**R. B. Woodward and the Logic of Testing in Organic Chemistry**

George Bor**g**

**Outline:**

*R. B. Woodward and the Logic of Testing in Organic Chemistry*

**Abstract**

A central question of philosophy of science is the nature of the knowledge achieved in science at its most successful. This paper starts from the premise that one of the best ways of addressing this question is to study how science turns data into evidence, for viewed from the inside science is an endeavor to do just that. Of particular interest from this perspective are transitional moments in which areas of research have gone from not being able to turn data into evidence effectively to being able to do so with great success. My thesis is that organic chemistry underwent just such a transition in the middle of the 20th century. During this period, chemists switched from turning the results of *chemical* experiments into evidence for hypotheses about molecular structure to turning *physical* data—absorption spectra, X-ray diffraction patterns, and so on—into evidence for such hypotheses. This switch changed the face of the discipline, for it allowed structural claims to become more secure. The question this paper will address is, what made this development possible? While historians have noted the importation of physical instrumentation into chemistry that occurred at this time, relatively little has been said from a historico-philosophical perspective concerning changes in how evidence was marshaled for structural hypotheses. Here I will defend four theses concerning the greater security of structural claims after the Instrumental Revolution, using the work of chemist R. B. Woodward in ultraviolet spectroscopy as the central example.

**1. Introduction**

In 1963, two years before winning the Nobel prize in chemistry, Robert Burns Woodward commented on the future of chemistry in the light of the recent solution of a problem that had mobilized chemists' efforts for decades:

> It will not be lost upon the reader—nor was it on at least some of the observers of the chemical scene in the late nineteen forties—that the almost simultaneous outcomes of the decades-long chemical degradative assault, and the incomparably shorter X-ray crystallographic investigations, presaged a future in which so singular an edifice as the chemical structure determination of strychnine was unlikely to find a parallel…This short history should give pause to those whose talent for despair is lavished upon an organic chemistry ornamented and supplemented—or as they fancy, burdened—by magnificent new tools which permit the establishment in days or weeks of enlightenments which once would have required months or years.[1]

The problem that had received converging solutions, one by chemical degradation techniques and the other by X-ray crystallography, was the elucidation of the molecular structure of strychnine, the investigation of which had begun in earnest in the late nineteenth century but which was considered solved only as of 1948. Though the structure of strychnine was one of the biggest challenges for classical chemistry, it was not the difficulty that made it significant for

---

[1] Woodward, R. B., Cava, C. M., W. Ollis, D. Hunger, A., Dafniker, H. U., & Schenker, K. (1963) The Total Synthesis of Strychnine. *Tetrahedron,* 19, 247-288 (p. 248).

Woodward, but rather that it was one of the last of its kind. As the "incomparably shorter X-ray crystallographic investigations" presaged, molecular structures would henceforth be solved in a time-frame of days or weeks rather than months or years. Chemistry could now "*begin* at the advanced vantage point of an established structure"[1]—begin, that is, the laboratory synthesis of the structures—rather than have to devote most of its efforts to establishing structure. What developments allowed chemists to establish structures so quickly, and, perhaps more importantly, with enough *security* to justify ambitious synthetic projects demanding the labor and resources of teams of chemists over many years?

Part of the answer is the generalization of instrumental techniques based on physical theory, such as X-ray crystallography, spectroscopy and mass spectrometry in the mid-20th century. The use of such techniques has led some historians to call these developments the "Instrumental Revolution."[2] Another part of the answer is the generalization of techniques for correlating features of molecular structure with features of spectroscopic data. These techniques allowed chemists to turn the data into evidence for claims about molecular structure. The history of these techniques begins in the late 19th-century, but they only reached maturity after the advent of quantum mechanics. Though many scientists were involved in the development of these techniques, some chemists and historians have highlighted Woodward's papers on UV spectroscopy in the early 1940s, papers that disclosed the so-called "Woodward rules," as a seminal contribution to this development. For example, Lord Todd, an eminent British chemist, wrote that

> [M]y attention was first drawn to him [Woodward] by a paper he published in 1941 (at the age of twenty-four) on the ultraviolet absorption spectra of α:β-unsaturated ketones; it seemed to me to herald a breakthrough in the use of spectroscopy in the study of molecular structure.[3]

---

[2] The term 'Instrumental Revolution' was coined by Dean S. Tarbell and Ann T. Tarbell in *Essays on the history of organic chemistry in the United States, 1875-1955*. Nashville: Folio Publishers, (p. 336). See also P. J. T. Morris (Ed.) (2002), *From Classical to Modern Chemistry: The Instrumental Revolution*. London: Royal Society of Chemistry.

[3] Cited in Slater, L. B. (2002). Instruments and Rules: R. B. Woodward and the tools of twentieth-century organic chemistry. *Stud. Hist. Phil. Sci., 33*, 1-33 (p. 17).

Slater (2002) notes that the Woodward rules "were among the first of this type of generalization," being "precise…and comprehensive in scope"[4] and that this "earlier work on UV spectra, and Arnold Beckman's 1941 introduction of a commercial UV instrument, pushed organic chemists to take note of advances in physical chemistry and instrumentation."[5] Benfey and Morris (2001) remark that

> the collating and generalizing abilities of the young Woodward still merit recognition as a major landmark in the use of physicochemical measurements in structural organic chemistry. In the three decades after 1941 such techniques as X-ray crystallography, infrared and ultraviolet spectroscopy, nuclear magnetic resonance, and mass spectroscopy gradually made chemical methods of structure determination almost completely extinct.[6]

This attention notwithstanding, in my opinion it has not been made sufficiently clear in what way the correlation techniques exemplified by the Woodward rules changed how chemists develop evidence for structural hypotheses. This paper is a contribution towards answering that question.

My basic claim is that the correlation techniques represent a change in the logic of testing applied to structure determination problems. This view contrasts with those such as Schummer's (2002) view that "the method of structure determination remained basically the same from a logical point of view, chemical and spectroscopic properties were both used in the same way" before and after the Instrumental Revolution.[7] 19th and early-20th century chemists had to make inferences about molecular structure without having empirical access to it. The classical approach consisted largely of measuring the bulk physical properties of compounds (e.g., melting point, empirical formula[8]), applying qualitative chemical tests for the presence of functional groups and converting unknown compounds into simpler, known ones in a process

---

[4] Slater (2002), p. 17.
[5] Slater (2002), p. 11.
[6] Benfey, O. T. and Morris, P. J. T. (Eds.) (2001), *Robert Burns Woodward: Architect and Artist in the World of Molecules*. Philadelphia: Chemical Heritage Foundation (p. 45). Additional remarks on Woodward's role in the instrumental revolution may be found in Morris, P. J. T. & Travis, A. S. The Rôle of Physical Instrumentation in Structural Organic Chemistry in the Twentieth Century. In Morris (2002), pp. 57-84.
[7] Schummer, J. The Impact of Instrumentation on Chemical Species Identity from Chemical Substances to Molecular Species. In Morris (Ed.) (2002), 188-211 (p. 196).
[8] The empirical formula of a compound expresses the ratio of the elements present. The molecular formula expresses the total number of each atom present and is the same as the empirical formula or some multiple of it.

known as 'chemical analysis'. In addition to the problem of gaining access, there was also the problem posed by the complexity of organic structures. These are not simple di- or tri-atomic molecules like $H_2$ or $H_2O$, but large ones containing dozens, even hundreds, of atoms. Classical structure determination dealt with the problems of access and complexity by applying chemical methods and a method of marshalling evidence derived from the basic hypothetico-deductive (H-D) model whose essential features were clearly articulated by Huygens.[9] An important limitation of this method is the necessity of conjecturing hypotheses, conjectures that often overreach the data and, as a consequence, lead to frequent back-tracking and revision of the hypotheses, as the examples will illustrate.

My claim is that the correlation techniques exemplified by the Woodward rules made possible a mode of hypothesis construction that overcomes some of the limitations of the H-D method. Specifically, I will argue that the "guess-and-check" method widely applied in classical structure determination was replaced by an experimentally-driven mode of hypothesis construction in which generic theory is used to build hypotheses about specific molecules from experimentally-determined answers to questions rather than by the confirmation of conjectures. This change made it possible for chemists to "argue more securely" (to use Newton's expression) for structural hypotheses. The question I will try to answer is: what made this change possible?[10]

It is tempting to try to answer this question by appealing to technological advances in instrumentation. But while these advances were obviously crucial, certain developments with respect to theory were also required in order to enhance the security of structural hypotheses. I will ask four questions concerning these developments. First, what is the logic of testing in modern structure determination, at least insofar as it follows Woodward's example? This question is the topic of sections 4 and 5. In the basic H-D model, conjectured hypotheses are

---

[9] Christiaan Huygens, Preface to *Treatise on Light*, 1690.
[10] The reader should note that while I will focus on the Woodward rules, I view the latter mainly as an example of a broader development; I am not trying to defend a priority claim for Woodward.

tested by experimental verification of observable consequences drawn from them. "Explanation" typically takes the form of one-way conditionals with the hypothesis (P) as antecedent and the phenomenon to be explained (Q) as consequent. In structure determination, for example, 'P' is typically a claim about the molecular structure of a chemical substance. A successful test of this claim consists of accurate prediction of observable phenomena deduced from it. In contrast to this "if P, then Q" logic, the Woodward rules have an "if R, then Q if and only if P" logic. Here, 'R' is a set of theoretical principles licensing a biconditional relationship between a theoretical quantity, P, and an observable quantity, Q. In contrast to the H-D method, the claim of interest, P, is here in the consequent. The significance of this change of position is that 'P' can be deduced from the value of 'Q' without having to be conjectured. The security of the deduction is underwritten by the reliability of the biconditional rather than the degree to which P is confirmed by observation. The reliability of the biconditional is established by comparison with a series of observations, which also provides H-D confirmation of the theoretical principles *en passant*. This biconditional logic is crucial for the aforementioned mode of hypothesis construction to become feasible, for it is what allows the scientist to avoid having to start with conjectures.

Second, spectroscopic measurements of parameters of molecular structure are indirect, and hence necessarily theory-mediated. A question one can ask of such measurements is whether the theoretical elements entering into them are playing a *constitutive* versus a merely *heuristic* role.[11] The significance of this question for the security of the results obtained by means of the measurements is that discarding the constitutive elements undercuts the results that presuppose them, whereas discarding the heuristic elements does not. I will argue that there were three principal constitutive elements entering into the Woodward rules: chemical structure theory, as elaborated largely in the 19th century; the chemical classification of bulk matter; and the

---

[11] For the distinction between constitutive and heuristic elements in scientific research, see Smith, G. E. (2010). Revisiting Accepted Science: The Indispensability of the History of Science. *The Monist, 93*, 545-579.

quantum-mechanical principle of the quantization of energy. In contrast, explanations of how chemical bonds are physically realized played a largely heuristic role, if any. This distinction is important for understanding the robustness of structure determination methods and their relative insensitivity to changes in background theories (theories of bonding, for example).

As hinted at in the preceding paragraph, quantum mechanics entered only selectively into the Woodward rules. Given that it is a fundamental theory for chemistry, on what grounds could such a piecemeal utilization of quantum mechanics be trusted? I will argue that there were three principal ways in which the Woodward rules, and similar measures used by chemists, were corroborated: cross-checking by measures based on different constitutive assumptions; chemical techniques, especially synthesis; and the "well-behavedness" of the measure, that is, its stability, convergence with other measures and amenability to increasing precision. The position argued for here is in contrast with arguments *à la* Duhem to the effect that what makes a theory-mediated measure reliable is holistic testing in the framework of a well-developed theory, for the Woodward rules were corroborated without constitutively presupposing such a theory. The second and third questions are the subject of section 6.

The question of corroboration brings us to the fourth question to be addressed here, that of how chemical structure theory is related to fundamental theory. One of the advantages of the correlation techniques is that they allow structural claims to be established without having to address the fundamental character of atoms and molecules. As we will see, the ability to use principles of quantum mechanics selectively allowed chemists to ignore the question of whether chemical structure theory, which was developed largely in the 19th century, is compatible with quantum mechanics. Chemists could thereby bring elements of quantum mechanics to bear on structure determination problems in the absence of a well-developed quantum theory of molecules, or at least of molecules in chemically-relevant environments. The fourth question will be addressed in section 7.

The changes in the logic of testing in structure determination that will be described here are of importance for historians and philosophers of science, as well as for chemists interested in method, for four reasons. First, it is an example of how theory-mediated measurement works, in practice, in the absence of a well-developed theory, in this case a theory of how relatively large molecules interact with light.  Second, it is an example of science that is driven by experiment rather than theory. The Woodward rules were not derived from an overarching theory. They had to be deduced empirically, and their further development was largely data-driven. A third reason is the methodological implications. On the one hand, many scientists "describe the role of evidence in their science as though it were just an application of hypothetico-deductive confirmation."[12] On the other hand, there is a tendency among philosophers of science to view the main challenge of science as one of developing correct theories and to treat the task of marshalling evidence for theoretical claims as secondary. Both views are distortions of actual scientific practice, and this study aims to contribute to correcting the distortions.

While these three reasons concern issues of importance for understanding scientific practice in general, the fourth reason concerns how to understand chemical practice in particular. The Instrumental Revolution made the discipline heavily reliant on physical instrumentation for identifying chemical substances. Getting clear on the logic of testing involved in the correlation techniques developed by Woodward and others should help historians and philosophers of science understand some of "the epistemic challenges, and advantages, that accompany reliance on these instruments."[13] I will argue that an advantage of the method championed by Woodward is that it allowed chemists to argue more securely for structural hypotheses. To show this, I will give a brief overview of the classical approach and then examine Woodward's strategy in detail

---

[12] W. L. Harper (2011), *Isaac Newton's Scientific Method: Turning Data into Evidence about Gravity and Cosmology*, Oxford: Oxford University Press, p. vii.
[13] Hendry, R. F., Needham, P., & Woody, A. I.  (2012). Introduction. In Hendry, R. F., Needham, P., & Woody, A. I. (Eds.) (2012), *Handbook of the Philosophy of Science. Volume 6: Philosophy of Chemistry*. Oxford: Elsevier B. V., p. 17.

as it is developed in his four seminal papers. I apologize in advance for having to go into chemical details that may be unfamiliar to the reader; while I have done my best to make the discussion accessible to a general audience, a certain degree of such detail is unavoidable. Scientific evidence really does lie in recondite details.[14]

## 2. The historical and scientific background of the Woodward Rules

As mentioned above, the fundamental problem for 19th and early-20th century chemists was how to draw inferences about molecular structure without having empirical access to it. The classical approach to structure determination consisted in inferring structure from experiments the design of which presupposed only macrophysical processes. For example, chemical analysis consisted of subjecting the compound to degradative conditions such as boiling in concentrated acid or caustic bases and comparing the properties of the products to those of known compounds. Further confirmation of a structural hypothesis could be obtained by reconstructing the target compound—'synthesizing' it—from known fragments and comparing the bulk properties of the synthetic product to the target. The arguments for a structural hypothesis could be quite complex, for they sometimes had to interrelate an enormous body of chemical data, as the chart in Figure 1 (taken from an early review of the strychnine work) illustrates.

---

[14] Please see the Appendix for a brief primer on structural formulae and the terminology employed in the following discussion.

