Assessment of Additive/Nonadditive Effects in Structure−Activity Relationships: Implications for Iterative Drug Design

Yogendra Patel,[†,⊥] Valerie J. Gillet,[†] Trevor Howe,[‡] Joaquin Pastor,[§,#] Julen Oyarzabal,*[,∥,#] and Peter Willett[†]

Department of Information Studies, University of Sheffield, Regent Court, 211 Portobello Street, Sheffield S1 4DP, United Kingdom, Department of Molecular Informatics, Johnson & Johnson Pharmaceutical R&D, Turnhoutseweg 30, 2340 Beerse, Belgium, Department of Medicinal Chemistry, Johnson & Johnson Pharmaceutical R&D, Jarama 75, 45007 Toledo, Spain, Department of Molecular Informatics, Johnson & Johnson Pharmaceutical R&D, Jarama 75, 45007 Toledo, Spain

Received August 28, 2008

Free-Wilson (FW) analysis is common practice in medicinal chemistry and is based on the assumption that the contributions to activity made by substituents at different substitution positions are additive. We analyze eight near complete combinatorial libraries assayed on several different biological response(s) (GPCR, ion channel, kinase and P450 targets) and show that only half-exhibit clear additive behavior, which leads us to question the concept of additivity that is widely taken for granted in drug discovery. Next, we report a series of retrospective experiments in which subsets are extracted from the libraries for FW analysis to determine the minimum attributes (size, distribution of substituents, and activity range) necessary to reach the same conclusion about additive/nonadditive effects. These attributes can provide guidelines on when it is appropriate to apply FW analysis as well as for library design, and they also have important implications for further steps in iterative drug design.

Introduction

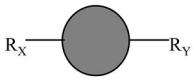
Medicinal chemistry is a complex universe in which the chemical structure space of biochemically active compounds, ligands or drugs, overlaps with the biological structure space formed by the target molecules.[1] Many attempts have been made to rationalize and quantify the complexity of medicinal chemistry space by considering chemical and biological systems, for example, by predicting ligand binding affinities; however, such analyses may lead to unreliable models because they are often based on flawed assumptions, such as additivity principles in biochemistry.[2] Indeed it has been appreciated for many years that additivity does not apply in complex systems, e.g., protein−ligand complexes.[2−4] However, the concept of additivity is a fundamental premise that is widely taken for granted[2] in different approaches in the drug discovery process, including: fragment-based ligand design, scoring functions and free energy calculations for ligand binding to proteins, and quantitative structure−activity relationships (QSARs[a]).

Fragment-based ligand design is a constructive approach to inhibitor design[5] in which different experimental approaches such as NMR,[6−8] tethering[9−13] and X-ray crystallography[14−17] are used to identify low-molecular-weight fragments that bind to a receptor. The fragments are then used as starting points to construct a lead compound under the assumption that the binding modes of the individual fragments are only minimally perturbed upon elaboration.[18] There are many cases where additivity in fragment-based ligand design has been demonstrated. For example, Stout et al.[15] deconstructed a natural substrate into fragments and used X-ray crystallography to demonstrate minimal perturbation in the binding of the fragments compared to the substrate as a whole. Saxty et al.[16] also used X-ray crystallography to iteratively elaborate an initial fragment hit into a nanomolar inhibitor and subsequently used a Free-Wilson analysis[19] to estimate the binding affinity of individual functional groups in order to direct further optimization. However, there are also many counter examples where additivity is not exhibited. For example, Babaoglu and Shoichet[17] found that none of the fragments resulting from substrate deconstruction bound in positions corresponding to their placement in the parent compound. Moreover, there are many examples where binding affinity only emerges once a critical number of functional groups are present with the molecular affinity far exceeding the sum of the component fragments: these include biotin,[20] several CDK inhibitors,[21] and stromelysin inhibitors.[22] In substrate recognition, too, distinct functional groups can combine with pronounced nonadditivity.[23]

Many of the scoring functions that are widely used to approximate the free energy of binding in protein−ligand docking also assume additivity, including those based on empirical and force field methods as found in DOCK,[24] CHARMm,[25] AutoDock,[26] GOLD,[27] LUDI,[28] FlexX,[29] ChemScore,[30] and GLIDE.[31] Force field methods approximate the binding free energy of a protein−ligand complex by a sum of van der Waals, electrostatics, and other contributions. The basic assumption is that different contributions to free energy of binding can be calculated separately and are additive; however, protein−ligand recognition is a concerted event and the thermodynamic quantities cannot be simply summed.[2] While these approaches are valid as prioritization tools in virtual screening and are useful for predicting experimental binding modes, none of the scoring functions has been shown to achieve useful estimates of experimental binding affinities.[32]

* To whom correspondence should be addressed. Phone: +34 91 7328000. E-mail: joyarzabal@cnio.es.
[†] Department of Information Studies, University of Sheffield.
[‡] Department of Molecular Informatics, Johnson & Johnson Pharmaceutical R&D.
[§] Department of Medicinal Chemistry, Johnson & Johnson Pharmaceutical R&D.
[∥] Department of Molecular Informatics, Johnson & Johnson Pharmaceutical R&D.
[⊥] Current address: Manchester Interdisciplinary Biocentre, The University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom.
[#] Current address: Spanish National Cancer Research Centre (CNIO), Experimental Therapeutics Programme, Melchor Fernandez Almagro 3, 28029 Madrid, Spain.
[a] Abbreviations: FW, Free-Wilson; SSE, scaled Shannon entropy; SAR, Structure−activity relationship.

10.1021/jm801070q CCC: $40.75 © 2008 American Chemical Society
Published on Web 11/14/2008



R_X = X1, X2,....
R_Y = Y1, Y2, ...

	Indicator variables					
Molecule ID	X1	X2	.....	Y1	Y2	.....
1	1	0	.....	1	0	.....
2	0	1	.....	1	0	.....
3	1	0	.....	0	1	.....
4	0	1	.....	0	1	.....
.....	.....	.....	.....	.....	.....	.....

$$y = C + a_1X1 + a_2X2 + a_3X3 + ...b_1Y1 + b_2Y2 + b_3Y3 + ...$$

$$y = C + \sum_{i=1}^{n} a_iXi + \sum_{j=1}^{m} b_jYj$$

Figure 1. FW analysis. X1, X2, Y1, and Y2 etc are possible R group substituents at two positions of variability on a common scaffold (R_X and R_Y respectively). Each molecule is represented by an indicator variable, one for each substituent: a value of 1 indicates that a molecule has that particular substituent at a particular position; 0 indicates that it does not.

In the absence of structural information for the receptor, quantitative predictions of ligand binding affinities can be obtained using QSAR techniques. However these techniques are also often based on the assumption of additivity. For example, Free-Wilson (FW) analysis can be applied to a series of analogues consisting of a common core structure and variable substituents, with the aim of deriving quantitative estimations of the contributions made by each substituent to overall activity.[19] The basic assumptions in FW analysis are that each substituent makes an additive contribution activity and that the substituent contributions are independent.[19] The FW method is limited to data sets of structurally similar ligands such as those explored in lead optimization. When applied in the right context, the FW approach (or its modified Fujita-Ban version[33]) can be very successful and it has been used extensively in medicinal chemistry, from the early days of QSAR through to the present day.[34−38] Moreover, there has been a resurgence of interest in FW analysis following the development of combinatorial chemistry techniques and fragment-based drug design. In combinatorial chemistry, a FW analysis can be used to predict the activity of all compounds in a full combinatorial matrix based on a cross-section of analogues that have already been synthesized. If a good model is obtained, then it is assumed that additive effects are present. In cases where activity depends significantly on the simultaneous presence or absence of two or more substituents and nonadditive substituent effects are present, then a FW analysis is not valid and nonlinear techniques such as artificial neural networks should be used.[39]

Given the complexity of protein−ligand interactions, and of medicinal chemistry space in general, the assumption of additivity should not be taken for granted because it can mislead future design efforts. Thus it is important to examine the SAR to determine if additivity is present. Furthermore, establishing the presence or absence of additive effects early in the lead optimization process also has clear efficiency implications for subsequent iterations of the design and test cycle: for data sets that are shown to be additive, a sequential synthetic approach can be followed based on the substituent contributions estimated by a Free-Wilson analysis; whereas for data sets that show nonadditive behavior, a fully combinatorial approach should be taken.[40,41]

Here we carry out a retrospective FW analysis of near-complete combinatorial libraries, tested over several biological assays, in order to: first, estimate whether additive or nonadditive effects are present for different combinations of chemical spaces and biological assays and, second, determine the minimum attributes of a data set (size, distribution of properties etc.) necessary to reach the same conclusion on additivity. The aim is to establish guidelines on the size and characteristics of compounds sets necessary to determine additivity in prospective iterative drug design due to its clear implications in this process.

