nature chemistry

PERSPECTIVE

https://doi.org/10.1038/s41557-019-0319-5

Towards the online computer-aided design of catalytic pockets

Laura Falivene[1], Zhen Cao[1], Andrea Petta[2], Luigi Serra[2], Albert Poater[1,3], Romina Oliva[4], Vittorio Scarano[2] and Luigi Cavallo[1*]

The engineering of catalysts with desirable properties can be accelerated by computer-aided design. To achieve this aim, features of molecular catalysts can be condensed into numerical descriptors that can then be used to correlate reactivity and structure. Based on such descriptors, we have introduced topographic steric maps that provide a three-dimensional image of the catalytic pocket—the region of the catalyst where the substrate binds and reacts—enabling it to be visualized and also reshaped by changing various parameters. These topographic steric maps, especially when used in conjunction with density functional theory calculations, enable catalyst structural modifications to be explored quickly, making the online design of new catalysts accessible to the wide chemical community. In this Perspective, we discuss the application of topographic steric maps either to rationalize the behaviour of known catalysts—from synthetic molecular species to metalloenzymes—or to design improved catalysts.

Humans have been unknowingly using natural catalysts—enzymes—for the production of food and beverages for thousands of years. More recently, modern chemistry has introduced synthetic catalysts, whose field of application is, in principle, infinite. Modern catalysis has the potential to address and solve some of the great challenges of the 21st century, especially those related to the increasing demands of sustainable energy[1,2]. However, whereas nature has had billions of years to develop highly efficient enzymes that can catalyse a finite set of reactions required for life (approximately 6,000 classified to date[3]), chemists have been working on catalyst development for less than a century—and to address the societal challenges outlined above they need to be able to create efficient catalysts on a much shorter timescale, namely decades, rather than millennia.

Both for nature and for chemists, developing a good molecular catalyst means designing a well-functioning catalytic pocket—the region of the enzyme or synthetic catalyst where the substrate binds and reacts—by properly arranging the atoms involved in the catalytic process in three-dimensional space[4] (Fig. 1). As an example, proteases, the enzymes that catalyse the cleavage of peptide bonds, have evolved into families defined by specific amino acids or metal atoms fundamental to their biological function, and they are all efficient in catalysing the same chemical reaction. Serine protease families have even evolved towards the same catalytic pocket through independent evolutionary pathways[5].

The fine tuning of the active site of proteases has provided these enzymes with a high degree of specialization by restricting their action to a selected number of peptide bonds located at specific sites in specific protein substrates[5]. This tuning has been the result of countless random mutations in their amino acid sequence and in the natural selection of the most effective solutions. In a synthetic chemistry laboratory, two main approaches are generally used to develop a new catalyst: one relies on the high-throughput screening of a large library of potential catalysts to explore as many candidates as possible[6–8], and the other consists of developing a catalyst with predefined properties in an informed way, possibly through a few iterations aimed at improving the catalyst's performance up to a satisfactory level.

One of the main challenges in catalyst design is tuning catalysts from poorly effective prototypes to systems fitting the stringent criteria for industrial applications. This is a daunting task, because even minimal differences in the active site of a catalyst can alter its catalytic performance. For instance, good selectivity usually requires favouring one reaction pathway over the others by approximately 2 kcal mol^{-1}, and reducing an activation barrier by the same amount can result in a multifold improvement in the catalytic activity. This energy difference is smaller than that associated with a hydrogen bond in water (approximately 5 kcal mol^{-1})[9], but is still sufficient to shuttle 95% of the reactants along the favoured reaction pathway.

Molecular descriptors for predictive catalyst design

Molecular descriptors are one of the pillars of predictive catalyst design[10]—they quantify properties of catalysts, enabling the correlation of experimental behaviour with structure[11–13]. Molecular descriptors were introduced by Hammett, who proposed an empirical equation based on two parameters to correlate rates and equilibrium constants of reactions involving aromatic compounds[14]. Since this seminal work, several other descriptors have been proposed to build quantitative structure–activity relationships within a given family of organometallic catalysts[10], in which a transition metal at the centre of the catalytic pocket is usually responsible for the catalytic action. The most common approach, introduced by Tolman[15] for phosphines, is to connect the experimental behaviour of catalysts to two different descriptors characterizing the ligands; one electronic descriptor capturing the electronic effects due to transmission along the chemical bonds, and one steric descriptor arising from forces (normally nonbonding) between parts of the molecule (Fig. 2a)[15].

[1]Catalysis Research Center, Physical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [2]Dipartimento di Informatica ed Applicazioni, University of Salerno, Fisciano, Italy. [3]Institut de Química Computacional i Catàlisi and Departament de Química, Universitat de Girona, Girona, Spain. [4]Department of Sciences and Technologies, University Parthenope of Naples, Naples, Italy. *e-mail: luigi.cavallo@kaust.edu.sa

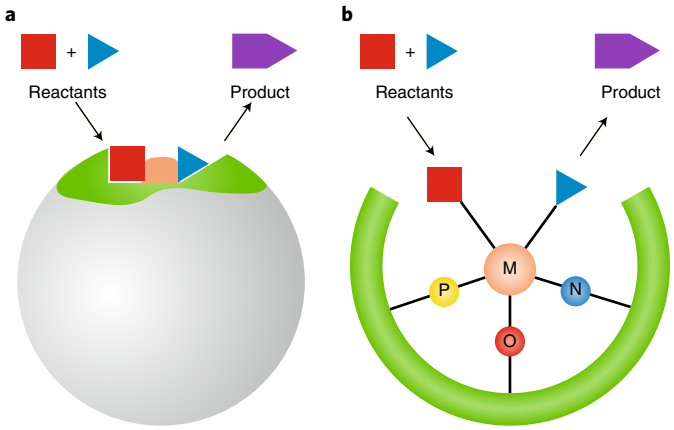

Fig. 1 | Catalytic pocket in enzymes and synthetic catalysts. a, The catalytic pocket of enzymes is characterized by a few residues directly involved in the catalytic action (bronze patch), and by residues indirectly involved, for example by altering the pK_a of a directly involved residue. Other residues may be responsible for binding the substrate (green patch), allowing the recognition and precise positioning of an enzyme's substrate in proximity to the chemically active catalytic residues. **b**, The catalytic pocket of a transition-metal-based synthetic molecular catalyst is usually characterized by a metal coordinated to one or more organic ligands. The metal is directly involved in the catalytic action, while the organic scaffold (green) plays an indirect catalytic role by modulating the electronic properties of the metal. The three-dimensional wrapping of the organic scaffold around the metal centre is responsible for the selectivity in substrate binding via shape complementarity, the formation of noncovalent binding interactions, or both. M, metal.

