Synthesis and Herbicidal Activity of Phenylproparginols

John J. Parlow* and Robert D. Clark

The Agricultural Group, Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167

A three-step/one-pot procedure was developed for preparing herbicidal fluorophenylproparginols from acetophenones by way of an enol phosphate intermediate. Side-by-side testing of 16 of the more active analogs showed that *trans*-2-methyl-1-[2-(4-methoxy-2,3,5,6-tetrafluorophenyl)ethynyl]-cyclohexanol (**21c**) was the most potent example. On the basis of the biological response data obtained for these first-tier analogs, a physiochemical "structural space" was defined. Twenty additional analogs were then targeted for synthesis in an experimental design set up to probe that structural space. Scaled rank sums analysis of herbicidal activity for these second-tier analogs produced a quantitative structure/activity relationship model indicating that **21c** bears a substitution pattern that is at or near the optimum for this class of chemistry.

Keywords: *Experimental design; phenylproparginol; QSAR; scaled rank sum*

INTRODUCTION

In the Monsanto Agricultural Group herbicide screening program, phenylproparginol **1** demonstrated modest herbicidal activity on economically important narrowleaf weeds. A synthetic program was then undertaken to identify those parts of **1** required for phytotoxicity and then to enhance the activity through systematic structural variation and elaboration.

The phenylproparginol nucleus (Figure 1) itself offers many possibilities for analog synthesis. We first approached the problem by considering the molecule as being composed of an alicyclic alcohol and a *para*-substituted aryl group with a bridging linker between them. Analogs were then prepared by modifying each component part separately. This allowed us to determine how much each component could be varied without compromising herbicidal activity. Once acceptable ranges of variation had been identified for each part, a multivariate experimental design (Brannigan and Duewer, 1991) was used to efficiently define the local "activity space" around the lead compound **1**.

Here we report the results of these efforts, which led ultimately to a 3-fold increase in unit activity after synthesis of 32 analogs.

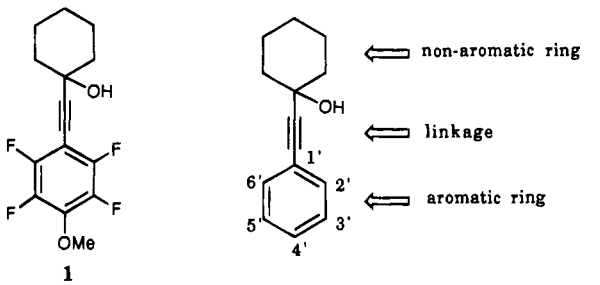
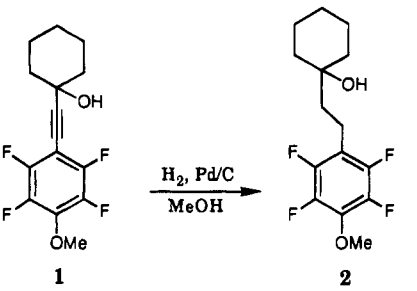
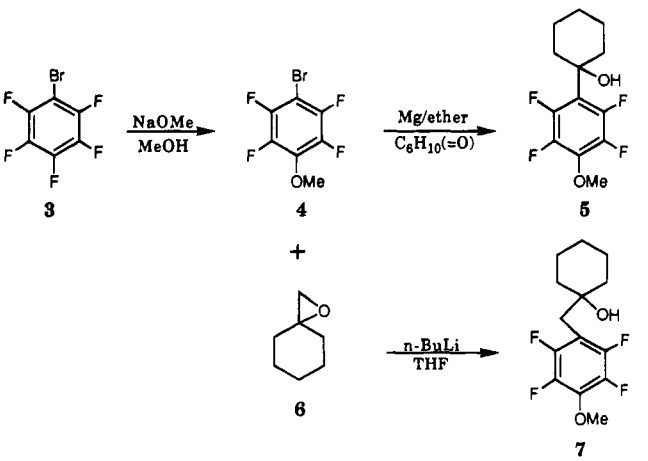


Figure 1. Phenylproparginol nucleus.

Scheme 1



Scheme 2



METHODOLOGY

Experimental Chemistry. Melting points were taken in open capillary tubes and are uncorrected. ^{1}H NMR, ^{19}F NMR, and ^{13}C NMR spectra were recorded on Varian XL-400 (400 MHz) or IBM AF-300 (300 MHz) spectrometers, and chemical shifts are expressed in parts per million relative to TMS, $CFCl_3$, and TMS, respectively. Chromatographic separations were performed on a Waters Prep 500A HPLC using silica gel columns. Elemental analyses were performed by Atlantic Microlab Inc., Atlanta, GA.

Synthesis. *1-[2-(2,3,5,6-Tetrafluoro-4-methoxyphenyl)ethyl]-cyclohexanol (2) (Scheme 1).* Compound **1** (1-[2-(2,3,5,6-tetrafluoro-4-methoxyphenyl)ethynyl]cyclohexanol) (0.75 g, 0.0025 mol) and 10% palladium on carbon were taken up in methanol (70 mL), and the mixture was shaken under hydrogen (50 psi over pressure) on a Parr hydrogenator until hydrogen uptake was complete. The slurry was filtered through Celite and the solvent removed to give the crude product, which was purified by HPLC (10% ethyl acetate–hexane) to give 0.65 g of **2** (85%) as a clear oil: ^{1}H NMR ($CDCl_3$) δ 1.46 (m, 2H), 1.68 (m, 9H), 1.84 (m, 2H), 2.92 (m, 2H), 4.17 (s, 3H).

1-Bromo-2,3,5,6-tetrafluoro-4-methoxybenzene (4) (Scheme 2). To bromopentafluorobenzene (**3**) (19.8 g, 0.080 mol) dissolved in methanol (100 mL) was added 95% sodium methoxide (5.0 g, 0.097 mol). After 3 h of stirring at room temperature, solvent was removed under reduced pressure and the residue

* Address correspondence to this author at Monsanto Co. U3H, 800 N. Lindbergh Blvd., St. Louis, MO 63167.

0021-8561/94/1442-2600$04.50/0

was dissolved in ether. The ether solution was washed with water and with brine and then dried over magnesium sulfate. Solvent was removed by evaporation, and the crude product obtained was purified by HPLC in hexane to give 17.6 g of **4** (85%) as a clear oil: ^{1}H NMR δ 4.07 (s, 3H).

1-(2,3,5,6-Tetrafluoro-4-methoxyphenyl)cyclohexanol (5) (Scheme 2). A few drops of 1,2-dibromoethane were added to magnesium turnings (0.30 g, 0.012 mol) stirring in anhydrous ether (40 mL). The mixture was brought to reflux, and **4** (1-bromo-2,3,5,6-tetrafluoro-4-methoxybenzene) (2.0 g, 0.0077 mol) was added dropwise as a solution in anhydrous ether (10 mL). The addition was carried out at a rate such that the reaction exotherm maintained the mixture at reflux. The solution was stirred for an additional 30 min at room temperature after addition was complete and then was cooled to 0 °C.

Cyclohexanone (0.85 g, 0.0087 mol) in anhydrous ether (10 mL) was added dropwise, with the addition controlled so as to keep the temperature below 10 °C. The reaction mixture was warmed to room temperature and stirred for 1 h and then was cooled to 0 °C and made acidic by addition of 2 N hydrochloric acid. Products were then extracted into ethyl ether. The organic layer obtained was washed with water and with brine and then dried over magnesium sulfate. Evaporation of solvent afforded a crude product which was purified by HPLC (8% ethyl acetate–hexane) to give 1.13 g of **5** (56%) as a yellow oil: ^{1}H NMR δ 1.12 (m, 1H), 1.22 (m, 2H), 1.64 (m, 3H), 1.96 (m, 4H), 2.35 (bs, 1H), 3.98 (s, 3H).

1-[(2,3,5,6-Tetrafluoro-4-methoxyphenyl)methyl]cyclohexanol (7) (Scheme 2). Compound **4** (1-bromo-2,3,5,6-tetrafluoro-4-methoxybenzene) (2.0 g, 0.0077 mol) in anhydrous tetrahydrofuran (20 mL) was treated at −78 °C with *n*-butyllithium (5.3 mL, 1.6 M in hexanes, 0.0085 mol). The addition was closely monitored to maintain the temperature below −68 °C, and the solution was stirred for an additional 15 min at −78 °C once addition was complete. This solution was transferred via cannula into a solution of **6** (1-oxaspiro[2.5]octane) (0.87 g, 0.0077 mol) in anhydrous tetrahydrofuran (15 mL). The mixture was allowed to warm to room temperature and stirred for 1 h, at which point 2 N hydrochloric acid was added until the solution was acidic. The acidified mixture was extracted with ether and the ether layer washed with water and then with brine and then was dried over magnesium sulfate. The solvent was removed by rotary evaporation, and the crude product obtained was purified by HPLC (10% ethyl acetate–hexane) to give 0.84 g of **7** (37%) as a white solid: mp 34–35 °C; ^{1}H NMR δ 1.31 (m, 2H), 1.54 (m, 9H), 2.80 (s, 2H), 4.03 (s, 3H).

