CCS CHINESE CHEMICAL SOCIETY · SIOC 1950 · ROYAL SOCIETY OF CHEMISTRY

RESEARCH ARTICLE

A unified predictive model for chiroptical sensing: a substrate-centric approach to predicting circular dichroism outputs across two chemically distinct organic classes

Lorenzo Cracco, [ID] †‡ Roberto Penasa, [ID] † Paolo Zardi, [ID] § Giulia Licini, [ID] Manuel Orlandi [ID] and Cristiano Zonta [ID] *

This study presents an innovative approach for rapid and reliable determination of circular dichroism outputs using a chiroptical supramolecular sensor based on an oxo-vanadium(v) complex. The research focuses on developing a unified predictive model capable of analyzing two distinct organic classes of chiral substrates: amides and carboxylates. Using a diverse library of forty-one substrates, molecular descriptors were calculated exclusively from the free-substrate structures, without considering the host–guest complex. This approach allowed for the construction of robust statistical models that correlate structural and electronic features of the substrates with the intensity of the induced circular dichroism (CD) signal. The resulting global model, based on only three terms, demonstrates good predictive capability for both substrate classes. This approach eliminates the need for individual calibration curves for each analyte, representing a significant step towards a universal platform for enantiomeric excess (ee) determination. This methodology opens new perspectives for high-throughput chiral analysis, with potential applications in fields such as asymmetric catalysis and drug discovery.

Introduction

High-throughput analysis (HTA) of stereochemical features is crucial especially in fields such as asymmetric catalysis and parallel synthesis for drug discovery, where large libraries of chiral molecules must be evaluated efficiently. Chromatographic methods, while accurate, often represent a significant bottleneck due to the time and resources required per analysis. Even with advances in chromatographic techniques,[1,2] there is an increasing interest in alternative strategies that can deliver rapid, cost-effective, and scalable results. In this context, optical spectroscopies, including UV-vis absorption and fluorescence, are promising for HTA applications.[3–5] Among these, chiroptical spectroscopies offer the advantage of direct enantioselective readouts without the need for derivatization or chiral stationary phases. Techniques such as electronic circular dichroism (ECD),[6–9] fluorescence-detected CD (FDCD),[10,11] Raman optical activity (ROA),[12] circularly polarized luminescence (CPL),[13,14] and vibrational CD (VCD)[15,16] have all gained attention. Between these choices, CD remains the most accessible and widely applied due to its simplicity and instrument availability. However, successful application of CD often requires the presence of strong chromophores.[17] This limitation has stimulated the development of supramolecular sensors designed to translate the chirality of an analyte into an observable CD signal.[8,9,18]

A particularly effective design strategy involves stereodynamic sensors, which are chromophoric systems that undergo rapid racemization and adopt a preferred stereochemical form upon binding a chiral guest.[19] Such sensors amplify the chiroptical signal through conformational bias, enabling sensitive and selective ee assessment. Despite their promising features, widespread adoption of chiroptical sensors in HTA is limited by the need to construct separate calibration curve for each analyte. This limitation introduces two main bottlenecks: (i) the necessity of a pure (or of known ee) reference compound and (ii) the time required to generate the corresponding calibration data. Recently, advances in computational chemistry and machine learning have begun to address these challenges.[20] Notably, Anslyn and co-workers have pioneered approaches that predict CD responses using molecular descriptors derived from probe-analyte adducts, enabling the

Department of Chemical Sciences, University of Padua, Via Marzolo 1, 35131 Padua, Italy. E-mail: cristiano.zonta@unipd.it

†These authors contributed equally to this work.

‡Present address: Université Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71 Avenue des Martyrs, 38044, Grenoble, France.

§Present address: Department of Chemical and Geological Sciences, University of Modena and Reggio Emilia, Via Campi 103, 41125, Modena, Italy.