Method

Free Wilson (FW) analysis is based on simple indicator variables that relate to different substituents on a common scaffold. Each molecule is described by a series of indicator variables, one for

Table 1. Details of the Various Data Sets Used

data set	property	no. of R_X groups	no. of R_Y groups	no. of compds in complete data set	no. of compds	percentage coverage of complete data set	min property value	max property value	range of property values
4R 1_136	**K channel**	**7**	**11**	**77**	**54**	**70.1**	**5.05**	**7.17**	**2.12**
4R 1_56	**Na channel**	**6**	**16**	**96**	**72**	**75.0**	**5.06**	**7.07**	**2.01**
4R 1_73	**class A GPCR**	**7**	**17**	**119**	**81**	**68.1**	**6.04**	**8.98**	**2.94**
D1 1_62	class A GPCR	14	15	210	178	84.8	5.01	7.86	2.85
D1 1_64	class A GPCR	14	15	210	178	84.8	5.16	8.99	3.83
D2 1_62	**class A GPCR**	**9**	**12**	**108**	**87**	**80.6**	**5.30**	**7.86**	**2.56**
D2 1_64	**class A GPCR**	**9**	**12**	**108**	**87**	**80.6**	**5.52**	**8.99**	**3.47**
D2 58_7	**class A GPCR**	**9**	**12**	**108**	**87**	**80.6**	**5.01**	**7.91**	**2.9**
D3 1_73	class A GPCR	22	5	110	98	89.1	5.46	8.98	3.52
D3 221_5	cellular metabolism	22	5	110	82	74.5	12.00	100.00	88.00
D4 100_152	**ion channel**	**13**	**7**	**91**	**69**	**75.8**	**5.03**	**7.98**	**2.95**
D5 100_193	class A GPCR	7	11	77	54	70.1	5.07	8.79	3.72
D6 76_97	**Ser/Thr kinase**	**19**	**6**	**114**	**96**	**84.2**	**6.79**	**8.90**	**2.11**
D7 221_5	cellular metabolism	12	5	60	42	70.0	0.00	91.5	91.5
D7 339_3	P450 3A4	12	5	60	40	66.7	12.01	85.26	73.25
D7 339_5	P450 2C9	12	5	60	40	66.7	16.81	110.79	93.98
D7 339_6	P450 2D6	12	5	60	40	66.7	11.35	54.23	42.88
D7 339_7	P450 1A2	12	5	60	40	66.7	3.93	102.75	98.82
D7 76_97	Ser/Thr kinase	12	5	60	42	70.0	7.18	8.87	1.69

Table 2. Characteristics of the Test Sets are Summarized over the Five Test Sets for Each Data Set and Include: The Maximum and Minimum Ranges of the Properties Being Predicted, the Maximum and Minimum SSE for Each R group, the Number of Compounds, and the Percentage of the Complete Data Set that is Represented by the Test Sets

data set	max property range	min property range	max R_x SSE	min R_x SSE	max R_y SSE	min R_y SSE	no. of compds	percentage coverage of complete data set
4R 1_136	**2.05**	**1.48**	**0.973**	**0.924**	**1.000**	**1.000**	**11**	**14**
4R 1_56	**1.89**	**1.43**	**0.949**	**0.910**	**1.000**	**1.000**	**16**	**17**
4R 1_73	**2.91**	**1.96**	**0.979**	**0.864**	**1.000**	**1.000**	**17**	**14**
D1 1_62	2.26	1.52	0.971	0.927	0.984	0.942	29−32	14−15
D1 1_64	3.65	2.59	0.968	0.950	0.974	0.951	30−32	14−15
D2 1_62	**2.56**	**1.77**	**0.973**	**0.953**	**1.000**	**1.000**	**12**	**11**
D2 1_64	**3.32**	**2.22**	**0.973**	**0.921**	**1.000**	**1.000**	**12**	**11**
D2 58_7	**2.74**	**1.57**	**0.973**	**0.953**	**1.000**	**1.000**	**12**	**11**
D3 1_73	3.33	2.42	0.995	0.995	0.991	0.870	23	21
D3 221_5	86.50	63.04	0.991	0.981	0.967	0.910	24−25	22−23
D4 100_152	**2.95**	**1.87**	**1.000**	**1.000**	**0.969**	**0.870**	**13**	**14**
D5 100_193	3.72	1.15	0.973	0.949	1.000	1.000	11	14
D6 76_97	**2.08**	**1.75**	**1.000**	**0.994**	**0.969**	**0.936**	**19−20**	**17−18**
D7 221_5	91.50	91.50	1.000	1.000	0.943	0.885	12	20
D7 339_3	73.25	73.25	1.000	1.000	0.960	0.885	12	20
D7 339_5	93.98	93.98	1.000	1.000	0.987	0.840	12	20
D7 339_6	42.88	42.88	1.000	1.000	0.987	0.916	12	20
D7 339_7	98.82	98.82	1.000	1.000	0.987	0.916	12	20
D7 76_97	1.68	1.68	1.000	1.000	0.970	0.912	12	20

Table 3. Characteristics of the Training Sets Including: The Number of Training Sets per Test Set, the Maximum and Minimum Ranges of the Properties Being Predicted, the Maximum and Minimum SSE for each R Group, the Maximum and Minimum Number of Compounds, and the Range of Training Set Sizes as a Percentage of the Number of Compounds Required for a Complete Data Set

data set	no. of training sets per test set	max property range	min property range	max R_X SSE	min R_X SSE	max R_Y SSE	min R_Y SSE	min no. of compds	max no. of compds	range of training set sizes as %
4R_1_136	**92**	**2.12**	**1.37**	**0.992**	**0.740**	**0.996**	**0.926**	**15**	**43**	**20−56**
4R_1_56	**102**	**1.93**	**1.45**	**0.991**	**0.712**	**0.996**	**0.929**	**21**	**56**	**22−58**
4R_1_73	**122**	**2.94**	**2.68**	**0.991**	**0.790**	**0.994**	**0.936**	**25**	**64**	**21−54**
D1_1_62	77	2.85	1.74	0.998	0.870	0.997	0.929	30	149	14−71
D1_1_64	77	3.83	2.18	0.997	0.872	0.996	0.933	29	148	14−71
D2_1_62	**77**	**2.56**	**1.58**	**0.998**	**0.898**	**0.996**	**0.925**	**28**	**78**	**26−72**
D2_1_64	**77**	**3.47**	**2.22**	**0.999**	**0.898**	**0.997**	**0.928**	**28**	**75**	**26−69**
D2_58_7	**97**	**2.90**	**1.57**	**0.998**	**0.842**	**0.996**	**0.927**	**22**	**78**	**26−72**
D3_1_73	62	3.52	2.59	0.999	0.921	0.999	0.924	43	75	39−68
D3_221_5	37	88.00	83.50	0.999	0.940	0.997	0.882	45	58	41−53
D4_100_152	**202**	**2.95**	**1.92**	**1.000**	**0.858**	**0.991**	**0.874**	**25**	**56**	**28−62**
D5_100_193	87	3.72	2.26	0.999	0.725	0.993	0.921	21	43	27−56
D6_76_97	**97**	**2.11**	**1.69**	**1.000**	**0.909**	**1.000**	**0.933**	**35**	**77**	**31−68**
D7_221_5	47	85.50	83.50	1.000	0.928	0.998	0.866	23	30	38−50
D7_339_3	92	70.79	67.11	0.986	0.928	0.998	0.871	23	28	38−47
D7_339_5	92	83.79	83.79	0.986	0.928	0.996	0.880	23	28	38−47
D7_339_6	92	35.44	35.06	0.986	0.928	0.998	0.885	23	28	38−47
D7_339_7	92	98.21	94.67	0.986	0.928	0.986	0.838	23	28	38−47
D7_76_97	47	1.33	1.33	1.000	0.928	0.996	0.838	23	30	38−50

each substituent, with a variable having the value of 1 if the substituent is present in the molecule and 0 if it is absent. Coefficients for the substituents are derived by applying multiple linear regression or partial least-squares (PLS) and indicate the contribution that each substituent makes to the property of interest, see Figure 1.