Conversion of a Methyl Ketone (9) to a Phenylproparginol (11a–e) via an Intermediate Enol Phosphate (10) (Scheme 3). Aryl methyl ketone **9** (0.010 mol) was dissolved in anhydrous tetrahydrofuran (20 mL) and chilled to −78 °C, and lithium diisopropylamide (0.011 mol) was added as a 2.0 M solution. The addition was closely monitored to maintain a reaction temperature below −68 °C; once addition was complete, the solution was stirred for an additional 15 min at −78 °C. A solution of diethyl chlorophosphate (0.011 mol) in anhydrous tetrahydrofuran (5 mL) was then added dropwise and the solution allowed to warm to room temperature.

The enol phosphate solution so obtained was added dropwise to a solution of 2.0 M lithium diisopropylamide (0.022 mol) in anhydrous tetrahydrofuran (20 mL) held at −78 °C. Once addition was complete, the solution was stirred for an additional 15 min at −78 °C. A solution of aliphatic ketone (0.011 mol) in anhydrous tetrahydrofuran (5 mL) was then added dropwise and the solution allowed to warm to room temperature. The mixture was made acidic by addition of 2 N hydrochloric acid and extracted with ether. The ether layer was washed with water and with brine and then dried over magnesium sulfate. Solvent was removed by rotary evaporation and the crude product obtained purified by HPLC in ethyl acetate–hexane.

Similarly, phenylproparginols **17** (Scheme 5), **20a,b,e,f** (Figure 2), **21a–f** (Figure 3), and **22, 23a, 24a, 25a, 26a, 27a, 28,** and **29** (Scheme 7) were prepared from the corresponding methyl ketone.

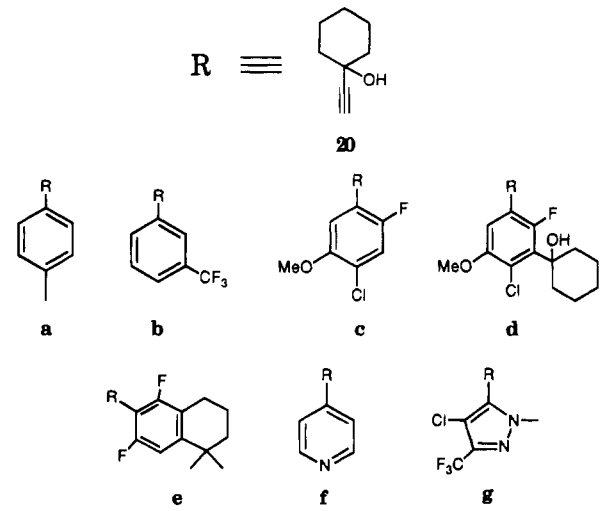


Figure 2. Proparginols with aryl variations.

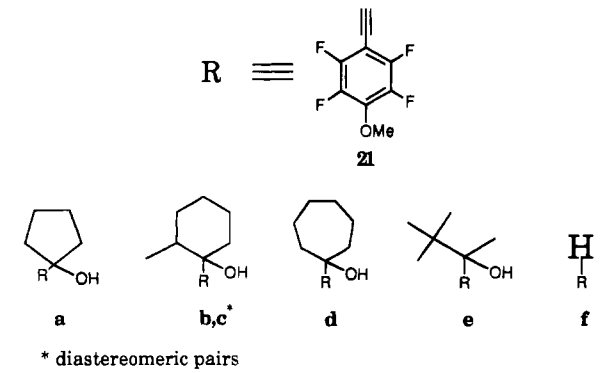


Figure 3. Alkyl and alicyclic phenylproparginols.

Preparation of Dibromoolefins (15) (Scheme 4). A solution of aldehyde **14** (0.010 mol) in dichloromethane (10 mL) was added dropwise to a cold (0 °C) solution of tetrabromomethane (0.020 mol) and triphenylphosphine (0.040 mol) in dichloromethane (30 mL). Once addition was complete, the solution was allowed to warm to room temperature, allowed to stir for 1–24 h, then filtered through Celite. The filtrate was washed with water and with brine and was dried over magnesium sulfate. Solvent was removed by evaporation, and the crude dibromoolefins obtained were purified by recrystallization or by HPLC in ethyl acetate–hexane.

Conversion of Dibromoolefins to Proparginols (16a–c) (Scheme 4). Dibromoolefin (**15**) (0.010 mol) was treated with 1.6 M *n*-butyllithium (0.022 mol) in anhydrous tetrahydrofuran (40 mL) at −78 °C. The addition was closely monitored to maintain a temperature below −68 °C. The solution was stirred for an additional 15 min at −78 °C after addition was complete, and then a solution of cyclohexanone (0.011 mol) in anhydrous tetrahydrofuran (10 mL) was added dropwise. The mixture was allowed to warm to room temperature and to stir for 1–24 h. The solution was acidified by addition of 2 N hydrochloric acid and then extracted with ethyl ether. The ether layer was washed with water and with brine and was dried over magnesium sulfate, and then solvent was removed under reduced pressure. The crude products were purified by recrystallization or by HPLC in ethyl acetate–hexane. Proparginols **20c,d,g** (Figure 2) were prepared from the corresponding dibromoolefins according to the same procedure.

Reaction of 17 with Nucleophiles To Afford Proparginols 18a–e (Scheme 5). Nucleophile (amine or sodium alkoxy/thioalkoxy salt) (0.011 mol) was added to a solution of **17** (1-[2-(pentafluorophenyl)ethynyl]cyclohexanol) (0.010 mol) in methanol or ethanol (20 mL). After stirring at reflux for 1–24 h, the solvent was removed under reduced pressure and the residue dissolved in ether. The ether solution was washed with water and brine and dried over magnesium sulfate. Solvent was removed by evaporation and the residue obtained

purified by recrystallization or by HPLC in ethyl acetate–hexane. In a similar fashion, proparginols **19a–d** (Scheme 6) and **23b,c, 24b–d, 25b,c, 26b** (Scheme 7) and **33** (Scheme 9) were prepared.

Bioevaluations. Tests were carried out in Teflon-lined 28 × 13 × 7.5 cm aluminum pans bearing 10 6-mm drainage holes. The soil used was Dupo silt loam which had been sieved (0.5-in. mesh) and sterilized by injection of steam (1 h at 85–90 °C). Soil was pressed into the pans to a depth of 6 cm using a grooved template, so that 10 10-cm furrows were formed into which seeds or propagules could be sown. Narrowleaf species tested included barnyardgrass (*Echinochloa crus-galli*; BYGR), yellow nutsedge (*Cyperus esculentus* L.; YENS), yellow foxtail (*Setaria glauca*; YEFT), seedling johnsongrass (*Sorghum halepense*; SEJG), large crabgrass (*Digitaria sanguinalis*; LACG) and downy brome (*Bromus tectorum*; DOBR).

Herbicides were applied as serial dilutions in acetone. A portion (8 mL) of each solution was sprayed in duplicate onto 550 g of soil, which was then thoroughly mixed and laid down over the seeded pans so as to give the application rates cited in the data tables. Overhead irrigation (0.6 cm) was applied immediately; thereafter, water was provided primarily by subirrigation, with light overhead watering applied as needed to keep the soil surface from drying out. Pans were maintained in greenhouses with supplemental lighting provided so as to afford a 14 h, 30 °C/10 h, 25 °C day/night cycle.

Analysis of Quantitative Structure/Activity Relationships (QSAR). Overall herbicidal effects were evaluated by inspection of aerial tissues 16 days after treatment. Characteristic symptoms observed for phenylproparginol-treated plants included stunting, apical inhibition, and necrosis. A growth reduction (GR) score between 0 (no injury) and 100% (no green tissue evident) was assigned by visual comparison with seedlings in check pans for which the soil had been treated with acetone alone. On this ordinal injury scale, 20% corresponds to a level of injury that would be commercially acceptable if the species were being raised as a crop, and 80% corresponds to a commercially acceptable level of weed control. The application rate required to obtain a growth reduction of 80% (ED_{80}) was estimated by interpolation or, where necessary, by logistic extrapolation. When negligible injury (GR <10%) was seen at any rate, the ED_{80} was arbitrarily assigned a value of 100 kg/ha. This no-effect "floor" was incorporated into calculations of potency (pED_{80}) so as to ensure uniformly positive values (Brannigan and Deuwer, 1991), setting $pED_{80} = \log_{10}(100/ED_{80})$.

Averaging ED_{80}s across species for QSAR is impractical, because usually there is some imprecisely titrated species for any particularly active or relatively inactive compound. Since these extreme activities are those which contribute most to the QSAR, analyses based on such averaged ED_{80}s tend to be poorly determined. One alternative, which is used here, is to average injury across all species at each rate for each compound and then estimate an ED_{80av} from these "average plant" GR titrations. This procedure minimizes variance in ED_{80av} but produces step-function titration curves. A more statistically robust approach is to convert GR scores across the entire test to ranks, which are summed across rates for each compound. The rank sums so obtained are then scaled from 0 (the minimum possible) to 100 (the maximum score). Such scaled rank sums (SRS) offer enhanced resolution over ED_{80}s in side-by-side tests using fixed titration ranges (Duewer and Clark, 1991). SRS statistics retain the robustness of rank statistics, but because they fall on an interval scale, they have many of the advantages of nominally continuous parametric measures of response such as ED_{80}.

