Electronic Supplementary Material (ESI) for Medicinal Chemistry Communications

A Matched Molecular Pair Analysis of *in-vitro* human microsomal metabolic stability measurements for methylene substitution or replacements - identification of those transforms more likely to have beneficial effects.

Alexander G. Dossetter..

AstraZeneca PLC, Alderley Park, Mereside, Cheshire, SK10 4TG

Phone +44 (0)1625517673. Fax +44(0)1625513624

Supplementary Information

Contents

1.0 *In-vitro* human liver microsomal screening protocol
Equipment and Reagents

All Acetonitrile, Formic Acid and Methanol solutions were HPLC grade, supplied by TheromFischer Scientific Ltd (Loughbrough, UK). NADPH was stored at -20 ºC when not in use and was supplied by Sigma Alderich Company Ltd (Dorset, UK). Collagenase A, supplied by Roche Diagnostics (Mannheim, Germany) and Trypsin Inhibitor, also supplied by Sigma, were both stored at 2-8 ºC prior to experimentation. Leibovitz solution supplied by Invitrogen (Paisley, UK) was also stored at this temperature. 0.4% (w/v) Tryphan Blue and DNase were ordered through Sigma and stored at room temperature and -20 ºC respectively.

All HLMs were provided by BD Biosciences and were made up by homogenising liver samples followed by a series of centrifugation steps in which the final pellet consisted of the microsomes. These were stored at -80 ºC. On delivery the HLMs have a protein concentration of 20 mg/ml. These were diluted to 1 mg/ml in 0.1 MINH Phosphate Buffer (pH 7.4) before experimentation. This buffer was made by the in-house Scientific Preparation Team (Alderley Park, AstraZeneca). Rat liver was delivered on ice from the Barriered Animal Breeding Unit (BABU) at AstraZeneca (Alderley Park, UK). These were obtained from 9 week old male Harlan Han-Wistar rats.

The Genesis Workstation 200 used in both assays was provided by Tecan UK Ltd (Reading, UK). Water used to wash the Tecan and in mobile phase mixtures was HPLC grade and supplied by in-house Ultra Purification Units (Marlow, UK). Both Sorval Legend RT (Newport Pagnell) and MSE Centaur 2 (London, UK) centrifuges were used.

1.1 Methods

This assay procedure was conducted using a Genesis Workstation 200 provided by Tecan. All compounds tested were dispensed by the Compound Management Group (CMG) at AstraZeneca at an initial concentration of 10 mM in dimethyl sulfoxide (DMSO) and initial volume of 5 µL. A total of 96 compounds could be tested in one assay, including 6 markers which are used to assess the accuracy and reliability of the generated data. The markers routinely used in this assay were: Phenacetin, Diclofenac, Diazepam, Dextromethorphan and Erythromycin. These specific markers were used because they are the substrates of some of

the major CYP enzymes involved in metabolism and thus provide a clear indication of viability of the HLMs used (Table 1). It should be noted that all marker solutions were pooled together before the assay and were tested as a cocktail.

When 96 compounds were run in any individual assay, two sets of markers were examined on the same plate - one at the front of the plate and one at the back. This was to ensure that there was no variability in the assay procedure throughout each experiment.

Compounds were first diluted 100 fold using acetonitrile followed by the transfer of 3 µL of diluted compound to 270 µL of HLMs. 27 µL of 11.1 mM NADPH was added to initiate the reaction. (HLMs were maintained on a "Te-shake" throughout the assay which was continually shaken at 9000 rpm after the addition of NADPH.) By this stage all compounds are at a concentration 10,000 fold lower than their initial concentration. (i.e. all compounds are now at a concentration of 1 µM). 25 µL of the reaction mixture was removed at time points; 2, 5, 10, 15, 20 and 30 minutes and added to 100 µL Internal Standard and acetonitrile quench solution, thus halting the reaction. Plates containing these samples were then spun at 3000 rpm for 15 minutes at 4 ºC.