We first apply FW analysis to (near) complete combinatorial libraries in order to classify them as *additive* or *nonadditive*: a library for which a good model is obtained is classified as additive, whereas a library for which the FW analysis fails is classified as nonadditive with the failure to generate a good model being indicative of interaction effects occurring between the substituents. For each library identified as additive, we then carry out a series of retrospective experiments in which compounds are progressively removed from the library and the FW analysis is applied to the reduced library in order to determine the minimum characteristics required to determine additive effects.

Data Sets. A total of 19 data sets of varying sizes was retrieved from the Johnson & Johnson Pharmaceutical R&D database comprising eight different sets of compounds (labeled 4R, D1, D2, D3, D4, D5, D6, and D7) with each compound set tested in from one to six assays that include different ion channels (pIC_{50} values),

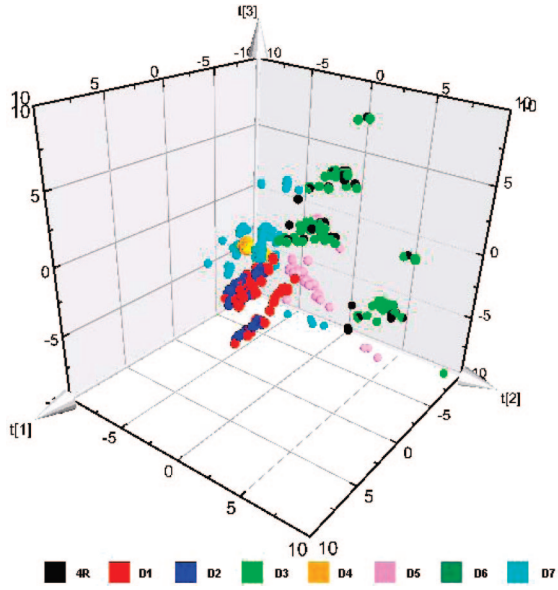


Figure 2. The three principal components (t[1], t[2] and t[3]) are derived from 213 1D and 2D descriptors (including electrotopological state keys, VSA descriptors and molecular property counts as implemented in Pipeline Pilot[42]) and cover 60.9% of the chemical space, defined by the 705 compounds in the data sets reported in Table 1.

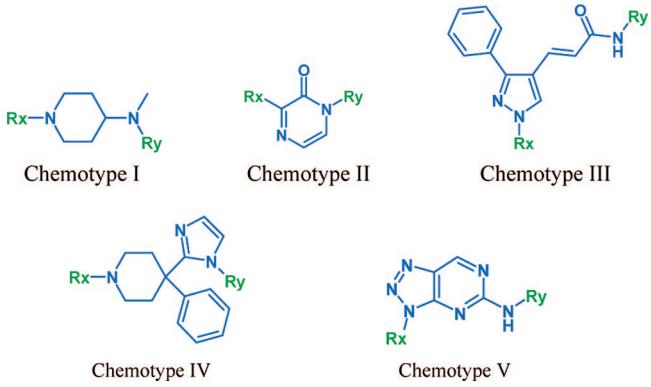


Figure 3. Compounds reported in Table 1 belong to these five scaffolds, and provide a good degree of diversity in chemical space.

class A GPCRs (pIC_{50} values), Ser/Thr Kinase (pIC_{50} values), cytochrome P450 isoforms (percentage of inhibition at 10 μM, as substrate concentration), and metabolic stability in human liver microsomes (percentage of parent compound after 15 min incubation). Taken together, these assays cover a broad biological space. All data sets contained two positions of variability. For three of the data sets (4R, D3, and D7) the numbers of compounds tested in each assay varies. Ideally, the data sets used in the analysis would be complete arrays with all compounds tested on all properties, however, no such data sets were available. The data sets were chosen as being as near complete as possible and vary from 67−89% complete. Table 1 gives summary details of the data sets and the assays they have been tested against (further details of all the data sets are provided as Supporting Information). The chemical space covered by all the compounds in the data sets is represented in Figure 2.

The eight compound sets reported in Table 1, and represented in Figure 2, belong to five different chemotypes. Each of these scaffolds bears two diversity points, R_x and R_y, as shown in Figure 3: (i) Chemotype I describes structures in the 4R (up to 81 compounds) and D3 (up to 98 compounds) data sets,[43] (ii) Chemotype II describes structures in the D1 (up to 178 compounds) and D2 (up to 87 compounds) data sets,[44] (iii) Chemotype III describes structures in the D4 (69 compounds) data set,[45] (iv) Chemotype IV describes structures reported in the D5 (54 compounds) data set,[46] and (v) Chemotype V describes structures in the D6 (96 compounds) and D7 (up to 42 compounds) data sets.[47]

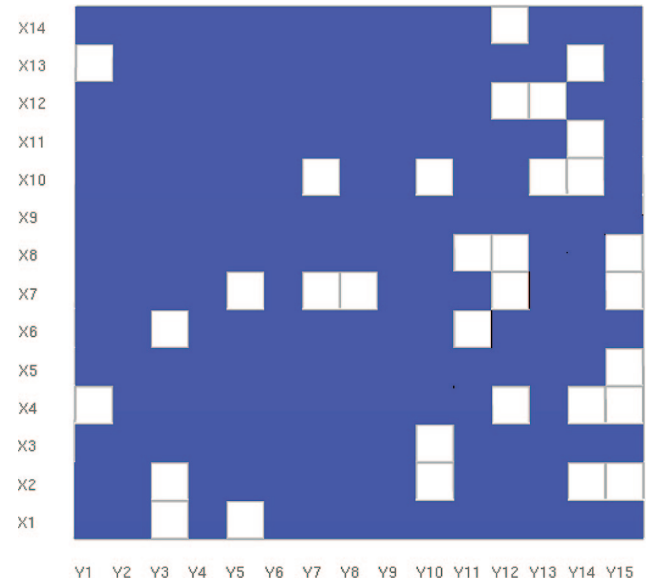


Figure 4. Distribution of substituents for the D1 data set. The dark squares indicate compounds present in the data set, with the white squares indicating compounds missing from the complete matrix.

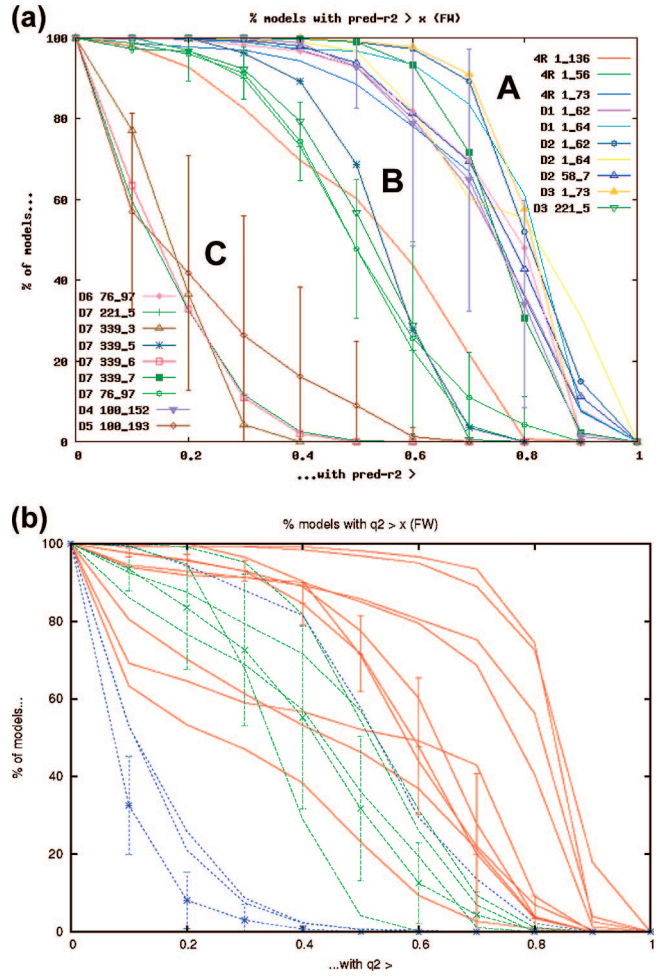


Figure 5. (a) Percentage of models with pred-$r^2 > x$. (b) Percentage of models with $q^2 > x$. The data sets classified as additive according to pred-r^2 are shown in red; those classified as partially additive are in green, and those classified as nonadditive are in blue.