Substituent parameters examined included Hammett's σ^*, the Hansch hydrophobicity parameter π, and the Charton–Taft steric coefficient E_{SV}. In those cases where some of the values for these parameters were not available in the existent literature, they were estimated by interpolation or extrapolation from values for homologous series. This included rings, which were algebraically "split" for analysis into polymethylene radicals. ED_{80}s were log-transformed to give potencies (pED_{80}) and offset by two decades to give uniformly positive values, as described above. Multiple regression was then carried out using the FIT MULTIPLE routine in RS/Explore from BBN Software Products Corp. to achieve a least-squares optimal fit to

$$pED_{80} = \log(100/ED_{80}) = c(1 + \sum a_i x_i + \sum b_i x_i^2) \quad (1)$$

A model including linear terms in all parameters was simplified by stepwise elimination of least-significant terms. Squared terms for all parameters were subsequently added into minimal linear models and then, again, eliminated stepwise.

Squared terms with positive coefficients correspond to nominal activity minima. Including them in a final model would grossly complicate the planning of further synthesis, so such terms were eliminated preferentially. First- and second-order coefficients a_i and b_i were then used to center parameters about their optima x_i^*, which were estimated from the quadratic terms ($x_i^{\ddagger} = -a_i/b_i$) (eq 2).

$$pED_{80} = \log(100/ED_{80}) = c'[1 + \sum b_i'(x_i - x_i^{\ddagger})^2] \quad (2)$$

RESULTS AND DISCUSSION

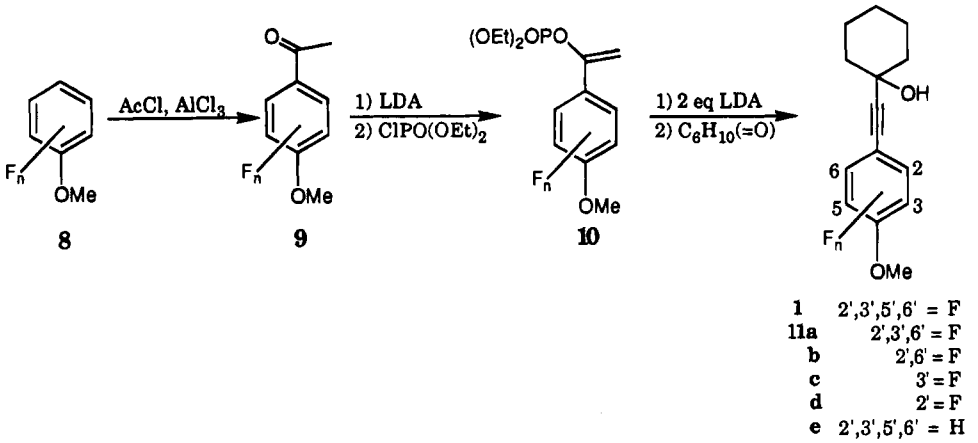
Initial Analog Syntheses and Evaluation. An initial set of analogs of **1** were prepared with the intention of evaluating the extent to which each segment could be independently varied without losing the desired biological activity. For the bridging acetylene in **1**, hybridization and chain length were both varied. The ethylene analog **2** was prepared by the reduction of **1** by catalytic hydrogenation (Scheme 1). The methylene-bridged and directly linked analogs were prepared from organometallic reagents derived from bromotetrafluoroanisole (**4**), as illustrated in Scheme 2. Condensation of the phenyl Grignard of **4** with cyclohexanone afforded the directly linked analog **5**, whereas lithium/halogen exchange and subsequent reaction with exocyclic epoxide **6** gave the methylene-linked analog **7**.

Aryl Variation. The second segment of the molecule to be investigated was the aromatic ring. Regiospecific effects of ring fluorination were investigated by preparing analogs of **1** from which aryl fluorines were systematically omitted. In general, the most efficient and easy way to prepare the desired phenylproparginols was a three-step/one-pot reaction starting with a substituted acetophenone (Scheme 3). Acetophenones **9** were prepared from anisoles **8** via Friedel–Craft acylation using acetyl chloride. The enolates of the acetophenones **9** were allowed to react with diethyl chlorophosphate to afford the enol phosphates **10**, which upon treatment with 2 equiv of lithium diisopropylamide eliminated phosphate to form the requisite acetylide anions. Quenching with cyclohexanone gave the desired proparginols **11**.

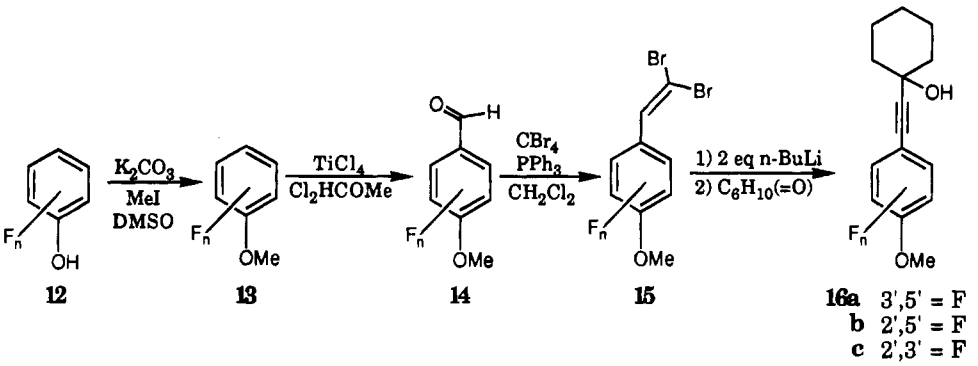
Unfortunately, the enol phosphate route produced a complex mixture of products when applied to difluoromethoxyacetophenones other than the 2′,6′ isomer. Evidently, abstraction of aryl hydrogen competes with vinyl deprotonation in these instances. The desired analogs were instead synthesized as shown in Scheme 4. The respective fluorophenols **12** were methylated to the anisoles **13**, which reacted with dichloromethyl methyl ether in the presence of titanium tetrachloride to yield the requisite benzaldehydes **14**. Condensation of the benzaldehydes **14** so obtained with carbon tetrabromide and triphenylphosphine afforded dibromoolefins **15**. Lithium acetylides were generated by metal/halogen exchange and double elimination and then quenched with cyclohexanone to afford the desired proparginols **16**.

The effect of the aryl *para* substituent was investigated by employing **17** as a substrate for nucleophilic

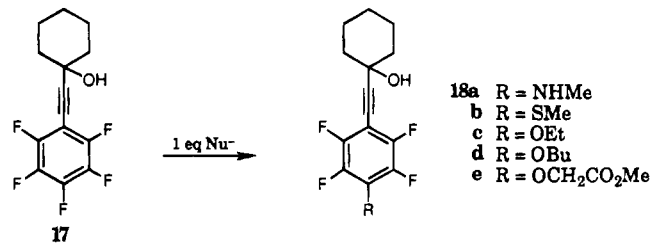
Scheme 3



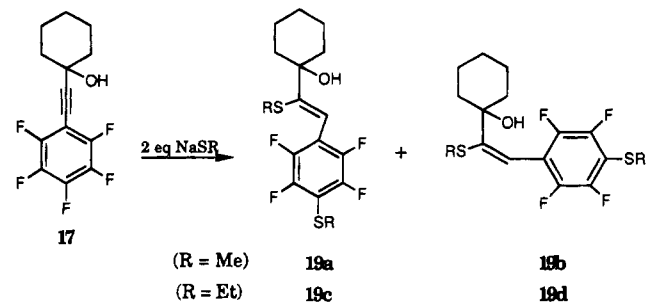
Scheme 4



Scheme 5



Scheme 6



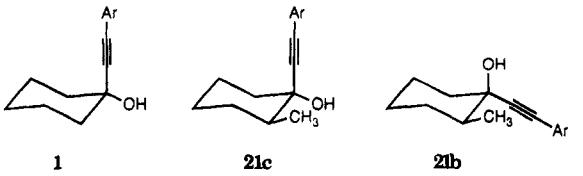
displacement. Somewhat surprisingly, it turned out that protection of the propargyl alcohol was generally not necessary: reaction of **17** with various nucleophiles cleanly afforded the desired *para*-substituted analogs **18** (Scheme 5). Displacement by alkoxides and amine proceeded smoothly, but reaction with excess thiolate was more complicated. A first equivalent of thiolate displaced the *p*-fluoride, but secondary attack occurred at the acetylenic bridge to give a mixture of *cis*- and *trans*-vinyl mercaptans **19**, as shown in Scheme 6.

A few other, more radical variations of the aromatic ring segment of the molecule were also tried. Each of these compounds **20** (Figure 2) was prepared by one of the procedures in Scheme 3 or 4.