**CHART VI**

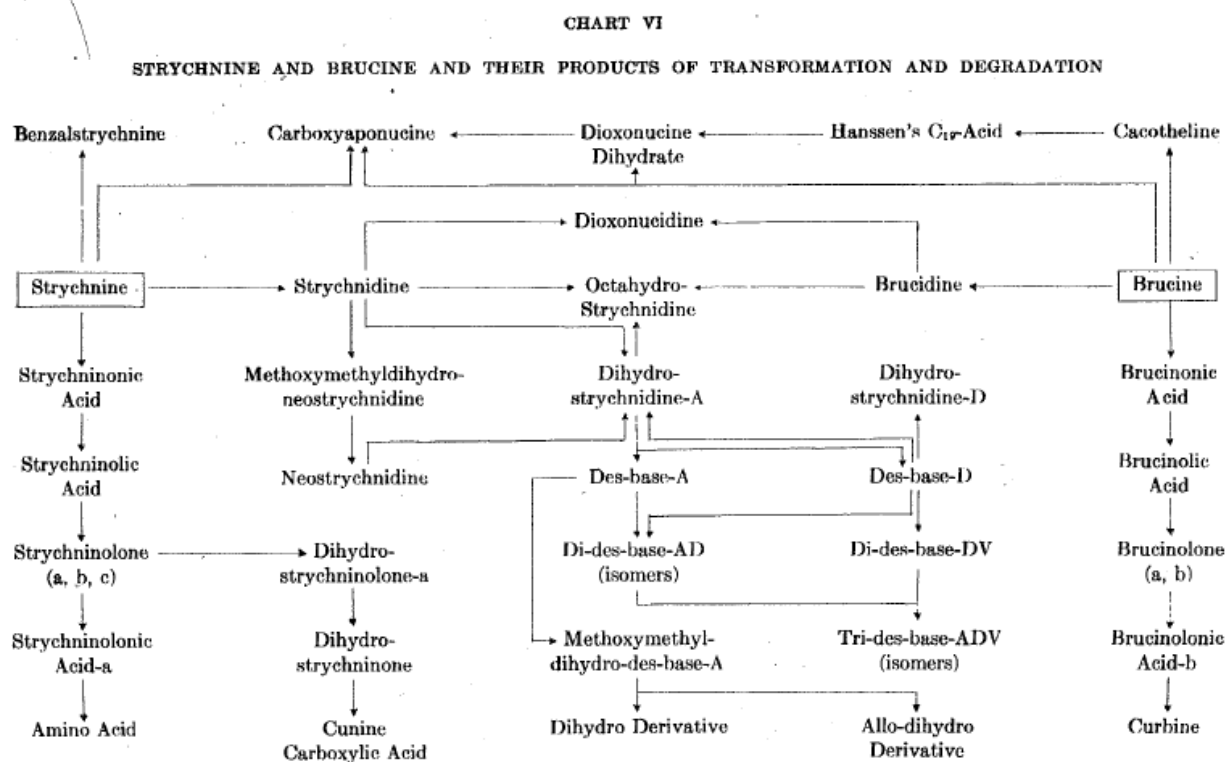STRYCHNINE AND BRUCINE AND THEIR PRODUCTS OF TRANSFORMATION AND DEGRADATION



*Figure 1.* The network of chemical transformations observed in the classical determination of the structure of strychnine and the closely related brucine [from H. L. Holmes, "The Strychnos Alkaloids," in *The Alkaloids: Chemistry and Physiology*, R. H. F. Manske and H. L. Holmes eds., vol. I, (New York: Academic Press, 1950), chapter 7, pp. 375-500.

While for simple compounds one could obtain high-quality evidence for structural hypotheses relatively easily, the determination of more complex structures was extremely time-consuming, often taking decades. Strychnine (**1**, Figure 2), for instance, was isolated in 1815, but the question of its structure was not definitively settled until 1948 despite intensive efforts on the problem: at least 245 papers were contributed to solving it from the time of strychnine's isolation to 1950.[15]

There were significant limitations to the classical approach to structure determination. The major technical limitation, of course, was that chemical means had to be used and these tend to be time-consuming, involving lengthy purifications and experimental processes. The major

---

[15] Slater, L. B. (2001). Woodward, Robinson, and Strychnine: Chemical Structure and Chemists' Challenge. *Ambix, 48*, 161-189. In contrast, complete structural determinations of molecules more complicated than strychnine are now routinely reported in single communications, most often with no subsequent revisions to the proposed structure.

logical limitation was that inherent in the hypothetico-deductive approach. Though I do not pretend to offer a survey of classical structural chemistry, I do propose that much classical structure determination work fit the following pattern. First, hypotheses were put forth about the structural subunits of the molecule. Second, broad, macroscopic chemical consequences were deduced from the hypotheses. Third, chemical experiments were conducted to test for these consequences. Lastly, a hypothesis for the overall structure of the molecule might be proposed as the best explanation for the set of experimental results obtained via the first three steps. Some examples from the strychnine work will be adduced to support this proposal.

As is well known, hypothesis confirmation tends to be tenuous in any case due to the possibility of confounding factors such as unnoticed alternative hypotheses or false background assumptions. As a result, much direct and indirect evidence must be accumulated before a complex hypothesis such as the structure of strychnine can be accepted with confidence. That is, a hypothesis about the overall structure of a molecule can easily reach far beyond the data providing evidence for it. In chemistry these problems were compounded by the lack of empirical access to molecular structure. All organic chemists had to go on in addressing questions about molecular structure was the coherence of the data as viewed in the light of structure theory. Hence hypotheses had to be accepted on the basis of an inference to the best explanation, with a continuing worry that new data would undermine the explanation. Indeed, each revision represented by the strychnine hypotheses in Figure 2 was forced by the acquisition of new data.[16,17]

Structure **A**, for example, was proposed by W. H. Perkin and Robert Robinson in 1910 on the grounds that it offered the simplest explanation, relative to competing hypotheses, of the

---

[16] For an account of the evolution of the proposed structures of strychnine, see Slater (2001).

[17] I here use the term 'chemical structure theory' in a broad sense to include the conventions for drawing structural formulae, the classification of functional groups, the theory of chemical bonding, the theories of electronic structure in atoms and molecules, and the theory of molecular shape. Later in the discussion I will specify a somewhat narrower sense.

available data: "the…expressions for strychninonic acid [a degradation product of strychnine] and strychnine…seem to us to be in every way satisfactory, since they account in a comparatively simple manner for all the known reactions of these substances."[18] A key consideration in tentatively accepting this structure was the size of ring 3, which contains six atoms, whereas the chief competing hypothesis considered by Perkin and Robinson had a seven-membered ring 3, the choice of a six-membered over a seven-membered ring being dictated by prior knowledge that six-membered rings form more readily than seven-membered. (Robinson would subsequently become a key player in the strychnine structural studies and a future Nobel prize winner). 20 years later, Robinson justified the adoption of hypothesis **B** in 1928 in terms of its ability to "illustrate" reactions of strychnine that had become known since 1910:

> The *iso*quinoline structure (Fawcett, Perkin, and Robinson, Part VII, J., 1928, 3082) was advanced mainly in order to illustrate a mode of symbolization of the reactions of strychnine which had been developed since 1912; especially, it was desired to emphasise the ether nature of the bases, the possession of one ethylenic double bond, and a theory…of the course of the degradations to strychninolone and curbine brought to light by Leuchs and his collaborators.[19]

Thus, structure **B** contained the new information that the ring 7 oxygen belongs to an ether (-C-O-C-), rather than a hydroxyl (-C-O-H) functional group; that the molecule contains only one double bond, on ring 7, rather than two on ring 4; and that strychnine possesses an isoquinoline rather than a quinoline ring system [see below (Figure 3) for the distinction between isoquinoline and quinoline ring systems; it later turned out that strychnine possesses neither]. Two years later he argued for a revision of that structure according to which ring 5 was now a bridged rather than a fused six-membered ring, and there was a return to a quinoline ring system (structure **C**). He framed his argument as an inference to the best explanation: "the new structure (VIII) [bearing a bridged ring] has many advantages over earlier suggestions, and can be justified experimentally in every detail…[seven points bearing on the experimental evidence follow]…[f]or the above

---

[18] Perkin, W. H., & Robinson, R. (1910). Strychnine, Berberine, and Allied Alkaloids. *J. Chem. Soc., 97,* 305-323.
[19] Menon, K. N., Perkin, W. H., & Robinson, R. (1930). Strychnine and Brucine. Part X. The Degradation of Dinitrostrycholcarboxylic acid: its Recognition as a Derivative of Quinoline and the Consequent Modifications of the Constitutional Formulae for the Strychnos Bases proposed in Part VII. *J. Chem. Soc., 133,* 830-843 (p. 832).

and other reasons, I consider that the structure (VIII) is almost certainly correct."[20] Subsequently, however, it became apparent that the quinoline system was actually a dihydroindole system, meaning that the ring fused to the benzene ring is five- rather than six-membered (see Figure 3). This, together with the identification of carbazole among the degradation products, the desire to include a tryptophan skeleton in order to relate the biogenesis of strychnine to that of other alkaloids, and the hypothesis of a blocked dihydroindole skeleton (assumed on experimental grounds) lead Robinson to propose structure **E** in 1932. Subsequent findings on the bromination of oxonucidine (a related compound; see the chart in Figure 1 above), however, seemed to contradict the blocked dihydroindole hypothesis because it was claimed (mistakenly, it turned out) that oxonucidine could be brominated at the 3-position of its dihydroindole moiety. These findings lead Robinson to propose an alternative structure (**D**) in the same year on the grounds that it "best expresses the whole of the known chemical behaviour of strychnine and its derivatives."[21] Structure **D** retained the dihydroindole substructure, but the latter was no longer blocked. Though he would revert to structure **E** in 1939, in 1946 Robinson proposed the correct structure of strychnine (**1**) because it "exhibits a more direct relation with cinchonine [a proposed biogenetic relative] than II [**E**], and a more direct explanation of the formation of β-collidine on degradation."[22] Thus Robinson proposed the correct structure in an attempt to provide a simpler explanation of a biogenetic relation as well as of an observed degradation product.[23]

---

[20] Menon, K. N. et al. (1930), pp. 833-834.

[21] Blount, B. K. & Robinson, R. (1932). Strychnine and Brucine. Part XX. Some Derivatives of Pseudostrychnine. Comments on a Recent Memoir of H. Leuchs—Über Strychnos-Alkaloide, LXVIII. *J. Chem. Soc., 135*, 2305-2310 (p. 2308).

[22] Briggs, L. H., Openshaw, H. T., & Robinson, R. (1946). Strychnine and Brucine. Part XLII. Constitution of the neo-Series of Bases and their Oxidation Products. *J. Chem. Soc.,* 903-908 (p. 907).

[23] Please see Figure 14 (Appendix) for structural formulae corresponding to the intermediates mentioned in this discussion.
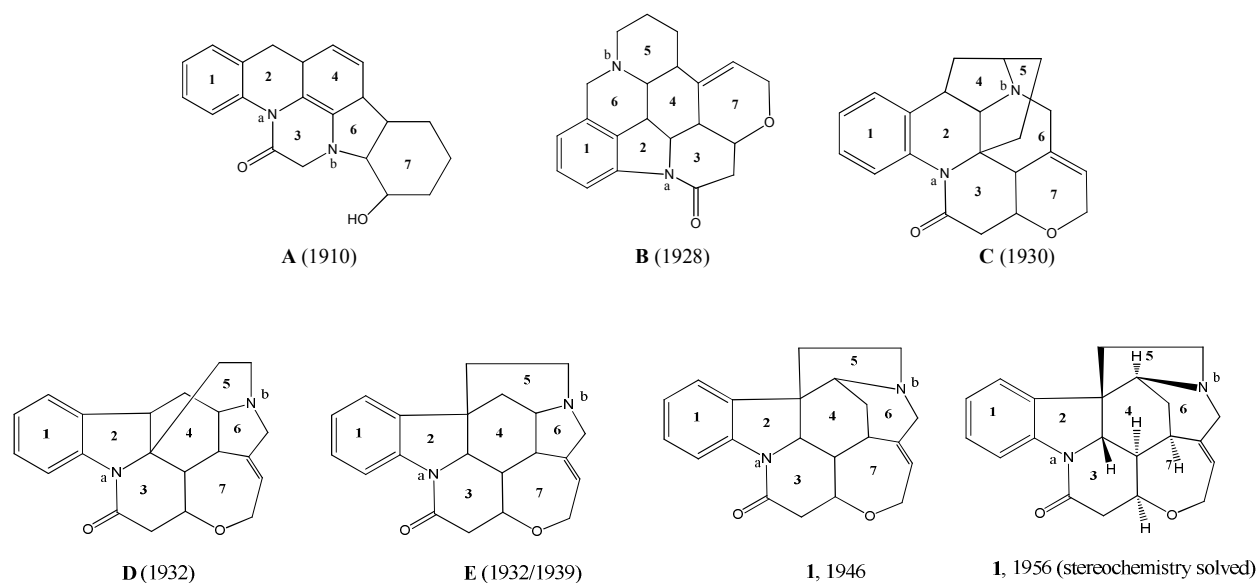
*Figure 2.* Strychnine hypotheses over the years. The numbering of **1** follows the scheme proposed in R. B. Woodward and W. J. Brehm, "The Structure of Strychnine. Formulation of the *Neo* Bases" (1948), *J. Am. Chem. Soc.* 70: 2107-2115  The rings of the other structures are numbered by analogy with this scheme.

As a further example I will briefly mention a puzzle in the strychnine saga that took 40 years to solve. The puzzle was the identity of the ring containing what was referred to as the "a-nitrogen" of strychnine. As the structures in Figure 2 show, up through 1930 it was believed that strychnine possessed what is known as a "quinoline" substructure (**G**, Figure 3), which consists of a benzene ring fused to a six-membered ring containing a nitrogen atom. The hypothesis that the a-nitrogen is contained in the quinoline substructure was first proposed by Julius Tafel in 1891 and 1898 and was based on analogies between the chemical behavior of strychnine and its derivatives and that of known compounds such as *N*-methyltetrahydroquinoline and dimethylaniline (Figure 3).[24] In 1928 further chemical evidence suggested that this hypothesis should be revised so that the b-nitrogen would now reside in the 6-membered ring, yielding an isoquinoline, and the a-nitrogen would reside in an adjacent 5-membered ring. In 1930 the hypothesis that the a-nitrogen resides in the quinoline system returned to favor, only to be discarded once more the following year in favor of the hypothesis that it is part of an "indole"

---

[24] a) Tafel, J. (1891). Über Strychnin. *Ann,. 264*, 33-84. b) Tafel, J.  (1898). Über Strychnin. Dritte Abhandlung. *Ann., 301*, 285-348. See also Robinson, R. et al. (1910).

system (Figure 3).[25] This latter hypothesis remained unchallenged, and as Figure 2 shows all subsequent proposals for the overall structure incorporated the indole substructure. Much of the controversy over the position of the a-nitrogen turned on the constitution of dinitrostrycholcarboxylic acid (**L**, Figure 3), a degradation product of strychnine that was alternately thought to be a quinoline-, isoquinoline- or indole-like system. From the standpoint of modern chemistry, it is astonishing that the elucidation of such a relatively simple structure would require 40 years.
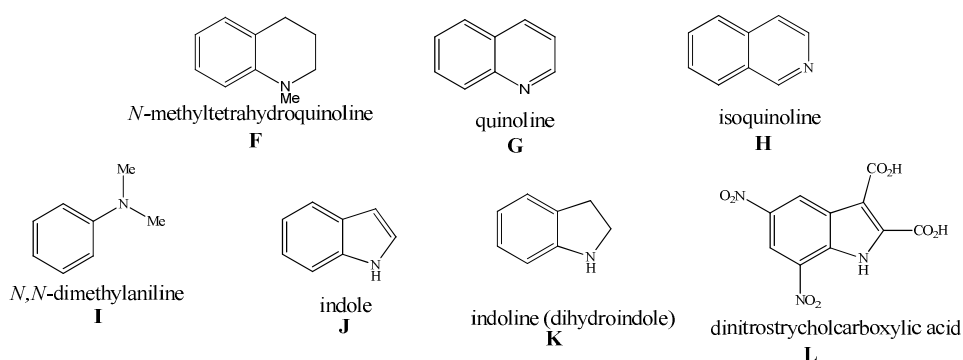


*Figure 3.* Quinoline, indole and related systems.