Marker	Major CYP Enzymes Responsible for Metabolism In HLMs
Phenacetin	CYP1A1, CYP1A2
Diclofenac	CYP2C9, other CYP2C and CYP3A4 enzymes
Diazepam	CYP3A4
Dextromethotphan	CYP2D6
Erythromycin	CYP3A4

Table 1 – The major CYP enzymes responsible for metabolism of the markers [20-25].

The reason why such low substrate concentrations were employed was to ensure that the assay procedure fitted the Michaelis-Menten model, assuming linear first order kinetics [3]. By utilizing substrate concentrations of 1 µM, believed to be well below the Michaelis Constant (K_m) the Michaelis-Menten equation simplifies to $v = (K_{cat}/K_m)[E][S]$, meaning that the rate of clearance is proportional to the concentration of substrate. The situation is avoided where a decrease in substrate has no effect on clearance as all available enzymes are still in use resulting in the maximum rate of clearance.

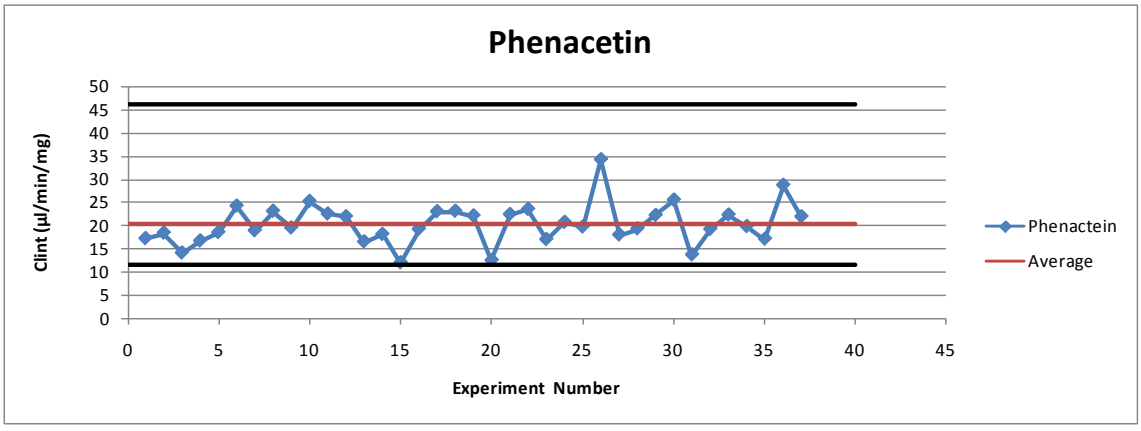
All samples were analysed by HPLC/MS/MS as detailed in section 3.5.