Alicyclic Ring Variations. The SAR for the alicyclic ring in **1** was explored by condensing tetrafluoromethoxyphenyl acetylide with various ketones to afford sundry alkyl and alicyclic phenylproparginols **21** (Figure 3). Reaction with 2-methylcyclohexanone gave both diastereomeric phenylproparginols **21b** and **21c**. The racemic diastereomers obtained were readily separated chromatographically, but we were unable to assign the relative stereochemistries on the basis of spectroscopic analysis.

SAR. Some of the first-tier analogs lacked the characteristic herbicidal activity on narrowleaf weeds in preliminary tests (data not shown). These included one of the racemic methylcyclohexanols, **21b**. A consideration of relative steric effects (Gordon, 1972; Bushweller *et al.*, 1970) suggests that the hydroxy group is probably equatorial in the principal conformer of **1**. If a tendency to keep the α-methyl group in **21** equatorial dominates the conformational equilibria of **21**, one diastereomer will be homologous to **1**, whereas the other will be quite different. On the basis of this consideration, we believe that the more active compound is probably *trans*-**21c** and that the less active compound is *cis*-**21b** (Figure 4).

Sixteen of the more active examples were tested side-by-side for pre-emergence activity, which yielded the results given in Table 1. Shortening the acetylenic bridge (**5** and **7**) attenuated potency but less so than did altering hybridization through reduction (**2**). The fact that all four isomeric difluoro analogs (**11b** and **16a–c**) exhibited nearly identical reductions in potency indicated that the extent of fluorination did much more



1 21c 21b

Figure 4. Relative stereochemistry of methylcyclohexyl analogs.

to determine herbicidal activity than did the position of fluorination.

Compound **21c** is 5 times as active as **21d**, which bears an isomeric alicyclic ring. *tert*-Butyl **21e** is actually more active than the cyclic **1**, which shows that an alicyclic ring is not necessary for activity. Incidentally, the potency of **21c** also indicates that the inactivity of **21b** is not due to enantiospecific congestion effects at the α carbons. Moreover, **1**, **17**, **18a**, and **18d** (which differ only in *para* substitution) show a 2-fold range in ED_{80}. On the other hand, modifying the acetylenic bridge or tetrafluorophenylene component of **1** sharply attenuated herbicidal activity. Taken together, these results point up the potential for enhancing potency by systematic variation of the α and *para* substituents. A second tier of analogs was then prepared to better define the SAR at these positions.

Second-Tier Analogs. The physicochemical "structural space" around the latter substituents was probed by applying the principles of multivariate experimental design (Brannigan and Duewer, 1991). Twenty target compounds were selected for synthesis so that substitution at the *para* and both α positions varied according to a Plackett–Burman design (Plackett and Burman, 1946). Such a design can accommodate up to 19 independent variables, allowing a 2-fold redundancy across three fundamental parameters at each of three positions. A two-level design balanced across the Charton–Taft steric constant E_{SV}, Hammett's electronic parameter σ^*, and the Hansch substituent lipophilicity π was then laid out. Seventeen of these analogs (Table 2) were readily synthesized from commercially available pentafluoroacetophenone (**30**), as illustrated in Scheme 7.

Two important classes of compounds, however, proved to be inaccessible using our approach. The design called for including phenyl groups as substituents R geminal to the alcohol, because they comprise one of the few tractable moieties which is bulky, hydrophobic, and electron withdrawing. Conversely, ethyl groups were desired as small, relatively nonhydrophobic and electron-donating aryl substituents Y. Synthesis of targeted α-benzylproparginols generally failed, however. Compound **29** was obtained at very low yield (3%) from benzyl propyl ketone, whereas α,α-diaryl ketone **32** was the sole product recovered from reaction of acetylide **31** with isobutyl benzyl ketone (Scheme 8). Evidently, enolate formation and subsequent *para* attack predominate in this case. Somewhat ironically, reaction with ethyl Grignard gave no *para* attack at all. Only addition to the triple bond was seen (Scheme 9). Note that although target analogs bearing α-phenyl or *p*-ethyl substituents were therefore "lost" from the experiment, it should be noted that those targeted compounds which were successfully prepared (Table 2) still make up an orthogonal basis set.

Results from a side-by-side test of these analogs are also given in Table 2. Some *p*-fluoro intermediates extraneous to the design were included in this test, along with **18c**, **21c**, and **21e**. The latter carbocylic compounds were included to evaluate the extent of test-to-test variability and to make it possible to compare results directly with the first-tier test. Again, **21c** was most active. Compounds **18c** and **21e** were also more active than was any analog in the experimental design.

QSAR. To obtain a good QSAR, no ED_{80} should be so high that a reproducible effect is not seen at any rate or so low that complete kill is seen at all rates. This consideration defines the useful dynamic range of the assay, which is about 0.03–30 kg/ha in this case. Large crabgrass and yellow foxtail ED_{80}s met this condition for all second-tier analogs, as did the weed average ED_{80}s.

Rank sum statistics are more robust, in that they are guaranteed to span an appropriate dynamic range if any species used in the ranking has an appropriate dynamic range and will still do so in many cases in which no individual species assayed can itself. Figure 5, which compares two SRS statistics with the weed average ED_{80}, shows that each exhibits a dynamic range in this test comparable to the parametric ED_{80}. Figure 5 also illustrates the log–linear relationship found between

Table 1. Herbicidal Activities of Phenylproparginol Analogs

	ED_{80} (kg/ha)							scaled rank sums[b]		
compd	BYGR	YENS	YEFT	SEJG	LACG	DOBR	av ED_{80}[a]	YENS	DOBR	BYGR
1	3.66	2.52	1.52	0.42	1.01	3.48	2.66	68	75	57
17	3.48	53.21	7.63	0.62	3.06	8.52	4.94	38	61	54
18c	1.96	3.39	0.64	0.36	0.55	4.02	2.33	65	72	68
18a	4.51	99.00	5.83	0.98	2.60	43.42	5.40	35	54	46
18d	3.24	37.63	3.00	0.85	1.18	53.21	4.32	35	45	38
16a	8.24	10.80	12.77	0.14	3.02	53.21	6.15	46	59	42
16c	6.46	99.00	11.40	1.32	2.55	53.21	6.51	38	41	37
16b	4.57	37.63	28.42	1.65	2.95	99.00	6.52	42	36	53
11b	28.42	11.21	5.21	0.31	2.58	75.19	5.82	50	41	32
3	25.07	18.06	6.02	0.79	7.37	99.00	8.49	50	33	33
21d	4.30	11.66	4.12	2.31	2.95	7.78	4.37	41	39	46
2	10.42	99.00	99.00	28.42	30.70	99.00	28.66	26	23	42
20e	8.16	11.67	99.00	0.14	7.07	43.37	10.68	44	32	40
21e	2.97	3.32	6.39	0.14	2.17	11.67	3.78	60	52	60
7	28.42	30.70	13.45	0.55	19.42	53.21	10.89	60	38	38
21c	0.51	2.39	0.40	0.18	0.32	3.23	0.87	82	83	83
SEM[c]	1.39*x*	1.80*x*	1.42*x*	1.93*x*	1.76*x*	1.69*x*	1.22*x*	±10.4	±6.8	±10.5

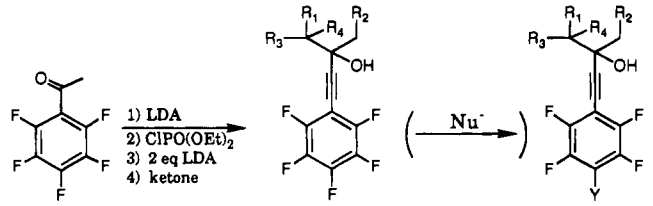
[a] ED_{80} for growth reduction averaged across all weeds. [b] Ranking criteria: **YENS**/YEFT/BYGR; **DOBR**/LACG/SEJG; and **BYGR**/LACG/YEFT/SEJG. [c] Calculated from pooled variances by analysis of variance for log(ED_{80}) or scaled rank sums.