118 | *Org. Chem. Front.*, 2026, **13**, 118–123

This journal is © the Partner Organisations 2026
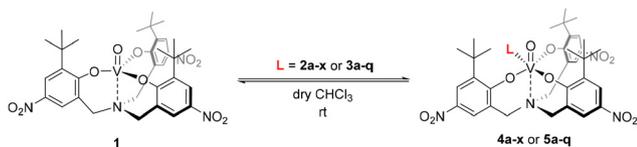
Scheme 1 Recognition process between the oxo-vanadium(v) amino-trisphenolate complex **1** and chiral Lewis bases.
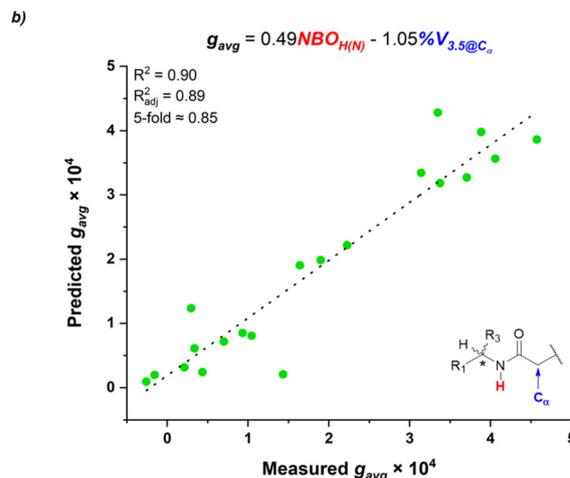

Fig. 1 (a) Enantiopure amides **2a–x** tested. (b) MLR modeling of g_{avg} of the amide adducts with the pivotal steric (blue) and electronic (red) molecular descriptors. The model was build using data of adducts **4a–t**.

theoretical construction of calibration curves.[21–23] This development marks a significant shift toward data-driven chiroptical sensing strategies.

Building on our group's recent work with stereodynamic metal-based sensors, we focused on vanadium complexes bearing tetradentate aminotrisphenolate ligands.[24] These complexes adopt a propeller-like conformation and exhibit enhanced chiroptical responses upon binding chiral Lewis bases. In particular, we reported an oxo-vanadium sensor that generates strong CD signals upon coordination (Scheme 1), including a UV-vis shift that enables concentration-independent ee determination *via* the anisotropy *g*-factor (the molar absorption coefficient, ε, is about 1.75×10^4 L mol^{-1} cm^{-1} at 450 nm, which corresponds to the absorption band that redshifts at 600 nm upon coordination of a Lewis base).[25]

Here, we present a novel predictive methodology, applicable to enantiopurity assessment, that operates across two chemically distinct classes of substrates, namely amides and carboxylates. Using a multiple linear regression (MLR) approach based on molecular descriptors computed solely from the substrate structures (*viz.* excluding the sensor), a unified calibration-free prediction model for CD response was constructed, offering a practical pathway toward universal HTA-compatible chiral probes.

Results and discussion

Amide model

To develop a theoretical model capable of predicting CD outputs, CD spectra of a structurally diverse library comprising twenty amides (Fig. 1a) in presence of the oxo-vanadium probe **1** were investigated. This library was carefully designed to provide a comprehensive representation of steric and electronic variations, thereby enabling a systematic evaluation of their influence on CD induction (see section S3 for synthetic details). As expected, the recorded CD spectra exhibited variable intensities depending on the nature of the substituents on the amide framework. For example, in the progression from amide (*R*)-**2a** to (*R*)-**2c**, where R$_1$ remains a phenyl group while R$_3$ increases in steric bulk from methyl to *tert*-butyl, a pronounced attenuation of the CD signal was observed (Fig. S73). To quantitatively analyse these variations, each compound underwent conformational analysis and geometry optimization, followed by computational parametrization to generate a set of thirty-five molecular descriptors (Table S3). These molecular descriptors were selected to comprehensively capture