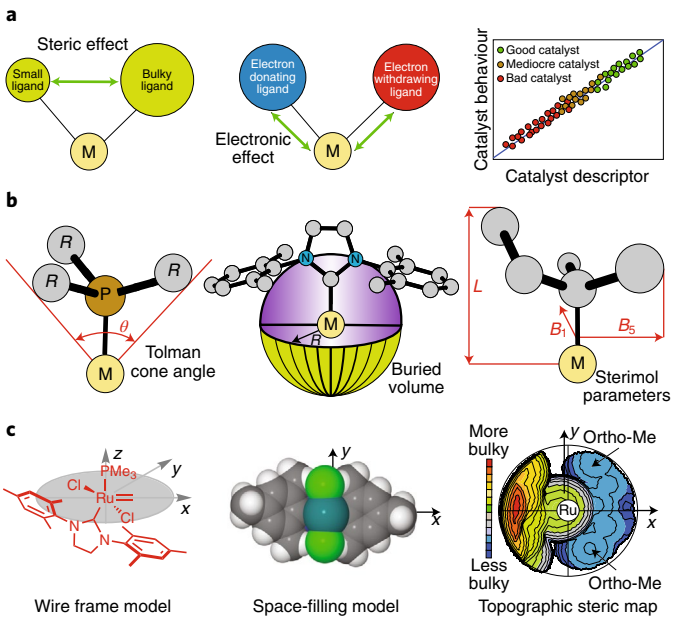

Fig. 2 | Schematic representation of descriptors used in catalysis. **a**, Steric and electronic molecular descriptors, capturing the hindrance and the electronic influence of ligands on the metal, can be used to correlate the experimental behaviour of a catalyst to its structure. **b**, Popular steric descriptors used in catalysis: the Tolman cone angle (θ), which measures the solid angle containing the whole ligand; the percentage of buried volume (%V_{Bur}), which measures the fraction of the first coordination sphere (of radius R) occupied by a ligand; and the Sterimol steric parameters, which measure the ligand length (L) and the minimal and maximal dimensions (B_1 and B_5, respectively) perpendicular to the ligand length. **c**, Wireframe, space-filling Corey–Pauling–Koltun, and topographic steric map representations of a ruthenium complex. M, metal. Adapted from ref. [36], American Chemical Society.

The steric descriptor introduced by Tolman, the cone angle θ (Fig. 2b)[15], is one of the most popular steric descriptors, and it can even be found in organometallic chemistry textbooks[16]. However, it has the disadvantage of being only applicable to phosphines. More general steric descriptors include those proposed by Taft[17], Charton[18,19] and Verloop[20,21], as well as the percentage of buried volume (%V_{Bur}), introduced by us to quantify the steric hindrance of N-heterocyclic carbenes (NHCs)[22–25]. As shown in Fig. 2b, the %V_{Bur} quantifies the fraction of the first coordination sphere around a metal centre that is occupied by the organic ligand. Because the %V_{Bur} focuses on space occupation around the metal, rather than on specific features of a given class of ligands, it can be used to build property–structure relationships for any class of catalysts and ligands. For example, this parameter has been used to quantify the steric effects of ligands in high oxidation-state metal catalysis[26] and, together with the Tolman cone angle, to explain the enhancement of nickel catalysis in cross-coupling reactions via remote steric effects[27].

Nevertheless, most of the proposed descriptors are limited by condensing the features of a given catalyst to single numbers, while the chemical behaviour is often more complex, because it is related to the three-dimensional (3D) shape of the catalytic pocket. This point becomes pivotal in the case of asymmetric synthesis, where the selectivity of the catalyst is related to effective de-symmetrization of the catalytic pocket. Single-value descriptors, such as the Tolman cone angle or the percentage of buried volume, are unable to capture these features. This limitation has inspired the search for other descriptors capable of capturing the 3D shape of a catalyst. Some of the proposed solutions are based on: (1) stereocartography, a mapping technique used to locate regions of maximum asymmetry and, consequently, stereo-induction around a chiral catalyst[28]; (2) the accessible molecular surface, a method for quantifying the intrinsic steric properties of a chelating ligand based on the solvent-accessible surface area of the metal centre[29]; and (3) Sterimol steric parameters, a set of parameters that can successfully capture the size and shape of a ligand on a metal by measuring the dimensions of the substituents along specific directions (Fig. 2b)[30,31]. These methodologies are powerful tools for quantifying differences in a catalytic pocket and quantitatively correlating the experimental behaviour of a catalyst with its structure. However, they cannot provide a 3D visualization of the shape of a catalytic pocket.

Visualizing the catalytic pocket

The concept of a catalytic pocket was developed in the field of enzymatic catalysis, and the importance of a properly shaped catalytic pocket has been well established since the introduction of the lock-and-key model by Fischer[32]. These concepts can easily be extended to molecular transition-metal catalysts, since high selectivity and good catalytic activity require shape complementarity between the substrate and the catalyst both in enzymatic and synthetic catalysis. However, the number of tools capable of visualizing the shape of the catalytic pocket in transition-metal complexes remains surprisingly low. The most common approach is to prepare a ball-and-stick or a space-filling Corey–Pauling–Koltun (CPK) visualization of the catalyst in a specific orientation[33,34], to maximize the exposure of the catalytic pocket (Fig. 2c). In addition to the difficulties associated with visualizing differences between related systems, these representations do not provide any quantitative description of the pocket surrounding the metal centre.

In this context, we introduced the concept of topographic steric maps[35,36] (Fig. 2c), which can be considered physicochemical analogues of physical maps in geography. Topographic steric maps

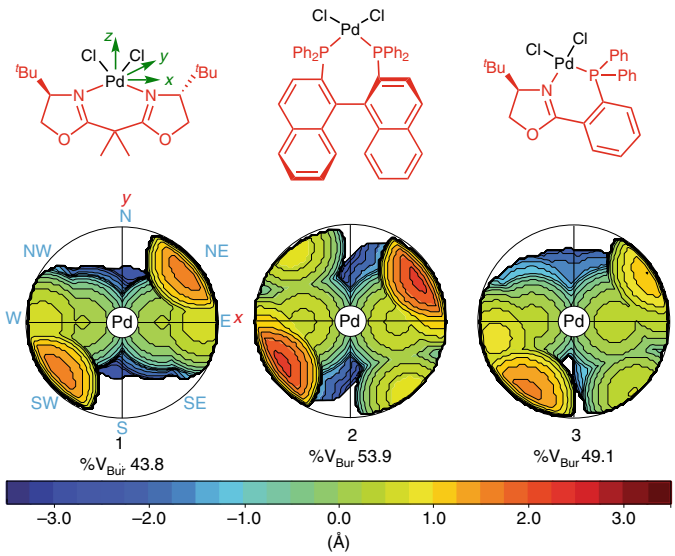

Fig. 3 | Topographic steric maps of transition metal complexes. Steric maps based on the density functional theory (DFT)-optimized structures of three related palladium complexes **1**–**3**. Only the ligands in red have been considered in the definition of the catalytic pocket. The steric maps are viewed down the z-axis; the orientation of the complexes is indicated for **1** and applies to the three systems. The isocontour scheme, in Å, is shown at the bottom. The red and blue zones indicate the more- and less-hindered zones in the catalytic pocket, respectively. Comparison of the steric maps allows differences to be identified in the shapes of the catalytic pockets of the three complexes. $\%V_{Bur}$, percentage of buried volume. Adapted from ref. [36], American Chemical Society.

provide an image of the surface of interaction between the catalyst and the substrate(s) that is shaped by the ligands in the complex, while the numerical array of points defining the surface can be used for quantitative analysis. Once the metal centre has been set to the equivalent of sea level and the whole complex has been oriented to maximize exposure of the metal centre from the top, altimetric contours offer a quantitative description of the catalytic pocket. Elevation from the metal is measured in a manner analogous to that used with physical maps for geographical features, and a colouring scheme from deep blue to dark red can be used to indicate regions in the catalytic pocket where there is space available to host the substrate below or above the 'sea level'.