Table 4. Average r^2, q^2, and pred-r^2 Values Obtained for Each Data Set[a]

data set	average r^2	average q^2	average pred-r^2	additivity class
D2 1_62	0.940	0.818	0.801	A
D3 1_73	0.911	0.449	0.794	A
D1 1_64	0.939	0.721	0.79	A
D2 1_64	0.939	0.815	0.772	A
D7 339_7	0.95	0.583	0.748	A
D2 58_7	0.891	0.594	0.742	A
D6 76_97	0.822	0.277	0.740	A
D1 1_62	0.927	0.69	0.722	A
D4 100_152	0.893	0.563	0.72	A
4R 1_73	0.934	0.421	0.716	A
D7 339_5	0.905	0.349	0.534	P
4R 1_136	0.908	0.479	0.518	P
D3 221_5	0.873	0.39	0.502	P
D7 76_97	0.916	0.401	0.500	P
D7 221_5	0.918	0.514	0.484	P
D5 100_193	0.742	0.082	0.191	N
D7 339_3	0.952	0.516	0.172	N
D7 339_6	0.795	0.13	0.155	N
4R 1_56	0.801	0.132	0.151	N

[a] Classification is based on pred-r^2 values where A = additive; P = Partially-additive; and N = non-additive.

Figure 4 shows the distribution of substituents in the D1 data set (similar plots for all data sets are available as Supporting Information).

Design of Training and Test Sets. For each data set, five different test sets were selected using a genetic algorithm (GA)[48] under the constraint that each test set contained at least one occurrence of each R_X and R_Y substituent; the compounds remaining after selection of a test set then formed a *training pool* from which training sets were selected. Multiple training sets were selected for each test set with the aim of varying training set attributes (training set size, range of values of the property to be predicted, and distribution of substituents in each R group), subject to the constraint that there was at least one occurrence of each substituent and ideally at least two occurrences (although the latter was not always possible).

The distribution of substituents for an R group was measured using the scaled Shannon entropy (SSE):

$$\mathrm{SSE} = \frac{\mathrm{SE}}{\log_2 N} \tag{1}$$

$$\mathrm{SE} = -\sum_i p_i \log_2 p_i \tag{2}$$

$$p_i = \frac{c_i}{\mathrm{total\_count}} \tag{3}$$

where SE is the Shannon entropy, N is the number of substituents in the R group, p_i is the probability of a substituent being the ith substituent, c_i is the number of occurrences of the ith substituent, and total_count is the total number of compounds in the training set. Each set of R group substituents (R_X and R_Y) was treated independently. If all of the substituents within an R group occur equifrequently, then the SSE is equal to one; if the training set contains only one substituent for one of the R groups (that is, all compounds occur in a single row (or column) of the matrix), then the SSE for that R group is equal to zero. However, because the subsets are constrained to contain at least one occurrence of each substituent, the SSE for an R group cannot be zero: for example, if there are 10 substituents, then the minimum SSE value that can be obtained is 0.774; if there are 5 different substituents, then the minimum SSE value is 0.809.

The characteristics of the test and training sets are summarized in Tables 2 and 3, respectively.

Classification of the Data Sets. A FW analysis (using the PLS module in MOE 2005.06) was carried out for each training set−test set combination and the resulting models were evaluated using r^2, q^2 (using leave-one-out (LOO) cross-validation), and predicted r^2 (pred-r^2). The resulting models generated for each test set were binned accumulatively according to pred-r^2 with bin values corresponding to: percentage of models with pred-r^2 > 0.1, percentage of models with pred-r^2 > 0.2,..., percentage of models with pred-r^2 > 0.9. The results are plotted in Figure 5a averaged over the five test sets for each data set. Standard deviations are shown for three of the data sets.

The data sets fall into three distinct bands: those where the majority of the models have high pred-r^2 (region A: > 50% of models have pred-r^2 ≥ 0.6), those where very few of the models show high pred-r^2 (region C: < 20% of models have pred-r^2 ≥ 0.4), and the intermediate case (region B: < 50% of models have pred-r^2 ≥ 0.6 and >20% of models have pred-r^2 ≥ 0.4). We conclude that pred-r^2 appears to provide a robust statistic for evaluation of additive effects and models falling in region A were classified as additive, those in region C as nonadditive, and those in region B as partially additive.

Table 4 shows the average r^2, q^2, and pred-r^2 values, obtained for all the models averaged over all training and test set combinations for a given data set, together with the classifications as defined above. While pred-r^2 appears to provide a robust statistic for evaluation of additive effects, in a prospective study, the aim is to determine additivity based on the compounds synthesized so far and the use of a test set is undesirable because it limits the compounds that are available for training. It would be more useful to be able to determine additivity using the q^2 statistic. Figure 5b shows the corresponding plot to Figure 5a based on q^2 (rather than pred-r^2), with the data sets colored according to the classifications in Table 4 (the additive data sets in red, the nonadditive data sets in blue, and the partially additive data sets in green). Unfortunately, although perhaps not surprisingly, the classification is less distinct when q^2 is considered. Although three of the four nonadditive data sets show poor q^2 values as expected, one of the nonadditive data

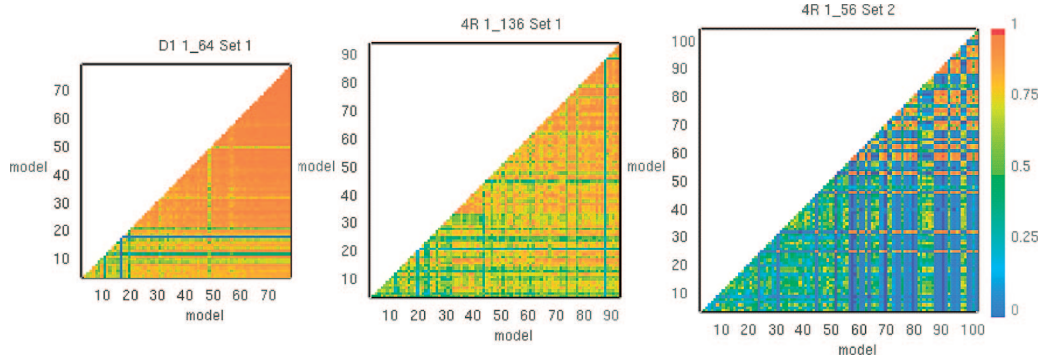


Figure 6. Correlation of coefficients in models generated for one test set for the additive data set D1 1_64 (left), the partially additive data set 4R 1_136 (middle), and the nonadditive data set 4R 1_56 (right). The average pairwise correlation coefficient, r^2, values are: 0.840 (D1 1_64), 0.725 (4R 1_136), and 0.331 (4R 1_56).

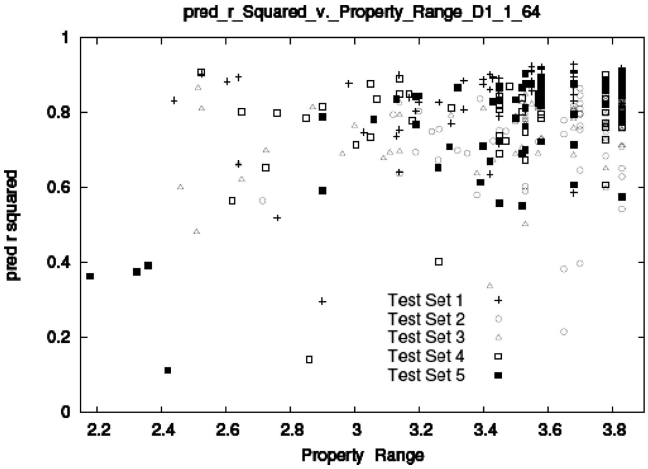


Figure 7. Variation of pred-r^2 with property range of the training set for an additive data set. The property range covered by each test set is as follows: test set 1: 3.65; test set 2: 2.59; test set 3: 3.32; test set 4: 3.55; test set 5: 3.13.

sets (D7 339_3) has relatively high average q^2. Furthermore, one of the additive data sets (D6 76_97) shows similar behavior to the nonadditive data sets with respect to its q^2 values and two of the additive data sets (D3 1_73 and 4R 1_73) are intermediate in average q^2 (<0.5). Thus, we conclude that classical q^2 using LOO cross-validation cannot be used to characterize the data sets as additive or otherwise. It is also evident from Table 4 that internal fit (r^2) does not provide a reliable indicator of additivity because a high r^2 is reported for all data sets.