As the strychnine story illustrates, when classical methods of structure determination are employed many chemical reactions have to be known before a complex structural hypothesis can become secure. Until then, the hypothesis is at the mercy of further facts that may lead to the discovery of hitherto unnoticed alternative hypotheses. The major difficulty had to do with the asymmetry of hypothesis confirmation that is characteristic of the HD method. As is well known, it is possible to falsify a hypothesis (relative to background assumptions) by deduction, but it is not possible to *confirm* one by deduction, even relative to background assumptions. The asymmetry of falsification and confirmation arise from the logic of hypothesis testing. Since the latter tests sentences of the form "if H, then O" where H is the hypothesis and O is an observable consequence, then the failure to observe O entails that H is false via the contrapositive "if not O,

---

[25] Menon, K. N., & Robinson, R. (1931). Strychnine and Brucine. Part XII. The Constitution of Dinitrostrycholcarboxylic Acid. *J. Chem. Soc., 134,* 773-777.

then not H." In contrast, no such entailment relation between H and O exists if O *is* observed, at least if all that theory can afford are one-way conditionals with the hypothesis as antecedent and the observation as the consequent.

If, on the other hand, theory can be used to derive conditionals with the observation as antecedent and the hypothesis as consequent, then the logical force of the observation is much stronger, for it now entails the hypothesis. Classical structure theory was largely unable to do this. Classical chemists had great difficulty using the theory at their disposal to derive structural hypotheses from the phenomena that served them as observations. For the most part, there was no way of deriving structure from the gross chemical and physical behaviors of compounds. As far as theory was concerned, for example, a given melting point was, in and of itself, compatible with an indefinite number of possible structures. The only way to reduce the number of candidates was to rely on the melting points of known compounds. The methods of measurement were developed independently of assumptions about molecular structure and largely pre-dated the rise of structure theory.

It is my contention that the importation of physical instrumentation into chemical structure determination enabled chemists to derive structural information from observation in a systematic and general way. The importance of these instruments in this regard is that they provided phenomena that could be systematically connected to structure by the joint application of structure theory and quantum mechanics. The papers disclosing the Woodward rules demonstrated how this could be done in a remarkably clear, simple and effective way. Before examining those papers, then, I will give a brief overview of the use of physical instrumentation in organic chemistry.

**3. The Development of UV Spectroscopy in Organic Chemistry**

*R. B. Woodward and the Logic of Testing in Organic Chemistry*

The instrumental methods imported into organic chemistry in the mid-20th century are collectively known as spectrometry.[26] Most spectrometric methods are forms of spectroscopy, the experimental evaluation of the manner in which a substance interacts with electromagnetic radiation. More specifically, spectroscopy measures the energy differences between allowed quantum states of a molecule by determining the wavelengths or frequencies of light absorbed or emitted by a sample (Figure 4). Absorption spectra are often characterized by the 'absorbance' *A* at a given wavelength; this quantity is defined by

$$A \equiv \log (I_0/I)$$

where *I* is the intensity of the light transmitted by the sample and $I_0$ is the intensity of light incident on the sample. When the sample is in solution and a cell must be used, $I_0$ is taken to be the intensity of light transmitted by the cell filled with pure solvent.

The absorbance is related to the path length *d* of the sample and the concentration *c* of absorbing molecules by the Beer-Lambert law,

$$A = \varepsilon c d$$

where the proportionality constant $\varepsilon$ is called the 'absorption' or 'extinction' coefficient. The absorption coefficient is a property of the absorbing material that varies with wavelength in a characteristic manner. The measurement of the absorbance along a continuum of wavelengths gives a characteristic curve, such as that shown for mesityl oxide in Figure 4. In the papers discussed here, the wavelength at which the absorbance reaches a maximum is called the 'wavelength of maximum absorption.'

The Woodward rules are concerned with UV spectroscopy in particular. The range of wavelengths of light of about 200-400 nm is called the ultraviolet; the visible region of the

---

[26] Another instrumental method that became important but that is not spectrometric is X-ray diffraction analysis. The latter identifies the location and nature of each atom in a molecule by studying the molecule's X-ray diffraction pattern. This technique is limited to crystalline substances, however, so modern structure determination generally relies more on spectrometry.

16

spectrum ranges from wavelengths of about 400 nm (violet light) to about 750 nm (red light). By the time of Woodward's research in this field, a significant amount of precedent existed for linking the UV spectrum of a compound with its structure. In the late 19th century, researchers studying the infrared spectra of organic molecules had correlated absorption bands with the presence of certain functional groups in the molecules.[27] In 1899, Hartley used UV spectroscopy to study the issue of tautomerism.[28,29] In the first two decades of the 20th century Henri had studied numerous natural products including strychnine, chlorophyll and cholesterol and had shown that the UV spectrum of an organic compound was characteristic of certain groups of atoms rather than a function of the whole molecule.[30] While he was able to correlate spectral features with functional groups and the spatial relations between functional groups, with one exception I am not aware of any work of Henri's correlating spectral features with the substitution pattern of functional groups—an important distinction, since it is the latter kind of correlation that gives access to individual bonds.[31] In 1940, immediately prior to the disclosure of the Woodward rules, Booker, Evans and Gillam published a classification correlating the spectral features of conjugated dienes with their substitution pattern.[32, 33]

---

[27] a) Abnery, W., & Festing, E. R. (1887). On the Influence of the Atomic Grouping in the Molecules of Organic Bodies on Their Absorption in the Infra-Red Region of the Spectrum. *Phil. Trans. 172*, 887-918; b) Julius, W. H. (1892). *Verhandl. Kanickl. Acad. Wetenschappen Amsterdam* I, 1.

[28] Tautomerism is the existence of a compound as an equilibrium mixture of isomeric forms. For example, aldehydes and ketones exist in solution as mixtures of the keto form and the enol form.

[29] Hartley, W. N., & Dobbie, J. J. (1899). A Study of the Absorption Spectra of Isatin, Carbostyril, and their Alkyl Derivatives in Relation to Tautomerism. *J. Chem. Soc., 75,* 640-661.

[30] Henri, V. (1919). *Études de Photochimie*, Paris: Gauthier-Villars.

[31] The exception is the UV absorption of linear alcohols of the form $CH_3(CH_2)_nOH$, which Henri showed varied with the value of *n*. See Henri (1919), p. 41ff.

[32] Booker, H., Evans, L. K., & Gillam, A. E. (1940). The Effect of Molecular Environment on the Absorption Spectra of Organic Compounds in Solution. Part I. Conjugated Dienes. *J. Chem. Soc.,* 1453-1463.

[33] As mentioned earlier, I am not trying to defend a priority claim for Woodward. Nevertheless, one respect in which I think his work goes beyond the Booker, Evans and Gillam work is his use of theoretical considerations to discover further features of molecular structure, a point that I will elaborate in the sequel. In contrast, Booker, Evans and Gillam offer a purely empirical classification and do not seem optimistic about relating it to fundamental theory: "[C]hange of intramolecular environment often has a very marked effect upon the absorption band due to a particular chromophore even when the chromophore itself remains unchanged. At present, these changes cannot be predicted from fundamental premises, although considerable knowledge of an empirical nature is accumulating" (Booker et al. (1940), pp. 1453-1454).

Also in 1940, Heinz Dannenberg at the Kaiser Wilhelm Institute for Biochemistry in Dahlem had published a monograph in which data for 208 steroids were collated and several numerical relationships were devised. For instance, he noted that an exocyclic double bond produced a slight displacement of the absorption band to the red.[34] In his discussion of the α,β-unsaturated ketones he noted that steroids with similar substituents on the double bond had similar maxima in their respective spectra and that the addition of a methyl group at the β position shifted the maximum absorption eleven microns toward the red. He also showed that a bromine atom attached to the double bond had a similar effect. While Dannenberg's work was a remarkable achievement, he did not seek to show the general nature of his results.[35] In contrast, Woodward showed how his rules are supported by many different compounds (though there is still a preponderance of steroids), and his correlations were further extended by Louis and Mary Fieser in the third edition of *Natural Products Related to Phenanthrene* (1949).
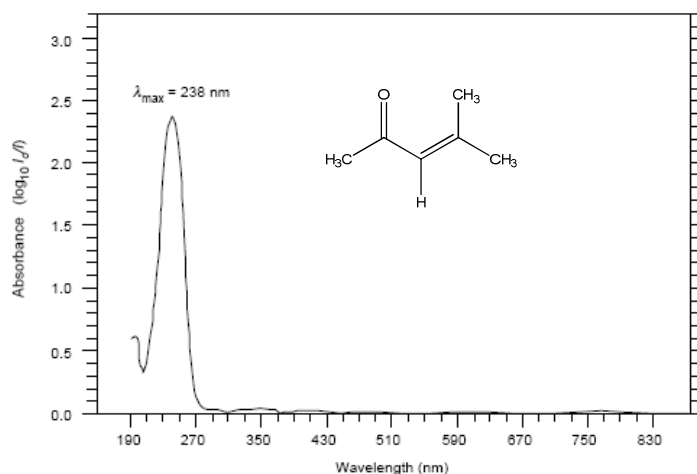


*Figure 4.* Ultraviolet spectrum of mesityl oxide.

UV spectroscopy is an important tool for the study of adjacent (conjugated) multiple bonds, e.g., for compounds having generic structures such as **2** and **3** below:

---

[34] Dannenberg, H. (1940). Über die Ultraviolettabsorption der Steroide. Monograph 21 for 1939 *Abhandlungen der preussischen Akademie der Wissenschaften.* Berlin : Verlag der Akademie der Wissenschaften.
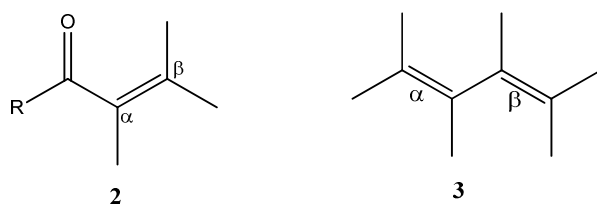[35] Such, in any case, is Benfey and Morris's (2001) assessment (p. 45).

*Figure 5*. A generic α,β-unsaturated ketone (**2**) and conjugated diene (**3**). 'R' denotes a functional group whose atom of attachment is carbon.

Structures such as **2** are called 'α,β-unsaturated ketones,' where 'α' and 'β' denote the position of the carbons in the carbon-carbon double bond. The substituents attached to the α and β carbons are represented here by lines.[36] Structures such as **3** are called 'conjugated dienes.' I have labeled each double bond in **3** also as either α or β for I will be dealing later with unsymmetrically substituted dienes. The Woodward rules concern the UV absorption spectra of molecules with structures **2** or **3**.

The basic principle of UV spectroscopy is that the electronic energy of a molecule can be changed if the molecule absorbs a quantum of light. The energies required for such electronic transitions are of the magnitude of bond strengths. The usual explanation for the absorption of UV light by conjugated molecules (though one neither explicitly affirmed nor denied by Woodward himself) is that the transition involves the excitation of an electron from a bonding to an antibonding state or 'molecular orbital' (Figure 6). The energy of light in the ultraviolet region is in the range of electrons occupying what are known as 'π'-molecular orbitals, and it is precisely these electrons that are involved in conjugated multiple bonds. The π-molecular orbitals of such conjugated systems extend over several atoms, and hence the energy differences between the orbitals are affected by the structure of the molecule. I will return to Woodward's own views later. In any case, both he and the usual explanation come to the same conclusion, that the  wavelength at which an absorption occurs carries information about molecular structure.

---

[36] For most of the cases examined by Woodward, the atom of attachment was carbon.
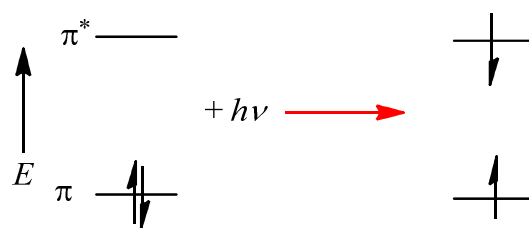
*Figure 6.* Ground and excited states. The single-headed arrows represent electrons.

UV spectroscopy was especially useful in the rapidly growing but highly complex field of steroid chemistry in the 1930s and 1940s. Chemists were having great difficulty determining the complete structures of steroids at the time, but the fact that steroids often contain structural motifs such as **2** and **3** meant that UV spectroscopy could be brought to bear in determining their structures. Woodward published his pioneering work on UV spectroscopy in 1941 and 1942. While the first paper, published in April, 1941, is entitled simply "Structure and the Absorption Spectra of α,β-Unsaturated Ketones", the following three papers bear the common title "Structure and Absorption Spectra" followed by a number (II-IV) and the subtitle of the specific paper. The four papers form a unit, with each one developing the method in a particular way. The common title hints at the possibility that the method developed here in the context of UV spectroscopy may be applicable to other forms of spectroscopy, as does Woodward's talk of "the power of the absorption method." In the following section I will go into a somewhat detailed analysis of these papers to support my argument that Woodward's method is novel. Those readers less interested in the technical details can skip to the summary on p. 28.

**4. The Woodward Rules**

The central idea of the April, 1941 paper is that one-to-one correlations can be established between the wavelength regions of maximum absorption of α,β-unsaturated ketones and the degree of substitution at the α and β positions.[37] Previous investigators had indicated that

---

[37] Woodward, R. B. (1941a). *J. Am. Chem. Soc., 63,* 1123-1126.

"in general the presence of an α,β-unsaturated carbonyl system in a given compound could be

correlated with intense selective absorption in the region 230-250 mµ."[38] Woodward continues:

> In this communication it is shown that the position *within this region* of the wave length [*sic*] of maximum
> absorption may be rigidly correlated with the extent of substitution of the carbon-carbon double bond in the
> α,β-unsaturated carbonyl system and that consequently the determination of this physical property throws a
> very considerable light upon the structure of the compound under examination.[39]

Woodward analyzes the literature values for the wavelength maxima of all the α,β-unsaturated

ketones of known structure whose absorption spectra had been measured to date, involving a

total of 45 sources. The data concern three subclasses of α,β-unsaturated ketones: α- or β-

substituted (**4** or **5** in Figure 7), αβ- or ββ-substituted (**6** or **7**), and αββ-subtituted (**8**). Since the

literature values had been measured in a variety of solvents, the data are corrected for solvent

effects so that the observed wavelength in any of the four solvents indicated in Figure 7 can be

"converted" into the wavelength in ethanol. The corrections are averages calculated from
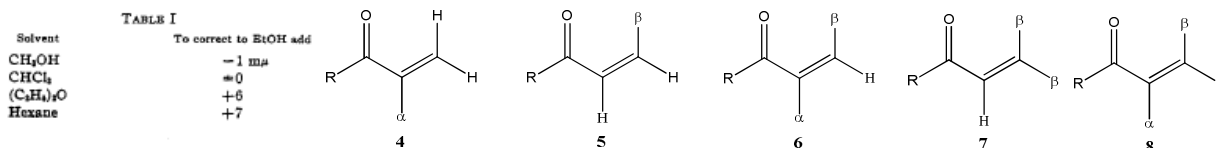
experimentally measured values.



| Solvent | To correct to EtOH add |
|---|---|
| CH₃OH | −1 mµ |
| CHCl₃ | ±0 |
| (C₂H₅)₂O | +6 |
| Hexane | +7 |

TABLE I

*Figure 7.* 'α' and 'β' represent substituents attached to the α and β carbons, respectively.

The solvent corrections are significant because they allow Woodward to establish the

one-to-one correlations. The previous investigators were able to discover a gross correlation

between the presence of an α,β-unsaturated carbonyl in a compound and intense selective

absorption in a region. In the absence of the corrections, however, it was difficult to detect any

patterns within this region. Correction for solvent effects generated new "data" that support a

refined correlation between position within this region and the degree of carbon-carbon double-

bond substitution.