Following a similar philosophy, a method for visualizing the hindrance around a catalytic centre has also been described by Houk, Montgomery and Liu, who used a graphical representation based on two-dimensional ligand steric contour maps, in which the contours describe the distance between the van der Waals surface of the ligand and the substrate, to explain the reversal of the regioselectivity in reductive couplings of alkynes and aldehydes catalysed by Ni–NHC complexes[37]. The same approach was used recently by Liu and Montgomery to rationalize the regio- and enantiocontrol in the same reaction when using a new class of exceptionally hindered, enantiopure NHC ligands[38].

Topographic steric maps of transition-metal complexes

In addition to offering an image of the catalytic pocket, topographic steric maps can also be considered fingerprints of transition-metal complexes. An illustration of this is presented in Fig. 3 for a set of three related, privileged chiral ligands[39]: bisoxazoline and binaphthyl ligands in C_2-symmetric complexes **1** and **2** and a phosphinooxazoline ligand in C_1-symmetric complex **3**. These complexes are known to induce enantioselectivity in a wide variety of textbook reactions, including Diels–Alder cycloadditions, Michael and Mukaiyama–Michael addition reactions, Heck reactions and asymmetric hydrogenations.

Simple visual inspections of the ball-and-stick or CPK structures of complexes **1**–**3** do not enable an immediate comparison of the various systems (see Fig. 2c for a representative comparison between different visualizations of a complex). Conversely, the steric maps in Fig. 3 identify the individual complexes and intuitively indicate how different ligands shape the catalytic pocket in different ways. The slim bisoxazoline ligand of **1** occupies space along the equator of the catalytic pocket, with the upward-oriented *t*butyl groups corresponding to the relatively soft bulges in the south-western and north-eastern quadrants[36]. The upward-facing phenyl groups of the binaphthyl ligand in **2** form a north-western to south-eastern groove similar to that in **1**. The main differences are in the naphthyl groups of **2**, which cover the bottom of the catalytic pocket in the north-western and south-eastern quadrants, and in the stronger hindrance provided by the phenyl groups of **2** in the south-western and north-eastern quadrants[36]. The catalytic pocket of **3** can be considered a hybrid of those of **1** and **2**, with the western hemisphere similar to that of **1** and the eastern hemisphere similar to that of **2**. Consistently, the $\%V_{Bur}$ of **2** is clearly larger than that of **1**, while the $\%V_{Bur}$ of **3** is intermediate between those of **1** and **2**.

Topographic steric maps in catalyst design

Topographic steric maps can be a useful guide for the design of new catalysts because they enable the determination of the effect of structural changes. For example, Bertrand and co-workers used topographic steric maps to design cyclic, six-membered (alkyl)(amino)carbene (CAAC-6) ligands capable of outcompeting their five-membered analogues (CAAC-5) in Pd-mediated α-arylations of ketones with aryl chlorides (Fig. 4a)[40]. Previous work on CAAC-5 Pd-complexes indicated that **4**, with unhindered ethyl substituents, was unable to promote the α-arylation reaction, while **5**, with a bulky menthyl substituent replacing the ethyl groups, showed high reactivity with the less sterically demanding aryl chloride, that is, chlorobenzene. However, **5** showed much lower reactivity with bulkier aryl chlorides, such as *o*-chlorotoluene or 2-chloro-*m*-xylene.

This knowledge led the authors to hypothesize that a CAAC ligand with large steric hindrance but more flexible than **5** would offer better catalytic efficiency and broaden the substrate scope to a large variety of aryl chlorides. To test this hypothesis, the CAAC-6 ligand of **6**, also having the advantage of being more electron donating, was selected. Buried volume analysis of the crystallographic structures of **4**, **5** and **6** confirmed the overall larger steric congestion around the Pd centre of **6** compared to both **4** and **5**. Further, analysis of the steric maps indicated that the additional steric hindrance in **5** is also placed around the north to south axis, which probably prevents coordination of bulky substrates, while in the steric map of **6** these are relatively unhindered zones, which possibly allows for coordination of bulkier substrates.

Subsequent comparison of the catalytic activities of **4**, **5**, and **6** in the α-arylation of propiophenone with different aryl chlorides indicated that CAAC-6 complex **6** outperforms CAAC-5 complexes **4** and **5**, and it offers excellent yields regardless of the structure of the aryl chloride (Fig. 4a), validating the catalyst design strategy[40].

Another example of catalyst design based on the comparative analysis of catalytic pockets in related complexes can be found in the development of chelating ligands for the Pd-catalysed copolymerization of ethylene with olefins that have a polar functional group (Fig. 4b). One of the drawbacks of the currently available catalysts is that only a few of them selectively generate high-molecular-weight linear polymers. Furthermore, the typical strategy used to increase molecular mass in olefin polymerization—the inhibition of the chain termination reaction by increasing the steric hindrance around the active centre of the catalytic system[42,43]—cannot

