? Does a minimal thiol surrogate, which is what makes an exhaustive site-and-face screen affordable, actually predict the behaviour of the real tripeptide? Does the thermodynamic ordering hold up once a higher level of theory is applied at every site? And is the thermodynamically preferred carbon also the kinetically accessible one? The last two answers differ from what the raw semi-empirical numbers suggest, and the split between thermodynamic and kinetic preference turns out to be the main result.

2 Computational methods

2.1 Molecular structures and stereochemistry

WFA was built from its isomeric SMILES string and checked against the PubChem reference entry, CID 265237, by exact International Chemical Identifier key (InChIKey) match (DBRXOUCRJQVYJQ-CKNDUULBSA-N) [12] before anything else was run. Structure handling, conformer embedding, stereoisomer enumeration and Cahn–Ingold–Prelog (CIP) labelling were all done in RDKit 2026.03.6 [13].

The nucleophile is modelled throughout as the deprotonated thiolate: methanethiolate (CH_3S^-) as the minimal probe, and glutathionate as the authentic confirmatory nucleophile, carrying a deprotonated thiol, deprotonated γ-glutamyl and glycine carboxylates, and a protonated α-amino group, for a net charge of −2.

Because the distinction is easy to blur, we set out here what is and is not being modelled, and hold to it in the Results. **Intrinsic thiolate reactivity** is the quantity this work computes: the energetics of RS^- adding to each electrophilic carbon, with every site treated identically so that the comparison between sites is internally consistent. **Glutathione speciation** is a separate matter. The thiol pK_a of glutathione lies between 8.7 and 9.2, so at pH 7.4 only some 2–5% of the pool is thiolate; the reactive fraction is small, and an absolute rate would require the computed thiolate reactivity to be weighted by it. Nothing here supplies that weighting, so no rate is predicted. **The non-enzymatic reaction at physiological pH** is the sum of those two factors, and the site ranking computed here transfers to it only to the extent that speciation affects

all four carbons equally, which it does, since the same nucleophile attacks each one. **The GST-catalysed reaction** lies outside the model altogether: an enzyme can lower the pK_a of the bound thiol, orient the two partners, and stabilise a transition state selectively, so neither the site ranking nor the relative accessibility computed here need carry over to it. Where the enzymatically derived C6 conjugate is discussed below, it is on those terms.

Modelling the thiolate rather than the thiol is the usual simplification for intrinsic-reactivity comparisons, since the elementary step is attack by RS^- and the rate scales as $[RS^-][\text{electrophile}]$.

Rather than hard-coding atom indices, the four electrophilic carbons were found by substructure matching and the resulting map checked against the expected connectivity: C3 as the β-carbon of the ring-A 2,3-en-1-one, that is the alkene carbon further from the C1 carbonyl; C24 as the β-carbon of the α,β-unsaturated side-chain δ-lactone; and C5 and C6 as the two epoxide carbons. Attack geometry is defined site by site. For the C2=C3 and C24=C25 alkenes the two prochiral faces are labelled α and β relative to the C19 angular methyl. The epoxide carbons admit only backside attack, anti to the C–O bond being broken, which fixes the approach direction and leaves one accessible geometry at each of C5 and C6.

Adducts were assembled with formal charges placed explicitly, atom by atom, instead of being inferred by resetting implicit hydrogen counts. The incoming sulfur becomes neutral and the −1 charge goes on the electrophile-side centre that should carry it: C2 for Michael addition at C3, the epoxide oxygen for ring opening at C6 or C5, and C25 for conjugate addition at C24. Every anionic adduct was then checked for exact molecular formula and mass balance against the expected product before any energy was computed.

This has a consequence at C24 worth spelling out. Addition there creates a new C–S stereocentre at C24 and generates a formal carbanion at C25. Since C25 carries the negative charge it is not a tetrahedral stereocentre, and RDKit's CIP implementation assigns it no configuration. The C24 adduct accordingly exists as two CIP-resolvable stereoisomers rather than four, and we label them R and S by the configuration at C24 itself. The coordinate files in the Supplementary Information (ESM_1) follow the same convention.

2.2 Conformer ensembles and thermochemistry

Each species began with 30–120 geometries embedded by ETKDGv3 and pre-optimised under MMFF94, then pruned at a 0.4 Å heavy-atom root-mean-square deviation (RMSD) and a 0.05 kcal mol^{-1} energy threshold. The 8–25 survivors were optimised with GFN2-xTB [8] in ALPB implicit water [9] using the ASE limited-memory BFGS implementation [14], after which the three or four lowest were re-optimised tightly and given a numerical Hessian (two-point central differences on analytic gradients, 0.005 Å displacement). Ensemble free energies are Boltzmann averages over those Hessian-verified conformers at 298.15 K and 1 bar, with vibrational entropy handled by Grimme's qRRHO interpolation [11] at a damping frequency of 100 cm^{-1}; the expressions themselves appear in ESM_1, Section S1.

Two reference-state conventions are worth flagging before these numbers are compared with experiment. Free energies use a 1 bar ideal-gas standard state with the ALPB solvation free energy added, not a 1 M solution standard state; for a bimolecular association that difference amounts to a uniform 1.9 kcal mol^{-1} offset at 298.15 K, which cancels in every site-to-site

comparison here and has not been applied. Reactants are also treated as separated and independently solvated, so the reference state contains no ion-pairing or encounter-complex stabilisation, consistent with the pose sampling in Section 3.8. Since everything reported is a difference between species computed under identical conventions, these choices move absolute values without disturbing the rankings that carry the conclusions.

Adduct formation free energies are quoted relative to the separated, independently solvated reactants:

$$\Delta G_{\mathrm{ads}} = G(\mathrm{adduct}) - G(\mathrm{WFA}) - G(\mathrm{Nu^-}) \qquad (1)$$

2.3 Reaction-coordinate scans and saddle-point searches

Methanethiolate was placed at each site and prochiral face along the Bürgi–Dunitz trajectory, 107° from the carbonyl axis for the enone and lactone and backside at 180° to the C–O bond for the epoxide carbons, starting from an S···C separation of 3.0 Å. The C–S distance was stepped down to 1.8 Å in 0.1 Å increments, giving 13 constrained points, with every other degree of freedom relaxed at each one.

Saddle-point searches were seeded from the scan's energy maximum, but only where that maximum fell at an interior point of the profile. When the profile falls monotonically from the first and most separated point, no interior barrier exists within the scanned range; the approach is then recorded as barrierless and no search is attempted. Searches that did run used the Sella eigenvector-following optimiser [15] with an internal-coordinate Hessian update, converged below a root-mean-square force of 0.02 eV Å^{-1}. Only a stationary point whose mass-weighted Hessian carried exactly one negative eigenvalue was accepted as a verified transition state. No stationary point located in this work met that test.

Three quantities have to be kept apart, and we give them distinct symbols and use them consistently below. The reaction free energy of Eq. 1 compares two minima and says nothing about the path between them. The scan maximum,

$$\Delta G_{\mathrm{scan}} = G(\text{highest sampled scan point}) - G(\mathrm{WFA}) - G(\mathrm{Nu^-}) \qquad (2)$$

is the free energy of the highest of the 13 constrained points on a relaxed profile, referenced to the separated reactants. It belongs to a discrete scan along a constrained coordinate rather than to a stationary point: it is neither variationally optimised nor Hessian-verified, and its position is limited by the 0.1 Å spacing of the scan. An activation free energy, ΔG‡, requires a Hessian-verified first-order saddle. Since none was located at any site, **no ΔG‡ is reported anywhere in this work**, and ΔG_scan should not be read as an estimate of one. Where a search reached a minimum or a higher-order saddle instead, that outcome is reported together with its imaginary-mode count.