It is evident from Table 4 that additive effects are dependent on both the structural space covered by the compounds and the biological target under consideration. For example, the D7 data set exhibits additive behavior with respect to P450 1A2 (339_7), however, nonadditive behavior is observed for P450 3A4 (339_3) and partially additive behavior is observed when the compounds are assayed on human liver microsomes (cellular metabolism: 221_5). On the other hand, different chemical spaces (data sets) exhibit different additivity effects for the same biological target. For example, while the D6 data set shows additive behavior for the Ser/Thr kinase (76_97) target, the D7 data set shows partially additive behavior on the same target. Finally some of the data sets show additive behavior for all biological responses analyzed, for example, D2 shows additive behavior for three different GPCR targets. Thus, the 19 cases represented in Table 4 that cover different chemical and biological spaces demonstrate the complexity of protein−ligand interactions, with only half of them showing fully additive behavior.

Analysis of Coefficients. An analysis of the model coefficients has been carried out in order to confirm the additivity classifications made thus far. It is expected that an additive data set will exhibit a greater degree of stability in the substituent coefficients generated over different training sets, than a nonadditive data set, i.e., the variation in the values of the coefficients a_1, a_2, b_1, b_2, etc., (see Figure 1) seen over multiple models will be less for an additive data set due to the increased reliability of each model relative to those generated for nonadditive data sets. Substituent stability has been assessed by carrying out a pairwise comparison of the models generated for each data set. For each model, the substituent coefficients are normalized to be in the range 0−1, with the minimum and maximum coefficient in each model set to 0 and 1, respectively (normalization is required because the absolute values can vary between models). The degree of correlation between two models is then measured by comparing the corresponding coefficients from two models using r^2. The pairwise correlation coefficients are plotted as heat maps in Figure 6 for an additive (left), partially additive (middle), and nonadditive (right) data set. In each case, the models are arranged according to increasing size

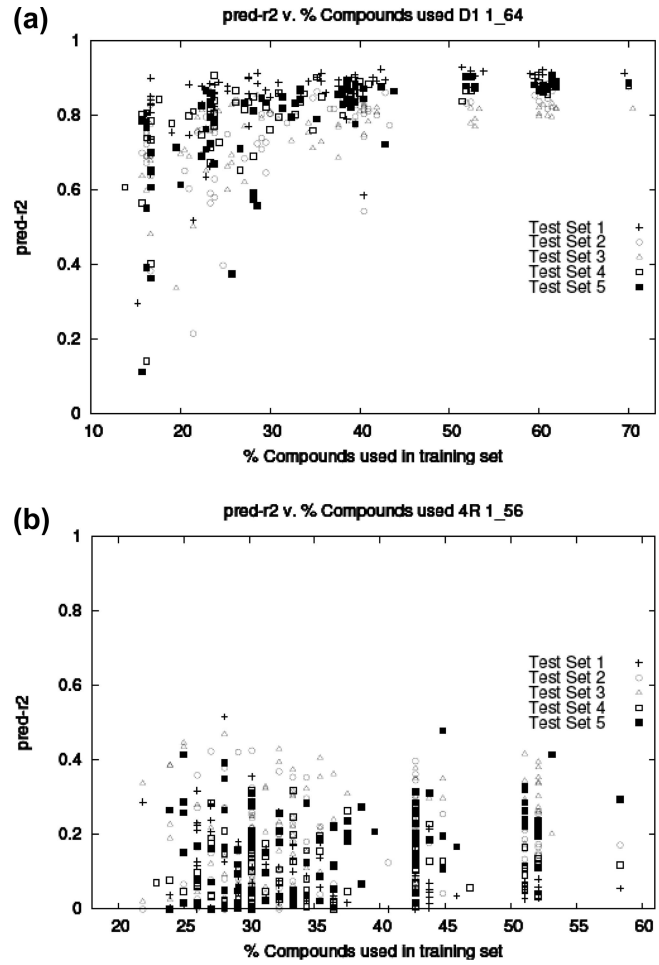


Figure 8. (a) Pred-r^2 vs training set size for the additive data set, D1 1_64. (b) Pred-r^2 vs training set size for the nonadditive data set, 4R 1_56.

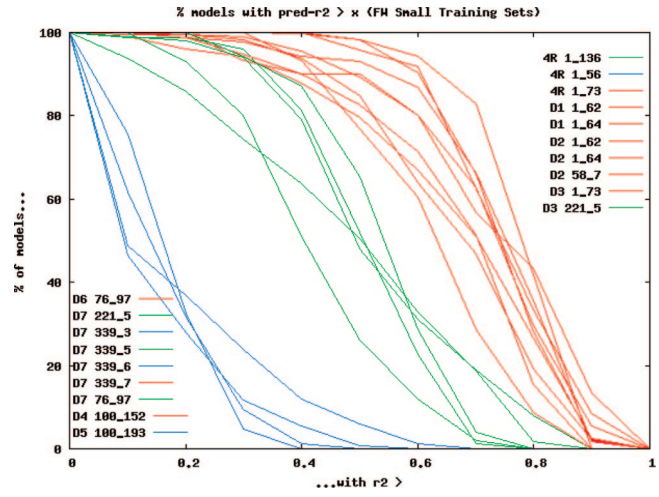


Figure 9. Percentage of models with pred-r^2 > x for models derived from small training sets (<30% data set) with additive models in red, partially additive in green, and nonadditive data set in blue.

of the training set. Each square indicates a pairwise comparison of models, with the color indicating the degree to which the coefficients are correlated: red indicates a correlation of 1.0 and blue indicates a correlation of 0.0. In all three cases, there is high correlation between the coefficients for the largest training sets. This is to be expected as the larger training sets are more similar: for example, if there are 70 compounds available and 60 are selected to form

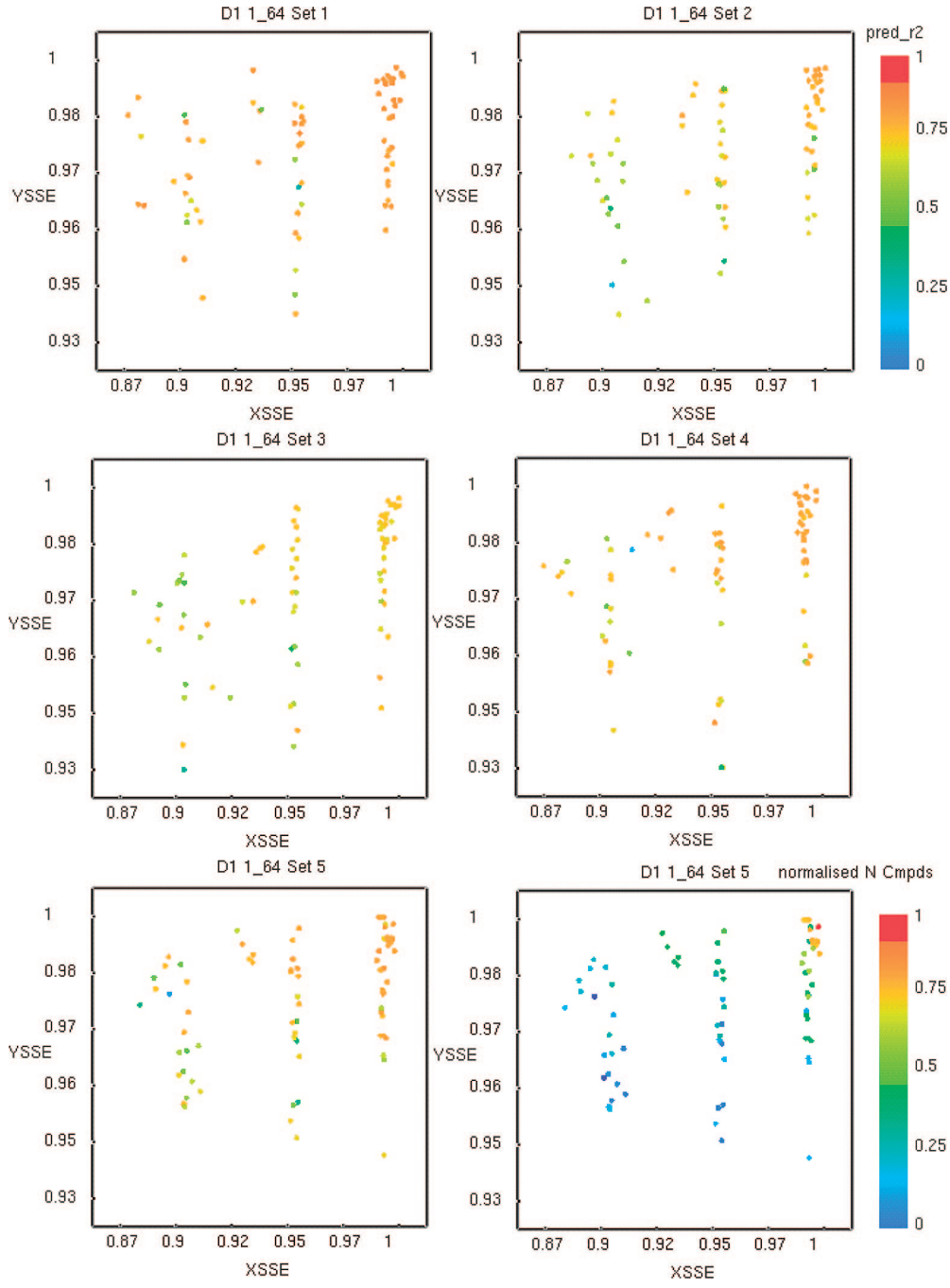


Figure 10. The first five plots show SSE values for the R_X (*x*-axis) and R_Y (*y*-axis) groups for training sets extracted from additive data set D1 1_64. Each data point represents one training set and is color coded to denote the pred-r^2 value (with red indicating pred-r^2 of 1 and blue indicating a pred-r^2 of 0). Each plot represents a different test set. The bottom right plot shows the same data points as the bottom left plot (the SSEs of the R_X and R_Y groups for the training sets corresponding to test set 5) with the data points now color coded according to the fraction of compounds in the training set (with the value 1 (red) representing all available compounds used in the training set).