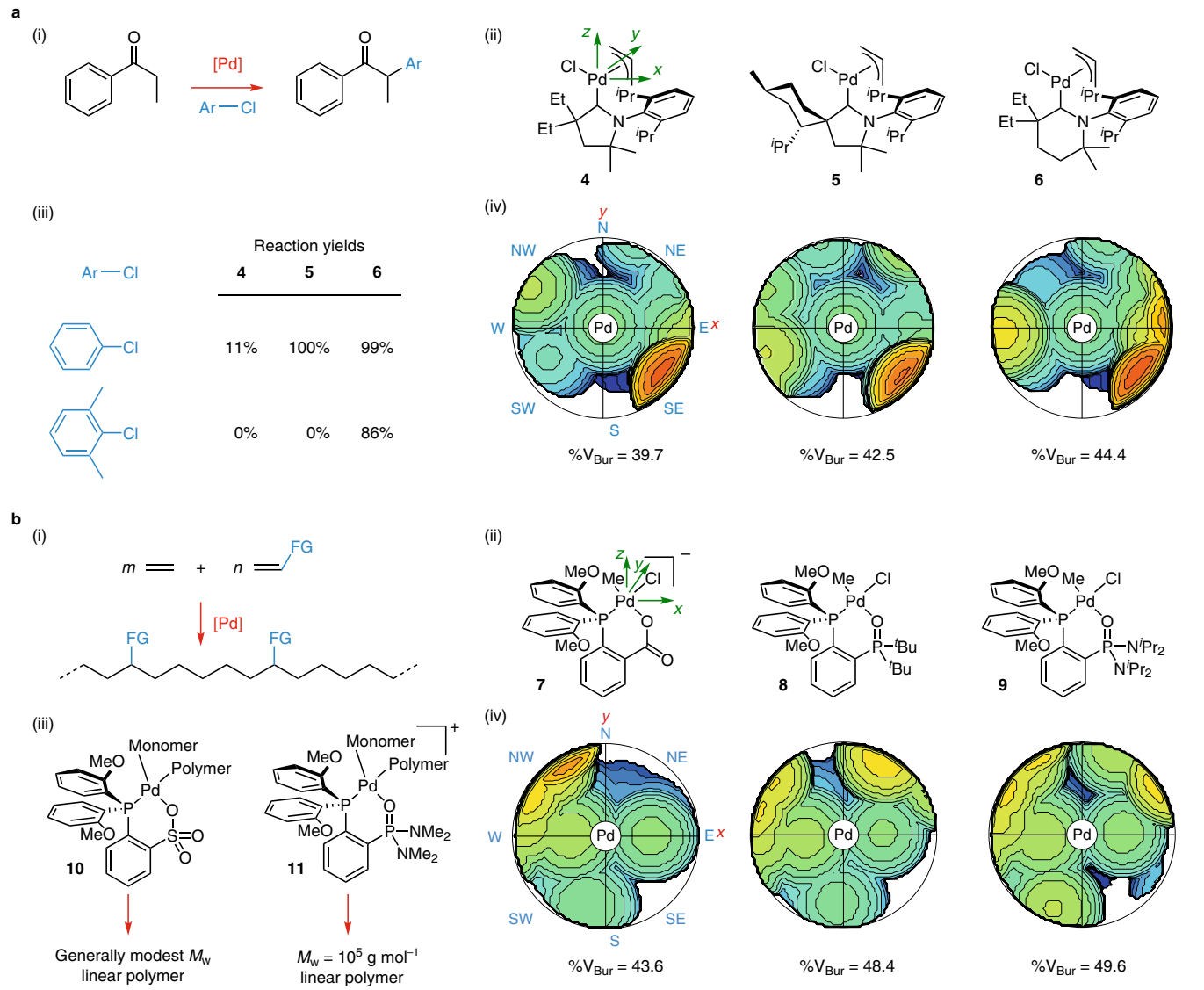

Fig. 4 | Application of topographic steric maps in catalyst design. a, Pd-catalysed α-arylations of ketones with aryl chlorides. (i) Reaction investigated by Bertrand and co-workers[40], α-arylation of propiophenone; (ii) Pd complexes compared by Bertrand and co-workers[40]; (iii) Performance of catalysts **4**–**6** in the α-arylation propiophenone; (iv) Steric maps of complexes **4**–**6**. A comparison of the steric maps of complexes **4**, **5** and **6** indicated that complex **6** was the most promising catalyst, because it has at the same time high steric hindrance and relatively free space along the north–south axis for coordination of bulky substrates. Subsequent experiments confirmed that **6** outperforms **4** and **5** in the α-arylation of propiophenone by aryl chlorides with different degrees of steric hindrance, such as chlorobenzene and 2-chloro-*m*-xylene. Adapted from ref. [40], American Chemical Society. **b**, Pd-catalysed synthesis of high-molecular-weight functional polyethylenes. (i) Reaction investigated by Carrow and colleagues[41], copolymerization of ethylene with functionalized 1-olefins; (ii) Pd complexes compared by Carrow and colleagues[41]; (iii) Performance of catalysts **10** and **11** in the synthesis of functionalized polyethylenes; (iv) Steric maps of complexes **7**–**9**. Comparison of the steric maps of complexes **7**, **8** and **9** indicated that the most hindered phosphine–phosphonic diamide ligand of **9** was the most promising skeleton for further structural variations. Subsequent experimental screening of a variety of different phosphine–phosphonic diamide ligands demonstrated that catalysts such as **11** are potent catalysts for the insertion copolymerization of ethylene with polar vinyl monomers to form linear, high-molecular-weight poly(ethylene-co-acrylate) materials. Adapted from ref. [41], American Chemical Society, 2018. In both **a** and **b**, the steric maps are viewed down the *z* axis; the orientation of the complexes is indicated for **4** and **7**, and this orientation applies to all of the systems. The isocontour scheme, in Å, is the same as that described in Fig. 3. $\%V_{Bur}$, percentage of buried volume.

be applied in a straightforward manner. For instance, increasing the steric hindrance of the substituents on the P atom of a phosphine-sulfonato Pd-ligand in the prototype Drent-type catalysts (**10**), for the copolymerization of ethylene with higher (and bulkier) olefins, drastically reduces the amount of incorporated co-monomer, as this steric hindrance is located in proximity of the coordinated monomer (Fig. 4b).

Carrow and co-workers resolved this dichotomy by designing Pd catalysts featuring a chelating P(v)–P(iii) ligand based on a phosphonic diamide-phosphine (PDAP) motif, such as in **9** (Fig. 4b)[41]. In addition to having different electronic properties, the steric hindrance of the PDAP skeleton can be tuned by changing the bulkiness of the substituents on the diamide N atoms, which is in positions remote from the coordinated monomer (Fig. 4b). This should preserve good insertion rates of the co-monomer.

The initial comparison of the steric hindrance of P,O-ligands in the crystallographic structures of Pd-complexes indicated that **9**, which has the prototype PDAP ligand with a diisopropyl,