Table 2. Test Results for Phenylproparginol Analogs Prepared Using Experimental Design

				ED_{80} (kg/ha)			scaled rank sums[b]		
compd	R_1	R_2	Y	YEFT	LACG	av[a]	YENS	DOBR	BYGR
22[d]	H	Et	F	1.65	2.63	4.76	53	46	44
23a	H	CH_2Ph	F	58.28	14.98	43.28	13	26	30
23b	H	CH_2Ph	NHBu	1.86	2.44	4.63	55	55	54
23c	H	CH_2Ph	OPe	3.33	3.00	6.09	63	48	49
24a	H	NEt_2	F	8.09	4.28	15.31	41	34	39
24b	H	NEt_2	OPe	3.06	4.04	8.11	40	45	40
24c	H	NEt_2	NHBu	6.33	5.07	9.89	48	32	32
24d	H	NEt_2	NMeEt	2.91	11.53	10.78	43	45	34
25a	H	OMe	F	2.25	6.42	8.30	64	37	37
25b	H	OMe	SBu	4.25	6.22	10.47	42	74	59
25c	H	OMe	NPr_2	4.16	59.43	15.04	28	23	23
26a	H	tBu	F	2.06	2.79	4.45	51	56	45
26b	H	tBu	NMeEt	5.24	2.51	6.87	49	51	57
27a	Et	iPr	F	7.71	2.54	7.24	39	71	49
27b	Et	iPr	NHBu	4.67	4.76	10.91	38	48	55
28	Et	Et	F	1.15	2.46	4.15	70	62	61
29	Et	Ph	F	3.20	7.85	12.30	45	33	34
18c[e]	CH_2	$(CH_2)_2$	OEt	1.68	0.69	2.15	76	70	78
21e[f]	Me	H	OMe	1.79	0.99	3.05	73	61	68
21c[e]	CH_2	$(CH_2)_2$	OMe	0.43	0.49	1.76	77	75	85
SEM[c]				±0.281	±0.204	±0.140	±9.2	±6.8	±7.2

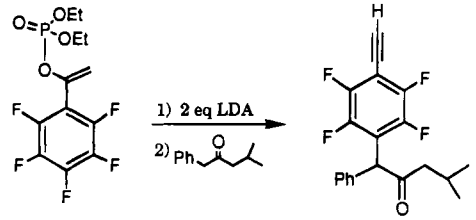
[a] ED_{80} for growth reduction averaged across BYGR, DOBR, LACG, SEJG, YEFT, and YENS at each rate. [b] Ranking criteria: **YENS**/YEFT/BYGR; and **BYGR**/LACG/YEFT/SEJG. [c] Calculated from pooled variances by analysis of variance for log(ED_{80}) or scaled rank sums. [d] R^3 and R^4 both H except as noted. [e] R^3 = Me. [f] R^3 = R^4 = Me.

Scheme 7



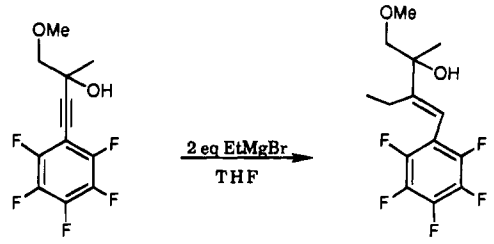
30 22-29 22-29

Scheme 8



31 32

Scheme 9



25a 33

weed ED_{80} and scaled rank sum statistics, with a slope here of approximately 45%/decade. This log−linearity is general, which suggests that SRS statistics can be a viable alternative to pED_{80} (or pED_{50}) values as response variables for any QSAR analysis based on linear free-energy relationships.

The underlying experimental design was explicitly "chiral", in that the α and α′ carbons were treated as independently variable in substitution. In fact, enantiomeric products were not separated, so that a total of 38 distinct analogs were actually included in the test (**18c** and **28** are each single compounds, because they are symmetrically substituted). For purposes of carrying out a numerical analysis of the data, the bulkier/more polar of the α substituents was arbitrarily identified with R_2 for the QSAR analysis (Table 2); the less bulky/less polar substituent was identified with R_1. This will be a valid simplification so long as one enantiomer dominates the biological activity of each racemic mixture. The diversity of physicochemical attributes in the substituents examined ensures that this is probably a safe assumption in our case. In terms of experimental design, this converts the underlying design from one at two levels for three parameters each across α and α′ substituents to a three-level design in σ^*, π, and E_{SV} for R_2 (hereafter designated σ_2^*, π_2, and E_2, respectively) and a truncated design balanced across one descriptive variable associated with R_1. Hansch's lipophilicity π for R_1 ($\equiv\pi_1$) was chosen as having the most convenient spread of values. Substitution of σ_1^* or E_1 would alter the absolute values of the coefficients in the regression models obtained, but not the significances or the relative magnitudes. It does mean that "optimal" parameter values are more properly identified with specific alkyl groups R_1 than with lipophilicity *per se*.

Compounds **21e** and **21c** were included in the analysis by adding indicator variables I_3 and I_4. The indicator variables I_3 (I_4) were set to 0 or 1 when R_3 (R_4) was H or Me, respectively. Quadratic terms were included to allow identification of parameter optima. In a few cases, second-order terms had significant positive coefficients, which would correspond to maximally "bad" values. These were disregarded because even if they are "real" and not analytical artifacts, they tend to obscure what direction to take in subsequent syntheses rather than make things clearer.

On the basis of the data in Table 2, only SRS_{BYGR} and pED_{80LACG} gave well-defined second-order QSAR eqs 3 and 4 when stepwise multiple regression analysis was applied. Standard errors of regression are cited par-