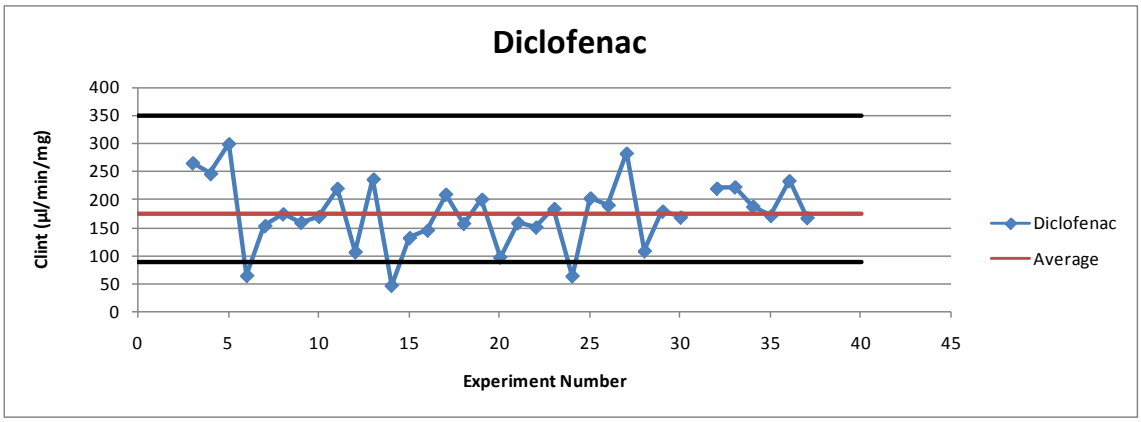
1.2 Calculations

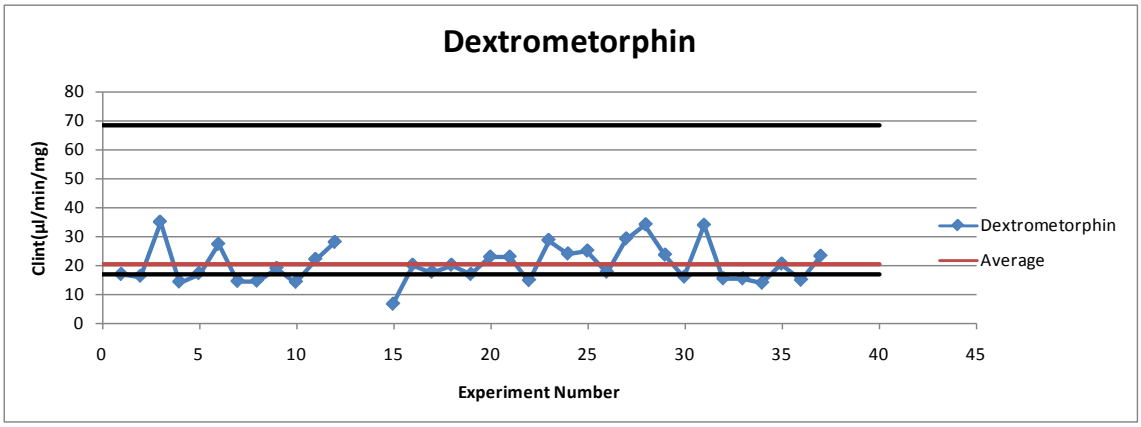
All incubations were analysed using QuanLynx^{Tm} browser. A "response" value was calculated for each compound at each time point by dividing the peak area of the Analyte peak on the chromatogram by the peak area of the Internal Standard [29]. These values were converted to percentages, with the highest response value equating to 100%.

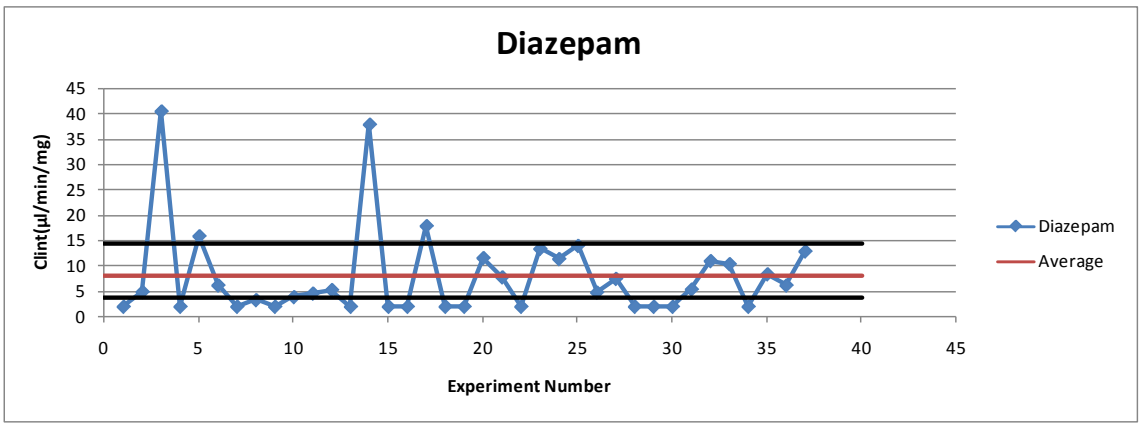
The Internal Standard was monitored throughout the course of each experiment, as large variations in Internal Standard may lead to inaccurate responses. For each compound a plot of "Percentage of Un-Metabolized Compound" vs Incubation time was plotted and the initial gradient determined (proportional to the initial rate of the reaction). This is equivalent to the negative form of the first order rate constant (k). This can also be calculated by: $LN(2)/T_{1/2}$. The rate constant was subsequently multiplied by 1000 as the concentration of protein in the microsomes was at 1 mg/mL. All Cl_{int} values therefore had units of µL/min/mg.

