Eur J Med Chem (1995) 30, 779–787
© Elsevier, Paris

New benzylpyrimidines: inhibition of DHFR from various species. QSAR, CoMFA and PC analysis

K-H Czaplinski[1], W Hänsel[1], M Wiese[2], JK Seydel[2]*

[1]*University of Kiel, Pharmaceutical Institute, 24118 Kiel;*
[2]*Institute for Experimental Biology and Medicine, Medicinal Pharmaceutical Chemistry, 23845 Borstel, Germany*

(Received 6 March 1995; accepted 8 June 1995)

Summary — To further analyze the structural requirements responsible for the enhanced activity of the newly developed, highly active benzylpyrimidine K-130, a series of trimethoprim analogues with various 4-anilinoalkoxy moieties has been synthesized and tested against dihydrofolate reductase (DHFR) derived from various species (*Mycobacterium lufu, Escherichia coli, Candida albicans* and rats). The importance of the secondary amino group for binding affinity could be shown by varying the substituent in the *para* position to the *sec*-amino moiety. This finding could be supported by multiple linear regression and comparative molecular field analysis (CoMFA). The polarized SO_2 group, which was thought to be responsible for the increased activity of K-130, seems not to be the only group important for receptor binding. Additionally, a high selectivity of the new compounds for DHFR derived from the various bacteria and *C albicans* compared with DHFR derived from rat liver is shown by PC analysis.

benzylpyrimidine / inhibition of DHFR / mycobacteria / *Escherichia coli* / *Candida albicans* / QSAR analysis / CoMFA / PC analysis

Introduction

X-ray crystallographic studies of the binary complex of dihydrofolate reductase (DHFR) from *Escherichia coli* and 3'-substituted analogues of trimethoprim (TMP) bearing a carboxylic acid moiety attached *via* an alkoxy side chain have demonstrated the ionic interaction of this carboxy group with the guanidinium moiety of arginine 57 (arg-57) of the active site of the enzyme [1].

Recent results of computer graphics supported the assumption of this additional binding site (arg-57) for a newly developed benzylpyrimidine K-130 [2]. This compound showed a 10-fold increase in inhibitory activity as compared with TMP for DHFR derived from *Mycobacterium lufu* [3]. Further changes in structure led to some inconsistencies with the assumption that the polarized SO_2-group of K-130 interacts with arg-57.

Replacement of the primary amino moiety of K-130 by an NO_2 group (compound K-122), which should cause a decrease in polarization of the SO_2 group, led to an equally active compound. Replacement by a methyl group led to the same result (K-150). However, the methylation of the *sec*-amino group (compound K-245), which should have a minor influence on the polarization of the SO_2 group, resulted in a significant decrease in activity, especially against *E coli*-derived DHFR.

Further information was obtained from K-130 derivatives, where the *sec*-amino group was replaced by a methylene group (KC-146). This structural change led to compounds that no longer showed an increase in inhibitory activity compared with TMP.

It was the aim of this work to scrutinize the mode of binding of K-130 and analyze the importance of the *sec*-amino group and the substituent in a *para*-position with regard to binding affinity. For this reason a series of 10 derivatives has been synthesized and tested against DHFR derived from various species [4].

Chemistry

All new compounds were prepared by alkylation of phenol 2,4-diamino-5-(3,5-dimethoxy-4-hydroxybenzyl)-pyrimidine (HDMP) [5] with the appropriate alkylarylbromide [6] with use of sodium trimethylsilanolate in dimethylformamide (DMF) and a following separation of the toluenesulfonyl group. Acidic hydro-

*Correspondence and reprints

lysis led to the compounds KC-1303, KC-1308 and KC-1311–KC-1318. The compounds KC-1307 and KC-1310 were obtained after treatment with lithium in liquid ammonia. Treatment of the appropriate alkylarylbromide with acid before alkylation of HDMP yielded KC-1300.

3D QSAR methods

The charges were calculated by the MNDO [7] method as implemented in SYBYL version 6.0 [8].

The comparative molecular field analysis (CoMFA) was performed using the QSAR option of SYBYL version 6.0 on a Silicon Graphics 4D/25 Personal Iris workstation. Unless specifically stated otherwise, default settings were used throughout. The steric and electrostatic potentials were generated using an sp^3 carbon probe with a 1+ charge. The grid used in the CoMFA study had a resolution of 2.0 Å and extended beyond the molecular dimensions by 4.0 Å in all directions. Cross-validation was performed using the leave-one-out method to determine the optimal number of components in the PLS analysis.