Table 3. Compounds 2–33 Prepared

product	molecular formula	calcd found C	calcd found H	^{1}H NMR ($CDCl_3$) δ[a]	yield[b]	mp[c] (°C)
2	$C_{15}H_{18}F_4O_2$ (306.30)	58.82 58.70	5.92 5.94	1.46 (m, 2H), 1.68 (m, 9H), 1.84 (m, 2H), 2.92 (m, 2H), 4.17 (s, 3H)	85	oil
3	$C_{13}H_{14}F_4O_2$ (278.24)	56.12 56.34	5.07 5.11	1.12 (m, 1H), 1.22 (m, 2H), 1.64 (m, 3H), 1.96 (m, 4H), 2.35 (bs, 1H), 3.98 (s, 3H)	56	oil
4	$C_7H_3F_4BrO$ (258.99)	32.46 32.62	1.17 1.20	4.07 (s, 3H)	85	oil
5	$C_{13}H_{14}F_4O_2$ (278.24)	56.12 56.34	5.07 5.11	(1.12 (m, 1H), 1.22 (m, 2H), 1.64 (m, 3H), 1.96 (m, 4H), 2.35 (bs, 1H), 3.98 (s, 3H)	56	oil
7	$C_{14}H_{10}F_4O_2$ (292.27)	57.53 57.68	5.52 5.51	1.31 (m, 2H), 1.54 (m, 9H), 2.80 (s, 2H), 4.03 (s, 3H)	37	34–35
11a	$C_{15}H_{15}F_3O_2$ (284.28)	63.38 63.25	5.32 5.79	1.14 (m, 1H), 1.55 (m, 7H), 1.94 (m, 2H), 2.05 (bs, 1H), 3.82 (s, 3H), 6.45 (m, 1H)	7	oil
11b	$C_{15}H_{16}F_2O_2$ (266.29)	67.66 67.50	6.06 6.27	1.18 (m, 1H), 1.59 (m, 7H), 1.97 (m, 2H), 2.05 (s, 1H), 3.72 (s, 3H), 6.39 (d, 1H)	26	69–71
11c	$C_{15}H_{17}F_1O_2$ (248.29)	72.56 72.14	6.90 7.04	1.15 (m, 1H), 1.57 (m, 7H), 1.93 (m, 2H), 2.47 (s, 1H), 3.81 (s, 3H), 6.79 (t, 1H), 7.09 (m, 2H)	47	76–78
11d	$C_{15}H_{17}F_1O_2$ (248.29)	72.56 72.51	6.90 6.99	1.18 (m, 1H), 1.61 (m, 7H), 1.93 (m, 2H), 2.05 (s, 1H), 3.72 (s, 3H), 6.55 (m, 2H), 7.24 (m, 1H)	41	oil
11e	$C_{15}H_{18}O_2$ (230.30)	78.23 78.00	7.88 7.81	1.18 (m, 1H), 1.60 (m, 7H), 1.93 (m, 2H), 2.25 (s, 1H), 3.71 (s, 3H), 6.74 (d, 2H), 7.28 (d, 2H)	61	56–57
16a	$C_{15}H_{16}F_2O_2$ (266.29)	67.66 67.28	6.06 6.22	1.21 (m, 2H), 1.54 (m, 7H), 1.88 (m, 2H), 1.89 (s, 1H), 3.93 (s, 3H), 6.88 (d, 2 H)	74	79–80
16b	$C_{15}H_{16}F_2O_2$ (266.29)	67.66 67.42	6.06 6.34	1.19 (m, 1H), 1.55 (m, 7H), 1.96 (m, 2H), 1.97 (s, 1H), 3.81 (s, 3H), 6.60 (dd, 1H), 7.03 (dd, 1H)	62	oil
16c	$C_{15}H_{16}F_2O_2$ (266.29)	67.66 67.22	6.06 6.10	1.18 (m, 1H), 1.32 (m, 7H), 1.93 (m, 2H), 2.15 (bs, 1H), 3.83 (s, 3H), 6.59 (dt, 1H), 7.04 (dt, 1H)	58	89–90
17	$C_{14}H_{11}F_5O$ (290.23)	57.94 57.85	3.82 3.90	1.38 (m, 1H), 1.67 (m, 7H), 2.04 (m, 2H), 2.14 (bs, 1H)	50	94–95
18a	$C_{15}H_{15}F_4NO$ (301.28)	59.80 59.67	5.02 5.01	1.20 (m, 1H), 1.62 (m, 7H), 1.94 (m, 2H), 2.25 (s, 1H), 3.05 (bs, 3H), 3.95 (bs, 1H)	81	87–88
18b	$C_{15}H_{14}F_4SO$ (318.33)	56.60 56.37	4.43 4.52	1.11 (m, 1H), 1.60 (m, 7H), 1.97 (m, 2H), 2.38 (s, 1H), 2.46 (s, 3H)	70	51–52
18c	$C_{16}H_{16}F_4O_2$ (316.29)	60.76 60.55	5.10 5.08	1.14 (m, 1H), 1.35 (t, 3H), 1.56 (m, 7H), 1.94 (m, 2H), 2.09 (s, 1H), 4.25 (q, 2H)	65	51–52
18d	$C_{18}H_{20}F_4O_2$ (344.35)	62.78 62.66	5.85 5.89	0.76 (t, 3H), 1.18 (m, 1H), 1.29 (sextet, 2H), 1.43 (m, 7H), 1.55 (quintet, 2H), 1.57 (m, 1H), 1.81 (bs, 1H), 4.04 (t, 2H)	72	oil
18e	$C_{17}H_{16}F_4O_4$ (360.30)	56.67 56.27	4.48 4.52	1.14 (m, 1H), 1.59 (m, 7H), 1.94 (m, 3H), 3.73 (s, 3H), 4.79 (s, 3H)	61	53–54
19a	$C_{16}H_{18}F_4O_1S_2$ (366.43)	52.45 52.50	4.95 4.95	1.20 (m, 1H), 1.63 (m, 7H), 1.80 (m, 2H), 1.96 (s, 3H), 2.05 (s, 1H), 2.46 (s, 3H), 6.66 (s, 1H)	56	64–65
19b	$C_{16}H_{18}F_4O_1S_2$ (366.43)	52.45 52.42	4.95 4.95	1.11 (m, 1H), 1.60 (m, 7H), 1.80 (m, 2H), 2.09 (s, 1H), 2.29 (s, 3H), 2.42 (s, 3H), 5.39 (s, 1H)	11	96–97
19c	$C_{18}H_{22}F_4O_1S_2$ (394.48)	54.81 54.96	5.62 5.61	0.99 (t, 3H), 1.17 (t, 3H), 1.18 (m, 1H), 1.64 (m, 7H), 1.80 (m, 2H), 2.03 (s, 1H), 2.34 (q, 2H), 2.88 (q, 2H), 6.70 (s, 1H)	52	oil
19d	$C_{18}H_{22}F_4O_1S_2$ (394.48)	54.81 54.93	5.62 5.62	1.19 (t, 3H), 1.28 (t, 3H), 1.46 (m, 1H), 1.60 (m, 7H), 1.90 (m, 2H), 2.71 (s, 1H), 2.79 (q, 2H), 2.88 (q, 2H), 5.60 (s, 1H)	11	oil
20a	$C_{15}H_{18}O_1$ (214.31)	84.07 84.11	8.47 8.48	1.59–1.64 (m, 7H), 1.97 (m, 3H), 2.05 (s, 1H), 2.33 (s, 3H), 7.09 (d, 2H, J = 7.8 Hz), 7.30 (d, 2H, J = 8.1 Hz)	95	85–86
20b	$C_{15}H_{15}O_2F_3$ (268.28)	67.16 66.78	5.64 5.76	1.25 (m, 1H), 1.60 (m, 7H), 1.96 (m, 2H), 2.10 (s, 1H), 7.34 (m, 1H), 7.48 (m, 2H), 7.60 (s, 1H)	45	63–64
20c	$C_{15}H_{16}O_2FCl$ (282.74)	63.72 63.87	5.70 5.84	1.18 (m, 1H), 1.62 (m, 7H), 2.00 (m, 3H), 3.86 (s, 3H), 6.89 (d, 1H, J = 6.1 Hz), 7.13 (d, 1H, J = 8.4 Hz)	55	84–85
20d	$C_{21}H_{26}O_3FCl$ (380.89)	66.22 65.89	6.88 6.87	1.30 (m, 2H), 1.62 (m, 14H), 2.10 (m, 6H), 3.84 (s, 3H), 6.86 (d, 1H, J = 5.8 Hz)	16	149–150
20e	$C_{20}H_{24}O_1F_2$ (318.41)	75.44 75.30	7.60 7.63	1.17 (s, 6H), 1.56–1.69 (m, 11H), 1.97 (m, 4H), 2.56 (t, 2H), 6.78 (dd, 1H)	38	78–79
20f	$C_{13}H_{15}ON$ (201.27)	77.58 77.40	7.51 7.56	1.47 (m, 1H), 1.80 (m, 7H), 2.14 (m, 2H), 3.21 (s, 1H), 7.41 (d, 2H), 8.70 (d, 2H)	26	103–104
20g	$C_{12}H_{14}F_3ClN_2O$ (306.71)	50.91 50.67	4.60 4.59	1.25 (m, 1H), 1.50 (m, 7H), 1.97 (m, 2H), 2.22 (s, 1H), 6.86 (s, 3H)	64	oil
21a	$C_{14}H_{12}F_4O_2$ (288.24)	58.34 58.38	4.20 4.31	1.79 (m, 5H), 2.01 (m, 4H), 4.03 (s, 3H)	22	oil
21b	$C_{16}H_{16}F_4O_2$ (316.29)	60.76 60.88	5.10 5.03	1.09 (d, 3H), 1.21 (m, 2H), 1.63 (m, 6H), 1.92 (s, 1H), 2.06 (m, 1H), 4.07 (s, 3H)	16	oil
21c	$C_{16}H_{16}F_4O_2$ (316.29)	60.76 60.94	5.10 5.06	1.09 (d, 3H), 1.28 (m, 2H), 1.59 (m, 6H), 2.10 (m, 1H), 2.25 (s, 1H), 4.08 (s, 3H)	23	36–37
21d	$C_{16}H_{16}F_4O_2$ (316.29)	60.76 60.85	5.10 5.04	1.60 (m, 8H), 1.93 (m, 3H), 2.13 (m, 2H), 4.08 (s, 3H)	50	75–76
21e	$C_{15}H_{16}F_4O_2$ (304.28)	59.21 59.32	5.30 5.34	1.08 (s, 9H), 1.53 (s, 3H), 2.06 (s, 1H), 4.08 (s, 3H)	43	oil
21f	$C_9H_4F_4O$ (204.12)	52.96 52.77	1.98 1.96	3.53 (s, 1H), 4.10 (s, 3H)	6	51–52
22	$C_{13}H_{11}F_5O$ (278.22)	56.12 56.25	3.99 4.02	0.90 (t, 3H), 1.50 (m, 2H), 1.52 (s, 3H), 1.66 (m, 2H), 2.13 (s, 1H)	36	oil

Table 3 (Continued)