2.4 Reactivity descriptors

Global electrophilicity follows the Parr formalism [16], using vertical ionisation potential (IP) and electron affinity (EA) taken from N ± 1 electron single points at the fixed solvated neutral geometry:

$$\mu = -\frac{IP + EA}{2} \qquad \eta = IP - EA \qquad \omega = \frac{\mu^2}{2\eta} \qquad (3)$$

Condensed Fukui functions for electrophilic attack come from Mulliken-type atomic populations q_k of the N and N + 1 electron species at that same geometry,

$$f^+(k) = q_k(N) - q_k(N+1) \qquad (4)$$

and are given alongside the fractional population of the lowest unoccupied molecular orbital (LUMO) on each atom. To anchor WFA on an external scale, six reference Michael acceptors (acrolein, methyl vinyl ketone, cyclohex-2-enone, methyl acrylate, acrylamide and *N*-ethylmaleimide) and an isolated C24-lactone fragment were computed by the same route.

2.5 DFT calibration

Truncated models were re-evaluated at M06-2X/def2-SVP [17] with density fitting and a domain-decomposition conductor-like screening (ddCOSMO) water continuum (ε = 78.3553) in PySCF 2.14.0 [18], run as single points on the GFN2-xTB geometry. Two separate decisions feed into this. The truncation routine is chosen by which calculation produced the source geometry, since conformer-ensemble adduct files and relaxed-scan trajectory files order their atoms differently; that is purely a file-format matter. Which atoms survive truncation is instead a chemical decision governed by the site: an A/B-ring-junction model for C3, C6 and C5, cutting the C6–C7 and C9–C10 bonds and capping the open valences with hydrogen link atoms, and a side-chain lactone model for C24. The calibration term is

$$\Delta\Delta E_{\mathrm{calib}} = \Delta E_{\mathrm{DFT}} - \Delta E_{\mathrm{xTB}} \qquad (5)$$

and serves to contextualise the full-molecule energetics rather than to rescale them after the fact. The same protocol was applied at six fixed points along the epoxide scan coordinate, which probes the shape of that profile independently of the saddle search.

The choice of level deserves comment, because the corrections it produces are several kcal mol^{-1} and they differ in sign between reaction classes. M06-2X was selected for its documented performance on main-group thermochemistry and on reactions involving charged organic species, which is the regime here. def2-SVP was selected on cost: the calibration spans six adducts and twelve scan points, and a larger basis was not affordable for the full set on the single core available. A split-valence basis of this size is, however, modest for quantifying energy differences of this magnitude, and basis-set incompleteness error does not cancel reliably between two reaction classes that differ as much as conjugate addition and strained-ring opening. We therefore repeated the calibration with def2-TZVP on the same geometries for the C3 and epoxide models, and report that comparison in Section 3.5. Every DFT number in this work is a single point on a GFN2-xTB geometry, so the corrections are estimates of a basis- and functional-dependent quantity rather than converged reaction energies.

2.6 Non-covalent pose sampling

Twelve WFA···GSH encounter complexes were constructed, six in random orientations and six directed at C3 from either face with initial S···C separations of 3.9–4.2 Å, then relaxed at the

same GFN2-xTB/ALPB(water) level to a force tolerance of 0.02 eV Å^{-1} inside 600 steps. The quantity reported is the electronic interaction energy,

$$\Delta E_{\mathrm{int}} = E(\mathrm{complex}) - E(\mathrm{WFA}) - E(\mathrm{GSH}) \qquad (6)$$

which contains no thermal correction and no explicit desolvation term, and is consequently not a binding free energy. Relaxed poses were screened for interfragment hydrogen bonding on a geometric criterion: donor–acceptor heavy-atom distance of 3.5 Å or less, hydrogen–acceptor distance of 2.5 Å or less, and a donor–hydrogen–acceptor angle of at least 120°.

2.7 Validation checks and software

Five checks were applied. The WFA structure reproduced the PubChem reference InChIKey exactly. Molecular formula and mass balance were verified for every adduct against the expected product before any thermochemistry. Each reported adduct minimum was confirmed free of imaginary frequencies, and every stationary point returned by a saddle search is quoted with its imaginary-mode count rather than assumed to be a transition state. The C6 relaxed scan was repeated in full after an interruption and reproduced the original 13-point profile to within 0.01 kcal mol^{-1} at every point, which establishes that the protocol is deterministic for a fixed starting geometry. Finally, each truncated DFT model was checked for electron-count parity and correct formula before submission, and the C3 calibration was re-derived independently when the calibration was extended to the remaining sites.

The calculations used RDKit 2026.03.6 [13], tblite 0.7.0 for the GFN2-xTB Hamiltonian and ALPB solvation, ASE 3.29.0 [14], Sella 2.6.0 [15] and PySCF 2.14.0 [18], all single-threaded, with no proprietary quantum-chemistry software at any stage. Fig. 1 summarises the workflow and Fig. 2 shows how geometries and energies move between the calculation stages; runtimes and hardware are listed in ESM_1, Section S2.

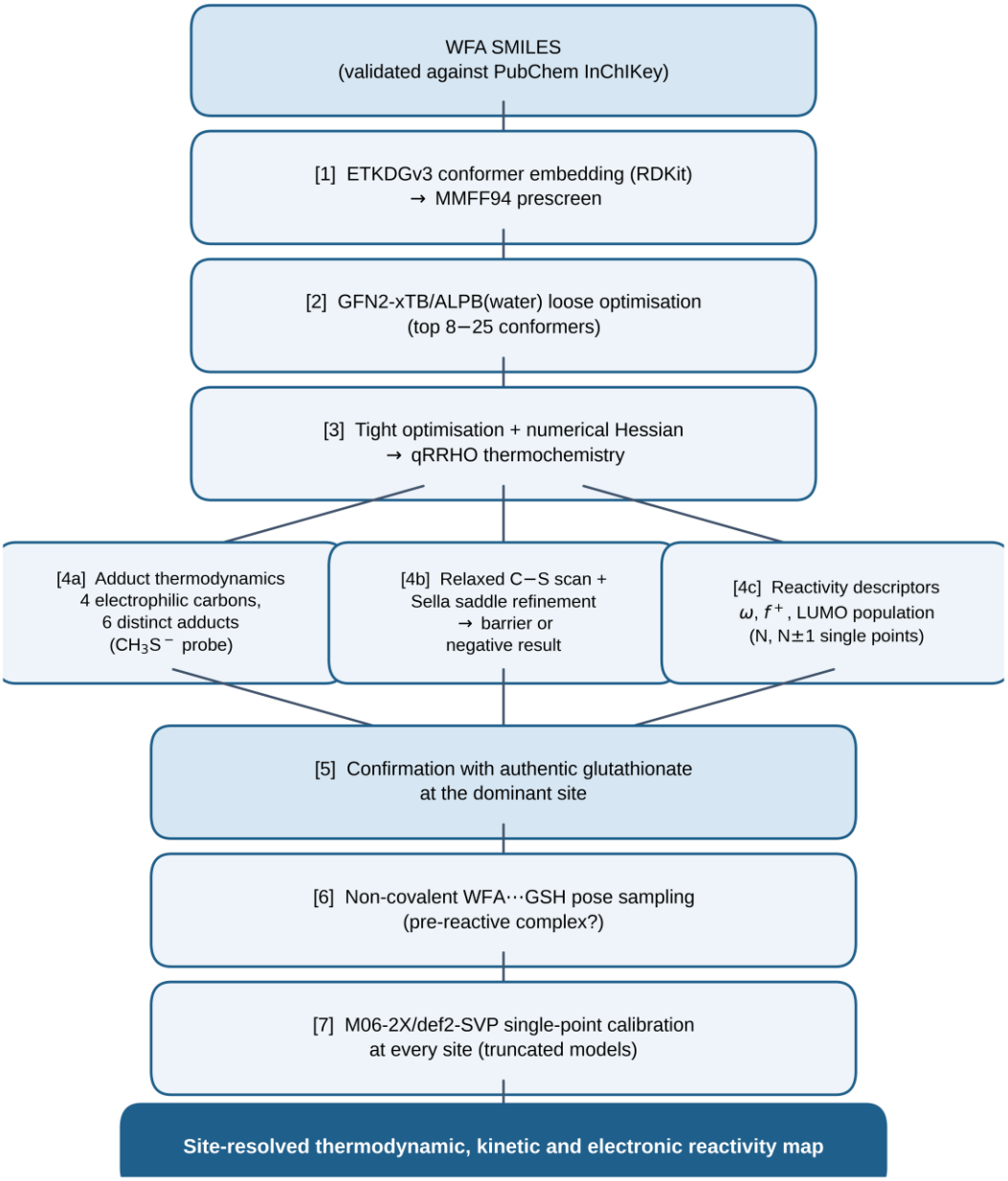


Fig. 1 Overall computational workflow, from structure validation through conformer-ensemble thermochemistry, reaction-coordinate scanning and reactivity descriptors to DFT recalibration