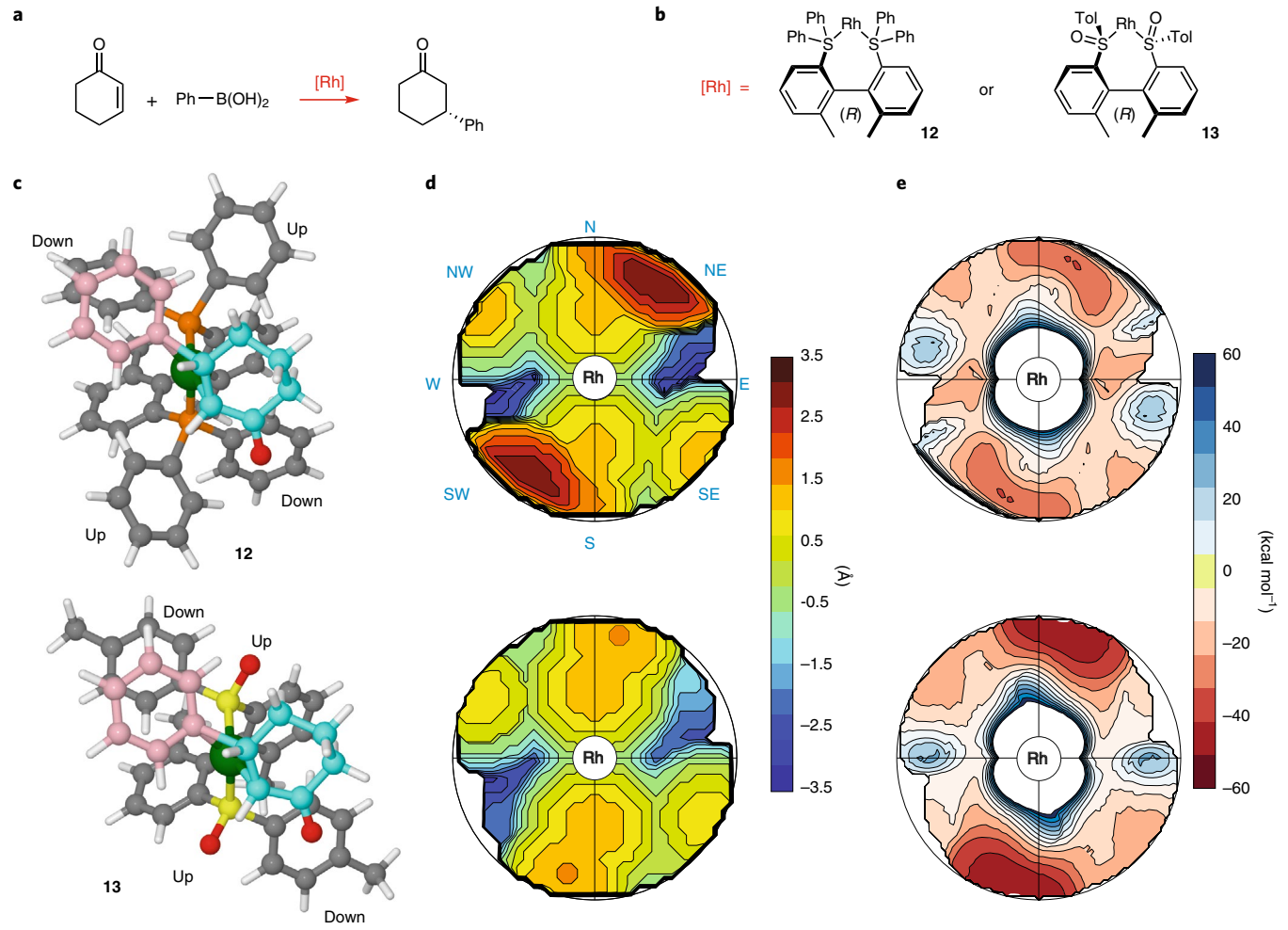

Fig. 5 | Rh-catalysed asymmetric addition of phenylboronic acid to 2-cyclohexenone. a, Scope of the reaction: enantioselective synthesis of 3-phenylcyclohexanone. **b**, Rh catalysts compared by Poater and co-workers[35]. **c**, Ball-and-stick representations of the favoured transition state geometries for the C–C bond formation steps between the six-membered rings of phenylboronic acid (pink) and 2-cyclohexenone (cyan), with **12** and **13**. **d**, Steric maps of the catalytic pockets of **12** and **13** in the transition-state geometries of the C–C bond formation step. The steric map of **12** is strongly asymmetric, with steric hindrance in the north-eastern and south-western quadrants due to the upward-oriented phenyl groups of **12**. This hindrance forms a northwest to southeast groove matching the relative disposition of the reacting groups in the favoured transition state. The steric map of **13** is flatter, with no zones of high steric hindrance. Therefore, **13** is not capable of enforcing a favoured orientation of the reacting groups in the C–C bond formation step through steric effects. **e**, Electrostatic potential maps of the catalytic pockets of **12** (top) and **13** (bottom) in the same transition states. The situation is reversed as compared to the steric maps: the electrostatic potential map of **13** is strongly asymmetric because of the highly negative electrostatic potentials in the northeast and southwest quadrants (due to the upward-oriented oxygen atoms of **13**), while the electrostatic map of **12** is flatter. Reproduced from ref. [35], Wiley.

phosphonic diamide group, displayed significantly higher steric hindrance when compared with other common ligands, such as the carboxylate-based ligand in **7** or the di-*tert*-butyl phosphine oxide ligand in **8**. In addition, visualization of the catalytic pocket indicated that the PDAP ligand in **9** hinders the O side of the complex (western hemisphere), while the impact on the P side, which is near the coordination position available for the incoming monomer in the eastern hemisphere, is minimal. This analysis led the authors to conclude that variations of the amine substituents in the PDAP ligand could be a good handle for tuning the polymerization behaviour of the catalyst.

Subsequent comparison of the catalytic copolymerization behaviour of Pd complexes bearing a series of differently substituted PDAP ligands allowed the authors to identify complexes, such as **11**, which are potent catalysts for the insertion copolymerization of ethylene with polar vinyl monomers to prepare high-molecular-weight copolymers (M_w up to 10^5 g mol^{-1}).

Physicochemical characterization of the catalytic pocket

As the steric maps identify the surface of the catalyst exposed to the substrate, this surface can be further analysed to characterize the catalytic pocket in terms of its physicochemical properties. For example, analysis of the catalytic pocket based on its steric properties alone is not sufficient for explaining the enantioselectivity in the asymmetric 1,4-addition of phenylboronic acid to 2-cyclohexenone, which leads to chiral 3-phenylcyclohexanone (Fig. 5a), promoted by Rh-catalysts **12** and **13** (see Fig. 5b). Both **12** and **13** are highly enantioselective, with enantiomeric excesses of 90% and 99%, respectively. While the good performance of **12** was expected, the high enantioselectivity achieved with **13** was surprising, as the enantioselective induction should be controlled by the small, upward-oriented S=O moieties (Fig. 5c). The characterization of the catalytic pocket in terms of its steric properties confirmed that the catalytic pocket of **12** is highly asymmetric[35], with a northwest to southeast oriented groove shaped by the upward-oriented *p*-tolyl groups (Fig. 5d). DFT calculations