both steric and electronic features of the amide. Steric effects were considered using Sterimol parameters and percent buried volumes, while electronic properties included both global descriptors (*e.g.*, HOMO–LUMO energy levels) and local electronic attributes such as natural bond orbital (NBO) charges on key atoms (see section S6.3 for a complete description of the employed parameters).[26] The chosen experimental observable was the average *g*-factor in the range 595–605 nm (g_{avg}). Because some adducts exhibited slightly noisy spectra, this average value was introduced to account for potential signal fluctuations at the wavelength of interest (600 nm). Earlier, it was observed that this is not the region with the best absorption of radiation, which would be at 340 nm (*e.g.*, $\varepsilon \approx 2.80 \times 10^4$ L mol^{-1} cm^{-1} upon coordination of (*R*)-**2a**), but it is the one where the free form of sensor **1** does not absorb and there is still good absorption by the adduct (*e.g.*, $\varepsilon \approx 1.45 \times 10^4$ L mol^{-1} cm^{-1} upon coordination of (*R*)-**2a**), enabling concentration-independent determination of ee.[25] To correlate g_{avg} with the structural features of the analytes, a statistical modelling protocol was applied (see section S6.4), obtaining a two-term statistical model (Fig. 1b). The first term represents an electronic parameter characterizing the nature of the amide proton (NBO$_{H(N)}$); the second term is a steric parameter, specifically the buried volume centred on the C$_\alpha$ atom relative

This journal is © the Partner Organisations 2026

Org. Chem. Front., 2026, **13**, 118–123 | **119**

to the carbonyl group ($\%V_{3.5@C_\alpha}$). Notably, as also highlighted by the relative magnitude of the parameters' coefficients, the steric parameter forms the core of the model, reflecting the spatial constraint imposed by substituents proximal to the carbonyl group, whereas the electronic term primarily modulates outliers within the amide-based dataset, bringing them closer to the main data points (Fig. S85 and S86). Indeed, analysis of the univariate correlation between $\%V_{3.5@C_\alpha}$ and g_{avg}, which is the dominant parameter in the model, revealed that the modulated outliers are all amides exhibiting a t-butyl or benzyl substituent on C_α (*i.e.*, (R)-**2d**, (R)-**2f**, (R)-**2r**, and (R)-**2s**) that should correlate with more or less half of the actually-displayed chiroptical signal in order to fall in the linear trend (Fig. S89). The model demonstrated strong predictive capability, with a coefficient of determination (R^2) of 0.90 and robust internal validation (5-fold = 0.85), where the label "5-fold" indicates the Q^2 obtained from k-fold cross-validation ($k = 5$). Furthermore, the high parameter-to-substrate ratio underscores the critical role of these identified features in determining CD intensity.

Carboxylate model

An extension of this predictive modelling approach was attempted for a second class of analytes, specifically carboxylic acids (Fig. 2a). These were employed in the presence of a non-coordinating base (Et_3N), to generate the corresponding anionic carboxylates. As with the amide series, a structurally variegated library of fifteen carboxylic acids was selected to probe the influence of steric and electronic factors on the induced CD response. For this series, thirty-two molecular descriptors (Table S4) were computed from geometry-optimized structures, drawing largely from the same parameters space employed in the amide analysis, with adjustments to account for the different coordination mode of the carboxylate group. Application of the multiple linear regression once again yielded a two-term model (Fig. 2b), though in contrast to the amide case, both terms in the carboxylate model correspond to electronic features. The first parameter corresponds to the IR stretching frequency of the carbonyl C=O bond, while the second represents the energy of the HOMO−3 orbital. As observed in the amide model, one descriptor contributes to the primary trend, while the secondary term improves model performance by accounting for the behaviour of outliers (Fig. S87 and S88), but a full understanding of the underlying reason remained elusive for the latter. Also in this situation, the regression R^2 parameter and the 5-fold cross-validation measure showed good values of 0.86 and 0.78 respectively.