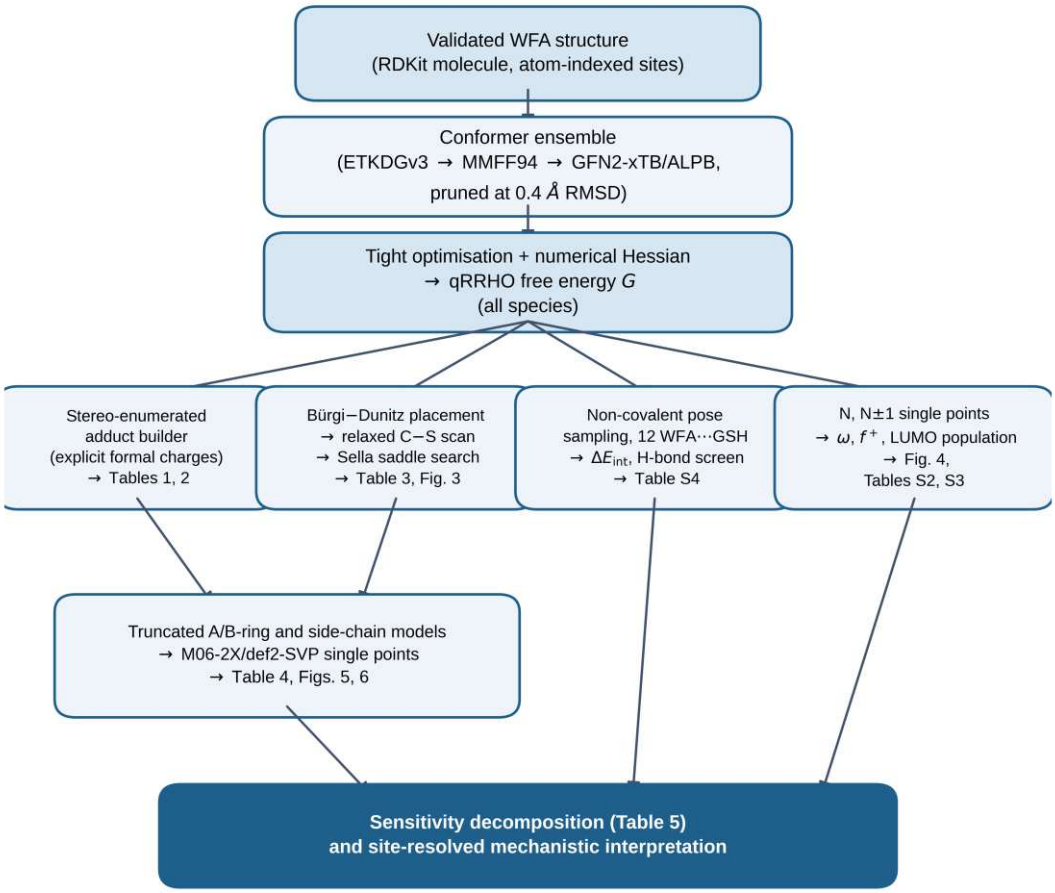


Fig. 2 Organisation of the calculation stages and the flow of geometries and energies between them

3 Results and discussion

3.1 Site-resolved thermodynamics with a methanethiolate probe

Adduct formation free energies for the minimal probe across all four electrophilic carbons are collected in Table 1. The C3 α-face Michael adduct is the most exergonic at −11.7 kcal mol⁻¹. Behind it come the two ring-opened epoxide adducts at −8.9 (C6) and −8.3 (C5) kcal mol⁻¹, then the C3 β-face adduct at −4.7. C24 trails well back, spanning +0.2 (R) to −4.4 (S) kcal mol⁻¹, and is the only site where one stereoisomer comes out net endergonic. None of the reported adducts has an imaginary frequency, so all are true local minima.

Table 1 Site-resolved adduct formation thermodynamics for the methanethiolate probe at the GFN2-xTB/ALPB(water) + qRRHO level. Bold indicates the most favourable entry within each site group

Site	Face or stereoisomer	ΔE (kcal mol⁻¹)	ΔG_{min} (kcal mol⁻¹)	ΔG_{ens} (kcal mol⁻¹)	n_{imag}
C3 (enone)	α	−25.72	−11.64	**−11.73**	0

Site	Face or stereoisomer	ΔE (kcal mol⁻¹)	ΔG_{min} (kcal mol⁻¹)	ΔG_{ens} (kcal mol⁻¹)	n_{imag}
C3 (enone)	β	−18.02	−4.51	−4.70	0
C6 (epoxide)	anti	−20.88	−8.78	**−8.90**	0
C5 (epoxide)	anti	−20.90	−8.07	−8.26	0
C24 (lactone)	R	−13.38	0.41	0.22	0
C24 (lactone)	S	−18.55	−4.35	**−4.36**	0

The 7.0 kcal mol⁻¹ α/β difference at C3 is consistent with the distinct steric environments of the two faces. Ring A is fused trans to ring B, and the angular C19 methyl occupies the β face. An α-face approach avoids this group and permits the new thioether to adopt a pseudo-equitorial position in a half-chair-like enolate; the β face approach is more crowded and favours a pseudo-axial arrangement. The preference is present in the electronic energy (7.7 kcal mol⁻¹) and remains similar after thermal and entropic corrections, which are 14.0 and 13.3 kcal mol⁻¹ for the two faces. This interpretation concerns the methanethiolate adducts and should not automatically be transferred to GSH. At C5 and C6, the calculated products differ mainly in which epoxide carbon bears sufur and which bears the alkoxide; their 0.6 kcal mol⁻¹ difference is too small to resolve confidently at this level.

Placing these numbers against the method's 2–6 kcal mol⁻¹ nominal accuracy, the 7–12 kcal mol⁻¹ that separates C3-α from the C24 stereoisomers is comfortably meaningful, as is the 7.0 kcal mol⁻¹ α/β preference within C3. The 2.8–3.4 kcal mol⁻¹ margin between C3-α and the epoxide adducts sits at the edge of the band. Anything under roughly 2 kcal mol⁻¹, the two epoxide carbons being the clearest case, is reported for completeness but should be read qualitatively.

3.2 Confirmation with authentic glutathionate

Whether the minimal probe transfers to the biologically relevant nucleophile was tested by recomputing both C3 diastereomeric adducts in full with authentic glutathionate (Table 2).