---

[38]Woodward (1941a), p. 1123. A 'mµ' or millimicron is a unit of length equivalent to 1 nanometer (nm).
[39] Woodward (1941a), p. 1123.
[41] Woodward (1941a), p. 1124.

The end result of the analysis is Table V, a concise listing of the correlations:

TABLE V

| Substitution | $\lambda_{max.,}$ m$\mu$ |
|---|---|
| $\alpha$ or $\beta$ | $225 \pm 5$ |
| $\alpha\beta$ or $\beta\beta$ | $239 \pm 5$ |
| $\alpha\beta\beta$ | $254 \pm 5$ |

*Figure 8.* The Initial Woodward Rules.

An important feature of the table is that there is no overlap in the $\lambda_{max}$ ranges, a fact that allows a one-to-one correlation to be established. Woodward was now in a position to specify what he meant by a "*rigid* correlation" in the passage quoted above: "the substitution for hydrogen of an alkyl group causes a shift toward the red of approximately 15 m$\mu$, and…the determination of $\lambda_{max}$ reveals unequivocally the extent of substitution of the carbon-carbon double bond in an $\alpha,\beta$-unsaturated ketone."[41] In other words, with these correlations the maximum can be deduced from the substitution pattern of the ketone and the pattern can be deduced from the maximum. Woodward was showing how the phenomena revealed by physical instruments could be used to deduce structural information from observations.

The rules presented in Table V are inductive generalizations, by which I mean that they generalize, from a necessarily limited set of examples, to all members of the class of which they are examples. It should be noted that inductively generalizing in the manner of Table V and the passage just quoted was a bold move, going far beyond the small sets of relatively simple cases used to support the generalizations, and also given the state of molecular quantum mechanics theory at the time.

In the second paper, published in October, 1941 and entitled "Structure and Absorption Spectra. II. 3-Acetoxy- $\Delta^5$-(6)-*nor*-cholestene-7-carboxylic Acid",[42] Woodward sets out to eliminate an apparent anomaly to the correlations established in the April paper. The anomaly was that the action of boiling pyridine on steroid **9** was reported to produce, in addition to

---

[42] Woodward, R. B., & Clifford, A. F. (1941b). *J. Am. Chem. Soc., 63,* 2727-2729.

compound **10**, three compounds with assigned structures **11a**, **b**, and **c** (Figure 9) all having $\lambda_{max}$ $\leq$ 230 nm, a value well outside the region predicted by Table V above for compounds with that substitution type ($\alpha\beta$). Since the correlations are supposed to be one-to-one, this anomaly poses an obvious challenge to Woodward's method. In response, he proposes an alternative structure, **12**, which has the same empirical formula and is consistent with the chemical tests conducted by the original researchers. Furthermore, he argues that the observed UV absorption spectrum is consistent with the reported values for $\alpha,\beta$-unsaturated carboxylic acids like **12**. He concludes his argument by presenting further chemical evidence in support of his proposal.
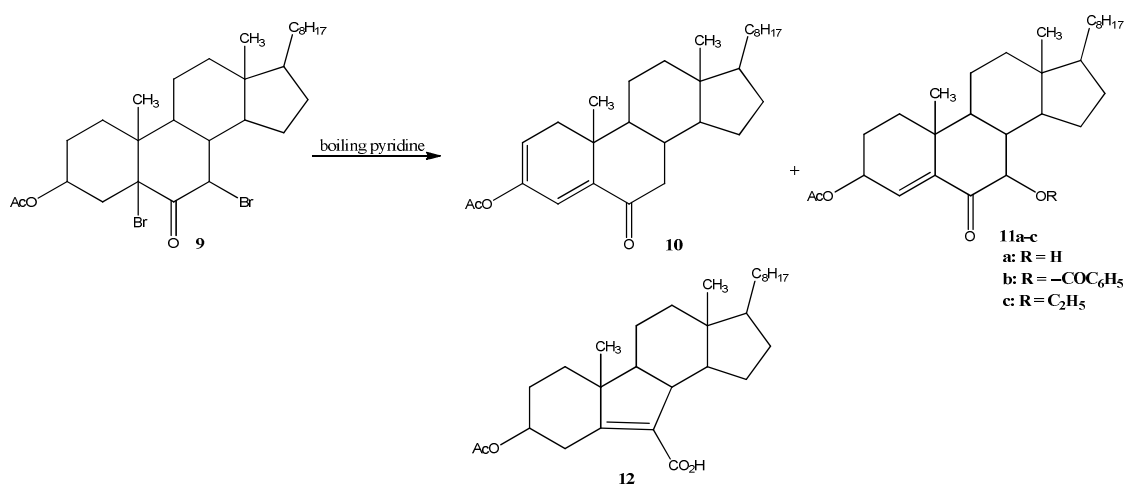


*Figure 9.*

Within the "Structure and Absorption Spectra" series, the second paper is significant because it highlights a key flaw in classical structure determination, namely, that the evidence supporting structural hypotheses tends to be of low quality because the hypotheses are greatly underdetermined by the data. Woodward in effect points out that the chemical data gathered by the original researchers was just as much evidence for structure **12**, which they failed to notice, as it was for structure **11**. This is a typical case of a conjectural hypothesis overreaching what the data call for. By contrast, if true, the correlations of the first paper rigorously exclude **11** as a possible candidate for they say that an $\alpha,\beta$-unsaturated ketone has the substitution pattern of **11** only if its maximum UV absorption lies within the range 239 $\pm$ 5 nm, a condition which the

compound in question does not satisfy.[43] On the other hand, the observed absorption spectrum was consistent with the presence of an α,β-unsaturated carboxylic acid. The point is that one can use the absorption method to make very precise inferences about substitution patterns on the basis of a single datum, $\lambda_{max}$. Hence the second paper seeks to mark the distance between the classical methods and the new method of structure determination that Woodward is championing.

The third part of the series, "Structure and Absorption Spectra. III. Normal Conjugated Dienes", published in January, 1942, extends the absorption method to a new class of compounds, conjugated dienes (**3** in Figure 5).[44] Woodward considers two major classes of such compounds, those in which both double bonds are present in a single ring (such as **13** in Figure 8) and those in which they are not (as in **14-16**). The former class absorbs in the region 255-290 nm, whereas the latter class absorbs in the region 215-250 nm. Woodward proposes that this difference is due to a "first order environmental effect." He then considers that the second class "may be considered normal" and proposes that variations within its absorption region "are certainly the result of second order environmental effects similar to those discussed in connection with α,β-unsaturated ketones in the first paper of this series."[45] The correlations of the first paper cannot be straightforwardly extended to this new class, however, for the data do not appear to support a one-to-one relationship between substitution type and wavelength region of maximum absorption: Some dienes of the same substitution type have significantly different values of $\lambda_{max}$, whereas other dienes of different substitution type have nearly identical maxima. For Woodward, this deviation from biunivocity signals that "positional factors are superimposed

---

[43] Since **12** has the αββ substitution pattern, it may look as if the $\lambda_{max}$ value should lie in the 254 $\pm$ 5 nm range, if Table V is correct. But the reader should note that **12** is a carboxylic acid, not a ketone, and Table V does not apply to the former functional group.
[44] Woodward, R. B., (1942a). *J. Am. Chem. Soc., 64,* 72-75.
[45] Woodward (1942a), p. 72.

upon the expected substitutive effect."[46] By a "positional factor" he means that the position of the double bond in the molecule makes a contribution to the observed maximum over and above the contribution due to the substitution type. Specifically, Woodward proposes that the positioning of a double bond exocyclic to a six-membered ring (as in **16**) increases $\lambda_{max}$ by 5 nm. He provides evidence for this claim by comparing calculated values of $\lambda_{max}$ to observed values for an array of dienes. The calculated values are obtained by taking butadiene (**14**) as the base value and adding 5 nm for each substituent (same effect, though smaller magnitude, as in the first paper)[47] and 5 nm for each exocyclic bond in the diene system.
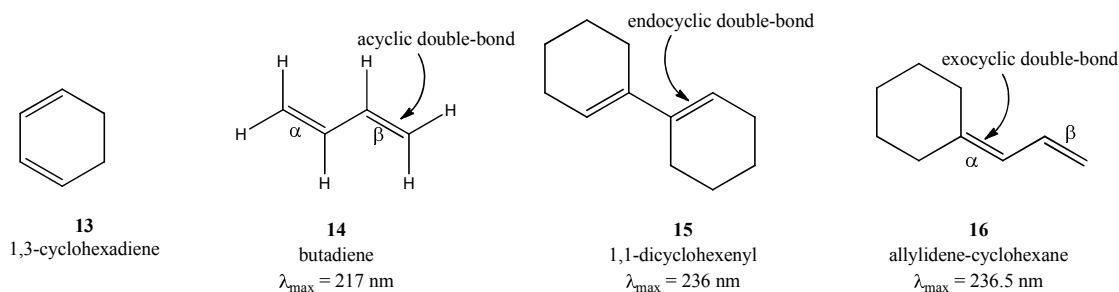


*Figure 10.* Some conjugated dienes

Woodward examines data for two classes of dienes. The first class contains dienes in which both double bonds are symmetrical with respect to the substitutive and positional factors, i.e., they have the same substitution pattern and they are both either exocyclic or not exocyclic (e.g., **14** and **15**). The second class contains dienes in which the double bonds are not symmetrical with respect to the substitutive and positional factors, i.e., they have different substitution patterns and one is exocyclic while the other is not (e.g., **16**). For both classes, the agreement between the calculated and the observed $\lambda_{max}$ is quite close, with an average deviation of 1 nm in the first case and of 3 nm in the second. Woodward also argues that the unsymmetrical dienes should be treated "as composite systems intermediate between two

---

[46] Woodward (1942a), p. 72.

[47] The magnitude is smaller because, in order to compare the conjugated dienes to the α,β-unsaturated ketones, the dienes are considered as averages of two ketones, each ketone bearing one of the two carbon-carbon double bonds of the diene. Hence the shift accompanying the introduction of a single substituent into a conjugated diene is halved.

symmetrical systems, since it is not possible at present to determine whether the small discrepancies in $\lambda_{max}$ (calcd.) are due to experimental variations, or whether…subtle constitutive factors shift the maximum slightly toward that characteristic of one or the other of the component systems,"[48] thus raising the possibility that improvements in measurement precision might reveal that the discrepancies have a clear signature, i.e., that they represent a real phenomenon rather than the vagaries of observation.

The important thing to note is the implicit logic of testing, which is that discrepancies between calculation and observation should be viewed as potentially telling us something about structural details making a difference in the observed spectra. This is typical of the feature of his method that is highlighted in the third paper. He begins the paper by noting a difference between the spectra of two classes of dienes and attributes the difference to a structural feature, namely whether or not both double bonds are present in a single ring. Focusing on one of the two classes, he notes that the correlations between structure and wavelength maxima for that class deviate from biunivocity, and uses this deviation to discover the effect of a further structural detail, the positions of the double bonds with respect to the ring(s) in the molecule. Combining these two features, he proposes a formula for predicting $\lambda_{max}$ for conjugated dienes. This sets him up for the further step of using the discrepancy between the predicted and the observed $\lambda_{max}$ to discover further "constitutive factors" making a difference in the absorption characteristics of dienes.

Two assumptions at work here are: 1) systematic deviations from a strict correlation between structure and wavelength maximum are to be viewed as containing information about structural sources of the deviations rather than as caused by factors exogenous to molecular structure, such as other natural causes or artifacts brought about by the use of highly approximate models of molecular structure; and likewise 2) discrepancies between calculation and

---

[48] Woodward (1942a), p. 75.

observation are clues to further significant structural details. There is a difference here with the basic HD method, for the point of using the Woodward rules to analyze spectra is not to make a prediction that could then be falsified by a comparison with observation, nor only to develop evidence for structures, but to determine whether other effects are superposed on those of the structural variables covered by the rules. Furthermore, the discovery of these second-order effects provides evidence that the initial correlation is not a mere curve-fit, for they show the discrepancies to be structural effects rather than mathematical artifacts. Thus the ability to find structural sources for the discrepancies provides reinforcing evidence for the "theory" being tested.

Woodward's use of discrepancies opens the door to a progressive refinement of the initial correlations, as demonstrated in the fourth paper. In "Structure and Absorption Spectra. IV. Further Observations on α,β-Unsaturated Ketones", published in the same issue as the third paper,[49] Woodward revisits the correlations proposed in the April, 1941 paper in the light of the effect of the exocyclic linkages on the wavelength maxima of dienes. From the data presented in that paper for ketones of substitution type αβ or ββ (Table V, Figure 8) he calculates an average value of $\lambda_{max}$ of 235 nm for those ketones whose carbon-carbon double bond is not exocyclic and an average value of 240 nm for those whose carbon-carbon double bond is. He concludes that α,β-unsaturated ketones are also subject to a positional effect whose magnitude is approximately the same as in the case of the dienes. Woodward uses the discovery of the positional effect in ketones to make the classification proposed in the first paper more precise. For αβ- or ββ- as well as for αββ-unsaturated ketones, the discovery of the positional effect generates two subclasses, a subclass with no exocyclic bond and a subclass with one exocyclic bond:[50]

---

[49] Woodward, R. B., (1942b). *J. Am. Chem. Soc., 64*, 76-77.
[50] While it is possible for monosubstituted ketones to have an exocyclic double bond, Woodward considers these too unstable for spectroscopic analysis and hence does not include them in the revised table.

TABLE I

| Substitution | | | Most probable $\lambda_{max.}$ |
|---|---|---|---|
| $\alpha$ or $\beta^a$ | | | 225 m$\mu$ |
| $\alpha\beta$ or $\beta\beta$ | i. | no exocyclic bond[b] | 235 |
| | ii. | one exocyclic bond | 240 |
| $\alpha\beta\beta$ | i. | no exocyclic bond | *ca.* 247 |
| | ii. | one exocyclic bond[c] | 252[d] |

*Figure 11.* The revised Woodward rules

Not only is the new classification more precise, it also allows Woodward to predict $\lambda_{max}$ for a subclass of then-unknown ketones, $\alpha,\beta,\beta$-substituted ketones having no exocyclic carbon-carbon double bonds ($\alpha\beta\beta$-i in the table).