Results and discussion

Inhibitor activity

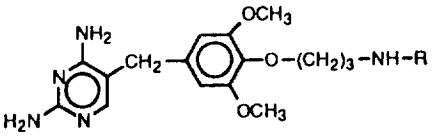
The IC_{50} values of the inhibitors were determined using enzyme extracts derived from four different species. *M lufu* was chosen as a Gram-positive organism, *E coli* as a Gram-negative one, *Candida albicans* as a representative of fungi, and rat as a vertebrate.

The IC_{50} data are shown in table I. Most of the new compounds in this series displayed a stronger activity than TMP and some were as active as K-130. This was the reason why we assumed an additional binding site, other than arginine 57, for these derivatives. The highly active KC-1311 with its CF_3 group instead of its SO_2 group strengthened this assumption and the possible importance of the *sec*-amino moiety. This is further supported by the loss of activity comparing K-130 and KC-146. In the latter the *sec*-amino moiety is replaced by a methylene group.

Quantitative structure–activity relationships

The most appropriate indicator of the electronic influence of the *para*-substituent on the NH group

Table I. Activity data (IC_{50}) of DHFR derived from *M lufu, E coli, C albicans* and rat liver.



Compound	*R*	*IC_{50}*				
		M lufu (μmol/L)	*C albicans (μmol/L)*	*Rat (μmol/L)*	*E coli (nmol/L)*	
					Obs	*Calc*[a]
KC-1300	Naphthyl-SO_2-Ph-CH_3	0.705	0.101	0.926	9.20	7.82[b]
KC-1303	$PhNO_2$	0.085	0.671	7.32	0.380	0.305
KC-1307	5,8-Dihydronaphthyl	0.319	0.643	13.46	0.728	0.629
KC-1308	$PhN(CH_3)_2$	0.200	3.460	33.21	5.400	6.68
KC-1310	$PhCH_3$	0.316	2.030	21.50	2.300	1.45
KC-1311	$PhCF_3$	0.034	0.498	4.44	0.383	0.615
KC-1315	PhCN	0.251	2.479	19.54	0.764	0.647
KC-1316	$PhCONHCH(CH_3)_2$	0.102	1.956	24.66	1.575	1.35
KC-1317	$PhCONH_2$	0.105	0.885	18.27	0.466	0.713
KC-1318	PhCOOH	0.029	0.512	4.38	0.370	0.439
KC-146[c]	$PhSO_2PhN(CH_3)_2$	0.366	0.781	–	2.800	–
TMP		0.312	30.360	190.89	1.590	–
K-122	$PhSO_2PhNO_2$	0.058	0.545	80.96	0.470	0.408
K-130	$PhSO_2PhNH_2$	0.040	0.266	24.19	0.550	0.670
K-150	$PhSO_2PhCH_3$	0.039	0.420	18.70	0.540	0.585

[a]Calculated with eq 5; [b]calculated with eq 4; [c]the *sec*-amino moiety has been changed into a methylene group.

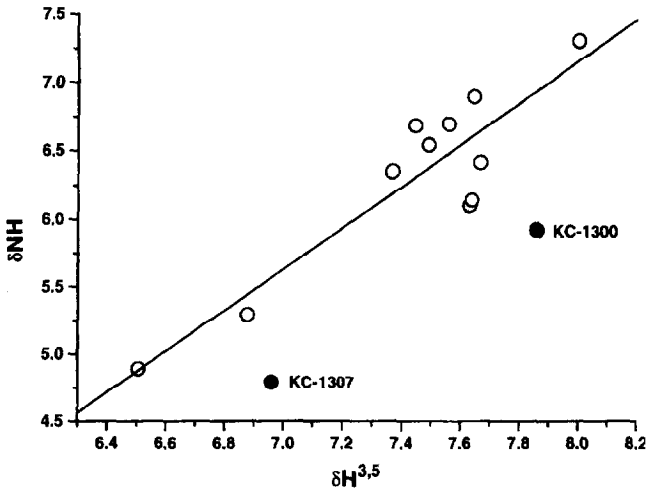


Fig 1. Correlation between the NMR chemical shifts of the NH group and the $H^{3,5}$ protons for the KC- and K-derivatives. The deviating dihydronaphthyl and naphthyl derivative are indicated.

seems to be the NMR chemical shift of the NH proton. This chemical shift is highly intercorrelated with the chemical shift of the $H^{3,5}$ protons (fig 1). The latter is more reliable due to less solvent and concentration dependency. The $H^{3,5}$ protons are on the phenyl ring in the *o*-position to the NH group. This descriptor together with other physicochemical parameters used in the final multiple linear regression analysis are listed in table II.

