En masse Analysis Yields Prioritized List of Building Blocks and Couplings for Lego-like Natural Product Synthesis

Andrea M. E. Palazzolo Ray and Martin D. Burke

General synthetic methods that broadly enable the preparation of functionally important organic small molecules can be exceptionally impactful. Identification of such methods has traditionally been done on an *ad hoc* basis. Conversely, a more systematic approach might reveal an impact prioritized list of unsolved methodology to access a given chemical space. Natural products are or have inspired many highly impactful human medicines, crop protectants, food preservatives, and biological probes. Thus, natural products represent a function rich chemical space for such an analysis. Here we report an *en masse* analysis of most linear natural products which has revealed an impact-prioritized list of building blocks and couplings that would collectively provide iterative synthesis-based access to greater than 75% of the chemical space occupied by all known linear natural products. Because this analysis was not restricted to known chemistry, many of the identified couplings represent an actionable list of methodological problems that currently lack such a generalized solution compatible with iterative synthesis. Moreover, for each identified coupling, deeper analysis of the corresponding local chemical environments to a breadth of two heavy atoms reveals the necessary substrate scope to access the functional space occupied by these linear natural products. We selected one of the identified top coupling problems that lacked a general solution, the stereoretentive cross-coupling of primary Csp^3 boronic acids to vinyl halides. Preliminary analysis into this coupling suggests that highly stereospecific couplings are accessible for *E* and *Z* olefins from the computationally pre-determined generalized substrate scope.