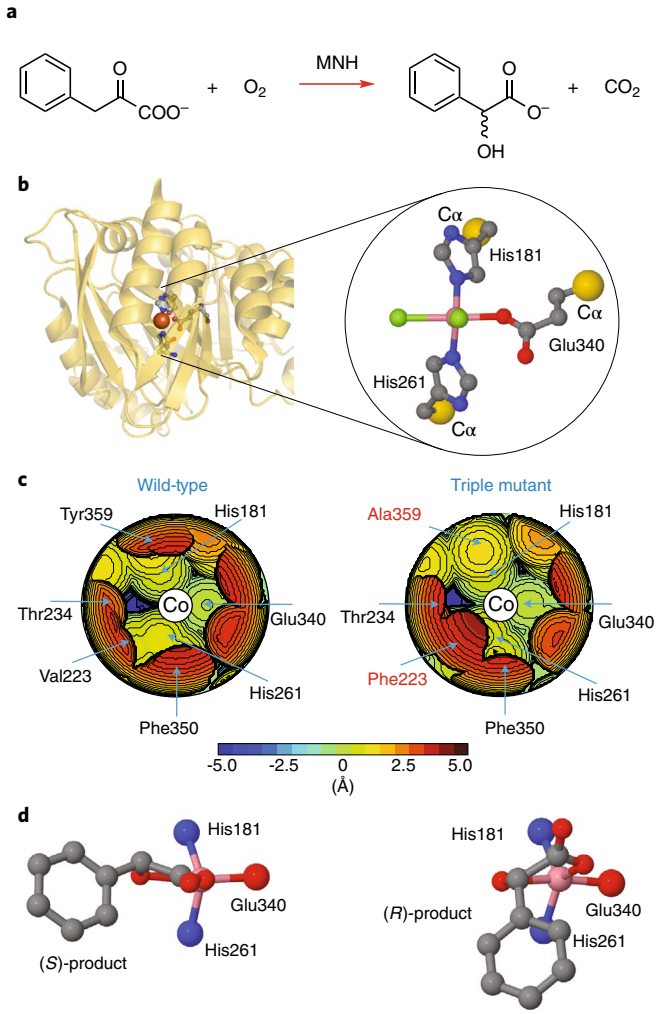

Fig. 6 | Steric maps of the catalytic pocket of metalloproteins. a, Scope of the reaction: oxidation of phenylpyruvate to mandelate. **b**, Ribbon representation of the structure of the S221M/V223F/Y359A mutant from *Streptomyces coelicolor* with the catalytic site in a ball-and-stick representation. The coordination positions available for catalysis are shown in green. **c**, Steric map of the active site in the crystallographic structure of the wild-type enzyme from *Amycolatopsis orientalis* and of the S221M/V223F/Y359A mutant from *Streptomyces coelicolor*. The metal is at the origin, the barycentres of the metal-coordinated nitrogen and oxygen atoms are on the *z* axis, and the metal-coordinated oxygen atom is on the *xz* plane. **d**, Ball-and-stick representation of the mandelate products bound in the crystallographic structure of the wild-type enzyme from *Amycolatopsis orientalis* and of the S221M/V223F/Y359A mutant from *Streptomyces coelicolor*. In both structures, Co(II) (pink) was used as a redox-inactive Fe(II) mimic. Adapted from ref. [36], American Chemical Society.

demonstrated that in the favoured transition state, the reacting groups are accommodated in this groove (Fig. 5c). In contrast, the steric map of **13** shows a substantially flatter catalytic pocket (Fig. 5d), suggesting that steric effects are not responsible for the high enantioselectivity exhibited by **13**.

Characterizing the surfaces of the catalytic pockets defined by the steric maps using the electrostatic potentials derived from DFT calculations provided a different perspective of the two catalysts (Fig. 5e). The electrostatic potential map of the catalytic pocket of **13** is remarkably asymmetric, with highly negative electrostatic potential hotspots located in proximity to the S=O moieties. In contrast, the electrostatic potential map of **12** is flatter. As a result, the electrostatic map of **13** has a shape similar to that of the steric map of **12**. The favoured transition state again places the reacting groups in the northwest to southeast groove, which is, in this case, shaped by the electrostatic potential on the surface of the catalytic pocket. In short, imaging catalytic pockets with the steric maps and characterizing their surfaces using a general property, such as the electrostatic potential, allowed us to explain the origin of the stereoselectivities of **12** and **13**.

Characterizing the catalytic pocket of metalloproteins

The application of topographic steric maps is not limited to synthetic molecular catalysts. For example, they can be used to characterize the catalytic pockets of natural or artificial metalloenzymes[44]. To illustrate how they can be used in the analysis and design of enzymes[45], in this section we discuss a comparison between the catalytic pocket of the wild-type mononuclear *p*-hydroxymandelate synthase from *Amycolatopsis orientalis* and that of an *in silico* designed triple mutant (S221M/V223F/Y359A) of a homologous enzyme from *Streptomyces coelicolor*.

The wild-type enzyme facilitates the conversion of phenylpyruvate to (*S*)-mandelate[46], while the mutant is designed to induce the opposite enantioselectivity and favour the formation of (*R*)-mandelate[46–48] (Fig. 6a). The active site of these proteins (Fig. 6b) presents a trigonal bipyramidal Fe centre with coordinating N-ε atoms of His181 and His261 in the equatorial plane and an oxygen from the carboxylate group of Glu340 along the main axis (according to the residue numbering of *Streptomyces coelicolor*). The other two coordination positions, coloured in green in Fig. 6b, are available for catalysis. In the X-ray structures of the wild-type enzyme (PDB ID: 2R5V)[49] and of the mutant (PDB ID: 3ZGJ)[48], these positions are occupied by the (*S*)-mandelate and the (*R*)-mandelate products, respectively. Co was used as a catalytically inactive mimic of Fe (refs. [48,49]) to intercept the intermediate in which the product is coordinated to the metal.

In both cases, the coordinating residues, His181, His261 and Glu340, are located at the bottom of the catalytic pocket[36]. As highlighted by the steric maps shown in Fig. 6c, in the catalytic pocket of the wild-type enzyme, the space above the metal is constrained by several residues (those in positions 223, 234, 261, 340, 350, and 359), and the aromatic ring of (*S*)-mandelate is thus placed in the south-western quadrant in an open subpocket between the metal and residues Thr234 and Val223. In contrast, in the mutant the aromatic ring of (*R*)-mandelate is placed in the north-western quadrant of the catalytic pocket. (*R*)-mandelate cannot be accommodated in the catalytic pocket of the wild-type enzyme, as its aromatic ring would clash with the large Tyr359 residue.

The impact of the V223F and Y359A mutations is evident from a comparison of the steric representations of the catalytic pockets: the V1223F mutation blocks the open space in the south-western quadrant, where the aromatic ring of (*S*)-mandelate would be placed, while the Y259A mutation opens space in the north-western quadrant, to accommodate the aromatic ring of (*R*)-mandelate.

Conclusions and outlook

Classifying the behaviour of catalysts using molecular descriptors that are unbiased towards the catalyst structure, such as the Sterimol steric parameters, topographic steric maps or the recently proposed average steric occupancy descriptors[50], can facilitate the rapid design of superior catalysts. Focusing on descriptors that are able to capture the shape of a catalytic pocket, topographic steric maps can be seen as fingerprints for characterizing transition metal catalysts and, based on preliminary results, metalloproteins, allowing their application in biocatalysis. Evolution of the concepts described in this perspective can be implemented by the catalysis community by using and modifying the source code, downloadable from the

SambVca web server, which calculates buried volumes and steric maps[51], under the GNU general public license[52].