Table 2 Cross-validation of the C3 site prediction with authentic glutathionate

Nucleophile	ΔG (C3-α) (kcal mol⁻¹)	ΔG (C3-β) (kcal mol⁻¹)	ΔΔG (α − β) (kcal mol⁻¹)
CH_3S^- (probe)	−11.7	−4.7	−7.0
Glutathionate (authentic)	−5.6	−18.9	+13.3

Both nucleophiles predict strongly that C3 addition is strongly exergonic, but they disagree on the preferred face. The compact probe prefers α by 7.0 kcal mol⁻¹; glutathionate prefers β by 13.3 kcal mol⁻¹. The two glutathionate adducts were built from equal numbers of retained unique conformers, 16 apiece, both are true minima, and both show sensible lowest vibrational frequencies of 12–13 cm⁻¹. Nor is the reversal small enough to hide inside the method's 2–6 kcal

mol⁻¹ uncertainty: it amounts to 13.3 kcal mol⁻¹ outright and 20.3 kcal mol⁻¹ relative to the probe's ordering, even considering the greater errors that a species with a double charge attracts, this is excessive. Sterics offer the simplest reading. Methanethiolate is small enough to reach either face on a near-ideal trajectory, whereas glutathionate has to fit its γ-glutamyl and glycine arms against the steroid, and that is evidently far easier on the convex β face. The result cautions against inferring facial selectivity for a bulky biological nucleophile from a minimal thiolate alone.

3.3 Kinetics of the approach coordinate

Table 3 collects the outcomes of the relaxed scans and saddle searches, with the corresponding free-energy profiles in Fig. 3.

Table 3 Outcomes of the relaxed scan and saddle search at each site. The final column is the scan maximum of Eq. 2, not an activation free energy; no Hessian-verified first-order saddle was obtained at any site, so no $\Delta G^{\ddagger}$ is reported

Site	n_{imag} at located stationary point	Imaginary frequency (cm^{-1})	d(C–S) at scan maximum (Å)	ΔG_{scan} (kcal mol^{-1})
C3-α	no search performed	—	—	no interior maximum in scan range
C3-β	no search performed	—	—	no interior maximum in scan range
C24 (face +, R)	2 (second-order saddle)	−101.7	2.82	4.8
C24 (face −, S)	2 (second-order saddle)	−68.2	3.69	3.6
C6 (epoxide)	0 (minimum)	16.4	3.41	−7.7
C5 (epoxide)	0 (minimum)	11.9	3.13	−11.2

Neither C3 face has an interior maximum. The highest of the 13 sampled points is the first, at 3.0 Å, after which the energy falls away steadily as the sulfur approaches carbon. There is accordingly no maximum from which to seed a saddle search. This is reported as the absence of a maximum within the 3.0 and 1.8 Å scan range — rather than as a barrier of zero. The calculations therefore did not locate a barrier under this protocol; they do not establish that the true free-energy profile is barrierless.

At C24, both faces the search converged on a genuine stationary point, but one carrying two negative Hessian eigenvalues. A second-order saddle is not a transition state, so neither qualifies as a verified barrier, and the scan maxima of 4.8 and 3.6 kcal mol⁻¹ are reported in their place under the convention of Eq. 2. Those values describe maxima on constrained profiles only and cannot be interpreted as verified barriers.

At both epoxide carbons (C5 and C6) the optimisation slid past the scan maximum into a downstream minimum, leaving the scan maximum as the only available quantity. Here, the reported values come out negative, at −7.7 and −11.2 kcal mol^{-1}. This is the clearest illustration of why ΔG_{scan} must not be read as $\Delta G^{\ddagger}$: a negative value is meaningless as an activation free energy. The likeliest cause is the qRRHO correction, which when applied point by point along a constrained scan need not vary monotonically with the electronic energy, so a single scan point's free energy is not obliged to track its electronic energy. The GFN2-xTB free-energy profile at the epoxide sites is therefore unresolved rather than measured, and Section 3.6 examines the same coordinate by a different route. No Hessian-verified first-order saddle was located anywhere in this study, at any site.

The lack of an interior peak at C3 might be partially due to the solvent model used. ALPB models the solvent as a continuum, omitting the explicit water shell surrounding the anionic thiolate. Consequently, as the nucleophile nears, the desolvation penalty is not addressed along the coordinate. This approach can lead to an underestimation of barriers in anion-neutral reactions. In experiments, the non-enzymatic addition of GSH to chalcone-type acceptors takes minutes to hours at physiological pH [7], which contradicts the interpretation of the calculated downhill profile as having no barrier. Table 3 should be employed for qualitative comparisions within a common model rather than for predicting absolute rates. While the C3 profile does not show an interior peak, the epoxide scans reveal significant maxima; incorporating explicit desolvation might also raise the epoxide barriers.

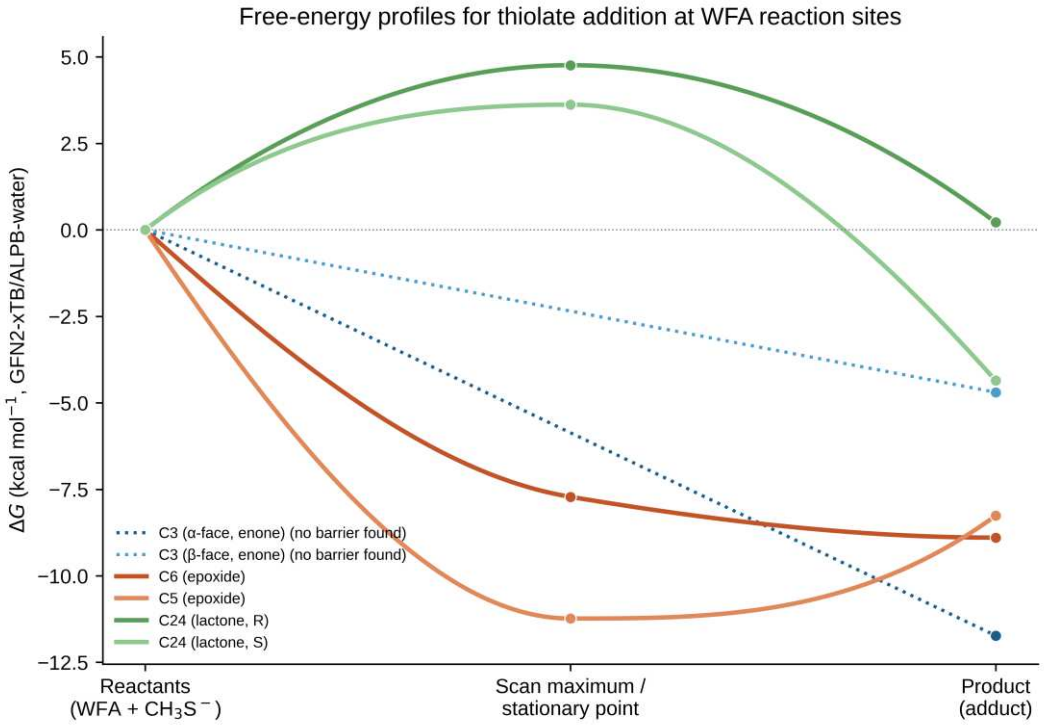


Fig. 3 Free-energy profiles for methanethiolate addition at the four candidate electrophilic carbons of WFA. The C3 traces (dotted) are monotonically downhill with no interior maximum; C24 shows a small apparent barrier at a second-order saddle; the epoxide points are scan-derived bounds discussed in Sections 3.3 and 3.6

3.4 Reactivity descriptors

WFA's intact ring-A enone returns an electrophilicity index of ω = 6.60 eV (Fig. 4a). That clears every simple reference acceptor tested by 2.1–2.8 eV, against acrolein at 4.53, methyl vinyl ketone at 4.40, cyclohex-2-enone at 4.36, methyl acrylate at 4.13 and acrylamide at 3.77 eV, and brings it close to *N*-ethylmaleimide at 6.03 eV, a bis-activated imide used as a fast and essentially irreversible thiol-alkylating reagent. The fused ring system rather than the enone group alone is responsible: the isolated C24-lactone fragment, computed identically, returns ω = 4.24 eV, which puts it among the acyclic references. This supplies an electronic-structure rationale for the 7–12 kcal mol^{-1} thermodynamic penalty that C24 carries in Table 1.