E coli DHFR

All the new derivatives were included in the first step of the QSAR analysis except the naphthyl derivative (KC-1300), which possesses an extended substitution pattern compared with the K-derivatives (eq 1).

$$\log(1/IC_{50}) = 0.74\ (\pm 0.26)\ \delta H^{3,5} \quad [1]$$
$$n = 9;\ r^2 = 0.54;\ s = 0.31;\ F = 7.1;\ r^2_{cv} = 0.25.$$

In all the equations here, the parameters are normalized for better comparison of the relative importance of the different descriptors used. Only 54% of the total variance in biological activity was explained using the NMR-chemical shift as descriptor. Inspection of the regression residuals shows that two compounds deviate, the isopropylamido (KC-1316) and the dihydronaphthyl (KC-1307) derivative. For the dihydronaphthyl derivative the same deviation from the regression line in figure 1 is observed as for the other derivative possessing a naphthalene ring system instead of a benzene ring. For the isopropropylamido derivative an additional steric effect can be assumed.

Omitting the two derivatives resulted in a highly significant correlation, explaining 87% of the variance. If the assumed steric effect is taken into account by the molar refraction (MR) as an additional descriptor, an equally significant equation is obtained including the two derivatives (eq 3).

$$\log(1/IC_{50}) = 0.93\ (\pm 0.16)\ \delta H^{3,5} \quad [2]$$
$$n = 7;\ r^2 = 0.87;\ s = 0.18;\ F = 32.5;\ r^2_{cv} = 0.72.$$

$$\log(1/IC_{50}) = 0.76\ (\pm 0.17)\ \delta H^{3,5} - 0.60\ (\pm 0.13)\ MR_{R4} \quad [3]$$
$$n = 9;\ r^2 = 0.90;\ s = 0.15;\ F = 27.9;\ r^2_{cv} = 0.83.$$

Table II. Data used for the multiple linear regression analysis.

Compound	*R*	$\delta H^{3,5}$	MR^a	$V_W{}^a$
KC-1300	Naphthyl-SO_2PhCH_3		37.82	77.29
KC-1303	$PhNO_2$	8.003	7.36	16.80
KC-1307	5,8-Dihydronaphthyl	6.960[b]	1.03	3.44
KC-1308	$PhN(CH_3)_2$	6.506	15.55	31.67
KC-1310	$PhCH_3$	6.877	5.65	13.67
KC-1311	$PhCF_3$	7.367	5.02	21.33
KC-1315	PhCN	7.446	6.33	14.70
KC-1316	$PhCONHCH(CH_3)_2$	7.631	23.76	53.68
KC-1317	$PhCONH_2$	7.638	9.81	22.02
KC-1318	PhCOOH	7.667	6.05	15.40
K-122	$PhSO_2PhNO_2$	7.646	39.53	80.42
K-130	$PhSO_2PhNH_2$	7.491	37.59	74.16
K-150	$PhSO_2PhCH_3$	7.560	37.82	77.29

[a]MR was calculated or taken from [13]; [b]δH^3 was used.

$\log(1/IC_{50}) = 0.62\ (\pm 0.10)\ \delta H^{3,5} - 0.96\ (\pm 0.10)\ MR_{R4}$ [4]
$n = 10$; $r^2 = 0.93$; $s = 0.15$; $F = 47.1$; $r^2_{cv} = 0.85$.

If the naphthyl derivative (KC-1300) is included, eq 4 is obtained with an identical statistical significance (see, however, CoMFA below). Similar results are obtained by replacing MR with V_w due to the fact that MR and the van der Waals volume (V_w) are highly correlated for this data set.

The assumed additional steric repulsion effect in the case of compound KC-1316 was substantiated by molecular modelling as discussed later.

Inclusion of the K-derivatives bearing as K-130 a substituted phenylsulfon moiety leads to the following equation (eq 5) (KC-1300 omitted).

$\log(1/IC_{50}) = 0.75\ (\pm 0.21)\ \delta H^{3,5}$ [5]
$n = 12$; $r^2 = 0.56$; $s = 0.26$; $F = 13.2$; $r^2_{cv} = 0.34$.