For example, one possible development could consist in exploring different ways of characterizing the surface of the catalytic pocket in terms of a series of physicochemical properties other than steric hindrance, as in the electrostatic potential maps of Fig. 5. In analogy to the characterization of protein surfaces in terms of the properties of the exposed amino acids[53], the catalytic pocket of catalysts could also be characterized in terms of their hydrophobic/hydrophilic, polar/apolar and aromatic patches; hydrogen bond acceptors and donors; or Lewis base and acid groups.

Another development could consist in the integration of the steric map source code with efforts to develop computational methods capable of real-time geometry optimization of molecular structures[54–57]. This would facilitate the implementation of a web-based 3D computer-aided design system[58], which could allow the interactive modification of an initial catalyst skeleton.

Finally, the digital version of the steric map, which is the array of points defining the surface in the Cartesian space, could be used as a digital steric descriptor within multilinear regression analysis[59–61], or could be embedded in a workflow for the high-throughput screening of new catalysts within machine learning approaches[50,62].

Code availability

The source code calculating buried volumes and steric maps is downloadable from the SambVca 2.1 web server, https://www.molnac.unisa.it/OMtools/sambvca2.1/index.html.

Received: 23 September 2018; Accepted: 25 July 2019;
Published online: 02 September 2019

References

1. *Chemistry for Tomorrow's World* (Royal Society of Chemistry, 2009); https://go.nature.com/31AWVup
2. *Basic Research Needs for Catalysis Science* (US Department of Energy, 2017); https://go.nature.com/2yU0iR1
3. Placzek, S. et al. BRENDA in 2017: new perspectives and new tools in BRENDA. *Nucleic Acids Res.* **45**, D380–D388 (2017).
4. Porter, C. T., Bartlett, G. J. & Thornton, J. M. The catalytic site atlas: a resource of catalytic sites and residues identified in enzymes using structural data. *Nucleic Acids Res.* **32**, D129–D133 (2004).
5. Neurath, H. Evolution of proteolytic enzymes. *Science* **224**, 350–357 (1984).
6. Holzwarth, A., Schmidt, H.-W. & Maier, W. F. Detection of catalytic activity in combinatorial libraries of heterogeneous catalysts by IR thermography. *Angew. Chem. Int. Ed.* **37**, 2644–2647 (1998).
7. Boussie, T. R. et al. A fully integrated high-throughput screening methodology for the discovery of new polyolefin catalysts: Discovery of a new class of high temperature single-site group (IV) copolymerization catalysts. *J. Am. Chem. Soc.* **125**, 4306–4317 (2003).
8. Taylor, S. J. & Morken, J. P. Thermographic selection of effective catalysts from an encoded polymer-bound library. *Science* **280**, 267–270 (1998).
9. Babin, V., Leforestier, C. & Paesani, F. Development of a "first principles" water potential with flexible monomers: Dimer potential energy surface, VRT spectrum, and second virial coefficient. *J. Chem. Theory Comput.* **9**, 5395–5403 (2013).
10. Todeschini, R. & Viviana Consonni, V. *Handbook of Molecular Descriptors* (Wiley, 2000).
11. Foscato, M., Occhipinti, G., Venkatraman, V., Alsberg, B. K. & Jensen, V. R. Automated design of realistic organometallic molecules from fragments. *J. Chem. Inf. Model.* **54**, 767–780 (2014).
12. Fey, N., Orpen, A. G. & Harvey, J. N. Building ligand knowledge bases for organometallic chemistry: Computational description of phosphorus(III)-donor ligands and the metal–phosphorus bond. *Coord. Chem. Rev.* **253**, 704–722 (2009).
13. Fey, N. The contribution of computational studies to organometallic catalysis: Descriptors, mechanisms and models. *Dalton Trans.* **39**, 296–310 (2010).
14. Hammett, L. P. The effect of structure upon the reactions of organic compounds. Benzene derivatives. *J. Am. Chem. Soc.* **59**, 96–103 (1937).
15. Tolman, C. A. Steric effects of phosphorus ligands in organometallic chemistry and homogeneous catalysis. *Chem. Rev.* **77**, 313–348 (1977).
16. Crabtree, R. H. *The Organometallic Chemistry of the Transition Metals* (Wiley, 2014).
17. Hansch, C., Leo, A. & Taft, R. W. A survey of Hammett substituent constants and resonance and field parameters. *Chem. Rev.* **91**, 165–195 (1991).
18. Charton, M. & Charton, B. Steric effects. v. barriers to internal-rotation. *J. Am. Chem. Soc.* **97**, 6472–6473 (1975).
19. Charton, M. Linear free-energy relationships. II. Proximity effects. *Chem. Tech.* **5**, 245–255 (1975).
20. Verloop, A., Hoogenstraaten, W. & Tipker, J. *Development and Application of New Steric Substituent Parameters in Drug Design*, Vol. 7, 165–207 (Academic Press, 1976).
21. Verloop, A. *The STERIMOL approach to drug design* (Marcel Dekker, 1987).
22. Poater, A. et al. SambVca: A web application for the calculation of the buried volume of N-heterocyclic carbene ligands. *Eur. J. Inorg. Chem.* https://doi.org/10.1002/ejic.200801160 (2009).
23. Cavallo, L., Correa, A., Costabile, C. & Jacobsen, H. Steric and electronic effects in the bonding of N-heterocyclic ligands to transition metals. *J. Organomet. Chem.* **690**, 5407–5413 (2005).
24. Poater, A. et al. Thermodynamics of N-heterocyclic carbene dimerization: The balance of sterics and electronics. *Organometallics* **27**, 2679–2681 (2008).
25. Hillier, A. C. et al. A combined experimental and theoretical study examining the binding of N-heterocyclic carbenes (NHC) to the Cp*RuCl (Cp* = η^5-C_5Me_5) moiety: insight into stereoelectronic differences between unsaturated and saturated NHC ligands. *Organometallics* **22**, 4322–4326 (2003).
26. Billow, B. S., McDaniel, T. J. & Odom, A. L. Quantifying ligand effects in high-oxidation-state metal catalysis. *Nat. Chem.* **9**, 837–842 (2017).
27. Wu, K. & Doyle, A. G. Parameterization of phosphine ligands demonstrates enhancement of nickel catalysis via remote steric effects. *Nat. Chem.* **9**, 779–784 (2017).
28. Lipkowitz, K. B., D'Hue, C. A., Sakamoto, T. & Stack, J. N. Stereocartography: a computational mapping technique that can locate regions of maximum stereoinduction around chiral catalysts. *J. Am. Chem. Soc.* **124**, 14255–14267 (2002).
29. Angermund, K. et al. Complexes ($_P$2)Rh(hfacac) as model compounds for the fragment ($_P$2)Rh and as highly active catalysts for C_O2 hydrogenation: The accessible molecular surface (AMS) model as an approach to quantifying the intrinsic steric properties of chelating ligands in homogeneous catalysis. *Chem. Eur. J.* **3**, 755–764 (1997).
30. Harper, K. C., Vilardi, S. C. & Sigman, M. S. Prediction of catalyst and substrate performance in the enantioselective propargylation of aliphatic ketones by a multidimensional model of steric effects. *J. Am. Chem. Soc.* **135**, 2482–2485 (2013).
31. Harper, K. C. & Sigman, M. S. Using physical organic parameters to correlate asymmetric catalyst performance. *J. Org. Chem.* **78**, 2813–2818 (2013).
32. Fischer, E. Einfluss der configuration auf die wirkung der enzyme. *Berichte Deutsch. Chem. Gesell.* **27**, 2985–2993 (1894).
33. Pauling, L. & Corey, R. B. The structure of fibrous proteins of the collagen-gelatin group. *Proc. Natl Acad. Sci. USA* **37**, 272–281 (1951).
34. Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl Acad. Sci. USA* **37**, 205–211 (1951).
35. Poater, A., Ragone, F., Mariz, R., Dorta, R. & Cavallo, L. Comparing the enantioselective power of steric and electrostatic effects in transition-metal-catalyzed asymmetric synthesis. *Chem. Eur. J.* **16**, 14348–14353 (2010).
36. Falivene, L. et al. SambVca 2. A web tool for analyzing catalytic pockets with topographic steric maps. *Organometallics* **35**, 2286–2293 (2016).
37. Liu, P., Montgomery, J. & Houk, K. N. Ligand steric contours to understand the effects of N-heterocyclic carbene ligands on the reversal of regioselectivity in Ni-catalyzed reductive couplings of alkynes and aldehydes. *J. Am. Chem. Soc.* **133**, 6956–6959 (2011).
38. Wang, H. et al. NHC ligands tailored for simultaneous regio- and enantiocontrol in Nickel-catalyzed reductive couplings. *J. Am. Chem. Soc.* **139**, 9317–9324 (2017).
39. Yoon, T. P. & Jacobsen, E. N. Privileged chiral catalysts. *Science* **299**, 1691–1693 (2003).
40. Weinstein, C. M. et al. Highly ambiphilic room temperature stable six-membered cyclic (alkyl)(amino)carbenes. *J. Am. Chem. Soc.* **140**, 9255–9260 (2018).
41. Zhang, W. et al. Electron-rich metal cations enable synthesis of high molecular weight, linear functional polyethylenes. *J. Am. Chem. Soc.* **140**, 8841–8850 (2018).
42. Deng, L., Woo, T. K., Cavallo, L., Margl, P. M. & Ziegler, T. The role of bulky substituents in Brookhart-type Ni(II) diimine catalyzed olefin polymerization: a combined density functional theory and molecular mechanics study. *J. Am. Chem. Soc.* **119**, 6177–6186 (1997).
43. Talarico, G., Busico, V. & Cavallo, L. "Living" propene polymerization with bis(phenoxyimine) group 4 metal catalysts: New strategies and old concepts. *Organometallics* **23**, 5989–5993 (2004).
44. Schwizer, F. et al. Artificial metalloenzymes: reaction scope and optimization strategies. *Chem. Rev.* **118**, 142–231 (2018).