In hard–soft acid–base language the ring-A enone is firmly a soft electrophile, a good match for a soft and polarisable thiolate and consistent with WFA's proteome-wide preference for cysteine over harder nucleophiles such as lysine ε-amino groups [2]. It is a global ground-state property of the intact molecule, evaluated before any bond forms, so it ranks intrinsic electrophilic character while remaining blind both to the steric accessibility of a particular carbon and to the strain-release driving force at the epoxide. Its role in this work is as an independent, externally anchored check on the C3-versus-C24 ordering that the explicit energetics already give, not as a replacement for them.

Condensed Fukui analysis (Fig. 4b) puts the largest f^+ values at C3 and the ring-A carbonyl oxygen, with a secondary maximum at C24. Both epoxide carbons and the epoxide oxygen show negligible f^+ and near-zero LUMO population. That fits the chemistry: epoxide reactivity comes from σ*-mediated strain release on backside attack rather than frontier-orbital electrophilicity, a mechanistically separate pathway from the two Michael acceptors. Full descriptor and Fukui data appear in ESM_1, Tables S2 and S3.

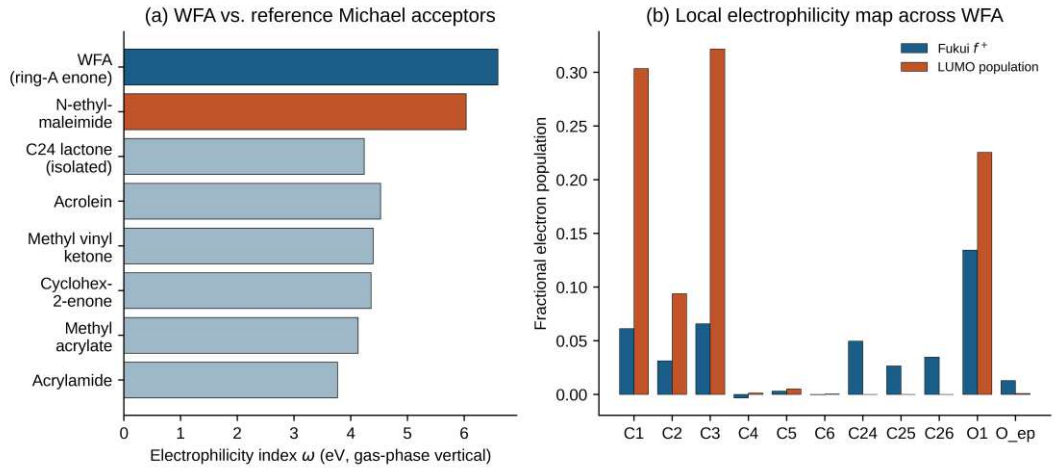


Fig. 4 a Gas-phase vertical electrophilicity index ω of WFA compared with six reference Michael acceptors and the isolated C24-lactone fragment. **b** Condensed Fukui f^+ and fractional LUMO population at each candidate electrophilic and adjacent atom of WFA

3.5 DFT calibration across all site types

Table 4 and Fig. 5 give the M06-2X/def2-SVP recalibration at every site, face and stereoisomer.

Table 4 Site-resolved DFT calibration of truncated-model reaction energies. All values are single-point M06-2X/def2-SVP on the GFN2-xTB/ALPB(water) geometry

Site (truncation model)	GFN2-xTB ΔE (kcal mol^{-1})	M06-2X/def2-SVP ΔE (kcal mol^{-1})	$\Delta\Delta E_{calib}$ (kcal mol^{-1})
C3 enone, α (A/B-ring)	−25.4	−13.2	+12.2
C3 enone, β (A/B-ring)	−19.9	−12.1	+7.8
C6 epoxide (A/B-ring)	−20.2	−27.6	−7.4
C5 epoxide (A/B-ring)	−20.1	−24.1	−4.0
C24 lactone, R (side-chain)	−13.9	−5.8	+8.1
C24 lactone, S (side-chain)	−18.3	−11.2	+7.1

At this stage, the correction does not exhibit a consistent bias. It is positive at both Michael-acceptor sites, with values reaching +7.8 and +12.2 kcal mol^{-1} at the two C3 faces, and ranging from +7.1 to +8.1 at C24. Conversely, it is negative at both epoxide carbons, measuring -7.4 and -4.0. This pattern, if taken literally, suggests a chemical interpretation, as conjugate addition and epoxide opening emphasise different aspects of a method: the former transforms a delocalised enone into a localised enolate, relying on the depiction of charge delocalization across the O-C1-C2 fragment while the latter changes a strained three-membered ring into an open alkoxide, depending on the accurate reproduction of ring strain in the reactant. However, the basis-set test outlined next indicates that this interpretation is unsustainable, and that the sign difference is primarily an aberration of def2-SVP.

Applying each site-specific correction to the full-molecule values places the epoxide adducts at −16.3 and −12.3 kcal mol^{-1}, below either C3 face at +0.5 and +3.1 and C24 stereoisomers at +2.7 and +8.4. The correction therefore reverses the uncorrected ranking. This result motivated the calibration at every candidate site rather than at only one representative reaction.

The reversed ranking depends on the calibration protocol and should not be treated as an independent prediction. The functional, basis set and single-point treatment can all affect the correction; none of the structures was optimised at the DFT level. Accordingly, the corrected values are described here as DFT-corrected estimates, not DFT-optimised reaction energies.

Basis-set sensitivity. Because def2-SVP is a modest basis for differences of this size, the calibration was repeated with def2-TZVP on the same geometries for the four adducts that carry the inversion. Table 5 gives both.

Table 5 Basis-set sensitivity of the calibration correction, in kcal mol^{-1}. All values are single points on the same GFN2-xTB geometry of the A/B-ring truncated model; only the basis set differs. Corr. is the correction $\Delta\Delta E_{calib}$ of Eq. 5; Est. applies it to that site's full-molecule value from Table 1

Site	xTB ΔE	SVP ΔE	TZVP ΔE	Corr. SVP	Corr. TZVP	Est. SVP	Est. TZVP

Site	xTB ΔE	SVP ΔE	TZVP ΔE	Corr. SVP	Corr. TZVP	Est. SVP	Est. TZVP
C3-α	−25.4	−13.2	−4.8	+12.2	+20.6	+0.5	+8.8
C3-β	−19.9	−12.1	−4.7	+7.8	+15.2	+3.1	+10.5
C6	−20.2	−27.6	−19.9	−7.4	+0.4	−16.3	−8.5
C5	−20.1	−24.1	−16.4	−4.0	+3.7	−12.3	−4.6

Increasing the basis set shifts each correction by +7.4 to +8.3 kcal mol⁻¹, a change comparable to the correction itself and at def2-TZVP all four corrections are positive. The sign difference between the reaction classes, and the mechanist interpretation based on that difference, therefore do not survive the basis set test. Absolute corrections from this protocol should be regarded as uncertain by several kcal mol⁻¹.

The basis set shift is also nearly uniform: its mean is +7.8 kcal mol⁻¹, with a spread of 0.9 kcal mol⁻¹ across four adducts. This common shift largely cancels in site-to-site comparisions. The ordering remains C6, then C5, then the C3 faces at both basis sets; the C6- minus-C3-α gap changes from 16.8 kcal mol⁻¹ to 17.3 at def2-TZVP. Thus, the reversed ranking is stable to this basis set change, even though the absolute corrections are not. The def2-SVP values are retained in Table 4 for consistency, with the estimates interpreted provisionally.

None of this removes the dependence on the functional or on the single-point protocol, and neither was tested. The thermodynamic ranking of these four sites should still be regarded as provisional pending full DFT optimisation of the adducts and their reactants.