product	molecular formula	calcd found C	calcd found H	^{1}H NMR ($CDCl_3$) δ[a]	yield[b]	mp[c] (°C)
23a	$C_{18}H_{13}F_5O$ (340.29)	63.53 63.76	3.85 3.89	1.57 (s, 3H), 2.01 (m, 2H), 2.06 (s, 1H), 2.85 (m, 2H), 7.16 (m, 5H)	42	60–61
23b	$C_{22}H_{23}F_4NO$ (393.42)	67.16 67.19	5.89 5.93	0.87 (t, 3H), 1.31 (m, 2H), 1.49 (m, 2H), 1.55 (s, 3H), 1.89 (m, 2H), 2.84 (m, 2H), 3.32 (m, 2H), 7.17 (m, 5H)	76	oil
23c	$C_{23}H_{24}F_4O_2$ (408.44)	67.64 67.30	5.92 6.34	0.89 (t, 3H), 1.41 (m, 4H), 1.64 (s, 3H), 1.77 (m, 2H), 2.10 (m, 2H), 2.90 (m, 2H), 4.24 (t, 2H), 7.25 (m, 5H)	42	oil
24a	$C_{15}H_{16}F_5NO$ (321.29)	56.08 56.09	5.02 5.01	1.00 (t, 6H), 1.47 (s, 3H), 2.46 (d, 1H, J = 13.3 Hz), 2.70 (m, 2H), 2.82 (m, 2H), 2.86 (d, 1H, J = 13.3 Hz)	14	oil
24b	$C_{20}H_{27}F_4NO_2$ (389.43)	61.68 62.05	6.99 6.90	0.85 (t, 3H), 1.08 (t, 6H), 1.32 (m, 4H), 1.47 (s, 3H), 1.70 (m, 2H), 2.43 (d, 1H, J = 13.3 Hz), 2.67 (m, 2H), 2.83 (m, 2H), 2.85 (d, 1H, J = 13.3 Hz), 4.16 (t, 2H)	50	oil
24c	$C_{19}H_{26}F_4N_2O$ (374.42)	60.95 60.68	7.00 6.91	0.92 (t, 3H), 1.03 (t, 6H), 1.35 (m, 2H), 1.50 (s, 3H), 1.55 (m, 2H), 2.45 (d, 1H, J = 13.2 Hz), 2.67 (m, 2H), 2.86 (m, 2H), 2.87 (d, 1H, J = 13.2 Hz), 3.37 (m, 2H), 3.91 (bs, 1H)	62	oil
24d	$C_{18}H_{24}F_4N_2O$ (360.39)	59.99 60.22	6.71 6.74	1.06 (t, 6H), 1.12 (t, 3H), 1.52 (s, 3H), 2.50 (d, 1H, J = 13.2 Hz), 2.73 (m, 2H), 2.85 (m, 2H), 2.85 (d, 1H, J = 13.2 Hz), 2.93 (s, 3H), 3.21 (q, 2H)	64	oil
25a	$C_{12}H_9F_5O_2$ (280.19)	51.44 51.54	3.24 3.29	1.49 (s, 3H), 3.31 (d, 1H, J = 9.3 Hz), 3.43 (s, 3H), 3.51 (d, 1H, J = 9.3 Hz)	41	oil
25b	$C_{16}H_{18}F_4SO_2$ (350.37)	54.85 54.70	5.18 5.01	0.88 (t, 3H), 1.36 (m, 2H), 1.52 (m, 2H), 1.56 (s, 3H), 2.93 (t, 2H), 3.43 (d, 1H, J = 9.3 Hz), 3.49 (s, 3H), 3.57 (d, 1H, J = 9.3 Hz)	28	oil
25c	$C_{18}H_{23}F_4NO_2$ (361.38)	59.83 59.58	6.42 6.38	0.78 (t, 3H), 1.40 (m, 4H), 1.48 (s, 3H), 3.08 (t, 4H), 3.40 (d, 1H, J = 9.3 Hz), 3.44 (s, 3H), 3.50 (d, 1H, J = 9.3 Hz)	12	oil
26a	$C_{15}H_{15}F_5O_2$ (306.27)	58.82 59.16	4.94 5.00	0.95 (s, 9H), 1.96 (s, 3H), 2.04 (s, 2H)	27	oil
26b	$C_{18}H_{23}F_4NO$ (345.38)	62.60 62.67	6.71 6.75	1.12 (s, 9H), 1.60 (t, 3H), 1.77 (s, 2H), 2.94 (bs, 3H), 3.22 (q, 2H)	76	oil
27a	$C_{16}H_{17}F_5O$ (320.30)	60.00 60.22	5.35 5.61	0.91 (t, 3H), 0.97 (d, 6H), 1.54 (m, 4H), 1.86 (bs, 1H), 1.96 (septet, 1H), 2.55 (s, 2H)	28	oil
27b	$C_{20}H_{27}F_4NO$ (373.43)	64.33 64.40	7.29 7.25	0.90 (t, 3H), 0.98 (t, 3H), 1.03 (d, 6H), 1.36 (m, 2H), 1.59 (m, 6H), 2.20 (m, 2H), 3.40 (t, 2H)	68	oil
28	$C_{15}H_{15}F_5O$ (306.27)	58.82 59.03	4.94 4.95	0.91 (t, 6H), 1.49 (m, 4H), 1.64 (m, 4H), 2.07 (s, 1H)	44	oil
29	$C_{19}H_{15}F_5O$ (354.32)	64.41 64.28	4.27 4.50	1.00 (t, 3H), 1.67 (m, 2H), 1.79 (m, 2H), 2.95 (d, 1H, J = 13.3 Hz), 3.09 (d, 1H, J = 13.3 Hz), 7.35 (m, 5H)	4	oil
32	$C_{20}H_{16}F_4O$ (348.34)	68.96 68.73	4.63 4.71	0.77 (d, 3H), 0.81 (d, 3H), 2.08 (septet, 1H), 2.29 (m, 2H), 3.50 (s, 1H), 5.11 (s, 1H), 7.24 (m, 5H)	17	oil
33	$C_{14}H_{15}F_5O_2$ (310.26)	54.20 54.10	4.87 4.85	0.81 (t, 3H), 1.31 (s, 3H), 2.00 (q, 2H), 2.50 (bs, 1H), 3.28 (d, 1H, J = 9.2 Hz), 3.34 (s, 3H), 3.46 (d, 1H, J = 9.2 Hz), 6.16 (s, 3H)	78	oil

[a] In $CDCl_3$ with TMS as internal standard; Bruker WM-360 and Varian XL-400 NMR spectrometers. [b] Yields are unoptimized with emphasis on purity of products rather than quantity. [c] Melting points were determined with a Mettler PF62 capillary melting point apparatus and are uncorrected.

enthetically below the relevant coefficients, with the residual standard deviation from regression given below the respective intercepts. Note that the constant term in each equation has been factored out to make it easier to compare coefficients and estimated errors. Note, too, that the absence of a term in σ_Y* results in a shift in the implied optimum $E_Y^{\ddagger}$ toward less hindered groups.

$$\underset{(\pm 7.66)}{SRS_{BYGR} = 74.50}\{1 + \underset{(\pm 0.093)}{0.389 I_3} - \underset{(\pm 0.200)}{1.022}(\pi_1 - 0.650)^2 - \underset{(\pm 0.384)}{1.456}(E_Y - 0.931)^2 - \underset{(\pm 0.054)}{0.158 E_2} + \underset{(\pm 0.075)}{0.240\sigma_Y^*}\} \quad (r^2 = 0.835) \quad (3)$$

$$\underset{(\pm 0.251)}{pED_{80LACG} = 2.11}\{1 + \underset{(\pm 0.098)}{0.221 I_3} - \underset{(\pm 0.313)}{0.848}(\pi_1 - 0.600)^2 - \underset{(\pm 0.154)}{0.676}(E_Y - 0.570)^2\} \quad (r^2 = 0.756) \quad (4)$$

The residual errors obtained are quite close to the SEM values obtained by analysis of variance across replicates (Table 1), which indicates that both models are about as accurate as can be expected given the underlying variability in biological response. Stable models could be obtained for the other response variables in Table 2, but only by centering about an optimal π_1 or E_Y value deduced from SRS_{BYGR} and pED_{80LACG} regressions. SRS_{YENS} allowed an independent direct estimation of the π_1 optimum, whereas both optima had to be constrained to obtain useful QSARs for SRS_{DOBR} and pED_{80av}. The results obtained are set out in eqs 5–7.

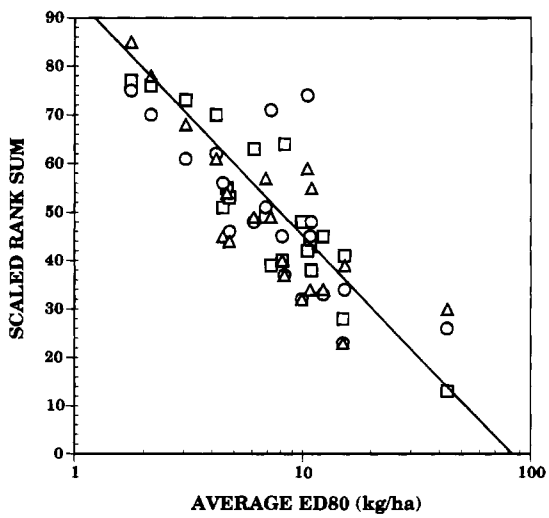


Figure 5. Scaled rank sums as a function of average weed ED_{80}: (○) SRS_{BYGR}; (△) SRS_{DOBR}; (□) SRS_{YENS}.

$$\mathrm{SRS_{YENS}} = 66.49\{1 + 0.359I_3 - 1.089(\pi_1 - 0.593)^2 - 1.013(E_Y - 0.931)^2 - 0.093E_2 + 0.203\sigma_Y^*\} \quad (r^2 = 0.458) \quad (5)$$
$$(\pm 14.17) \quad (\pm 0.196) \quad (\pm 0.578) \quad (\pm 0.793) \quad (\pm 0.111) \quad (\pm 0.155)$$

$$\mathrm{SRS_{DOBR}} = 69.12\{1 + 0.324I_3 - 0.912(\pi_1 - 0.650)^2 - 1.856(E_Y - 0.931)^2 - 0.204E_2 + 0.327\sigma_Y^*\} \quad (r^2 = 0.653) \quad (6)$$
$$(\pm 10.88) \quad (\pm 0.142) \quad (\pm 0.306) \quad (\pm 0.588) \quad (\pm 0.083) \quad (\pm 0.115)$$

$$\mathrm{pED_{80av}} = 1.54\{1 + 0.295I_3 - 0.730(\pi_1 - 0.650)^2 - 0.623(E_Y - 0.931)^2 - 0.080E_2 + 0.098\sigma_Y^*\} \quad (r^2 = 0.486) \quad (7)$$
$$(\pm 0.267) \quad (\pm 0.157) \quad (\pm 0.337) \quad (\pm 0.648) \quad (\pm 0.091) \quad (\pm 0.126)$$