The significance of the fourth paper within the series is not merely that it refines the initial classification. It showcases an important feature of the classification constituted by such correlations, namely its amenability to a process of successive approximations. Woodward takes the correlation from previous research to be true and then uses it to discover new structural details that make a difference in the absorption spectra. These details are incorporated into the classification, extending it to new compound classes and adding more fine-grained correlations to the ones he started with, as well as solving structural problems. The process can then begin again using the refined and extended classification, generating a pattern of cumulative progress in which new correlations are continually being discovered and existing ones become ever more ramified (Figure 12). His procedure amounts to moving beyond the basic hypothetico-deductive method, for what is being tested here is not agreement between theory and observation but rather whether discrepancies between theory and observation carry information about as yet unknown chemical sources. Woodward's correlational "theory" would not be falsified by a discrepancy with observation, but only by a failure to find a robust chemical source for it.

| Substitution | Most probable $\lambda_{max}$ | | Substitution | Most probable $\lambda_{max}$ | | Substitution | | Most probable $\lambda_{max}$ |
|---|---|---|---|---|---|---|---|---|
| | | | α or β | $225 \pm 5$ nm | | α or β | | 225 nm |
| α,β-unsaturated ketones | 130-250 nm | → | α,β or ββ | $239 \pm 5$ nm | → | α,β or ββ | i. no exocyclic bond<br>ii. one exocyclic bond | 235 nm<br>240 nm |
| | | | αββ | $254 \pm 5$ nm | | αββ | i. no exocyclic bond<br>ii. one exocyclic bond | *ca.* 247 nm<br>252 nm |

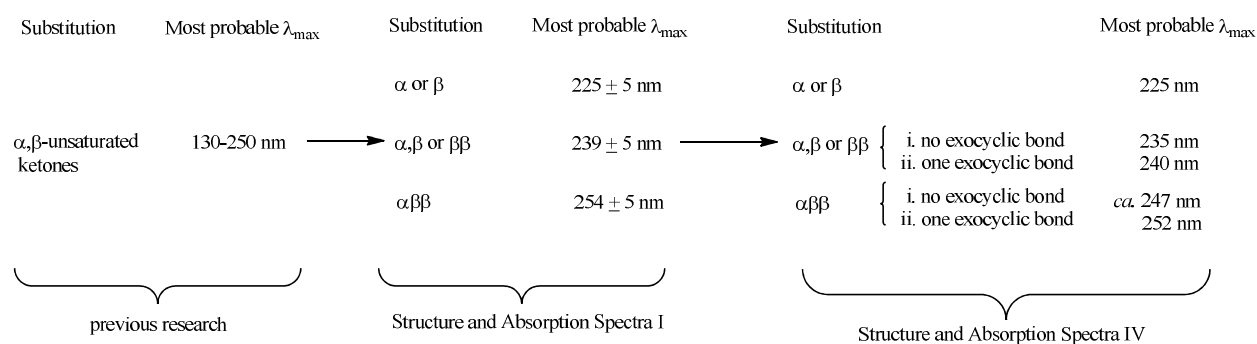previous research  —  Structure and Absorption Spectra I  —  Structure and Absorption Spectra IV

*Figure 12.* Woodward's expanding classification

In summary, the four papers make the following methodological contributions. The first paper shows that the absorption method can be used to gain empirical access to individual bonds. It does so by refining a gross correlation, between absorption wavelength and a type of molecular structure, to afford a set of correlations that can be used to determine whether that structure has one, two or three substituents bound to it. The second paper emphasizes a weakness of the classical method of developing evidence for a hypothesis—hypothesize a structure and then deduce observable consequences from it—by  showing how the evidence thus obtained tends to be unreliable due to the possibility of confounding factors such as unnoticed alternative hypotheses. It also shows how the new method can be used to weed out alternatives. The third paper demonstrates how the Woodward rules can be used to turn data that appear to refute them into evidence for them, by showing how deviations from the rules can be chemically significant. The fourth paper (and also the third) highlights the expansionary logic of the absorption method, showing how an initial classification afforded by the correlation technique can be progressively refined, precisely by pinning down chemical sources of the deviations. Underpinning all four papers is the idea that the data carry information about structure, and can be used as the starting point for deducing structural hypotheses. They thus offer the promise of greatly reducing the need for conjecture.

## 5. Structure Determination Using the Woodward Rules

The Woodward rules have the basic logical form

*R. B. Woodward and the Logic of Testing in Organic Chemistry*

1. If R, then $\Delta M$ when $C_1 \& \ldots \& C_m$ if and only if S

where R consists of background assumptions such as that molecules have well-defined structures, assumptions from quantum mechanics, the theories presupposed in the design of the spectrometer and chemical structure theory; $\Delta M$ is the measured effect (e.g., absorption in a wavelength region); the $C_i$'s are the experimental conditions (e.g., of UV spectroscopy); and S is a structural subunit, say, an exocyclic double bond.

Since more than one structural subunit may correspond to a given wavelength region, the rules are more accurately represented as:

2. If R, then $\Delta M$ when $C_1 \& \ldots \& C_m$ if and only if $S_1$ or $S_2$ or $\ldots$ $S_n$. For example, in the rule that an $\alpha,\beta$-unsaturated ketone absorbs at $225 \pm 5$ nm if and only if it is $\alpha$- or $\beta$-substituted, '$\Delta M$' is the absorption value, '$C_1 \& \ldots \& C_m$' are the conditions of the UV spectroscopic experiment, and the disjunction of '$S_i$'s' is '$\alpha$- or $\beta$-substitution.'

This logical framework gives rise to a form of progress that consists of reducing the number of possible $S_i$'s that can be correlated with a particular measurement by increasing measurement precision, improving the physical theory presupposed in the design of the instrument or removing sources of systematic error. Conversely, new $S_i$'s can be discovered as resolution is improved and sources of systematic error are removed. A second form of progress consists in multiplying the different types of measurement. If one measurement cannot be used to decide between alternative structural subunits, another one that is sensitive to the difference between them can be employed. Alternative measurements can also be used to produce converging

evidence for a structural hypothesis. Thus the different measurements made possible by physical instrumentation can be used to cross-check and complement each other.

A very important feature of the techniques pioneered by Woodward is that they allow for the correlation of spectral values not only with structural subunits but also with connections and interactions *between* subunits. For example, the positional effect studied in the third and fourth papers allows one to determine whether the carbon-carbon double bonds of conjugated systems are connected to rings. Thus the correlation techniques also produce biconditionals of the form:

3.  If P, then $\Delta M$ when $C_1 \& \ldots \& C_m$ if and only if $S_1$—$S_2$ or $S_3$—$S_4$ or $\ldots S_m$—$S_n$, where the em dash is a shorthand for something like 'and is connected to a.' For example, in the rule that an $\alpha,\beta$-unsaturated ketone absorbs at 240 nm if and only if it is $\alpha,\beta$- or $\beta,\beta$-substituted with one exocyclic bond, the '$S_m$—$S_n$' term can be spelled out as 'the molecule is an $\alpha,\beta$-disubstituted $\alpha,\beta$-unsaturated ketone, the alkene component of which is connected exocyclically to a ring.'

This is the form of correlation proposed in Table I, Figure 9 above. Taken together, these rules allow chemists to determine structure from spectra via a "layering" approach. First, hypotheses of structural units are deduced from spectra via rules (2). Second, the connections between these units are deduced with the help of rules (3). The molecular formula together with rules for assessing the relative stabilities of proposed structures (such as the "octet rule") greatly constrain the set of plausible candidate structures that can be generated from the application of (2) and (3). Though there are many compounds for which UV spectroscopy does not provide useful structural information, with more general methods, such as nuclear magnetic resonance, it is often possible to determine the structure of a compound from a single spectrum and the molecular formula.

In what way do the modern methods of structure determination allow contemporary chemists to argue more securely for structural claims than their classical counterparts? As shown above, one risk in the classical approach was that the content of conjectural hypotheses about molecular structures could reach well beyond the data providing evidence for them. Woodward attempts to eliminate this risk by constructing a "theory" that consists of law-like relationships between variables arrived at by inductively generalizing over the data. Since these relationships are biconditionals, the inference of structural features from spectral features, and vice-versa, follows deductively. His approach therefore represents an attempt to confine risk to the inductive generalizations.[53] When this approach is applied to the determination of molecular structure, it has the effect of limiting the content of structural hypotheses to no more than the data clearly call for. The chemist can build up a hypothesis for the overall structure of the molecule piece by piece, with each piece deduced from observation via one of the correlations.

A contemporary example of the layering approach is shown in Figure 13. In this case, the researchers were able to "assemble" the structure of menthol from fragments deduced from nuclear magnetic resonance (NMR) data. The subunits of each fragment were deduced using correlations between absorption frequencies and structural subunits, and then the connections between the subunits and, later, the fragments, were deduced from correlations between the frequencies of the subunits.

---

[53] Newton arguably took the same approach to confining risk in the *Principia*. For a discussion, see Smith, G. E. (2012). How Newton's *Principia* changed physics. In Janiak, A., & Schliesser, E. (Eds), *Interpreting Newton.* Cambridge, U.K.: Cambridge University Press, 360-396.
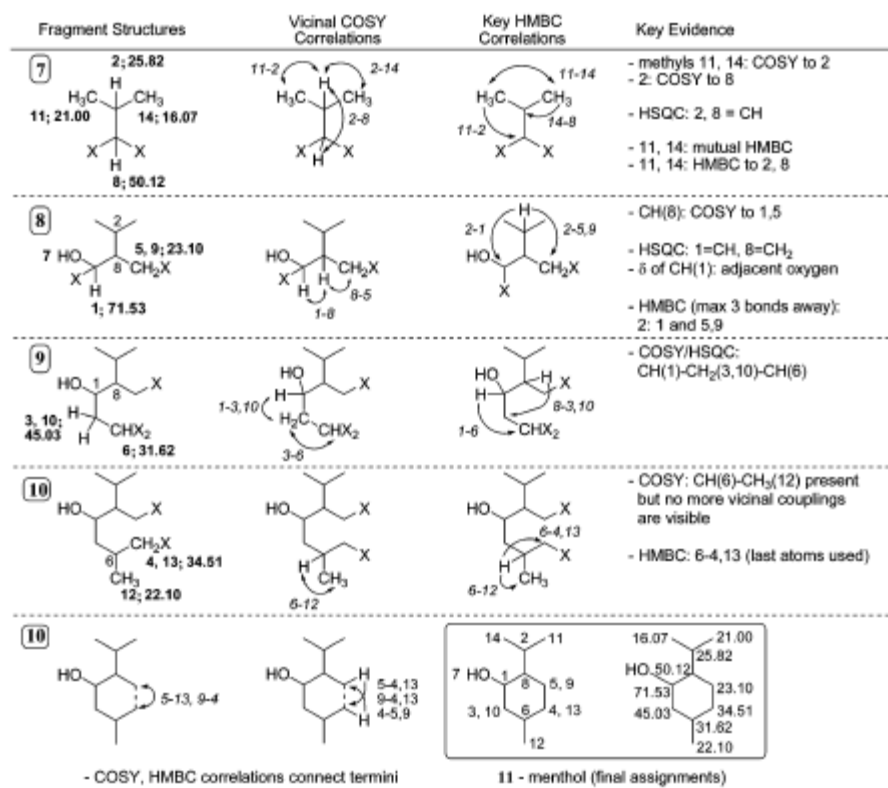
Fragment assembly for menthol. Numbers denote proton identifiers and carbon chemical shifts.

*Figure 13.* Structure elucidation of menthol using NMR spectroscopy. Reproduced from Kwan, E. E. & Huang, S. G. (2008). Structural Elucidation with NMR Spectroscopy: Practical Strategies for Organic Chemists. *Eur. J. Org. Chem.*, 2671-2688 (p. 2677).

The final structure is constructed out of answers provided by the spectroscopic experiments listed in the "Key Evidence" column to questions about the components of the molecule and the connectivities between them. At no point in the argument was recourse made to gross chemical and physical behavior to justify the assignments.

This progress was, of course, made possible by the development of tools of measurement whose design presupposes processes occurring at the micro- rather than the macrophysical level. The new instrumentation did not, however, in and of itself entail a change in the logic of testing. For example, spectra could have been used merely as "fingerprints" for the comparison of newly isolated compounds with known ones, much as bulk physical properties were used.[54] The use of phenomena that measure parameters of the hypotheses that explain them is a vital ingredient for

---

[54] Early practitioners of spectroscopy used it in precisely this way. See Morris & Travis (2002), p. 65.

a successive approximations approach such as that shown in Figure 12, for only on this condition do deviations from theory carry information that can be exploited in developing a more accurate successor theory.

The main risk with Woodward's approach is the discovery of a clear counter-example to the initial inductive generalization. The correlations for mono- and trisubstituted α,β-unsaturated ketones of the first paper were based on relatively small samples, yet in formulating his rules Woodward assumes that the compounds in his data set are pure enough examples of the substitution types of interest to represent those types as a whole.[55] The data sets were quite limited, both quantitatively—six and nine compounds were used to establish the correlations for the αβ/ββ- and αββ-substituted ketones, respectively, in the first paper, 11 and 18 for the symmetrical and unsymmetrical dienes in the third paper—and qualitatively, that is, in terms of diversity: the generalizations were based on relatively simple compounds, many of them steroids. One might legitimately ask why Woodward did not try to expand the data set before generalizing and applying the correlations to new classes of compounds. The risk was that a counter-example would show the correlations to be artifacts of limited data.

Such a discovery would undermine at least part of the evidential reasoning predicated on the correlations. For example, if a conjugated diene were discovered that violated the one-to-one relation between substitution type and wavelength maximum but did not have a compensating secondary effect, then the reasoning that supported the hypothesis of the positional effect would be undermined because it would no longer be certain whether that effect was what was giving rise to the observed deviations from biunivocity. The hypothesis of the one-to-one relationship itself would also be undermined (at least in the case of conjugated dienes) because the positional hypothesis was part of the evidence for it. It follows that all solutions to structural problems that

---

[55] The confidence he places in the generalizations is evident in his remarks when confronted with apparent counterexamples. See for instance Woodward (1941a), p. 1125 and (1941b), pp. 2727-2728.

had been predicated on the correlation would have to be re-evaluated. Hence there was a risk that all the conclusions concerning dienes that were predicated on the Woodward rules might collapse like a house of cards.

Woodward deals with the risk entailed by the inductive generalizations by (a) limiting them to data that are well-behaved. By "well-behaved" I mean a measure that is stable, convergent and amenable to increasing precision. A measure is stable if it gives the same result when the variables are varied. Data are "convergent" when different measures of a quantity give the same value. The significance of amenability to increasing precision is that it is evidence that the measurements are zeroing in on a real underlying quantity. In the case at hand, the variables were the substituents and the solvent. Had the values of $\lambda_{max}$ for each substitution type been less stable as the substituents were varied, one would have been entitled to wonder whether the

---

[58] In 1949, Louis and Mary Fieser presented an overview of the total body of data supporting the Woodward rules as well as well as expanded, refined and revised versions of the rules themselves. Dienes and $\alpha,\beta$-unsaturated ketones not included in the original data sets studied here were shown to fit the Woodward rules, thus supporting Woodward's inductive generalizations. The rule for calculating the $\lambda_{max}$ of conjugated dienes based on the reference value of a parent diene, discussed earlier in the context of the third paper, was refined by basing the reference value on data from several steroids, rather than on the single and overly simple case of butadiene used by Woodward. The rule was also extended to cover trienes, tetraenes, and enol acetates. An analogous rule for calculating the $\lambda_{max}$ of $\alpha,\beta$-unsaturated ketones based on the reference value of a parent compound was proposed and shown to be applicable not only to $\alpha,\beta$-unsaturated ketones but also to the corresponding dienones and 1,4-enediones. Fieser and Fieser also summarize data from studies of the effects of varying the nature of the substituents attached to $\alpha,\beta$-unsaturated ketones. In addition, they discuss successful parallel efforts to correlate infrared absorption characteristics with specific features of steroid structures. In short, the results summarized by Fieser and Fieser either supported the Woodward rules or at least supported the basic approach of correlating spectral characteristics with structural features. Fieser, L. F. &Fieser, M. (1949). *Natural Products Related to Phenanthrene.* New York: Reinhold Publishing Corporation (pp. 184-219).

correlations were picking out a single dominant effect. Furthermore, the ability to eliminate solvent-induced variations by applying a correction provided an additional degree of security, for it provided evidence for a structural regularity underlying the apparent variations in wavelength maximum. Corroboration of the Woodward rules was provided by the use of different spectrometric techniques (though not by Woodward himself) as well as the use of synthesis and traditional chemical techniques. Though Woodward did not try to increase the precision of the measurements he used, in the long term the refinement of the correlations depends on reducing the number of possible $S_i$'s that can be correlated with a particular measurement value, which can be achieved by increasing measurement precision. In general, the history of molecular spectroscopy reveals a tendency towards increasing precision across techniques.

Perhaps counter-intuitively, he also deals with the risk by (b) immediately predicating further research on the correlations rather than trying to improve the initial data set. By assuming that the correlations are rigid, Woodward made them easy to disconfirm in subsequent experiments, thus reducing the likelihood of a garden path. The results documented in the four "Structure and Absorption Spectra" papers constitute a sequence predicated on the hypothesis that the initial correlation between the α,β-unsaturated ketone functional group and absorption wavelength, provided by previous research, is true. This sequence consisted in extending the correlations to new classes of compounds, increasing their precision and eliminating apparent anomalies. Subsequent research by others confirmed and extended Woodward's rules.[58]   Had the initial Woodward rules been mere artifacts of limited data, it is unlikely that this body of research would have succeeded to the extent that it did. In short, the quality of Woodward's data provided an initial test of the generalizations, and these were then retested by the subsequent research predicated on them.