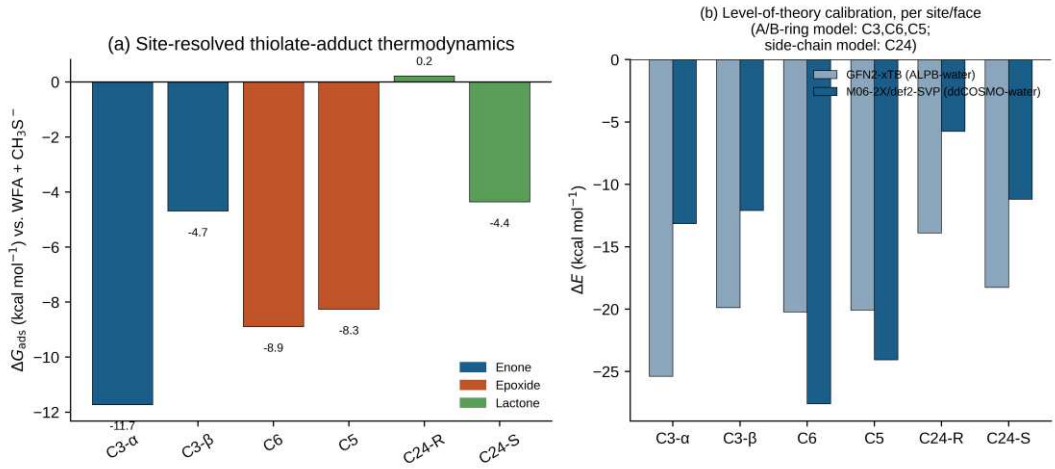


Fig. 5 a Site-resolved adduct formation free energies for the methanethiolate probe, coloured by electrophilic motif. **b** GFN2-xTB compared with M06-2X/def2-SVP single-point reaction energies for the truncated model at each site and face

3.6 Independent DFT test of the epoxide barrier

The GFN2-xTB scans did not provide usable free-energy barriers at either epoxide carbon (Section 3.3). Single-point M06-2X/def2-SVP calculations were run at six fixed points along the same relaxed-scan coordinate, at d(C–S) = 3.00, 2.30, 2.20, 2.10, 2.00 and 1.80 Å, and compared with GFN2-xTB at identical geometries (Fig. 6).

This comparison tests whether the rise along the sampled epoxide coordinate persists when the electronic-structure method changes while the geometry is held fixed. It cannot establish a transition state, but it can show whether the scan feature is unique to GFN2-xTB.

Both methods place the maximum at the same C–S distance, 2.10 Å, at both sites. The feature is therefore geometrically localised and reproduced across two very different electronic-structure treatments, which is what one expects of a real feature of the potential-energy surface associated with S_N2-like backside epoxide opening rather than a quirk of the semi-empirical Hamiltonian or of the scanning protocol. DFT also places the maximum higher, not lower: measured against each profile's own 3.00 Å point, +43.5 against +23.1 kcal mol^{-1} at C6, and +51.8 against +40.8 kcal mol^{-1} at C5.

These four numbers are single-point electronic energies at fixed scan geometries. They are not activation energies. They carry no thermal or entropic correction, the geometries were optimised at a different level from the one that evaluates them, the maximum is located only to the 0.1 Å resolution of the scan, and no saddle point was sought or found. What they establish is narrower than a barrier height but still useful: along this coordinate, at this geometry, both methods agree that the energy rises substantially before the adduct forms, and DFT makes the rise larger than GFN2-xTB does. A quantitative activation energy for epoxide opening would require a DFT-optimised, Hessian-verified transition state.

One further consistency check is available at the product-like end of the C6 scan, where the DFT-minus-xTB gap of −9.9 kcal mol^{-1} at 1.80 Å matches in sign and rough magnitude the independently derived full-adduct correction of −7.4 kcal mol^{-1} in Table 4. The two come from different geometries, so exact agreement was never expected.

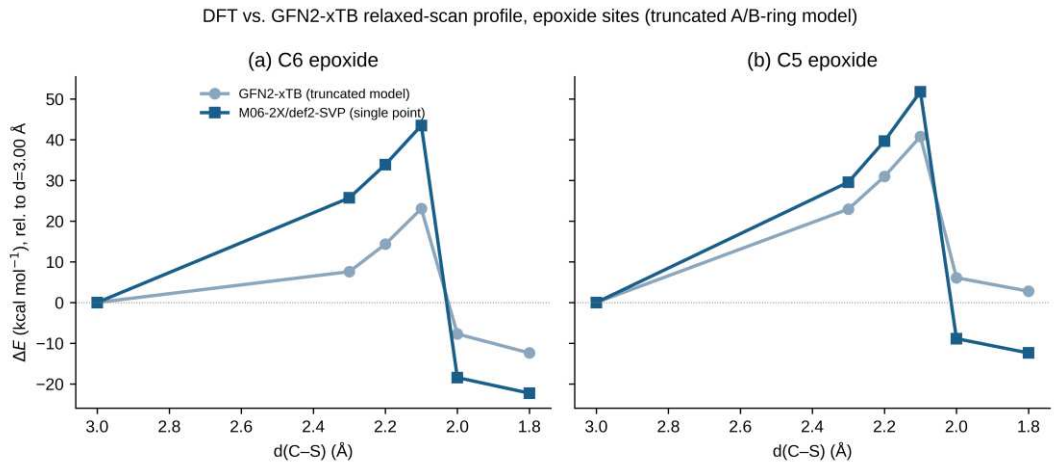


Fig. 6 GFN2-xTB compared with M06-2X/def2-SVP single-point electronic energies along the same relaxed C–S approach coordinate for the truncated A/B-ring model at **a** the C6 and **b** the C5 epoxide carbons, each referenced to its own 3.00 Å point

3.7 Sensitivity to methodological choices

Table 6 breaks down how each adduct free energy responds to the thermal correction, the conformational average, the entropy treatment and the level of theory.

Table 6 Component-wise sensitivity of ΔG_{ads} (kcal mol^{-1}). The final column applies each site's own correction from Table 4 to that site's full-pipeline value and is a DFT-corrected estimate, not a DFT-optimised reaction energy

Adduct	Full pipeline ΔG_{ens}	Bare ΔE	ΔG_{min} only	Harmonic entropy only	DFT-corrected estimate
C3 enone, α	−11.73	−25.72	−11.64	−15.90	+0.5
C3 enone, β	−4.70	−18.02	−4.51	−9.05	+3.1
C6 epoxide	−8.90	−20.88	−8.78	−13.01	−16.3
C5 epoxide	−8.26	−20.90	−8.07	−11.98	−12.3
C24 lactone, R	0.22	−13.38	0.41	−3.77	+8.4
C24 lactone, S	−4.36	−18.55	−4.35	−8.46	+2.7

The thermal and entropic correction is large, at 13–16 kcal mol^{-1}, but it varies little between sites and so leaves relative rankings intact. Conformational averaging beyond the single lowest conformer moves ΔG by less than 0.3 kcal mol^{-1} anywhere, which suggests the conformer protocol is adequate for this system. Replacing qRRHO with a purely harmonic entropy shifts every value by 4–5 kcal mol^{-1} in the same direction. These tests describe sensitivity within the GFN2-xTB protocol; they do not address uncertainty in the DFT calibration.

Among the four components examined, then, only the level of theory changes the ordering. The thermal, conformational and entropy choices are internal to the GFN2-xTB workflow and had limited effects on the ranking. By contrast, the DFT corrections come from one single-point protocol and determine the revised order. The sensitivity analysis therefore locates the uncertainty rather than removing it, and it puts that uncertainty squarely on the calibration.

3.8 Non-covalent encounter complexes

Eleven of the twelve sampled WFA···GSH poses converged within the specified tolerance. The twelfth, which as it happens formed two geometric hydrogen bonds, did not converge and is left out of the summary statistics, though its geometry is retained in ESM_1. Over the eleven converged poses ΔE_{int} ranges from −11.3 to +8.2 kcal mol^{-1} with a mean of −4.1, and not one satisfies the hydrogen-bond criterion of Section 2.6. The six directed poses relax to final S···C3 separations between 1.95 and 4.17 Å; the three closest are plausibly sampling the early attractive stretch of the same approach coordinate examined in Section 3.3 rather than any distinct encounter minimum. Per-pose values are listed in ESM_1, Table S4.