Examining the coefficients in the two models, it is clear that a strong relationship exists between steric hindrance at C_α and g_{avg} for the amides, and between carbonyl stretching frequency and g_{avg} for the carboxylates. This is probably related to stereoinduction in the tripodal ligand, as showed also by the crystal structure of the adduct.[25] For the amides, the bulkiness of the substituent on C_α is an encumbrance towards the two adjacent t-butyl groups on top of the probe's helix and with the oxo group on the vanadium atom. This is reflected on the intensity of the optical signal. As an example, (R)-**2e** bears a phenyl substituent on C_α and the adduct exhibits a small g_{avg} due to a low energy differentiation between the two diastereomeric structures, since both are sterically hindered on top of the helix, as opposed to the high value of the adduct given by (R)-**2a** which displays a smaller methyl substituent. For the carboxylates, an extensive explanation is less straightforward. Generally, it was observed that the highest values of g_{avg} are exhibited by analytes with both an aromatic moiety and a heteroatom on C_α. Therefore, the carbonyl stretching frequency presumably takes into account the electronic features that those substituents bring in the system.

Unified model

A key feature of this modelling strategy is that all molecular descriptors were computed exclusively on the substrate, independently of the vanadium probe. To the best of our knowledge, this substrate-centric approach has not been previously reported in the context of chiroptical sensing. This idea of focusing solely on the intrinsic structural and electronic features of the free substrate allows to deal with prejudicial choices, while significantly reducing the computational cost associated with model construction. This innovative strategy thus enhances both the generality and practical applicability of the predictive framework.
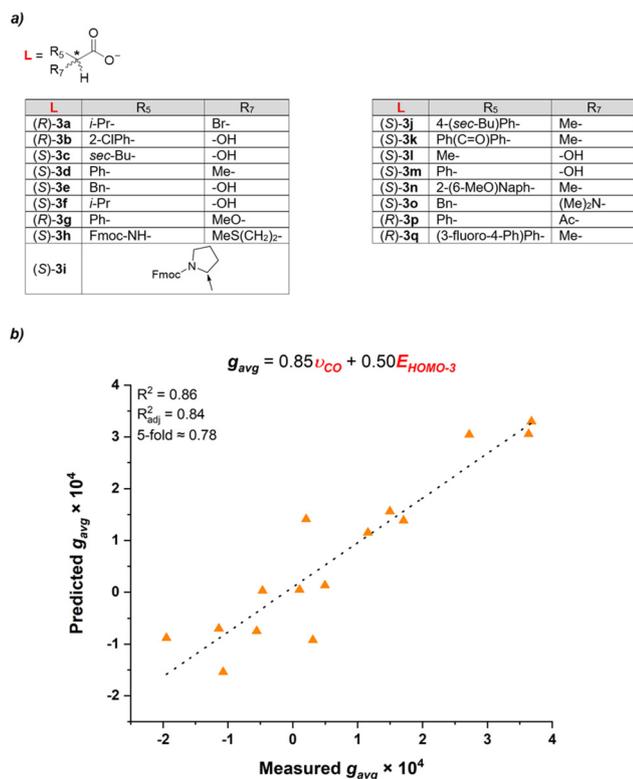

Fig. 2 (a) Enantiopure carboxylic acids **3a–q** (here already specified as carboxylates) tested. (b) MLR modeling of g_{avg} of the carboxylate adducts with the pivotal molecular descriptors. The model was build using data of adducts **5a–o**.