A major part of the variance is still explained by the electronic substituent effect, with an almost identical regression coefficient. The inclusion of MR indicated no significant contribution. This could be due to an overcompensation of the negative steric effect by an additional binding site for the extended structures of K-122, K-130 and K-150. The results of a nonlinear fit using MR and MR^2 seem to support this hypothesis (eq 6).

$\log(1/IC_{50}) = 0.76\ (\pm 0.13)\ \delta H^{3,5} - 0.60\ (\pm 0.17)\ MR_{R4} + 0.70\ (\pm 0.17)\ MR^2_{R4}$ [6]
$n = 12$; $r^2 = 0.87$; $s = 0.16$; $F = 18.5$; $r^2_{cv} = 0.68$.

Interestingly, a highly significant nonlinear dependency of the IC_{50} values on $\delta H^{3,5}$ can also be derived as found in the CoMFA analysis omitting compound KC-1316 (eq 7).

$\log(1/IC_{50}) = 0.68\ (\pm 0.15)\ \delta H^{3,5} - 0.37\ (\pm 0.15)\ (\delta H^{3,5})^2$ [7]
$n = 11$; $r^2 = 0.88$; $s = 0.14$; $F = 30.0$; $r^2_{cv} = 0.79$.

Surprisingly, the naphthyl derivative KC-1300 was found to be an extreme exception in this data set, its activity being very low and much less than that of TMP. This may indicate strong steric repulsion of the naphthyl ring moiety and/or the attached phenylsulfon group (see eq 3 and 4).

DHFRs derived from other sources

Multiregression analysis using the same descriptors as in eq 1–6 did not lead to equations with satisfying significance for the IC_{50} values determined for DHFRs derived from *M lufu*, *C albicans* and rat liver. However, electronic substituent effects again explain a major fraction of the variance. Equations were found for *M lufu* (eq 8 and 9) and *C albicans* (eq 10 and 11).

$\log(1/IC_{50}) = 0.58\ (\pm 0.33)\ \delta H^{3,5} - 0.021\ (\pm 0.332)\ MR_{R4}$ [8]
$n = 9$; $r^2 = 0.34$; $s = 0.37$; $F = 1.5$; $r^2_{cv} = 0.26$.

$\log(1/IC_{50}) = 0.75\ (\pm 0.21)\ \delta H^{3,5} - 0.29\ (\pm 0.34)\ MR_{R4} - 0.077\ (\pm 0.34)\ MR^2_{R4}$ [9]
$n = 12$; $r^2 = 0.47$; $s = 0.33$; $F = 2.4$; $r^2_{cv} = 0.07$.

$\log(1/IC_{50}) = 0.54\ (\pm 0.26)\ \delta H^{3,5} - 0.57\ (\pm 0.27)\ MR_{R4}$ [10]
$n = 9$; $r^2 = 0.59$; $s = 0.24$; $F = 4.3$; $r^2_{cv} = 0.32$.

$\log(1/IC_{50}) = 0.42\ (\pm 0.21)\ \delta H^{3,5} - 0.24\ (\pm 0.27)\ MR_{R4} + 0.78\ (\pm 0.26)\ MR^2_{R4}$ [11]
$n = 12$; $r^2 = 0.68$; $s = 0.31$; $F = 5.6$; $r^2_{cv} = 0.41$.

Differences in binding sites of the various DHFRs can also be conducted from the results of principal component and CoMFA analysis for the IC_{50} values determined for the DHFR of the four different species.

Descriptors beyond $\delta H^{3,5}$ and MR, such as total lipophilicity (log P), did not improve the correlations. Only poor correlations were found between the IC_{50} values of rat liver with all chemical descriptors used.

CoMFA analysis

Two different alignments of the molecules were used in the CoMFA study. In both cases the torsion angles of the benzylpyrimidine moiety (the common pharmacophore of the derivatives) were set to the values observed in the trimethoprim–DHFR complex [9, 10] and for the flexible alkyl chains an extended conformation was chosen. In the first alignment all derivatives were superimposed in a way to give the most similar conformations. The naphthyl derivative KC-1300, which is conformationally mostly restricted, served as a template on which the other derivatives were fitted so that the NH group, the phenyl rings and the SO_2 group were exactly overlayed. This led to conformations for the K-derivatives bearing a phenyl ring instead of the naphthyl ring which are energetically accessible but not optimal. In the second alignment the energetically most favourable conformations obtained from a systematic search and subsequent energy minimization were used. The naphthyl derivative was superimposed by a least squares fit of the NH and SO_2 groups and the dihydronaphthyl derivative by aligning the NH group. This led to a spatially different orientation of the rings for these two compounds.