45. Rothlisberger, D. et al. Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
46. Choroba, O. W., Williams, D. H. & Spencer, J. B. Biosynthesis of the vancomycin group of antibiotics: involvement of an unusual dioxygenase in the pathway to (*S*)-4-hydroxyphenylglycine. *J. Am. Chem. Soc.* **122**, 5389–5390 (2000).
47. Hubbard, B. K., Thomas, M. G. & Walsh, C. T. Biosynthesis of L-*p*-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. *Chem. Biol.* **7**, 931–942 (2000).
48. Pratter, S. M. et al. Inversion of enantioselectivity of a mononuclear non-heme iron(II)-dependent hydroxylase by tuning the interplay of metal-center geometry and protein structure. *Angew. Chem. Int. Ed.* **52**, 9677–9681 (2013).
49. Brownlee, J., He, P., Moran, G. R. & Harrison, D. H. Two roads diverged: the structure of hydroxymandelate synthase from *amycolatopsis orientalis* in complex with 4-hydroxymandelate. *Biochem.* **47**, 2002–2013 (2008).
50. Zahrt, A. F. et al. Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science* **363**, eaau5631 (2019).
51. Falivene, L. et al. SambVca 2. A web tool for analyzing catalytic pockets with topographic steric maps. *Organometallics* **35**, 2286–2293 (2016).
52. *GNU v.3* (Free Software Foundation, 2007).
53. Hubbard, S. J., Campbell, S. F. & Thornton, J. M. Molecular recognition: conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors. *J. Mol. Biol.* **220**, 507–530 (1991).
54. Hostaš, J. & Řezáč, J. Accurate DFT-D3 calculations in a small basis set. *J. Chem. Theory Comput.* **13**, 3575–3585 (2017).
55. Brandenburg, J. G., Bannwarth, C., Hansen, A. & Grimme, S. B97-3c: A revised low-cost variant of the B97-D density functional method. *J. Chem. Phys.* **148**, 064104 (2018).
56. 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.* **15**, 1652–1671 (2019).
57. Mühlbach, A. H., Vaucher, A. C. & Reiher, M. Accelerating wave function convergence in interactive quantum chemical reactivity studies. *J. Chem. Theory Comput.* **12**, 1228–1235 (2016).
58. Wu, D., Rosen, D. W., Wang, L. & Schaefer, D. Cloud-based design and manufacturing: a new paradigm in digital manufacturing and design innovation. *Comput. Aided Des.* **59**, 1–14 (2015).
59. Santiago, C. B., Guo, J.-Y. & Sigman, M. S. Predictive and mechanistic multivariate linear regression models for reaction development. *Chem. Sci.* **9**, 2398–2412 (2018).
60. Sigman, M. S., Harper, K. C., Bess, E. N. & Milo, A. The development of multidimensional analysis tools for asymmetric catalysis and beyond. *Acc. Chem. Res.* **49**, 1292–1301 (2016).
61. Harper, K. C. & Sigman, M. S. Three-dimensional correlation of steric and electronic free energy relationships guides asymmetric propargylation. *Science* **333**, 1875–1878 (2011).
62. Ahneman, D. T., Estrada, J. G., Lin, S., Dreher, S. D. & Doyle, A. G. Predicting reaction performance in C–N cross-coupling using machine learning. *Science* **360**, 186–190 (2018).

Acknowledgements

L.C. thanks the King Abdullah University of Science and Technology (KAUST). This research used resources of the Core Labs and of the KAUST Supercomputing Laboratory. A.P. is a Serra Húnter fellow and thanks the Spanish MICINN for the project PGC2018-097722-B-I00. R.O. thanks University Parthenope 'Finanziamento per il Sostegno alla Ricerca Individuale di Ateneo – Annualità 2017' for funding.

Author contributions

L.C. conceived and designed the project. L.F. and A.P. provided the DFT calculations and the buried volume and steric maps analyses. Z.C. wrote the SambVca source code. R.O. provided the analysis of the biomolecules. A.P., L.S. and V.S designed and implemented the SambVca web application. All authors contributed to the discussion, L.C. and R.O. wrote the manuscript and all authors commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to L.C.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.