120 | *Org. Chem. Front.*, 2026, **13**, 118–123

This journal is © the Partner Organisations 2026

The interesting results obtained from the construction of the models for both amide and carboxylate substrates prompted an investigation into the feasibility of constructing a unified model that could account for CD responses across both organic classes. However, such a goal posed a distinct challenge, primarily due to the intrinsic structural and electronic differences between the two substrate families, which complicates the identification of a common set of molecular descriptors (see section S6.3). Nonetheless, by carefully choosing a minimal yet representative subset of parameters (Table S5), it was possible to develop a cross-class model capable of distilling the key factors governing CD induction in this system. Fig. 3 illustrates the global model derived from multiple linear regression encompassing all thirty-five substrates employed until this point. The model featured a good correlation (R^2 = 0.81) and a positive internal validation measure (5-fold = 0.77). Furthermore, an external validation was conducted to evaluate the ability of the model to predict an employed test set composed of four amides (**2u–2x**) and two carboxylates (**3p–3q**). It was found that the model reliably predicted the CD signals of the test set (test R^2 = 0.80). As anticipated, the increased chemical diversity and sample size necessitated a broader descriptor space. Nevertheless, only three terms were required to achieve a good correlation. These include two electronic descriptors and one second-order term that integrates both steric and electronic contributions. Consistent with the individual models, the IR stretching frequency of the carbonyl bond plays a key role, reflecting its critical role as the substrate's coordination site with the vanadium probe. A second electronic parameter, $E_{\mathrm{HOMO}-4}$, acts as a modulator for the outlier substrates; however, the rationale remains unclear, as with the carboxylates. The third parameter, which is a second-order term, likely captures a combined steric and electronic influence centred around the α-carbon relative to the carbonyl group. Nonetheless, even if on first approximation the steric and electronic contributions are highlighted by the fitting equations, it is difficult to correlate each parameter directly to specific physical-organic features. These promising results in constructing a predictive model for the CD outputs of two substrate classes open new avenues for developing more sophisticated approaches. Such advancements could extend the substrate scope and significantly accelerate the ee assessment process.

Conclusions

The development of fast and reliable methods for enantiomeric excess (ee) assessment is crucial across various scientific disciplines. This article reports a novel approach that addresses this need by developing a unified predictive model for circular dichroism (CD) outputs across two chemically distinct organic frameworks: amides and carboxylates. The combined parametrization of these two organic classes is facilitated by the presence of the common C=O moiety. In this regard, inclusion of other classes would be more challenging, but not impossible if new escamotages will be devised. This substrate-centric approach, which computes molecular descriptors exclusively from free substrates without including the host–guest complex, offers a conceptually simple yet informative strategy about the physical-organic chemistry of the process. This method not only yields good predictive models but also eliminates the need for separate calibration curves, marking a significant step towards a universal ee-sensing platform. The ability to construct a unified model across chemically diverse analytes demonstrates the potential for broader applicability and higher efficiency in chirality assessment. Furthermore, by focusing on the intrinsic properties of free substrates, this approach overcomes challenges associated with modelling noncovalent assemblies and reduces computational costs. These advancements pave the way for expanding substrate scope, improving model interpretability, and integrating this approach into automated or high-throughput workflows for asymmetric reaction screening. As such, this work marks the first step towards extending this unified approach to other classes of analytes, opening new avenues for the development of more sophisticated methods for ee determination, potentially revolutionizing high-throughput analysis in areas such as asymmetric catalysis and drug discovery.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data that support the findings of this study are available in the supplementary information (SI). Supplementary infor-
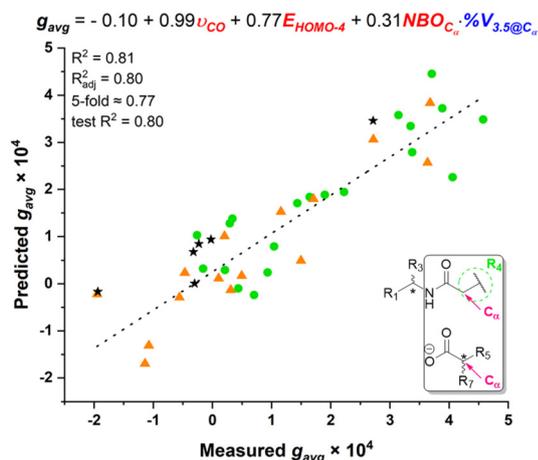

$$g_{avg} = -0.10 + 0.99\,\nu_{CO} + 0.77\,E_{HOMO-4} + 0.31\,NBO_{C_\alpha}\cdot\%V_{3.5@C_\alpha}$$

R² = 0.81
R²_adj = 0.80
5-fold ≈ 0.77
test R² = 0.80

Fig. 3 MLR modeling of g_{avg} of adducts **4a–t** (green dots) and **5a–o** (orange triangles) with the pivotal steric (blue) and electronic (red) molecular descriptors. The black stars correspond to adducts **4u–x** and **5p–q**, which were employed as external validation set. The bottom right specifies how the analytes have been described independently of the organic classes to which they belong.