1.3 The Markers









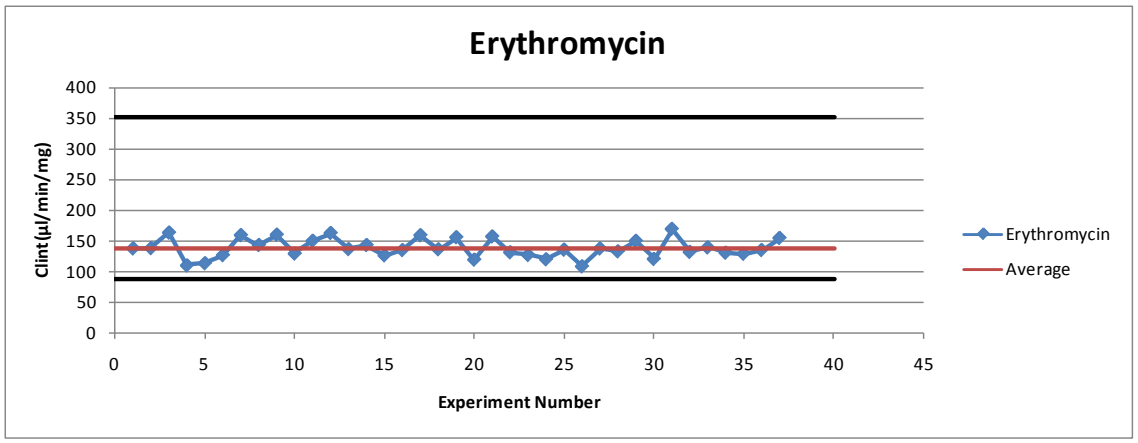


Figure - Graphs showing the variation in Cl_{int} for the markers used in HLM experiments: (a) Phenacetin (b) Diclofenac (c) Dextrometorphin (d) Diazepam (e) Erythromycin, over the course of the experimental procedure. The red line present on each graph indicates the average Cl_{int} of that maker. The black lines present on each graph indicate a 2 fold range either side of the expected Cl_{int}.

The 5 markers, Phenacetin, Diclofenac, Dextrometorphin, Diazepam and Erythromycin were run routinely with each HLM assay and their Cl_{int}s monitored over the experimental procedure. Each experiment was only considered to pass when the 3 out of 5 markers fell within 2 fold to the expected Cl_{int} for that marker. The expected Cl_{int}s were calculated from IBIS by calculating the mean of all compounds tested in previous experiments.

Database searches used to find matched pairs.

Searches were performed on the AstraZeneca PLC database of global HLM Clint results; 135K at the time. SMIRK reaction strings were used to find the substrate (starting methylene) and product (substituted methylene or replacement). Multiple strings were used for some of the transforms group to avoid logic operators and most specifically to avoid finding terminal methylene groups i.e. $–CH_2CH_3$ as the CH_3 matches a non defined carbon. Thus specifically CH or CH_2 were defined and the results combined. The column [No.] refers to the transform number in tables 1 and 2, where no matches were found in the HLM database this has been indicated. Where low matched pair numbers were found these have been indicated by a transform number closest to that in Table 1 and 2, and the value in parentheses was the

number of matched pairs found. These were examined statistically and if they added no conclusion were not included in the paper.