The underlying question is whether WFA and GSH associate in a defined, bound orientation before the covalent step, with a preferred orientation that could influence facial or regiochemical selectivity. These data do not support such a complex, and they do not rule one out either. Interaction energies are widely spread and mixed in sign, the most favourable poses are not the ones sitting closest to the electrophilic carbon, and no pose shows the interfragment hydrogen bonding that would hold a specific orientation in place.

The pose screen is inconclusive because it uses electronic energies without thermal corrections or explicit desolvation and includes only twelve locally relaxed starting poses. It shows only that this sampling protocol did not identify a well-defined, stable encounter complex. The results cannot establish that association is non-directional or that no complex exists. Broader sampling and an explicit-solvent free energy treatment would be needed to assess geometry-dependent interactions and hydrogen bonding.

3.9 Mechanistic interpretation and an experimental test

This approach scans and corrected product energies give different indications. C3 is the only site without an interior maximum along the scanned coordinate, whereas the DFT-corrected estimates favour the epoxides over C3 by approximately 13–19 kcal mol^{-1}. This pattern is consistent with kinetic rather than thermodynamic control of the non-enzymatic reactions: C3 may be more accessible even though the calculated epoxide products are lower in energy.

This is an interpretation, not a demonstration of kinetic control. A quantitative comparison would require activation free energies at C3 and the epoxide sites, but none is available. The absence of a C3 scan maximum is a result from one method and solvent model, not evidence of a zero barrier; the epoxide maxima are single points on unoptimised geometries, not barrier heights. The calculations establish a qualitative difference between the sampled approach profiles, opposite to the corrected thermodynamic ordering, DFT-optimised, Hessian-verified transition states at both site types are needed to test the interpretation.

The C6 connectivity reported for enzymatically produced CR-777 [6] is compatible with this interpretation but does not validate it. GST could alter thiol ionization, orient the reactants or stabilize a transition state in ways not represented in the present calculations. Because the enzymatic reaction lies outside the model, no direct conclusion about its mechanism follows from these results.

The prediction is directly testable, and the discriminating experiment is regiochemical rather than mass-based. C3 and C6 conjugates are constitutional isomers of identical elemental composition, so both appear at *m/z* 778.4 $[M+H]^+$ (monoisotopic 777.35 Da = WFA 470.27 + GSH 307.08) and the precursor mass assigns neither. Three observables separate them. In ^1H NMR, C3 addition saturates the 2,3-en-1-one and removes the characteristic H2/H3 olefinic doublets near δ 6.2 and 6.7 ppm, replacing them with upfield aliphatic resonances while leaving the H6 epoxide proton near δ 3.2 ppm untouched; C6 addition inverts that pattern. In the ultraviolet, C3 addition destroys the enone chromophore at 225–230 nm while C6 addition preserves it. In tandem mass spectrometry both regioisomers lose the diagnostic glutathione neutral fragments at −129 Da (pyroglutamate) and −75 Da (glycine), but third-stage fragmentation (MS^3) of the steroid core distinguishes them, since retro-Diels–Alder cleavage of ring A carries the thioether substituent only in the C3 conjugate.

A proposed experiment would incubate WFA with excess GSH in aqueous buffer at pH 7.4 and ambient temperature, monitor absorbance at 228 nm, and quench aliquots for chromatographic separation. Loss of the 228 nm band would track consumption of the C3 enone, while NMR and MS would be used to assign the isolated conjugate. A time course could also test the proposed reversibility difference: if the C3 Michael adduct undergoes retro-Michael elimination while the C6 epoxide-derived adduct does not, a product mixture initially enriched in C3 that shifts toward

C6 over hours or days would support the kinetic control interpretation. This outcome remains a prediction to be tested, not an observed result.

On these criteria, the non-enzymatic reaction of WFA with GSH in buffer would yield mainly the C3 conjugate. A C6 product would remain compatible with the results; the corrected thermodynamics in Section 3.5 make it a plausible minor or slow-forming component.

3.10 Limitations

Five limitations bear on how these numbers should be used, and the first two are the ones that constrain the central claims.

Every DFT value quoted is a single point on a GFN2-xTB geometry of a truncated model, not a DFT-level optimisation. Consequently, the corrections estimate a quantity that depends on the functional, the basis set and the geometry at which it is evaluated. The basis-set test in Section 3.5 quantifies one of those dependencies, since the def2-SVP to def2-TZVP shift is large but nearly uniform and so cancels the site ranking. It does not test functional dependence or geometry relaxation. Because the functional was not varied and no DFT geometry optimisation was performed, the reversed thermodynamic ordering remains provisional until reactants and adducts are optimised at DFT level.

No Hessian-verified first-order saddle was located at any site, so no activation free energy is reported. Every quantity in Table 3 is a maximum on a constrained scan, sampled at 0.1 Å resolution, and is negative in two cases, and the DFT epoxide values in Section 3.6 are single points at fixed geometries. The kinetic-control reading in Section 3.9 rests on a qualitative contrast between two profiles rather than on comparison of verified barriers, and is offered on that basis.

The conformer ensembles keep only three or four Hessian-verified geometries from an initial embedding of 30–120. Agreement between ensemble and minimum values in Table 6 suggests this suffices, but an unsampled lower-energy conformer cannot be excluded, least of all for the flexible tripeptide, and no multi-seed statistical replication was attempted.

ALPB implicit solvation under-describes short-range, orientation-specific hydrogen bonding and provides no explicit desolvation penalty for an approaching anion. This bears on the pose sampling of Section 3.8, on the absence of a C3 scan maximum discussed in Section 3.3, and more loosely on the glutathionate face preference of Section 3.2.

Finally, the thiolate model supports conclusions about intrinsic relative reactivity between sites. No absolute rate is predicted, the speciation weighting described in Section 2.1 is not applied, and nothing here extends to the GST-catalysed reaction.

4 Conclusions

This study compared thiolate addition to the four electrophilic carbons of withaferin A using GFN2-xTB/ALPB(water) conformer-ensemble thermochemistry with qRRHO corrections, relaxed C–S approach scans, reactivity descriptors and M06-2X/def2-SVP recalibration at every site.

At the uncorrected semi-empirical level, C3 is the most favourable site, at $\Delta G = -11.7$ kcal mol^{-1} for the α face, followed by C6 at −8.9 and C5 at −8.3, and well clear of the C24 lactone at +0.2 to −4.4. Applying site-specific M06-2X/def2-SVP corrections inverts that ordering, placing the epoxide adducts at −16.3 and −12.3 kcal mol^{-1} while both C3 faces are slightly endergonic. Repeating the calibration with def2-TZVP moves every correction by about +7.8 kcal mol^{-1}, enough to remove the difference in sign between the two reaction classes and with it any mechanistic reading of that difference. Because the shift is nearly uniform, however, the site ranking is unchanged at both basis sets, and the gap between C6 and C3-α changes only from 16.8 to 17.3 kcal mol^{-1}. The inverted ordering is therefore robust to basis set while the absolute corrections behind it are not, and we present it as a DFT-corrected estimate. Functional dependence and the single-point protocol remain untested, so the thermodynamic ranking should be treated as provisional until the adducts are optimised at the DFT level.

The kinetic evidence is less definitive. No Hessian-verified first-order saddle was located at any site, so no activation free energy can be reported. The C3 approach coordinate has no interior maximum at either face, while at the epoxide carbons (C5 and C6) both GFN2-xTB and M06-2X place a pronounced maximum at the same C–S distance, 2.10 Å, with DFT the higher of the two at +43.5 (C6) and +51.8 (C5) kcal mol^{-1} relative to the 3.00 Å point. These are single-point scan maxima on unoptimised geometries, not activation barriers. The contrast between the C3 and epoxide profiles is consistent with kinetic control of the non-enzymatic reaction, but it does not establish it. Verified transition states at both site types remain necessary.