This journal is © the Partner Organisations 2026

Org. Chem. Front., 2026, **13**, 118–123 | **121**

mation: compounds syntheses and characterizations, CD measurements, additional description of the computational and statistical modelling, and cartesian coordinates of the most stable conformers. See DOI: https://doi.org/10.1039/d5qo01281c.

Acknowledgements

References

1 C. De Luca, S. Felletti, F. A. Franchina, D. Bozza, G. Compagnin, C. Nosengo, L. Pasti, A. Cavazzini and M. Catani, Recent developments in the high-throughput separation of biologically active chiral compounds via high performance liquid chromatography, *J. Pharm. Biomed. Anal.*, 2024, **238**, 115794.

2 C. L. Barhate, L. A. Joyce, A. A. Makarov, K. Zawatzky, F. Bernardoni, W. A. Schafer, D. W. Armstrong, C. J. Welch and E. L. Regalado, Ultrafast chiral separations for high throughput enantiopurity analysis, *Chem. Commun.*, 2017, **53**, 509–512.

3 D. S. Hassan, F. S. Kariapper, C. C. Lynch and C. Wolf, Accelerated Asymmetric Reaction Screening with Optical Assays, *Synthesis*, 2022, 2527–2538.

4 S. Sheykhi, L. Mosca, M. Pushina, K. Dey and P. Anzenbacher, Exploiting fluorescent zinc(II) and copper (II) complexes for enantiomeric excess determination of hydroxycarboxylates, *Chem. Commun.*, 2020, **56**, 8964–8967.

5 M. Pushina, S. Farshbaf, E. G. Shcherbakova and P. Anzenbacher, A dual chromophore sensor for the detection of amines, diols, hydroxy acids, and amino alcohols, *Chem. Commun.*, 2019, **55**, 4495–4498.

6 J. S. S. K. Formen, J. R. Howard, E. V. Anslyn and C. Wolf, Circular Dichroism Sensing: Strategies and Applications, *Angew. Chem., Int. Ed.*, 2024, **63**(19), e202400767.

7 E. Nelson, J. A. Bertke, F. Y. Thanzeel and C. Wolf, Organometallic Chirality Sensing via "Click"–Like η6-Arene Coordination with an Achiral Cp*Ru(II) Piano Stool Complex, *Angew. Chem., Int. Ed.*, 2024, **63**(26), e202404594.

8 M. Quan, X. Y. Pang and W. Jiang, Circular Dichroism Based Chirality Sensing with Supramolecular Host–Guest Chemistry, *Angew. Chem., Int. Ed.*, 2022, **61**(23), e202201258.

9 G. Pescitelli, L. Di Bari and N. Berova, Application of electronic circular dichroism in the study of supramolecular systems, *Chem. Soc. Rev.*, 2014, **43**, 5211–5233.

10 R. Penasa, F. Begato, G. Licini, K. Wurst, S. Abbate, G. Longhi and C. Zonta, Fluorescence detected circular dichroism (FDCD) of a stereodynamic probe, *Chem. Commun.*, 2023, **59**, 6714–6717.

11 A. Prabodh, Y. Wang, S. Sinn, P. Albertini, C. Spies, E. Spuling, L.-P. Yang, W. Jiang, S. Bräse and F. Biedermann, Fluorescence detected circular dichroism (FDCD) for supramolecular host–guest complexes, *Chem. Sci.*, 2021, **12**, 9420–9431.