---

[61] Duhem, P. (1991). *The Aim and Structure of Physical Theory* (P. P. Wiener, Trans.). Princeton: Princeton University Press (p. 182). (First published 1906).

**6. The Constitutive Assumptions of the Woodward Rules and Their Corroboration**

In this section I will address the questions posed in the introduction concerning what theoretical elements entered constitutively into the Woodward Rules and how the piecemeal use of quantum-mechanical theory was corroborated. The claim that the method of marshaling evidence for molecular structure in organic chemistry went significantly beyond the H-D method starting in the mid-20th century raises the more general point that the logic of testing in a science is to a large degree dependent on decisions and options concerning measurement. Pierre Duhem made this point forcefully in *The Aim and Structure of Physical Theory* when he contrasted the logic of an experiment in "sciences close to their origins," such as physiology and "certain branches of chemistry," with that of an experiment in physics. According to Duhem, an experiment in physics is more complicated than an experiment in physiology because every experiment in physics presupposes a theory of measurement that is itself a physical theory, for physical theory is necessarily presupposed in the use of instruments and in the interpretation of the data.[61] As a result, Duhem argues, the test of a physical theory is always also a test of the theory presupposed in the instruments. Duhem makes the further point that while the physiologist and the chemist do appeal to theories when they use physical instruments, they appeal to physical theories, not theories in their own domains. They can therefore treat the theories justifying the use of the instruments as beyond the scope of their tests, whereas the physicist cannot.

What is the significance of physics having to provide its own theory of measurement? As is well known, for Duhem it was significant because, he claimed, it entails a holistic logic of testing, the famous thesis that "an experiment in physics can never condemn an isolated hypothesis but only a whole theoretical group." An alternative interpretation, however, is that since the measurements used to support physical theories are themselves mediated by physical theory, success in theory-mediated measurement is a primary form of evidence for physical

theory.[62] More precisely, if the theory of measurement shares assumptions with the theory being tested, then successful measurement—in the sense of providing well-behaved data—is a form of evidence for the theory being tested. A corollary is that successful theory-mediated measurement permits a form of bootstrapping, in which the ability to obtain well-behaved values for theoretical parameters allows evidence to be accumulated for a theory even when the latter is in a rudimentary state, before "a whole theoretical group" is available.

The quantum-mechanical theory of organic molecules was in such a rudimentary state when organic chemists began to analyze spectra for information about structure, and I will argue that Woodward's work is a case of such bootstrapping. Before doing so, however, I will illustrate the alternative interpretation of theory-mediated measurement by considering a related case from physics, namely, the method of analysis of the band spectra of diatomic and simple polyatomic molecules that was developed in the 1930s and 1940s in the wake of the emergence of quantum mechanics.[63] With this method, frequencies of absorption or emission are related to energetic states of the molecule under study through quantized models of molecular motion. From these models equations are deduced representing the allowed transitions between different electronic, vibrational or rotational levels of the electronic states participating in the transition.[64] The equations together with the spectral data permit the deduction of various theoretical quantities, such as the energetic position of the vibrational levels (via the transition frequencies), the

---

[62] Smith (2012), pp. 368-369.

[63] In what follows I will discuss the method as it is illustrated in Gerhard Herzberg's 1950 compendium *Molecular Spectra and Molecular Structure* (second edition). Malabar: Krieger Publishing Company.

[64] For example, the frequencies representing transitions between vibrational levels of two electronic states of a diatomic molecule (neglecting rotational effects) are given by the equation

3.  $v = v_{00} + \omega_0'v' - \omega_0'x_0'v'^2 + \omega_0'y_0'v'^3 + \cdots - (\omega_0''v'' - \omega_0''x_0''v''^2 + \omega_0''x_0''v''^3 + \cdots)$

where $\omega_0'$ and $\omega_0''$ are the vibrational frequencies of the two vibrational levels involved, $\omega_0'x_0'$ and $\omega_0''x_0''$ are anharmonicity constants arising from the fact that a diatomic molecule does not vibrate exactly like a harmonic oscillator, $v'$ and $v''$ are the vibrational quantum numbers associated with the vibrational levels, $v$ is the calculated frequency for a transition between vibrational levels with quantum numbers $v'$ and $v''$ and $v_{00}$ is the frequency of the transition when the quantum numbers are both 0. The quantities related by the equation can be assigned values by comparison with spectra.

[68] Herzberg (1950), pp. 466-472.

vibrational frequencies, terms reflecting the anharmonicities of the vibrations, the force constants, and the electronic energy levels of molecules. Similar analyses can be applied to the rotational frequencies to obtain information about the geometrical arrangement of the nuclei in the molecule, for example very precise values of the internuclear distances. Once calculated, these theoretical quantities can be used to deduce phenomena in other fields, such as thermodynamics, chemistry and astronomy. The deduction of thermodynamic quantities, such as the enthalpy, heat capacity, entropy and free energy of gases, corroborate this application of quantum mechanics because the thermodynamic quantities can often be measured independently of the band spectra by means of direct thermal measurements.[68]

All of this analysis is made possible by applying the Born-Oppenheimer (BO) approximation to molecular quantum mechanics. The BO approximation exploits the large mass difference between electrons and nuclei to separate the nuclear and electronic components of the molecular wavefunction, allowing each component to be solved independently.[69] The approximation was disclosed in 1927, and the ensuing decades of research into molecular structure based on it were remarkably successful. For example, the 1939 edition of the diatomic molecule volume of Gerhard Herzberg's *Molecular Spectra and Molecular Structure*, which became the compendium of the research, references 734 articles, almost all of them post-BO. In the final printing of the 1950 edition, the number had grown to 1611. At the end of the volume there is a 79-page Table 39 that lists structural properties for more than 250 distinct diatomic molecules.

The deduction of theoretical quantities from the observed frequencies was made possible by the theory of measurement employed, which includes the principles that (i) the energy of light is quantized and is measured by the frequency; (ii) the energetic states of a molecule are quantized; (iii) the energies of the light absorbed or emitted by a molecule measure the energies

---

[69] Born, M., & Oppenheimer, R. (1927). Zur Quantentheorie der Molekeln. *Ann. Physik, 84*, 457-484.

of the transitions between these states; (iv) the allowed transitions are governed by selection rules derived from wave mechanics; and (v) the BO approximation is valid. These principles are or involve principles of quantum mechanics, of course, and so the ability to obtain well-behaved values for the theoretical quantities provides a form of "bootstrapping" evidence for quantum theory. More broadly, the success of this research constitutes a form of evidence for the theoretical assumptions involved in the analysis of the band spectra, the reasoning here being that if the hypotheses were false it is unlikely that a research process predicated on them would be successful. "Success" here involves both H-D confirmation of predictions, the ability to obtain well-behaved measurements of properties of interest and the ability to use those results to obtain well-behaved measurements of further theoretical quantities.

If it is plausible that physics having to provide its own theory of measurement allows it to develop evidence through success in theory-mediated measurements, what can be said about other sciences? In particular, what can be said about the methods of organic structure determination discussed in this paper? Duhem claims that "special sciences" like physiology differ from physics in not *having* to employ terms whose connections to the observable world are mediated by assumptions that fall within the scope of their respective theories. But in speaking of physiology and certain branches of chemistry as being close to their origins, he seems to suggest the possibility that these sciences can and perhaps should employ such terms. In the context of this paper a natural question to ask is whether chemists have appealed to chemical theories in the experimental procedures they use to determine the structures of molecules, and if so, should they have?

I will argue that $20^{th}$ century chemists such as Woodward started to presuppose assumptions and principles from chemical structure theory in their measurements (and not just

quantum mechanics, for example). Classical chemical structure theory was developed in the 19th century, in the work of Kekulé, Frankland, Brown, van t'Hoff and others, and has continued to inform the practice of synthetic chemists to this day. Woodward applies it to the study of absorption spectra. At the core of structure theory is what Hendry has identified as the "structural view" of the chemical bond.[72] According to this view, bonds are material parts of the molecule which are responsible for spatially localized and directional submolecular relationships between individual atomic centers.[73] According to this definition, bonds are defined in terms of their theoretical role rather than in terms of how they are physically realized (e.g., by shared pairs of electrons). Their theoretical role is to provide links in the topological structure of molecules and to constrain the spatial arrangements of atoms within them (thereby providing a foundation for stereochemistry). The question of how bonds are physically realized is left to "empirical and theoretical investigation," for example in quantum mechanics, to answer.[73] Some of the basic principles of structure theory are:

4. A bond is individuated by the atomic centers it links.[73]

5. Atoms in molecules can only participate in a fixed number of bonds ('valency').

6. Atoms are assigned the valence characteristic of their element. A structural formula should saturate all the available valences. For example, since carbon has a valence of four, each carbon atom in a structural formula should form four bonds.

7. The network of bonds constituting the structure of a molecule, or a substructure thereof, is invariant (with appropriate exceptions made for such phenomena as fluctuations from equilibrium positions, tautomerism, etc.)

---

[72] Hendry, R. F. (2010). The Chemical Bond: Structure, Energy, and Explanation. In Suárez, M., Dorato, M., & Rédei, M. (Eds.) *EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association.* Berlin: Springer, 117-127.
[73] Hendry (2010), p. 125.

It is the network of bonds that is the primary analytic and synthetic objective of organic chemists. As a result, the latter are more interested in determining which atoms are connected to which, what is the bond order of those connections (single, double, etc.) and what stereochemical features arise from the spatial relationships between bonds than in developing theories of how bonds are realized or in measuring fundamental constants and structural quantities, like bond lengths, to high precision (in contrast to the work of physicists such as Herzberg, for example). In order to gain empirical access to the bonding network, the chemist needs at least three components of a theory of measurement: (i) a parameter that carries information about the bonding network, especially individual bonds; (ii) a way of relating that parameter to an observable property; and (iii) some assumptions about the entities to which access is desired that make the parameter in (i) carry the desired information.

In the "Structure and Absorption Spectra" papers (i) is provided by the energy of the electronic transition. As I will show later, there is evidence that Woodward conceptualizes the transition as that of an electronic oscillator. The observable property in (ii) is provided by the wavelength of light absorbed. The way of relating the energy to the wavelength is provided by principles of quantum mechanics. (iii) is provided by a tacit "working hypothesis": the extent of substitution of the carbon-carbon double bond in the α,β-unsaturated carbonyl system makes a contribution to the energy of the electronic transition. This working hypothesis presupposes that each additional bond attached to the α,β-unsaturated carbonyl system makes a distinct contribution to the energy, or in other words that the contributions are cumulative and hence can be distinguished by measurement. It further presupposes the principles of structure theory discussed above. The provisional acceptance of this hypothesis licenses Woodward's careful study of the data from the 45 cited sources in the first paper, a study that results in his claim that the position of the wavelength of maximum absorption may be rigidly correlated with the extent

of substitution of the carbon-carbon double bond. For a reason I will explain shortly, Woodward's working hypothesis is more controversial than it might appear at first glance. Moreover, it is historically significant because it allows the absorption spectrum of an organic molecule to provide *empirical access to individual bonds—the basic unit of structure theory—* something that had not yet been achieved for molecules of this complexity with a similar degree of generality.[74]

The utility of the working hypothesis for structure determination depends on the numerical values associated with the different substitution patterns, for if the relative values were within the limits of precision or if there were significant overlap between compound classes the correlation between wavelength and structure would not help to distinguish structural differences empirically. In the absence of *ab initio* methods for predicting them, the numerical values must be determined by experiment. So a question that could only be answered by experiment was whether the energetic contribution of the substitution pattern is large enough to make a detectable difference in the absorption spectra.

With the experimentally-determined values in hand, structural features could be deduced. Both results were incorporated into subsequent structure determination efforts as I have described above. There are two ways in which evidence from the success of continuing research is being accumulated here. One is the success of future applications of the correlations, and structural information deduced by means of them, to other substances. Another is the success of

---

[74] The importance of distinguishing the contributions of individual structural subunits to the properties of molecules is recognized by contemporary research in structure theory. For example, R. W. F. Bader, theorist of the influential "atoms in molecules" approach to modeling chemical structure, writes that

> each atom in a system makes an additive contribution to the average value of every system property. This is the principle underlying the cornerstone of chemistry—that atoms and functional groupings of atoms make recognizable contributions to the total properties of a system. In practice, we recognize a group and predict its effect upon the static and reactive properties of a system in terms of a set of properties assigned to the group. In those limiting situations wherein a group is essentially the same in two different systems, one obtains a so-called additivity scheme for the total properties, for in this case the atomic contributions, as well as being additive over each system, are transferable between the systems.

> *Atoms in Molecules*, (Oxford: Clarendon Press, 1994 [first published 1990]), p. 137.

other methods that also presuppose them. Synthesis is particularly effective in this regard. Since each step in a synthesis (save the first) presupposes that the previous step produced the desired compound, the determination of the structure of any of the intermediates corroborates the structural assignments of all the intermediates upstream of it. An incorrect structural assignment of an intermediate will very likely reduce the chances for a successful synthesis. Moreover, since the number of possible targets is practically infinite, there is a practically infinite number of syntheses that could corroborate previous structural assignments and correlations. A consequence is that the acceleration of synthetic efforts brought about by the application of these correlational techniques has also accelerated the corroboration *of* those techniques.

The working hypothesis is amenable to an open-ended extension by the addition of further concepts from chemical theory. In its initial form it only presupposes that substitution will have an effect. Later on, Woodward complicates the hypothesis by supposing that positional factors play a role and bringing concepts such as ring strain, reactivity and chemical equilibrium to bear in explaining deviations from biunivocity.[75] Thus what the wavelength ends up being a measure *of* becomes increasingly fleshed out in terms of structure theory. Moreover, recall that Woodward found it necessary to correct for solvent effects in order to observe the correlations. The correction was another way that chemical theory was entering into the theory of measurement. For the solvent correction presupposes the distinction between a mixture and a compound and the related concepts of solution, solvent and solute. Furthermore, Woodward explains the solvent's effect on absorption in terms of the polarity of the solvent. This suggests a hypothesis that electrostatic interactions between the solvent and the solute are making a difference to the way the solute molecules absorb light, and hence how the composition of the sample mixture can affect the measured property.

---

[75] Woodward (1942a), p. 72.

In summary, Woodward's work shows how both structure theory and concepts from the chemical classification of matter can enter into the theory of measurement that is brought to bear in structure determination using spectroscopy, in addition to the physical theory already presupposed in the design of the instruments. In contrast, the classical chemists were much more like Duhem's physiologist in that chemical theory, especially structural, was largely absent from their methods of observation.

I now turn to the issue of the constitutive assumptions of the Woodward rules, and in particular the issue of what assumptions were borrowed from quantum theory. This issue brings me to the controversial aspect of Woodward's working hypothesis I alluded to earlier. The hypothesis presupposes that a conception of molecular structure developed in the 19th century can be integrated with the use of instruments that presuppose quantum theory. A risk in accepting the hypothesis is that the structural conception will turn out to be incompatible with that theory. As has been pointed out by others,[76] it is far from clear whether, according to quantum mechanics, molecules have localized bonds, static structures, or the kinds of shapes and directional features that structural formulae represent. Furthermore, at the time of the "Structure and Absorption Spectra" papers the application of quantum formalisms to the analysis of molecular structure was still being worked out and could hardly be considered very secure. Another potential point of conflict is that, as I have shown in the preceding paragraphs, the model likely assumed in these papers ignores wave mechanics, and moreover is false from the standpoint of the latter, for the model treats electrons as charged point-masses oscillating in accordance with Hooke's law. So there was a risk that the classificatory apparatus erected on the

---

[76] For a sample of the literature, see: a) Weisberg, M., Needham, P., & Hendry, R. F. Philosophy of Chemistry. *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/win2011/entries/chemistry/ and references cited therein. b) Woody, A. I. Concept Amalgamation and Representation in Quantum Chemistry. In Woody, A. I. et al. (2012), 427-466.

basis of this "amalgamation"[77] of structure theory and quantum theory would have to be entirely reinterpreted, for it would no longer be clear what the spectra were telling chemists about the constitution of molecules. A consequence is that the evidence for the reactivity of substances, as determined by these methods, would have to be reevaluated.