Methanethiolate reproduced glutathionate's site preference but not its facial preference: it favoured α over β by 7.0 kcal mol^{-1}, whereas glutathioate favoured β over α by 13.3. Since minimal-thiol probes are what make exhaustive site and face screens affordable, that failure is worth weighing in any study that relies on one.

The calculations also illustrate two broader methodological points. First, the DFT corrections differ between Michael addition and epoxide opening, so calibrating a semi-empirical method at one motif may not transfer to another motif in the same molecule. Multi-site electrophiles require site-appropriate calibration. Second, the probe-tripeptide comparison in Section 3.2 suggests that a minimal thiolate can reproduce site ranking while failing to predict facial preference. Authentic nucleophile calculations are therefore important when facial selectivity is a key outcome.

WFA's electrophilicity index 6.60 eV is close to *N*-ethylmaleimide and far above simple enones, which fits its behaviour as a competent cysteine alkylator. Further work should prioritise a fully DFT-optimised transition state for the epoxide pathway, an explicit-solvent treatment of the encounter complex, and the regiochemical experiment set out in Section 3.9, which would separate the predicted C3 conjugate from the known enzymatic C6 product on NMR, ultraviolet and MS^3 criteria rather than on precursor mass.

Acknowledgements

The authors used Claude (Anthropic) to draft and edit the text of this manuscript and its Supplementary Information, and to assist in preparing the analysis scripts and figures. The authors ran and checked all calculations, verified the scientific content, interpreted the results, and take full responsibility for the work in its published form.

Declarations

Funding This research received no external funding.

Competing interests The authors declare no competing financial or non-financial interests.

Ethics approval Not applicable. This study is entirely computational and involved no human participants, animals or biological materials.

Data availability All input structures, optimised geometries, absolute energies, raw calculations, analysis scripts and figure-generation code underlying every table and figure are provided as Supplementary Information (ESM_1). The pipeline depends only on open-source software.

Author contributions Bhargava Reddy: conceptualisation, methodology, software, investigation, formal analysis, writing – original draft. Hari Prasad: methodology, validation, writing – review and editing. P.V.Suresh: conceptualization, supervision, resources, writing – review and editing. All authors read and approved the final manuscript.

References

1. Bailly C. Covalent binding of withanolides to cysteines of protein targets. Biochem Pharmacol. 2024;226:116405. https://doi.org/10.1016/j.bcp.2024.116405

2. Nie HJ, Fu YJ, Long S, Wang JY, Zhao WS, Zhai LH, Yang YL, Chen XH. Chemoproteomics reveals proteome-wide covalent and non-covalent targets of withaferin A. Acta Pharmacol Sin. 2025;46:1782–93. https://doi.org/10.1038/s41401-024-01468-5

3. Bargagna-Mohan P, Hamza A, Kim YE, Khuan Abby Ho Y, Mor-Vaknin N, Wendschlag N, Liu J, Evans RM, Markovitz DM, Zhan CG, Kim KB, Mohan R. The tumor inhibitor and antiangiogenic agent withaferin A targets the intermediate filament protein vimentin. Chem Biol. 2007;14:623–34. https://doi.org/10.1016/j.chembiol.2007.04.010

4. Bargagna-Mohan P, Deokule SP, Thompson K, Wizeman J, Srinivasan C, Vooturi S, Wendschlag N, Liu J, Mohan R. Withaferin A effectively targets soluble vimentin in the glaucoma filtration surgical model of fibrosis. PLoS One. 2013;8:e63881. https://doi.org/10.1371/journal.pone.0063881

5. Nicholas AW, Rosazza JP. Reactions of withaferin-A with model biological nucleophiles. Bioorg Chem. 1976;5:367–72. https://doi.org/10.1016/0045-2068(76)90021-3

6. Rabhi C, Arcile G, Le Goff G, Da Costa Noble C, Ouazzani J. Neuroprotective effect of CR-777, a glutathione derivative of withaferin A, obtained through the bioconversion of Withania somnifera (L.) Dunal extract by the fungus Beauveria bassiana. Molecules. 2019;24:4599. https://doi.org/10.3390/molecules24244599

7. Kenari F, Pintér Z, Molnár S, Borges ID, Camargo AJ, Napolitano HB, Perjési P. (E)-2-benzylidenecyclanones: part XIX. Reaction of (E)-2-(4′-X-benzylidene)-1-tetralones with cellular thiols: comparison of thiol reactivities of open-chain chalcones and their six- and seven-membered cyclic analogs. Int J Mol Sci. 2024;25:7773. https://doi.org/10.3390/ijms25147773

8. Bannwarth C, Ehlert S, Grimme S. GFN2-xTB—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. J Chem Theory Comput. 2019;15:1652–71. https://doi.org/10.1021/acs.jctc.8b01176

9. Ehlert S, Stahn M, Spicher S, Grimme S. Robust and efficient implicit solvation model for fast semiempirical methods. J Chem Theory Comput. 2021;17:4250–61. https://doi.org/10.1021/acs.jctc.1c00471

10. Bedart C, Vergoten G, Bailly C. Withangulatin A identified as a covalent binder to Zap70 kinase by molecular docking. Computation. 2025;13:207. https://doi.org/10.3390/computation13090207

11. Grimme S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. Chem Eur J. 2012;18:9955–64. https://doi.org/10.1002/chem.201200497

12. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE. PubChem 2023 update. Nucleic Acids Res. 2023;51:D1373–80. https://doi.org/10.1093/nar/gkac956

13. RDKit: open-source cheminformatics software, version 2026.03.6. https://www.rdkit.org. Accessed 28 Sep 2026.

14. Larsen AH, Mortensen JJ, Blomqvist J, Castelli IE, Christensen R, Dułak M, Friis J, Groves MN, Hammer B, Hargus C, Hermes ED, Jennings PC, Jensen PB, Kermode J, Kitchin JR, Kolsbjerg EL, Kubal J, Kaasbjerg K, Lysgaard S, Maronsson JB, Maxson T, Olsen T, Pastewka L, Peterson A, Rostgaard C, Schiøtz J, Schütt O, Strange M, Thygesen KS, Vegge T, Vilhelmsen L, Walter M, Zeng Z, Jacobsen KW. The atomic simulation environment—a Python library for working with atoms. J Phys Condens Matter. 2017;29:273002. https://doi.org/10.1088/1361-648X/aa680e

15. Hermes ED, Sargsyan K, Najm HN, Zádor J. Sella, an open-source automation-friendly molecular saddle point optimizer. J Chem Theory Comput. 2022;18:6974–88. https://doi.org/10.1021/acs.jctc.2c00395

16. Parr RG, Szentpály LV, Liu S. Electrophilicity index. J Am Chem Soc. 1999;121:1922–4. https://doi.org/10.1021/ja983494x

17. Zhao Y, Truhlar DG. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements. Theor Chem Acc. 2008;120:215–41. https://doi.org/10.1007/s00214-007-0310-x

18. Sun Q, Zhang X, Banerjee S, Bao P, Barbry M, Blunt NS, Bogdanov NA, Booth GH, Chen J, Cui ZH, Eriksen JJ, Gao Y, Guo S, Hermann J, Hermes MR, Koh K, Koval P, Lehtola S, Li Z, Liu J, Mardirossian N, McClain JD, Motta M, Mussard B, Pham HQ, Pulkin A, Purwanto W, Robinson PJ, Ronca E, Sayfutyarova ER, Scheurer M, Schurkus HF, Smith JET, Sun C, Sun SN, Upadhyay S, Wagner LK, Wang X, White A, Whitfield

JD, Williamson MJ, Wouters S, Yang J, Yu JM, Zhu T, Berkelbach TC, Sharma S, Sokolov AY, Chan GKL. Recent developments in the PySCF program package. J Chem Phys. 2020;153:024109. https://doi.org/10.1063/5.0006074

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.docx