Table 1 SMARTS used to find matches

SMIRK	TRANSFORM	No.
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][C:2]([H])([CH3])[C:3]	racemic_aryl-CH2-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][C:2]([H])([CH3])[C:3]	racemic_aryl-CH2-alkyl	9
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][C@:2]([H])([CH3])[C:3]	Rchiral_aryl-CH2-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][C@@:2]([H])([CH3])[C:3]	Schrial_aryl-CH2-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][C@:2]([H])([CH3])[C:3]	Rchiral_aryl-CH2-alkyl	15
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][C@@:2]([H])([CH3])[C:3]	Schrial_aryl-CH2-alkyl	No Matches
[c:1][C:2]([H])([H])[c:3]>>[c:1][C:2]([H])([CH3])[c:3]	racemic_aryl_CH2_aryl	10
[c:1][C:2]([H])([H])[c:3]>>[c:1][C@:2]([H])([CH3])[c:3]	Rchiral_aryl-CH2-aryl	13c (7)
[c:1][C:2]([H])([H])[c:3]>>[c:1][C@@:2]([H])([CH3])[c:3]	Schrial_aryl-CH2-aryl	13d (6)
[c:1][C:2]([H])([H])[N:3]>>[c:1][C@:2]([H])([CH3])[N:3]	Rchiral_aryl-CH2-N	5
[c:1][C:2]([H])([H])[N:3]>>[c:1][C@@:2]([H])([CH3])[N:3]	Schrial_aryl-CH2-N	6
[c:1][C:2]([H])([H])[N:3]>>[c:1][C:2]([H])([CH3])[N:3]	racemic_aryl-CH2-N	4
[c:1][C:2]([H])([H])[O:3]>>[c:1][C@:2]([H])([CH3])[O:3]	Rchiral_aryl-CH2-O	7a (10)
[c:1][C:2]([H])([H])[O:3]>>[c:1][C@@:2]([H])([CH3])[O:3]	Schrial_aryl-CH2-O	7b (9)
[c:1][C:2]([H])([H])[O:3]>>[c:1][C:2]([H])([CH3])[O:3]	Racemic_chrial_aryl-CH2-O	7
[c:1][C:2]([H])([H])[S:3]>>[c:1][C@:2]([H])([CH3])[S:3]	Rchiral_aryl-CH2-S	8a (14)
[c:1][C:2]([H])([H])[S:3]>>[c:1][C@@:2]([H])([CH3])[S:3]	Schrial_aryl-CH2-S	8b (3)
[c:1][C:2]([H])([H])[S:3]>>[c:1][C:2]([H])([CH3])[S:3]	Racemic_chrial_aryl-CH2-S	8
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][C:2]([C])([C])[C:3]	aryl-gem-di-M-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][C:2]([C])([C])[C:3]	aryl-gem-di-M-alkyl	11
[c:1][C:2]([H])([H])[c:3]>>[c:1][C:2]([C])([C])[c:3]	aryl-gem-di-M-aryl	11a (3)
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][C:2]([F])([F])[C:3]	aryl-CF2-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][C:2]([F])([F])[C:3]	aryl-CF2-alkyl	12
[c:1][C:2]([H])([H])[c:3]>>[c:1][C:2]([F])([F])[c:3]	aryl-CF2-aryl	13
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][C:2]([F])([F])[C:3]	alkyl-CF2-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C:2]([F])([F])[C:3]	alkyl-CF2-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C:2]([F])([F])[C:3]	alkyl-CF2-alkyl	18
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][C:2]([C])([C])[C:3]	alkyl-gem-di-M-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C:2]([C])([C])[C:3]	alkyl-gem-di-M-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C:2]([C])([C])[C:3]	alkyl-gem-di-M-alkyl	16
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][O:2][C:3]	alkyl-O-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][O:2][C:3]	alkyl-O-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][O:2][C:3]	alkyl-O-alkyl	19