training sets, then the maximum number of compounds that can differ between the two training sets is 10, or 16.6%. For the additive data set, there is a good correlation between the coefficients developed for the substituents for all training sets except the smallest (indicated by a large area of red/orange). For the partially additive data set, the region of low correlation is extended (i.e., there is more green in the heat map). The nonadditive data set shows little correlation between the coefficients, except for the very large training sets. Similar plots are seen for the different test sets for each data set and each plot is characteristics of other data sets in the same additivity classification. These findings support the earlier classification of the data sets.

Identifying Nonadditive R-Groups. The stability of the substituents across different models has also been used in an attempt to identify nonadditive substituents in data sets classified as partially additive. If such substituents can be identified, their removal from the test sets should lead to improved prediction performance. The substituents in a model are first ranked according to the relative sizes of their coefficients (rankings are used because, as indicated above, the absolute values of the coefficients can vary widely from one model to another). The variation in ranking across the models is then determined and an unstable (nonadditive) substituent is defined as one whose standard deviation in ranking is greater than one-fifth of the number of substituents in that R-group, with each

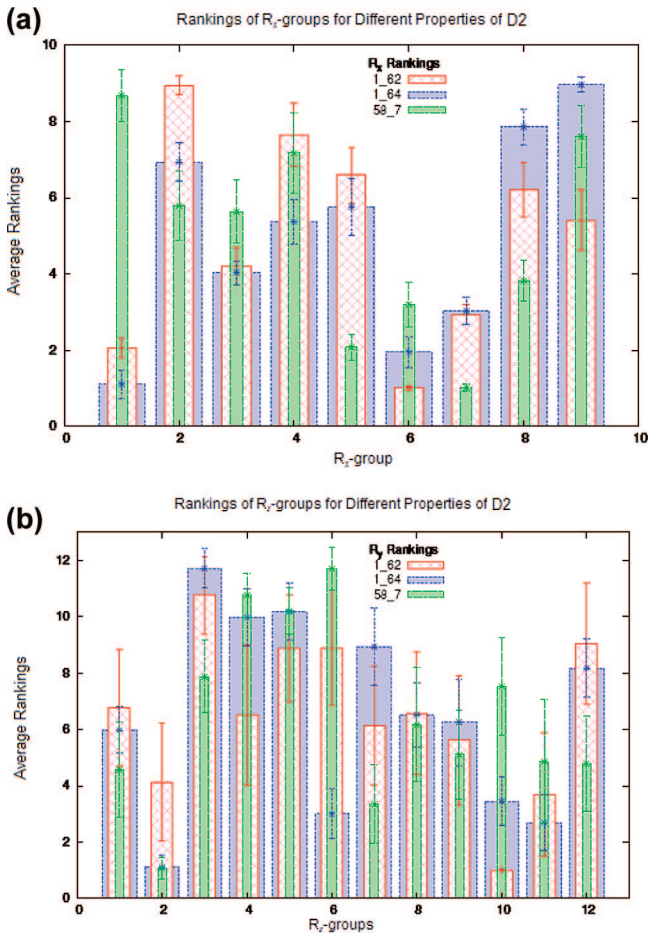


Figure 11. (a) Substituent profile for three end points for the R_x group of data set D2. Each substituent is plotted along the *x*-axis and the average ranks of the substituents are plotted on the *y*-axis. (b) Substituent profile for three end points for the R_y group of data set D2. Each substituent is plotted along the *x*-axis, and the average ranks of the substituents are plotted on the *y*-axis.

Table 5. Selected Compounds and Their Properties for Data Set D2

R_x group	R_y group	1_62	1_64	58_7
X1	Y1	6.74	8.06	5.64
X1	Y2	6.68	8.94	5.98
X1	Y6	6.76	8.61	5.01
X1	Y7	6.96	7.69	6.06
X1	Y9	6.81	7.77	5.77
X1	Y8	6.79	8.06	5.59
X1	Y10	7.77	8.84	5.10
X1	Y11	6.85	8.55	5.17
X1	Y12	6.81	7.78	5.23
X7	Y2	6.88	8.42	7.91

R-group (R_X and R_Y) considered independently. The partially additive data set, D7 221_5, is used to illustrate the change in pred-r^2 when unstable substituents (4 out of 12 substituents) are removed from it. Initially, when all compounds are considered, only 4% of the models have pred-$r^2 \geq 0.6$, but after removing the nonadditive substituents, most of the models (82% over all test sets) have pred-$r^2 \geq 0.6$ (plot showing this case is available as Supporting Information, Figure S10).

Training and Test Set Attributes. Following the classification of the data sets as additive, partially additive, or nonadditive, the next step was to determine the minimal requirements for training and test sets in order to obtain a reliable assessment about additive effects. Training and test set attributes may provide key guidelines for medicinal chemists to design new libraries, not only to obtain a SAR but also to be able make an assessment on additive effects for those structures against the target(s) under analysis. Thus, the determation of these attributes could have an impact on the initial design and prioritization of compounds for synthesis and on the definition of the best strategy (sequential or combinatorial) for further iterations of the drug discovery process.

The attributes of the five test sets selected for each data set are summarized as follows (Table 2): (i) test set size, between 11% and 23% of the complete data set; (ii) range of activity values, between 31% and 100% of the range covered by data set; and (iii) the distribution of substituents within the R-groups (as measured by the SSE) is very high, larger than 0.95 in most cases; this is mainly due to the constraint that each test set should contain at least one occurrence of each R_x and R_y substituent. A large number of training sets, between 47 and 202, was generated for each test set (Table 3). Thus, a more detailed analyses of training set attributes was performed compared to test set attributes, however, the conclusions for desirable training attributes are based on a consensus response across all test sets, therefore, we suggest that any combination of the test set attributes used is valid. The attributes of the training sets that were investigated are: the range of activity values of the compounds in the training set, training set size, and the distribution of substituents within the R-groups (as measured by the SSE).

Activity Range. Figure 7 shows the variation of pred-r^2 with the range of activity values in the training sets for the additive data set, D1_1_64. The pred-r^2 values are all relatively high and show no correlation with the range of activity values covered. Similar plots are seen for all other data sets, regardless of their classification as additive, nonadditive, or partially additive. Thus we conclude that activity range does not appear to be a factor in the ability to generate good predictive models.

Training Set Size. Figure 8a shows pred-r^2 plotted against training set size for the additive data set, D1 1_64. Training set size is measured as the percentage of compounds in the complete combinatorial data set, and it should be noted that training sets containing 100% of the compounds were not possible due to the use of a test set. It can be seen that successful models (pred-$r^2 \geq 0.6$) are still generated when only 30% of the complete data set is used. The corresponding plot for a nonadditive data set ((4R 1_56) is shown in Figure 8b, where all models have poor predictive performance (pred-$r^2 < 0.5$) regardless of training set size. The two plots are characteristic of data sets in the same additivity classification.

The relationship between pred-r^2 and training set size was investigated further by comparing the average pred-r^2 for training sets containing <30% of the compounds (for a small number of cases where training sets with <30% did not exist, the minimum size was increased to 40% or 50%) with the average pred-r^2 for the remaining training sets, that is, those containing >30% of compounds. The maximum difference between the average pred-r^2 of the large and small training sets is 0.163 over all data sets (the average difference is 0.07), indicating that there is very little difference between the best models of the large and small training sets. Pred-r^2 values for the small training sets (<30% of the compounds in the data set) have also been extracted from Figure 5a and are shown in Figure 9, where the clear distinction between the additive (red), partially additive (green), and nonadditive (blue) data sets is still apparent. The data sets are colored according to the classification reported in Table 4.