12 Y. Tian, G. Fang, F. Wu, J. G. Kauno, H. Wei, H.-Y. Hsu, F. Li, G. Xu and W. Niu, Raman spectroscopic technologies for chiral discrimination: Current status and new frontiers, *Coord. Chem. Rev.*, 2025, **526**, 216375.

13 J. Gong and X. Zhang, Coordination-based circularly polarized luminescence emitters: Design strategy and application in sensing, *Coord. Chem. Rev.*, 2022, **453**, 214329.

14 N. A. Carmo dos Santos, E. Badetti, G. Licini, S. Abbate, G. Longhi and C. Zonta, A stereodynamic fluorescent probe for amino acids. Circular dichroism and circularly polarized luminescence analysis, *Chirality*, 2018, **30**, 65–73.

15 C. Bravin, G. Mazzeo, S. Abbate, G. Licini, G. Longhi and C. Zonta, Helicity control of a perfluorinated carbon chain within a chiral supramolecular cage monitored by VCD, *Chem. Commun.*, 2022, **58**, 2152–2155.

16 T. P. Golub, T. Kano, K. Maruoka and C. Merten, VCD spectroscopy distinguishes the enamine and iminium ion of a 1,1′-binaphthyl azepine, *Chem. Commun.*, 2022, **58**, 8412–8415.

17 N. Berova, L. Di Bari and G. Pescitelli, Application of electronic circular dichroism in configurational and conformational analysis of organic compounds, *Chem. Soc. Rev.*, 2007, **36**, 914.

18 L. You, D. Zha and E. V. Anslyn, Recent Advances in Supramolecular Analytical Chemistry Using Optical Sensing, *Chem. Rev.*, 2015, **115**, 7840–7892.

19 C. Wolf and K. W. Bentley, Chirality sensing using stereodynamic probes with distinct electronic circular dichroism output, *Chem. Soc. Rev.*, 2013, **42**, 5408.

20 W. L. Williams, L. Zeng, T. Gensch, M. S. Sigman, A. G. Doyle and E. V. Anslyn, The Evolution of Data-Driven Modeling in Organic Chemistry, *ACS Cent. Sci.*, 2021, **7**, 1622–1637.

21 J. R. Howard, J. R. Shuluk, A. Bhakare and E. V. Anslyn, Data-science-guided calibration curve prediction of an MLCT-based ee determination assay for chiral amines, *Chem*, 2024, **10**, 2074–2088.

22 J. R. Howard, A. Bhakare, Z. Akhtar, C. Wolf and E. V. Anslyn, Data-Driven Prediction of Circular Dichroism-Based Calibration Curves for the Rapid Screening of Chiral Primary Amine Enantiomeric Excess Values, *J. Am. Chem. Soc.*, 2022, **144**, 17269–17276.

23 J. J. Dotson, E. V. Anslyn and M. S. Sigman, A Data-Driven Approach to the Development and Understanding of Chiroptical Sensors for Alcohols with Remote γ-Stereocenters, *J. Am. Chem. Soc.*, 2021, **143**, 19187–19198.

24 G. Licini, M. Mba and C. Zonta, Amine triphenolate complexes: synthesis, structure and catalytic activity, *Dalton Trans.*, 2009, 5265.

122 | *Org. Chem. Front.*, 2026, **13**, 118–123

This journal is © the Partner Organisations 2026

25 P. Zardi, K. Wurst, G. Licini and C. Zonta, Concentration-Independent Stereodynamic g-Probe for Chiroptical Enantiomeric Excess Determination, *J. Am. Chem. Soc.*, 2017, **139**, 15616–15619.

26 C. B. Santiago, J.-Y. Guo and M. S. Sigman, Predictive and mechanistic multivariate linear regression models for reaction development, *Chem. Sci.*, 2018, **9**, 2398–2412.

This journal is © the Partner Organisations 2026

Org. Chem. Front., 2026, **13**, 118–123 | **123**