To make this point clearer, I will compare the status of structural formulae in the 19th century to their status in contemporary chemistry. There is textual evidence that late 19th-century chemists were divided on the nature of the information embodied in structural formulae.[78] Some apparently thought they were merely tools summarizing the chemical behavior of substances, whereas others seem to have leaned towards a more realist interpretation according to which the formulae embodied microstructural information. This ambiguity is understandable given the experimental evidence used to identify substances at the time, which consisted largely of observations of the bulk chemical properties of substances. In the event that physical theory were to show that a realist interpretation of structure theory was untenable, the instrumentalist interpretation of structural formulae would be safe, for the observations of the chemical properties would not be altered. Not only would the instrumentalist interpretation be safe, but so would much of the chemical knowledge that had been accumulated about substances. As Schummer puts it,

> chemical properties…differ from physical properties in that they establish relations between chemical substances, between reaction partners and reaction products connected by chemical transformations. Systematically arranging all chemical knowledge yields a network structure in which every substance is related to every other substance by direct or indirect links. In such a network, the identity of each substance corresponds to its network location defined by its characteristic relations to other substances…classical chemical structure theory reproduces that network on a theoretical and sophisticated level, such that the relations between substances correspond to relations between chemical structures.[79]

---

[77] Woody (2012).
[78] Cf. Hendry (2010), pp. 120-121 and Weisberg, M. et al. (2011).
[79] Schummer, J. The Impact of Instrumentation on Chemical Species Identity from Chemical Substances to Molecular Species. In Morris (Ed.) (2002), p. 196.

Spectroscopic properties, on the other hand, "do not describe relations between different chemical substances, but the response of an isolated material sample to electromagnetic fields." I would argue that this difference between chemical and spectroscopic properties affects the status of structural formulae. Imagine a substance characterized solely in terms of spectrometric properties,[80] and that these properties of the substance are assigned to a structure. Now imagine that physical theory shows structure theory to be radically wrong. All chemical information would immediately vanish; one would be left with nothing more than a "fingerprint." The reason for this is that the structural formula supports an indefinite number of inferences concerning the chemical behavior of the substance. Because it is "known" what functional groups a structure contains and what the spatial relationships between the groups are, many chemical properties may be inferred without doing a single experiment. In the absence of a *chemical* characterization, the structural formula acquires the theoretical function of mediating the (theoretical) relations between the chemical substance whose structure it represents, and every other chemical; in other words, the substance's chemical properties are inferred from its structure. Many predictions can be made, often quite reliably, without having done a single chemical experiment. When a substance is characterized in terms of its chemical properties, on the other hand, the situation is reversed: the structure is inferred from its chemical properties. In the latter case, the theoretical function of the structural formula becomes less important because the observations of the chemical properties speak for themselves: chemical methods provide a source of knowledge of chemical behavior that is independent of structure theory. The downside is that many experiments might have to be performed before reliable predictions about chemical behavior can be made.

    If we now imagine that a significant fraction of substances were characterized exclusively by their spectrometric properties, then in the event that structure theory had to be discarded there

---

[80] I write "spectrometric" so as to include mass spectral as well as spectroscopic properties.

would be a large number of substances whose chemical behavior was largely a mystery. Even though it is possible to correlate features of "fingerprints" with functional groups, many predictions about reactivity hinge on the spatial relationships between functional groups, and these predictions necessarily presuppose a determinate structure. Thus, basing substance identification on spectrometric properties came with the risk that developments in quantum theory would undermine the predictive power of structural formulae.[81]

Balancing the risk was significant evidence for the promise of this approach. This evidence included the prior evidence linking spectral patterns with structural features such as those Woodward cites at the beginning of the first paper. There, he mentions three key pieces of evidence as background to his own research: (i) the measurement by Henri and coworkers of the absorption spectrum of mesityl oxide; (ii) the discovery by Scheibe and coworkers that the exact position of the wavelength of maximum absorption is solvent-dependent; (iii) the generalization proposed by Menschick and coworkers that the presence of an α,β-unsaturated carbonyl group could be correlated with intense selective absorption in wavelength region 230-250 nm. In addition to this directly relevant evidence, there was also the wider body of evidence, discussed above, linking spectral data with structural subunits.

Another source of evidence was the analysis of band spectra presupposing the Born-Oppenheimer approximation. While the two research programs might seem only distantly related, they both presuppose a rigid molecular backbone, as does classical chemical structure theory.[82] So the evidence supporting the BO approximation could, to some extent, be taken as evidence for chemical structure theory, and vice-versa. So even though the rationale for the BO

---

[81] In reality, of course, most substances are partially chemically characterized, if only because it is known how they are produced. The point remains, however, that much chemical information would be lost were structure theory discarded.

[82] In the Born-Oppenheimer treatment the electronic energies are calculated, for each internuclear distance, as if the nuclei were fixed. This step in the approximation amounts to treating the nuclei like classical particles with definite trajectories rather than as "wave-particles" whose location and motion are characterized by a wavefunction. These energies are then incorporated into the potential within which the nuclei vibrate. The minimum in the potential well represents the equilibrium bond length.

approximation may have been insufficient at the time it was accepted, as some have argued,[83] its basic presupposition has been retested repeatedly both by physicists studying band spectra and chemists engaged in structure determination.

With respect to the Woodward rules specifically there were three other safeguards besides the corroborating evidence from the band spectra. One was that the measurements would become increasingly well-behaved with time, in the sense that they would become increasingly stable, converging and amenable to increasing precision in the ways indicated earlier. Another was the corroborating evidence from chemical methods of analysis and synthesis. A third comes from the fact that the quantum-theoretic assumptions in his work divide into two types, constitutive and heuristic. In brief, heuristic assumptions may be important for guiding research but can subsequently be discarded without jeopardizing the results obtained therein, whereas constitutive assumptions can be discarded only at the cost of *prima facie* undercutting the results that presuppose it.[84]

It seems clear that Woodward is assuming basic, qualitative principles of quantum mechanics, such as that energy is quantized. These principles allow him to turn his rules into more than mere empirical correlations, for they allow him to turn deviations from his correlations into evidence for structural causes. For example, the claim that "the introduction of strain into a given chromophore will labilize the electronic system with a consequent shift toward the red of the absorption band associated with the system" implies a relationship between strain energy and absorption wavelength that Woodward uses to explain the absorption characteristics

---

[83] a) Sutcliffe, B. T., & Woolley, R. G. (2012). Atoms and Molecules in Classical Chemistry and Quantum Mechanics. In Woody, A. I. et al. (Eds.) (2012), pp. 388-426. b) Sutcliffe, B. T., & Woolley, R. G. (1977). Molecular Structure and the Born-Oppenheimer Approximation. *Chem. Phys. Lett., 45,* 393-398.
[84] On the distinction between the constitutive and heuristic elements entering into scientific research, see G. E. Smith, (2010), pp. 545-547.

of conjugated dienes.[85] It seems reasonable, therefore, to impute to him some version of the Bohr frequency rule, which involves the proportionality

8.  $E_1-E_2 \propto 1/\lambda_{max}$

where $E_1$ and $E_2$ are the energies of the molecular states involved in the absorption. (8), of course, entails that if $\lambda_{max}$ has the value *a*, then $E_1-E_2$ has a value proportional to *1/a*, and vice-versa. In other words, any absorption wavelength can be correlated in a one-to-one fashion with a transition energy. By means of additional assumptions relating the energy to structure, the absorption wavelength can be correlated in a one-to-one fashion to structural subunits. Some of the key assumptions are:

$A_1$: The absorption wavelength measures a quantity of energy that is a parameter of molecular structure.

$A_2$: The energy absorbed by a sample varies (inversely) as the wavelength of light. I.e., the energy of light is quantized.

$A_3$: (from $A_1$) The molecule is confined to discrete energy states.

$A_4$: The quantity of energy measured by the absorption wavelength includes contributions from the substitution pattern of the α,β-unsaturated ketone.[86]
$A_{1-4}$ are necessary for making differences in the absorption wavelengths of different compounds carry information about structural differences between them. Together, these four assumptions allow the expectation that the value of the absorption wavelength will vary systematically with the substitution pattern. This systematic dependency makes the wavelength carry information about structure and allows wavelength to be predicted from a given structure.

---

[85] Woodward (1942a), p. 72.
[86] Other important background assumptions not discussed in the main text include:
 $A_5$: The ultraviolet absorption characteristics of an organic compound are the same for all molecules of that compound (a necessary assumption for relating measurements on bulk sample to microstructure)
 $A_6$: Similarity of spectra between substances can be taken as evidence of similarity of structure ("same effect, same cause")
 $A_7$: The known structures used to support the generalization in $A_6$ were assigned correctly
 $A_8$: The first law of thermodynamics (to ensure that the energy of the light absorbed or emitted by a molecule equals the energy absorbed or lost by the molecule)

It is worth noting how crucial quantization is to the possibility of developing evidence from molecular spectra. If all energies were accessible to a molecule, then molecular structures could not be distinguished according to how much energy they absorbed, for they would all be capable of absorbing the same energies. And if light energy were not quantized, then the wavelength of light could not serve as a measure of molecular energies. What is remarkable here is not so much the use of quantization *per se*, but rather how much work is being done by a simple, qualitative principle, work that does not presuppose the much more elaborate theory erected on the basis of the principle.

The only textual clue as to how Woodward might have been conceptualizing the interaction between light and molecular structure is a footnote stating that "[p]ossibly the electrons of the single bond which is α,β to each of the double bonds of the diene chromophore are sufficiently labilized to permit oscillations approximating those of a conventional six-electron system such as that of the trienes."[87] From the footnote and a reference to a paper by G. N. Lewis and Melvin Calvin,[88] a plausible conclusion is that Woodward was conceptualizing the absorption spectra as the effect of transitions between states of an electronic oscillator, together with the assumptions of "simple quantization." Key among these are, of course, assumptions $A_{1-3}$ and, in addition, the further "structural" assumption that the molecular energy levels are determined by the constitution of the molecule, which allows the analysis of the characteristic energies of the absorbed radiation to be interpreted in terms of that constitution. What is not

---

[87] Woodward (1942a), p. 72.

[88] In that paper, Lewis and Calvin attempt to explain the absorption spectra of dyes in terms of a "quasi-classical" model. Starting from the observation that "it appears unlikely that in the near future these means [i.e., quantum-mechanical methods] alone will permit the interpretation of the absorption of light by such complex structures as the organic dyes," Lewis and Calvin propose a model based on the idea that "in the absorption of light the energy is taken up by electronic oscillations." The oscillations are considered as "analogous to classical oscillations, but subjected to the rules of simple quantizations." They point to qualitative agreement with experimental data as evidence for their model. See Lews, G. N. & Calvin, M. (1939). The color of Organic Substances. *Chem. Rev., 25*, 273-328.

mentioned in Woodward's papers is quantum mechanics itself—the wave-like character of matter, the Schrödinger equation, the uncertainty principle, orbitals and so on.

Indeed, an interesting feature of $A_{1-4}$ is that they posit a relationship between wavelength and energy, and between energy and molecular structure, independently of whether they can be derived from or explained by an overarching theory. Furthermore, these relationships are posited without containing terms from or presupposing *any specific* theory of the energetics of molecular structure, such as molecular quantum mechanics. Presupposing $A_1$ and $A_2$ has two principal effects. First, as evidenced by the third paper, it allows Woodward to go beyond a purely empirical correlation between wavelength and structure by allowing him to explain deviations from the initial correlations in terms of energetic effects caused by structural differences. Second, it provides a margin of security from the vicissitudes of theory change by grounding his approach in background assumptions that are minimally theory-laden. For example, were the 'molecular orbital' theory of UV spectra alluded to in section 3 to turn out to be false, the structures identified using the Woodward rules would not necessarily have to be revised, for their identification does not rely on heavily theory-laden concepts such as 'molecular orbitals.' What they do rely on is the "theory-thin" hypothesis that the wavelength of maximum absorption of a compound measures a quantity of energy that is characteristic of the structure of the compound, and that the value of this quantity is partially determined by the substitution pattern of the compound. The effect is to insulate substance identification from possible failures in our explanations, thus providing a relatively secure foundation for synthesis.

To return to the heuristic/constitutive distinction: the assumptions of "simple quantization" discussed above entered constitutively into Woodward's work, whereas the model of the electronic oscillator was entering only heuristically: Woodward only brings it up to make the claim that steric factors can have an effect on the absorption characteristics of chromophores seem plausible, but his subsequent use of that claim to distinguish dienes with endocyclic and exocyclic double bonds does not depend on that bonding model. Simple quantization, on the other hand, ensures that the absorbed wavelengths are measuring a parameter of molecular structure, its energy, and in general is the basis of the usefulness of spectroscopy for exploring the structures of molecules. On the other hand, it is at least *prima facie* compatible with different theories of the chemical bond (as oscillating electrons, as shared pairs of electrons, as molecular orbitals, as energetic stabilization of the molecule, etc.)[.] So if the model of the electronic oscillator were later discarded, the results obtained with Woodward's correlational techniques would not be undercut. That said, a safeguard is obviously not a guarantee, and there was always the possibility that the weight of evidence in favor of an interpretation of quantum theory that was incompatible with classical structure theory would force a radical revision of the latter.

## 7. A Comment On the Relation of Organic Structure Determination to Quantum Mechanics

---

[107] For a discussion of evidence from theoretical support, see Hempel, C. G. (1966). *Philosophy of Natural Science*. Prentice Hall: Upper Saddle River, p. 38.

So far this paper has made two basic claims. The first is that spectroscopic methods allow chemists to make more secure arguments for claims about the microscopic structure of substances than they could by conjecturing hypotheses. The second is that classical structure theory was presupposed in the theory of measurement underwriting this use of spectroscopy. But this use of spectroscopy also brings up a new problem for the security of claims about molecular structure. The theory of measurement combines classical structure theory with quantum theory. As noted above, it has been suggested that the latter theory might be incompatible with the former. In fact the *Stanford Encyclopedia of Philosophy* entry on "Philosophy of Chemistry" presents the issue as a rather stark opposition between the weight of experimental evidence in favor of something like the classical structure theory on the one hand, and the ontology entailed by quantum theory on the other:

> Do molecules have the kinds of shapes and directional features that structural formulas represent? Given the history I have discussed so far it seems like the answer is obviously yes. Indeed, a number of indirect experimental techniques including x-ray crystallography, spectroscopy, and product analysis provide converging evidence of not only the existence of shape, but specific shapes for specific molecular species.

> Despite this, quantum mechanics poses a challenge to the notion of molecular shape. In quantum mechanical treatments of molecular species, shape doesn't seem to arise unless it is put in by hand.