Distribution of Substituents. Figure 10 shows a typical relationship between the distribution of substituents in the R_X (*x*-axis, XSSE) and R_Y groups (*y*-axis, YSSE) and model performance (color coding is used to indicate pred-r^2, with red representing pred-r^2 of 1 through to blue for pred-r^2 of 0) for an additive data set. The first five plots represent different test sets, and the final (bottom right) plot shows the same training sets as the bottom left but with the color coding used to represent training set size rather than pred-r^2. Although pred-r^2 increases as the SSE of the training set increases, the final plot indicates that this can be attributed to the increase in the number of compounds in the training set as the SSE tends toward one.

Similar plots were obtained for the small (<30% of the full data set) training sets extracted from the same data set, where it can be seen that the relationship between pred-r^2 and the distribution of the substituents in the training set has been lost (the plot showing this case is available as Supporting Information, Figure S11). Thus, we conclude that the distribution of substituents within the training set has little effect on the predictive ability of a model.

Substituent Profiling: Impact on Lead Optimization. For a data set that exhibits additive behavior over several different assays or end points, an analysis of the substituent coefficients for the different assays can provide valuable information. For example, substituents that make a positive contribution to multiple assays can be distinguished from those that confer selectivity by making a positive contribution to one end point and a negative contribution to another. Such an analysis assumes that the substituents are stable across the multiple models and is therefore only applicable to additive data sets. A substituent profile is constructed by first ranking the coefficients in each model with the substituent with maximum coefficient at rank one; the substituent ranks are then averaged over all models for the data set. A graphical representation of a substituent profile can be used as a visualization tool to enable qualitative decision-making and allows medicinal chemists to discuss and rationalize the SAR in a multifactorial manner, which, therefore, helps to inform future directions in lead optimization.

Among the data sets analyzed and classified according to the additive/nonadditive effects in the SAR, data set D2 is highlighted because it shows additive effects for three different targets, class A GPCRs: 1_62, 1_64, and 58_7. Therefore, we can use substituent profiling to identify and rationalize the impact of the different substituents in a multifactorial optimization and look for those combinations of substituents with the desired activity and selectivity. Substituent profiles for data set D2 are shown in Figure 11a,b and indicate that compounds containing substituent X1 (which represents *p*-methoxy phenyl) should have higher activity for targets 1_62 and 1_64 compared to 58_7. This is supported by the activity values of the compounds that contain this substituent, which are shown in Table 5. The profiles also indicate that a combination of X7 and Y2 should correspond to high activity for 58_7, low activity for 1_62, and intermediate activity for 1_64. This compound does indeed have low activity for 1_62 and high activity for 58_7; however, it has higher activity for 1_64 than anticipated. This could be because the X7 group, for which all data sets have a similar ranking, has more influence on activity than the Y2 group.

This visualization tool allows the relative substituent coefficients across multiple biological end points to be superimposed, as demonstrated here, and provides easy-to-interpret clues on the importance of individual substituents in a multifactorial optimization, whether the aim is to maximize the contribution to activity for all the end points or to optimize on selectivity. In addition to this qualitative analysis, the FW models also provide quantitative estimates of substituent contributions, which can be used to make more accurate predictions, for example, between substituents that may have identical rankings but different weights in the QSAR models such as X7 and Y2 described above. The combination of qualitative and quantitative analyses can be used to prioritize compounds for the next iteration in the synthesize−test cycle, based on their multifactorial profiles.

Conclusions

FW analysis was carried out for eight near-complete combinatorial libraries assayed on a total of 19 biological responses, and pred-r^2 and LOO q^2 were calculated in an attempt to classify the libraries as additive, partially additive, and nonadditive. While pred-r^2 provided a clear distinction between the data sets, enabling a classification to be made, no clear distinction was seen using q^2: with data sets identified as additive using pred-r^2 ($\geq$0.6) having low q^2 and, conversely, data sets identified as nonadditive exhibiting high q^2 values. In addition, substituent stability analyses confirmed the additivity classifications made using the pred-r^2 criterion. Thus, we conclude that pred-r^2 appears to be a good indicator of additivity effects in FW analysis, whereas q^2 using LOO cross-validation is less reliable. This finding is consistent with previous studies that have questioned the reliability of LOO q^2.[49] Furthermore, and perhaps unsurprisingly, we found that r^2 did not provide a reliable indicator of additive effects. Only half of the data sets (10 out of 19 in Table 4) examined showed clear additive behavior for their biological responses, which leads us to question the concept of additivity that is widely taken for granted in many approaches to drug discovery process. The assumption of additive effects, when the biological response is dependent on interaction between the substituents, could lead to nonproductive directions being pursued in an iterative lead-optimization process. For an additive data set, a sequential approach to synthesizing new compounds can be followed based on estimated R-group contributions calculated during FW analysis; conversely, a combinatorial approach should be followed for data sets that show nonadditive behavior, and in this scenario, a FW analysis is not valid. We have also seen that the relationship between chemistry space and biological space is complex, as might be expected. Thus, a library may exhibit additive behavior for one biological response and nonadditive behavior for a different response, and different libraries may demonstrate different additivity behavior when assayed on the same biological target.

We then carried out a retrospective analysis of the additive data sets in which compounds were progressively removed from the training sets in order to determine the minimum requirements for assessing additivity. The aim was to determine guidelines for assessing additivity that could be used prospectively in a lead optimization project and, as consequence, for the design of libraries of compounds to be synthesized. If the presence of additive effects can be established by analyzing a small fraction of a virtual library, then a sequential approach to synthesis is justified based on the quantitative estimations of substituent contributions. Furthermore, if additive effects are seen across multiple properties, then a multifactorial analysis can be followed based on substituent profiles. The combination of qualitative visualization tools and quantitative QSAR analyses can be used to rationalize SAR and prioritize compounds for the next iteration in the synthesize−test cycle based on their multifactorial profiles. If a library is assessed as being nonadditive, then a combinatorial approach to synthesis should be followed.

The retrospective analysis demonstrated that pred-r^2 remained robust as an indicator of additive effects when as few as 30% of the compounds were present in the training set, with the constraint that there is as least one occurrence of each substituent present. For training sets with fewer than 30% of the full matrix, the FW analysis began to fail even for additive data sets. No clear relationship was found for the other training set attributes investigated, namely property range and distribution of substituents. The test sets used in the analysis contained from 11−23% of the full-matrix, and were also constrained to contain at least one occurrence of each substituent.

The use of pred-r^2 to assess additivity is undesirable because it requires additional compounds to be synthesized and assayed in order to form the test set; in total, between 41−53% of the full matrix is required considering both training and test sets. However, our investigations so far have indicated that the use q^2 can be misleading. Future work will consider more sophisticated implementations of cross-validation, for example, leave-group-out (LGO) and constrained cross-validation in which

similar constraints to those applied to the test sets are used to ensure that all substituents in the leave-out sets are still represented in the "training" set. It is hoped that such an analysis may lead to a reduction in the number of compounds that need to be synthesized in order to determine additivity.

Acknowledgment. We thank Chemical Computing Group for software support and Johnson & Johnson, The Royal Society, and the Wolfson Foundation for funding support.