Some nonsignificant coefficients ($t < 2, P > 0.10$) have been retained in eqs 5–7. This was done in part to avoid distorting the values for better-determined coefficients and in part to indicate how close to zero the coefficient in question was predicted to be. A regression coefficient of $+0.010 \pm 0.020$ for E_2, for example, would not in and of itself be significant but would conflict seriously with the value of -0.158 ± 0.054 from eq 3. The value of -0.093 ± 0.111 obtained in eq 5, on the other hand, does not conflict with the value in eq 3. The only major qualitative discrepancy among the models defined by eqs 3–7 is in the optimal value for E_Y. Unfortunately, the experimental design we used precludes a meaningful direct estimation of error for estimates of the substitution parameter optima. The sensitivity of the models to the differing optima was instead examined by a perturbation technique. The E_Y optimum found using $\mathrm{pED_{80LACG}}$ (eq 4) was cross-substituted into the analysis for $\mathrm{SRS_{BYGR}}$. The resulting weakened model obtained is defined by eq 8. Conversely, eq 9 gives the result of centering the data for $\mathrm{pED_{80LACG}}$ on the optimum obtained for $\mathrm{SRS_{BYGR}}$; in addition, note that having shifted the E_Y downward causes the term in σ_Y^* to drop out. In the regression for $\mathrm{pED_{80LACG}}$ the shift in optimum actually affords a small improvement in fit to the data. Moreover, eq 9

$$\mathrm{SRS_{BYGR}} = 82.58\{1 + 0.193I_3 - 0.760(\pi_1 - 0.650)^2 - 0.350(E_Y - 0.570)^2 - 0.138E_2\} \quad (r^2 = 0.740) \quad (8)$$
$$(\pm 9.29) \quad (\pm 0.092) \quad (\pm 0.221) \quad (\pm 0.157) \quad (\pm 0.060)$$

$$\mathrm{pED_{80LACG}} = 1.71\{1 + 0.440I_3 - 1.196(\pi_1 - 0.600)^2 - 1.441(E_Y - 0.931)^2 - 0.132\sigma_2^* + 0.288\sigma_Y^*\} \quad (r^2 = 0.779) \quad (9)$$
$$(\pm 0.255) \quad (\pm 0.143) \quad (\pm 0.404) \quad (\pm 0.578) \quad (\pm 0.057) \quad (\pm 0.114)$$

includes more well-determined terms than does eq 4, which generally makes it a more useful guide for further synthesis.

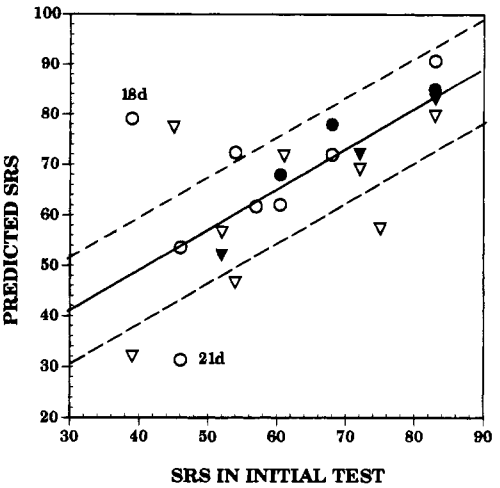


Figure 6. Predicted scaled rank sums (SRS) as a function of actual values in the initial test. Solid symbols (●, ▼) indicate predictions based on individual compounds included in both tests, whereas open symbols (○, ▽) indicate predictions based on regression models: (●, ○) $\mathrm{SRS_{BYGR}}$; (▼, ▽) $\mathrm{SRS_{DOBR}}$. The solid line corresponds to eq 8; the dashed lines correspond to an error range of ±10.

Surprisingly, including global lipohilicities (log P) failed to improve any of the models examined. This was true whether terms in log P were included initially or added in to the simplified models.

Predictivity. Residual deviation from regression and r^2 are valid measures of a model's internal consistency and goodness-of-fit, but the ability to predict the activity of extraneous compounds is of greater interest in most QSAR applications. In our case, those first-tier analogs which fit the template shown for Table 1 were used to test the model's predictivity. Predicted values for $\mathrm{SRS_{BYGR}}$ and for $\mathrm{SRS_{DOBR}}$ are plotted in Figure 6 as a function of their actual scores in our initial side-by-side test (Table 1); solid symbols show the relationship between actual scaled rank sums for the three compounds common to both tests.

The correlation is only fair, with Spearman rank correlations (Spearman, 1904) between predicted and observed values of $\varrho = 0.347$ for $\mathrm{SRS_{BYGR}}$ (eq 3), 0.524 for $\mathrm{SRS_{DOBR}}$ (eq 6), and 0.395 for $\mathrm{pED_{80LACG}}$ (eq 4); the respective r^2 values were 0.294, 0.298, and 0.359. Applying alternative eq 8 to predict $\mathrm{SRS_{BYGR}}$ lowers the value of the predictive ϱ to 0.240 but raises the predictive r^2 slightly to 0.333. Using eq 9 instead of eq 4 to predict $\mathrm{pED_{80LACG}}$, on the other hand, increases both ϱ (to 0.575) and r^2 (to 0.423). In all cases, the major outliers are **21d** and **18d**. That the former analog is more active than predicted presumably reflects extraordinary conformational constraints imposed by its cycloheptyl ring. The pentyloxy **18d**, on the other hand, bears a more extended Y group than did any example in the experimental design set upon which the model was based. That it is much less active than predicted may reflect some binding site "wall" which the models examined here do not adequately accommodate.

Final Model. There is overall a roughly linear relationship between the predicted and observed scaled rank sums, which is to be expected since different titration levels and activity ranges were included in the two tests. Inspection of the data for those compounds common to both tests suggests the relationship defined by eq 10, which corresponds to the solid line in Figure 6. This defines a median line, such that half of the data points fall above it and half below it.

$$SRS_{pred} = 0.8(SRS_{initial}) + 17 \quad (10)$$

An experimental error in measurement in SRS of approximately ±10 (Tables 1 and 2) gives an anticipated range set off in Figure 6 by the dashed lines, which is clearly consistent with the data obtained. Equation 10 was then used to transform data from the initial test for inclusion in a composite QSAR. Application of stepwise multiple regression to data pooled from both tests then gave the model defined by

$$SRS_{BYGR} = \underset{(\pm 8.22)}{76.55}\{1 + \underset{(\pm 0.081)}{0.172}I_3 - \underset{(\pm 0.111)}{0.505}(\pi_1 - 0.697)^2 - \underset{(\pm 0.190)}{0.568}(E_Y - 0.720)^2 - \underset{(\pm 0.055)}{0.143}E_2\} \quad (r^2 = 0.718) \quad (11)$$

This provides an additional indication of how sensitive eq 3 is to incidental changes in the data from which it was derived. The composite model implies that the optimal substituent R_1 is a methyl group ($\pi = 0.65$) or part of a ring ($\pi \approx 0.8$ for methylene); somewhat counterintuitively, the models all suggest that the smaller R_2 is, the better. The optimal substituent Y is one for which E_{SV} falls somewhere between 0.6 and 0.9, and the positive coefficients in σ_Y* evident in eqs 3, 5, and 6 suggest that electronegative groups at this position tend to enhance activity somewhat. Note again that the lower E_Y optimum counterbalances the loss of the σ_Y* term. Isopropoxy ($E_{SV} = 0.75$, $\sigma^* = 1.51$) should be optimal. This E_{SV} range encompasses several other moderately hindering nucleophilic groups, but ethylthio ($E_{SV} = 0.94$, $\sigma^* = 1.44$), butylamino ($E_{SV} = 0.70$, $\sigma^* = 1.08$), and methylthio ($E_{SV} = 0.60$, $\sigma^* = 1.47$) analogs, for example, are all expected to be appreciably less active.

By far the most effective way to increase herbicidal activity, however, is to add a methyl group R_3 at the α carbon ($I_3 = 1$), which confers a 20–40% jump in SRS, which corresponds to a 2-fold reduction in ED_{80} (*cf.* Figure 5). Further substitution at this carbon (*e.g.*, $I_4 = 1$), however, confers no significant increase in potency. In fact, since E_{SV} is necessarily greater than or equal to 0, eq 11 indicates that the maximum attainable SRS_{BYGR} is about 90, corresponding (Figure 5) to an ED_{80av} somewhat greater than 1 kg/ha. This is less than twice the potency (half the ED_{80}) of **21c** in the second test, which was the most active analog in both tests. Since even a "perfect" analog would therefore not be active enough to commercialize, this synthesis project was discontinued.

LITERATURE CITED

Brannigan, L. H.; Duewer, D. L. Experimental Design in the Development of Biological Active Compounds. *Pharmacochem. Libr.* **1991**, *16*, 553–556.

Bushweller, C. H.; Beach, J. A.; O'Neil, J. W.; Rao, G. U. The A Value of Hydroxyl Determined by the Nuclear Magnetic Resonance Peak Area Method at −83°. *J. Org. Chem.* **1970**, *35*, 2086–2087.

Duewer, D. L.; Clark, R. D. Rank-Order Analysis for Robust Multiresponse, Multiblock Comparisons: Evaluation of Herbicide Interactions. *J. Chemom.* **1991**, *5*, 503–521.

Gordon, A. J. *The Chemist's Companion*; Wiley: New York, 1972; pp 156–157.

Plackett, R. L.; Burman, J. P. The Design of Optimum Multifactorial Experiments. *Biometrika* **1946**, *33*, 305–328.

Spearman, C. The Proof and Measurement of Association Between Two Things. *Am. J. Psychol.* **1904**, *15*, 72–101.

Received for review April 14, 1994. Accepted August 15, 1994.[⊗]

[⊗] Abstract published in *Advance ACS Abstracts,* October 1, 1994.