Indeed, since much of the evidence presupposes quantum mechanics, one may wonder why the former is not nullified by the latter, that is, shown not to be evidence for molecular structure after all. In "Must a Molecule Have a Shape?", a paper that sparked debate on this subject, R. G. Woolley argues that size and shape are not intrinsic properties of molecules but rather reflect the influence of a many-body environment. Woolley raises the possibility that the idea that molecules have a determinate shape reflects a parochialism in the evidence used to support the idea: namely, a parochialism arising from the fact that conventional spectroscopic and chemical measurements are conducted on bulk samples rather than isolated molecules. If this is correct, an appropriate response might be, not to discard the evidence for classical structure theory, but rather to view it as evidence for a theory of non-isolated molecules that holds within limits of

precision that are specified by the resolving power of the measuring techniques used. For example, UV and NMR experiments are typically performed on bulk samples and therefore have much less resolving power than, say, the molecular beam technique and laser radiation. The problem might then be, not so much the potential incompatibility of quantum theory with structure theory, but rather with specifying the limits of precision within which the latter holds. A plausible scenario is therefore that even if it is true that, according to quantum theory, isolated molecules do not have determinate structures, this conclusion need not invalidate the evidential reasoning supporting structure theory or entail that there is something illusory about molecular structures. The attribution of determinate microscopic structures to substances can still be justified, only with appropriate qualifications, such as that the degree to which the nuclei may be considered "fixed" depends on the resolving power of the experimental methods used.

Whether or not such a solution is feasible, one moral of Woolley's argument is that focusing on the nature and limits of measurement procedures can inform our understanding of apparent conflicts that arise when one theory is claimed to be more general or fundamental than another. This point is important, for there is a tendency in philosophy of science to assume that the process of constructing theories in the empirical sciences can be understood independently of the process of marshalling evidence for them. The debate over whether or not shape is an intrinsic property of molecules highlights the fact that the two processes mutually condition each other: structure theory was required to turn the data into evidence for claims about molecular structure, and, conversely, the nature of the data may have had an important impact on the development of structure theory.

## 8. Concluding Remarks

In summary, Woodward's work exemplified how the data from physical instrumentation, introduced into chemistry during the "Instrumental Revolution," could be successfully marshaled into evidence for molecular structure. The new methods permitted a more sophisticated logic of

testing to be adopted than was usually possible using classical methods. Their successful application required only a piecemeal utilization of the quantum theory underlying the new technology. The evidence provided by the new methods could be corroborated in three ways: by converging spectrometric measurements, by traditional chemical techniques, and by the well-behavedness of the measures. Furthermore, the techniques could be applied even though the quantum theory of molecules was still being worked out.

One lesson of this episode is that the answer to what aspects of theory are "confirmed" by experiment should always be arrived at by a careful study of the logic of testing in the research process. The research documented in Woodward's "Structure and Absorption" papers is an example of how theory-mediated measurements can work, in practice, in the absence of a well-developed theory; it is an example of exploratory experimental science. The Woodward rules were local relationships that were not derived from an overarching theory. Yet some overarching theory was crucial to producing the knowledge they represent, for it was used to turn spectral data into evidence for the rules in the ways discussed in this paper. In the study of such research, interpretations of the role of theory in science that reduce it to one of either representing or explaining miss an important part of what is going on in this kind of science. Another question one can ask concerning the role of theory in exploratory experimental science, beyond whether it represents or explains observations, is how it is used to marshal the data into evidence for the claims to knowledge produced by the research. Answering this question can sometimes lead to surprising insights into what aspects of theory are being tested in a particular research process. In the application of spectroscopy to organic chemistry, it turns out that an elaborate molecular quantum-theoretical apparatus was unnecessary for the process to get started. This conclusion is surprising in light of the foundational role that quantum mechanics is supposed to have with respect both to chemistry and to the understanding of spectroscopy.

Though not derived from a theory, the Woodward rules were nonetheless "theoretical" in that they related measurements to theoretical entities, generalized the relations far beyond the narrow sample data on which they were based, and implied that only the theoretical entities invoked make a difference to the measurements. Such local relationships are often necessary in the early stages of theory construction, where the evidence for them cannot come from the theoretical support derivations provide but must come, at least initially, from agreement with the data and, subsequently, from the success of the research predicated on them.[107] Subsequent investigations not only extended Woodward's initial results but also generalized his approach to other forms of spectrometry and to a vast array of compound classes. Today, standard textbooks are filled with tables listing correlations like those proposed by Woodward.[108] By 1963, structure elucidation had progressed to such an extent that Woodward declared its replacement by *de novo* synthesis as the main task confronting organic chemistry.[109]

## 9. Appendix

The diagrams in Figures 1-3 are examples of structural formulae, diagrams used by organic chemists to represent molecular structures. They are based on a symbolism developed by G. N. Lewis. Simple examples of what are called "Lewis structures" are shown below:



| methane | ammonium ion | water | carbon dioxide | carbon dioxide (incorrect) |
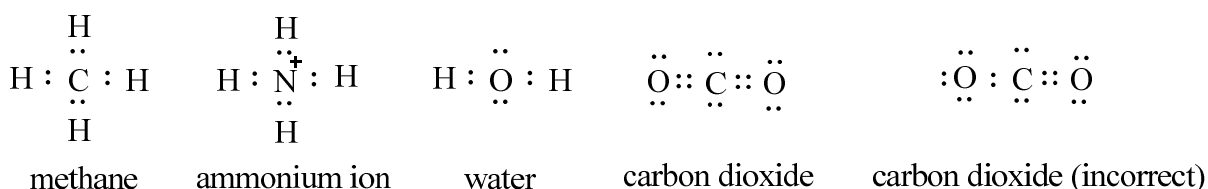
As Streitweiser, Heathcock and Kosower put it, "such structures not only provide simple and convenient representations of ions and compounds but are also valuable in providing an accurate

---

[108] See, for example, Silverstein, R. M.,  Webster, F. X., & Kiemle, D. J. (2005). *Spectrometric Identification of Organic Compounds*. Hoboken: John Wiley & Sons, Inc.
[109] Woodward *et al.* (1963), p. 248.

accounting for electrons. They form an important basis for predicting relative stabilities."[110]

Briefly, suitable structures can be derived using three simple rules:[111]

1. Show all and only valence electrons (the only electrons involved in bonding)

2. Each element should, to the greatest extent possible, have eight valence electrons, what is called a complete "octet." In the case of hydrogen, the octet rule is replaced by a "duet" rule that requires each hydrogen atom to have, to the greatest extent possible, two valence electrons.

3. Formal charges are assigned by dividing each bonding pair of electrons equally between the bonded atoms. The number of electrons assigned to each atom is compared with the neutral atom and appropriate positive or negative charges are assigned. Lone pairs are assigned to single atoms.

These rules are illustrated in the examples above. For example, the structure of methane has eight electrons, corresponding to one from the valence shell of each hydrogen atom and four from that of the carbon atom. All the structures have atoms with eight valence electrons (or two in the case of hydrogen) save for the incorrect structure of carbon dioxide, which assigns ten electrons to carbon. Application of rule 3 results in neutral molecules with the exception of the ammonium ion. In this case four electrons are assigned to nitrogen, and since the neutral atom has five electrons in its valence shell, a positive charge is assigned to it in the Lewis structure. The reason for rule 2 is that species with incomplete octets or duets are usually unstable or highly reactive. As shown in the structures, multiple bonds are represented simply by assigning pairs of electrons corresponding to the bond order (two pairs for double bonds, three pairs for triple bonds).

---

[110] Streitweiser, A., Heathcock, C. H., & Kosower, E. M. (1992). *Introduction to Organic Chemistry.* New York: Macmillan Publishing Company. p. 8.
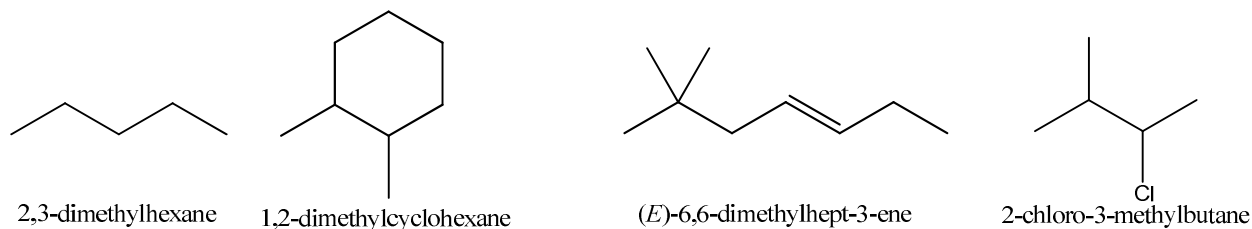[111] Streitweiser et al. (1992), pp. 8-9.

Further simplification can be achieved by using what are known as Kekulé structures. In this convention, each electron-pair bond is replaced by a line. Electron pairs not involved in bond formation are generally omitted, unless they are needed to call attention to particular properties of the molecule. The Kekulé structures corresponding to the Lewis structures above are:



methane          ammonium ion          water          carbon dioxide          methoxide
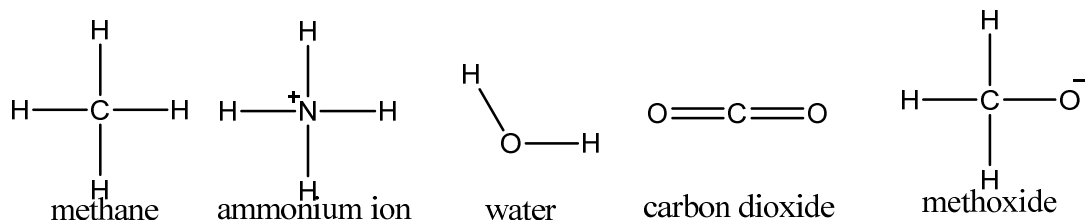
In these structures, the lone pairs are understood to be present and their presence is signified by appropriate formal charges, as shown for methoxide.

Though Lewis and Kekulé structures are useful for simple molecules, for larger molecules they can be awkward and so further streamlining of the drawing conventions is introduced in the form of line structures. Each line represents a chemical bond and the intersections of the bond lines are occupied by carbon atoms unless otherwise noted. Such formulae also contain implicit carbon-hydrogen bonds, for it is assumed that each carbon atom makes four bonds and hence every carbon atom that has fewer than four explicit bonds is understood to be bound to as many hydrogens as necessary to bring the total number of its bonds up to four. A few simple examples are shown below; more examples are shown in the text:



2,3-dimethylhexane          1,2-dimethylcyclohexane          (*E*)-6,6-dimethylhept-3-ene          2-chloro-3-methylbutane

The basic topology of molecular structures in organic chemistry includes chains, rings and functional groups. A functional group is a group of atoms that shows a relative constancy of

properties when attached to different carbon chains.[112] Table 1 lists a number of functional groups that appear in the text:
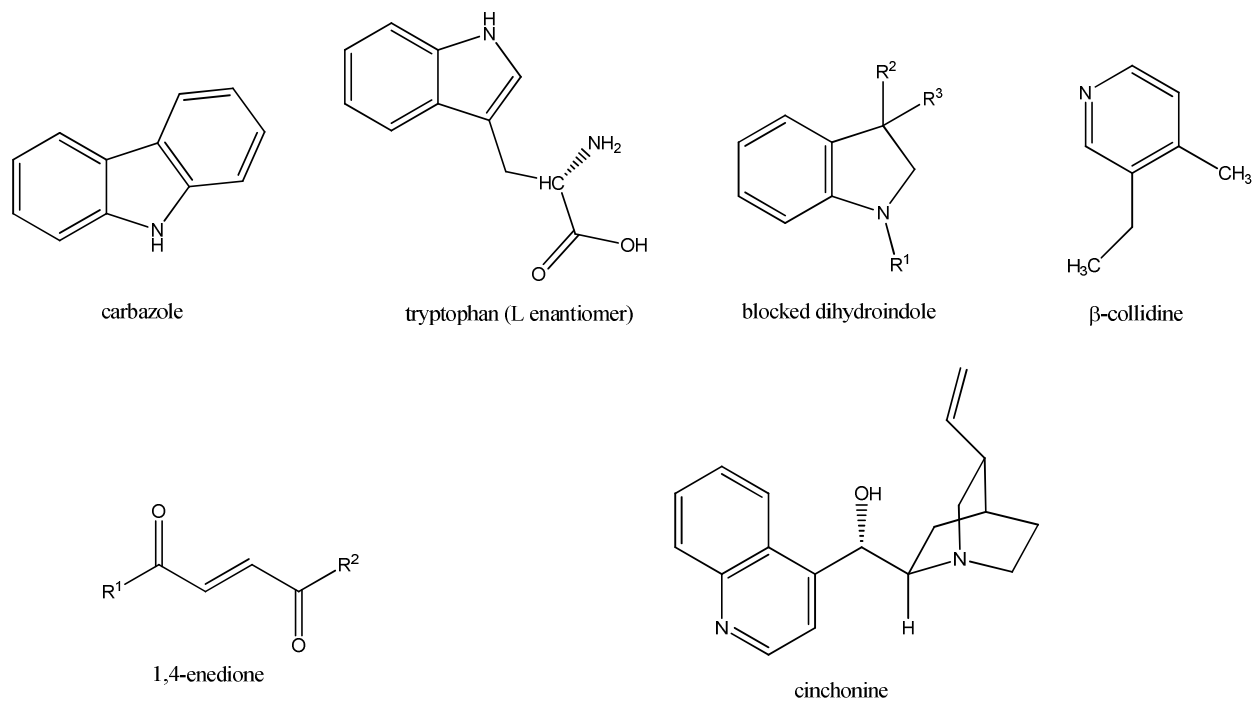
| Class | General Structure | Characteristic Functional Group | Example | Class | General Structure | Characteristic Functional Group | Example |
|---|---|---|---|---|---|---|---|
| alkanes | R-H | none | $CH_4$ | amines | | | |
| alkenes | | | $H_3CHC{=}CH_2$ | primary amines | $R{-}NH_2$ | $-NH_2$ | $CH_3NH_2$ |
| aromatic ring | | | | secondary amines | $R{-}N(H){-}R'$ | | $CH_3NHCH_3$ |
| | | | | tertiary amines | $R{-}N(R''){-}R'$ | | $CH_3NCH_3$ / $CH_3$ |
| alkyl halides | RF, RCl, RBr, RI | F, Cl, Br, I | $CH_3Cl$ | aldehydes | $R{-}C({=}O){-}H$ | | $H_3C{-}C({=}O){-}H$ |
| alcohols | $R{-}OH$ | $-OH$ | $CH_3OH$ | ketones | $R{-}C({=}O){-}R'$ | | $H_3C{-}C({=}O){-}CH_3$ |
| ethers | $R{-}O{-}R'$ | $-O-$ | $CH_3OCH_3$ | carboxylic acids | $R{-}C({=}O){-}OH$ | $-C({=}O){-}OH$ | $H{-}C({=}O){-}OH$ |
| | | | | esters | $R{-}C({=}O){-}OR'$ | | $H_3C{-}C({=}O){-}OCH_3$ |
| | | | | amides | $R{-}C({=}O){-}NR_2'$ | | $H_3C{-}C({=}O){-}NH_2$ |
| | | | | nitro compounds | $R{-}NO_2$ | $-NO_2$ | $H_3C{-}NO_2$ |

*Table 1.* Some functional groups. Partial reproduction of Table 3.1 from Streitweiser et al. (1992).

A 'substituent', a term I will be using later, is a functional group attached to a chain of *n* carbon atoms, where *n* can be any positive integer. The attachment of functional groups to carbon chains is termed 'substitution' because the substituents are conceived as replacing hydrogen atoms. Hence the number of such replacements on a chain is referred to as the 'degree of substitution' and the configuration of substituents on a chain is referred to as the 'substitution pattern.' The pattern can be indicated in the name of a compound by using numbers or Greek letters (α, β, γ etc.) according to rules maintained by the International Union of Pure and Applied Chemistry (IUPAC).

Some miscellaneous compounds that are mentioned in the text are shown in Figure 14:

---

[112] Streitwesier et al. (1992), p. 36. An alternative definition in terms of consistency of chemical reactivity from molecule to molecule is also proposed (p. 39).

carbazole       tryptophan (L enantiomer)       blocked dihydroindole       β-collidine

1,4-enedione       cinchonine

*Figure 14.*