Supporting Information Available: Additional details of the data sets. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Wess, G. How to escape the bottleneck of medicinal chemistry. *Drug Discovery Today* **2002**, *7*, 533–535.
(2) Dill, K. A. Additivity principles in biochemistry. *J. Biol. Chem.* **1997**, *272*, 701–704.
(3) Szwajkajzer, D.; Carey, J. Molecular and biological constraints on ligand-binding affinity and specificity. *Biopolymers* **1997**, *44*, 181–198.
(4) Jencks, W. P. On the Attribution and Additivity of Binding Energies. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4046–4050.
(5) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* **2004**, *47*, 3463–3482.
(6) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531–1534.
(7) Hajduk, P. J.; Gomtsyan, A.; Didomenico, S.; Cowart, M.; Bayburt, E. K.; Solomon, L.; Severin, J.; Smith, R.; Walter, K.; Holzman, T. F.; Stewart, A.; McGaraughty, S.; Jarvis, M. F.; Kowaluk, E. A.; Fesik, S. W. Design of adenosine kinase inhibitors from the NMR-based screening of fragments. *J. Med. Chem.* **2000**, *43*, 4781–4786.
(8) Szczepankiewicz, B. G.; Liu, G.; Hajduk, P. J.; Abad-Zapatero, C.; Pei, Z. H.; Xin, Z. L.; Lubben, T. H.; Trevillyan, J. M.; Stashko, M. A.; Ballaron, S. J.; Liang, H.; Huang, F.; Hutchins, C. W.; Fesik, S. W.; Jirousek, M. R. Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. *J. Am. Chem. Soc.* **2003**, *125*, 4087–4096.
(9) Erlanson, D. A.; Wells, J. A.; Braisted, A. C. Tethering: Fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 199–223.
(10) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. Site-directed ligand discovery. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9367–9372.
(11) Braisted, A. C.; Oslob, J. D.; Delano, W. L.; Hyde, J.; McDowell, R. S.; Waal, N.; Yu, C.; Arkin, M. R.; Raimundo, B. C. Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *J. Am. Chem. Soc.* **2003**, *125*, 3714–3715.
(12) Maly, D. J.; Choong, I. C.; Ellman, J. A. Combinatorial target-guided ligand assembly: Identification of potent subtype-selective c-Src inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2419–2424.
(13) Hindi, S.; Deng, H. T.; James, L.; Kawamura, A. Selective photolabeling of Lck kinase in complex proteome. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5625–5628.
(14) Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H. Fragment-based lead discovery using X-ray crystallography. *J. Med. Chem.* **2005**, *48*, 403–413.
(15) Stout, T. J.; Sage, C. R.; Stroud, R. M. The additivity of substrate fragments in enzyme-ligand binding. *Struct. Folding Des.* **1998**, *6*, 839–848.
(16) Saxty, G.; Woodhead, S. J.; Berdini, V.; Davies, T. G.; Verdonk, M. L.; Wyatt, P. G.; Boyle, R. G.; Barford, D.; Downham, R.; Garrett, M. D.; Carr, R. A. Identification of inhibitors of protein kinase B using fragment-based lead discovery. *J. Med. Chem.* **2007**, *50*, 2293–2296.
(17) Babaoglu, K.; Shoichet, B. K. Deconstructing fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2006**, *2*, 720–723.
(18) Hajduk, P. J. Puzzling through fragment-based drug design. *Nat. Chem. Biol.* **2006**, *2*, 658–659.
(19) Free, S. M.; Wilson, J. W. Mathematical Contribution to Structure−Activity Studies. *J. Med. Chem.* **1964**, *7*, 395–399.
(20) Green, N. M. Avidin. *Adv. Protein Chem.* **1975**, *29*, 85–133.
(21) Congreve, M. S.; Davis, D. J.; Devine, L.; Granata, C.; O'Reilly, M.; Wyatt, P. G.; Jhoti, H. Detection of ligands from a dynamic combinatorial library by X-ray crystallography. *Angew. Chem., Int. Ed.* **2003**, *42*, 4479–4482.
(22) Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, G. M.; Marcotte, P. A.; Severin, J.; Walter, K.; Smith, H.; Gubbins, E.; Simmer, R.; Holzman, T. F.; Morgan, D. W.; Davidsen, S. K.; Summers, J. B.; Fesik, S. W. Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. *J. Am. Chem. Soc.* **1997**, *119*, 5818–5827.
(23) Miller, B. G.; Wolfenden, R. Catalytic proficiency: The unusual case of OMP decarboxylase. *Annu. Rev. Biochem.* **2002**, *71*, 847–885.
(24) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411–428.
(25) Momany, F. A.; Rone, R. Validation of the General-Purpose Quanta(R)3.2/Charmm(R) Force-Field. *J. Comput. Chem.* **1992**, *13*, 888–900.
(26) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
(27) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(28) Bohm, H. J. Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 309–323.
(29) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
(30) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions 0.1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.
(31) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
(32) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.
(33) Fujita, T.; Ban, T. Studies on Structure−Activity Relationship 0.3. Structure−Activity Study of Phenethylamines as Substrates of Biosynthetic Enzymes of Sympathetic Transmitters. *J. Med. Chem.* **1971**, *14*, 148–152.
(34) Paliakov, E.; Henary, M.; Say, M.; Patterson, S. E.; Parker, A.; Manzel, L.; MacFarlane, D. E.; Bojarski, A. J.; Strekowski, L. Fujita-Ban QSAR analysis and CoMFA study of quinoline antagonists of immunostimulatory CpG-oligodeoxynucleotides. *Bioorg. Med. Chem.* **2007**, *15*, 324–332.
(35) Vasanthanathan, P.; Lakshmi, M.; Rao, A. R. QSAR study on pyridinone derivatives as HIV-1 non-nucleoside reverse transcriptase inhibitor: a mixed approach. *Med. Chem.* **2007**, *3*, 227–232.
(36) Pissurlenkar, R. R. S.; Malde, A. K.; Khedkar, S. A.; Coutinho, E. C. Encoding type and position in peptide QSAR: Application to peptides binding to class I MHC molecule HLA-A*0201. *QSAR Comb. Sci.* **2007**, *26*, 189–203.
(37) Globisch, C.; Pajeva, I. K.; Wiese, M. Structure−activity relationships of a series of tariquidar analogs as multidrug resistance modulators. *Bioorg. Med. Chem.* **2006**, *14*, 1588–1598.
(38) Lee, S. G.; Chmielewski, J. Rapid synthesis and in situ screening of potent HIV-1 protease dimerization inhibitors. *Chem. Biol.* **2006**, *13*, 421–426.
(39) Schaper, K. J. Free-Wilson-type analysis of non-additive substituent effects on THPB dopamine receptor affinity using artificial neural networks. *Quant. Struct.−Act. Relat.* **1999**, *18*, 354–360.
(40) Sehon, C.; McClure, K.; Hack, M.; Morton, M.; Gomez, L.; Li, L. N.; Barrett, T. D.; Shankley, N.; Breitenbucher, J. G. Pyrazole CCK1 receptor antagonists. Part 2: SAR studies by solid-phase library synthesis and determination of Free-Wilson additivity. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 77–80.
(41) Ge, N. X.; Cho, S. J.; Hermsmeier, M.; Poss, M.; Shen, C. F. Testing non-additivity of biological activity in a combinatorial library. *Comb. Chem. High Throughput Screening* **2002**, *5*, 147–154.
(42) *Pipeline Pilot*; Scitegic, 9665 Chesapeake Drive, Suite 401, San Diego, CA 92123-1365. Available from SciTegic Inc. at http://www.scitegic.com.
(43) Bosmans, J. P. R. M. A.; Love, C. J.; Van Lommen, G. R. E. Preparation of 2,4-diaminopyrimidines as dopamine D4 antagonists. Patent WO9743279 A1, 1997.
(44) Andres-Gil, J. I.; Alcazar-Vaca, M. J.; Linares de la Morena, M. L.; Martinez Gonzalez, S.; Oyarzabal Santamarina, J.; Pastor-Fernandez, J.; Vega Ramiro, J. A.; Drinkenburg, W. H. I. M. Preparation of

substituted pyrazinone derivatives as α2c-adrenoreceptor antagonists. Patent WO07071639 A1, 2007.
(45) Brotherton-Pleiss, C. E.; Dillon, M. P.; Gleason, S. K.; Jen Lin, C. J.; Schoenfeld, R. C.; Villa, M.; Zhai, Y. Preparation of heterocyclic inhibitors of P2 × 3 useful for treating pain, genito-urinary, gastrointestinal, and respiratory disorders. Patent US 2007037974 A1, 2007.
(46) Steckler, T. H. W.; Janssens, F. E.; Leenaerts, J. E.; Fernandez-Gadea, F. J.; Gomez-Sanchez, A.; Meert, T. F. Preparation of 4-phenyl-4-(imidazol-2-yl)piperidine derivatives as selective non-peptide δ-opioid agonists for treatment of depression and anxiety, Patent WO04089372 A1, 2004.
(47) Freyne, E. J. E.; Love, C. J.; Cooymans, L. P.; Vandermaesen, N.; Buijnsters, P. J. J. A.; Willems, M.; Embrechts, W. C. J. Preparation of triazolopyrimidines as glycogen synthase kinase 3 inhibitors. Patent WO05012304 A2, 2005.
(48) Holland, J. H. *Adaptation in Natural and Artificial Systems*; The University of Michigan Press: Ann Arbor, MI, 1975.
(49) Golbraikh, A.; Tropsha, A. Beware of q2! *J. Mol. Graphics Modell.* **2002**, *20*, 269–276.

JM801070Q