[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][O:2][C:3]	aryl-O-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][O:2][C:3]	aryl-O-alkyl	No Matches
[c:1][C:2]([H])([H])[c:3]>>[c:1][O:2][c:3]	aryl-O-alkyl	20
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][S:2][C:3]	alkyl-S-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][S:2][C:3]	alkyl-S-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][S:2][C:3]	alkyl-S-alkyl	22
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][S:2][C:3]	aryl-S-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][S:2][C:3]	aryl-S-alkyl	No Matches
[c:1][C:2]([H])([H])[c:3]>>[c:1][S:2][c:3]	aryl-S-alkyl	23
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][S:2](=O)(=O)[C:3]	alkyl-S(=O)(=O)-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][S:2](=O)(=O)[C:3]	alkyl-S(=O)(=O)-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][S:2](=O)(=O)[C:3]	alkyl-S(=O)(=O)-alkyl	24
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][S:2](=O)(=O)[C:3]	aryl-S(=O)(=O)-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][S:2](=O)(=O)[C:3]	aryl-S(=O)(=O)-alkyl	No Matches
[c:1][C:2]([H])([H])[c:3]>>[c:1][S:2](=O)(=O)[c:3]	aryl-S(=O)(=O)-alkyl	24a (1)
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][N:2][C:3]	alkyl-N-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][N:2][C:3]	alkyl-N-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][N:2][C:3]	alkyl-N-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][N:2][C:3]	aryl-N-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][N:2][C:3]	aryl-N-alkyl	No Matches
[c:1][C:2]([H])([H])[c:3]>>[c:1][N:2][c:3]	aryl-N-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][N:2]([C])[C:3]	alkyl-N(CH3)-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][N:2]([C])[C:3]	alkyl-N(CH3)-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][N:2]([C])[C:3]	alkyl-N(CH3)-alkyl	21
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][N:2]([C])[C:3]	aryl-N(CH3)-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][N:2]([C])[C:3]	aryl-N(CH3)-alkyl	No Matches
[c:1][C:2]([H])([H])[c:3]>>[c:1][N:2]([C])[c:3]	aryl-N(CH3)-alkyl	21a (2)
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][C@:2]([H])([C])[C:3]	Ralkyl-CH(CH3)-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C@:2]([H])([C])[C:3]	Ralkyl-CH(CH3)-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C@:2]([H])([C])[C:3]	Ralkyl-CH(CH3)-alkyl	15
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][C@@:2]([H])([C])[C:3]	Salkyl-CH(CH3)-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C@@:2]([H])([C])[C:3]	Salkyl-CH(CH3)-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C@@:2]([H])([C]	Salkyl-CH(CH3)-alkyl	14

)[C:3]		
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][C@:2]([H])([F])[C:3]	Ralkyl-CH(F)-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C@:2]([H])([F])[C:3]	Ralkyl-CH(F)-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C@:2]([H])([F])[C:3]	Ralkyl-CH(F)-alkyl	17
[$([CH]):1][C:2]([H])([H])[$([CH]):3]>>[C:1][C@@:2]([H])([F])[C:3]	Salkyl-CH(F)-alkyl	No Matches
[$([CH]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C@@:2]([H])([F])[C:3]	Salkyl-CH(F)-alkyl	No Matches
[$([CH2]):1][C:2]([H])([H])[$([CH2]):3]>>[C:1][C@@:2]([H])([F])[C:3]	Salkyl-CH(F)-alkyl	17
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][C@:2]([H])([F])[C:3]	Raryl-CH(F)-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][C@:2]([H])([F])[C:3]	Raryl-CH(F)-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH]):3]>>[c:1][C@@:2]([H])([F])[C:3]	Saryl-CH(F)-alkyl	No Matches
[c:1][C:2]([H])([H])[$([CH2]):3]>>[c:1][C@@:2]([H])([F])[C:3]	Saryl-CH(F)-alkyl	No Matches