If the naphthyl derivative KC-1300 was included no predictive CoMFA models could be obtained for the activities against any of the different DHFRs using either the whole data set or varying subsets. Therefore this derivative was excluded from further CoMFA analysis.

Using the first alignment and the default values (cut-off = 30 kcal, minimum sigma = 2 kcal), CoMFA

Table III. Influence of the field cut-off value on cross-validated r^2-values for the CoMFA analysis of the derivatives analyzed in eq 3 (n = 9) and eq 6 (n = 12) using the first alignment rule.

Cut-off[a] *(kcal)*	*n*	*E coli*		*M lufu*	
		Electrostatic only	*Steric + electrostatic*	*Electrostatic only*	*Steric + electrostatic*
3	9	0.740 (5)	0.603 (2)	0.186 (4)	0.201 (1)
30	9	0.225 (2)	0.340 (2)	0.284 (4)	0.240 (1)
3	12	0.706 (4)	0.621 (3)	0.401 (1)	0.425 (2)
30	12	0.232 (3)	0.354 (4)	0.422 (5)	0.474 (2)

The optimal number of components as determined from cross-validation is given in parentheses. [a]Minimal sigma values of 0.2 and 1 kcal for cut-off values of 3 and 30 kcal, respectively.

models with low predictivity were obtained. In order to study the influence of the various CoMFA parameters on the predictivity of the derived models, both the field cut-off and the minimum sigma value were varied. While the influence of different minimum sigma values in the range of 0.2 to 2 kcal proved to be of minor importance, the reduction of the field cut-off value resulted in a drastic increase of r_{cv}^2 for *E coli*-derived DHFR (table III). In addition the influence of the steric field changed; its inclusion no longer resulted in an increase of r_{cv}^2.

Using the second alignment both the field cut-off and minimum sigma value were of minor importance, but the data set chosen showed a strong influence on the derived CoMFA model. For the activities against *E coli*- and *C albicans*-derived DHFR, low cross-validated r^2-values were obtained if the dihydronaphthyl (KC-1307) and the isopropylamido (KC-1316) derivatives were included in the analysis (table IV), while for DHFR from *M lufu* results similar to those obtained with the first alignment were obtained (data not shown).

The reason why the compounds KC-1300, KC-1307 and KC-1316 are outliers can be rationalized as follows. If the first alignment is used the highly active K-derivatives are forced into the same orientation as the low active KC-1300 derivative. For the second alignment using optimized conformations these derivatives possess unique steric properties encountered only once in the data set.

1) The isopropyl group of KC-1316 occupies a region in space which is different from all other derivatives.
2) The phenyl ring of the naphthyl derivative KC-1300 shows an orientation which is different from the K-derivatives independently of whether the NH and SO_2 group or the phenyl/naphthyl ring were aligned.
3) The orientation of the rings for the dihydronaphthyl and the naphthyl derivative is different from the other compounds as they have energy minima at deviating C-N-phenyl torsion angles.

The results obtained from the various CoMFA can be summarized as follows. The activities of the derivatives against *E coli*-derived DHFR can be described by the electrostatic field at the *para*-substituent, probably indicating a dipole interaction with the receptor site. The importance of the *sec*-amino group for enhanced activity is indicated by CoMFA, provided additional derivatives in which this group is replaced by a methylene group are included (*eg*, compounds like KC-146). Otherwise the differences

Table IV. Influence of the data set used on cross-validated r^2-values for the CoMFA analysis of the derivatives analyzed in eq 1 (n = 7), eq 3 (n = 9), eq 1 plus the three K-derivatives (n = 10) and eq 6 (n = 12) using the second alignment (field cut-off: 30 kcal; minimum sigma: 1 kcal).

n	*E coli*		*C albicans*	
	Electrostatic only	*Steric + electrostatic*	*Electrostatic only*	*Steric + electrostatic*
7	0.850 (3)	0.394 (1)	0.697 (1)	0.418 (1)
9	0.234 (1)	0.312 (2)	0.219 (2)	0.166 (1)
10	0.859 (5)	0.584 (2)	0.776 (1)	0.688 (2)
12	0.279 (1)	0.329 (3)	0.426 (2)	0.406 (2)

The optimal number of components as determined from cross-validation is given in parenthesis.

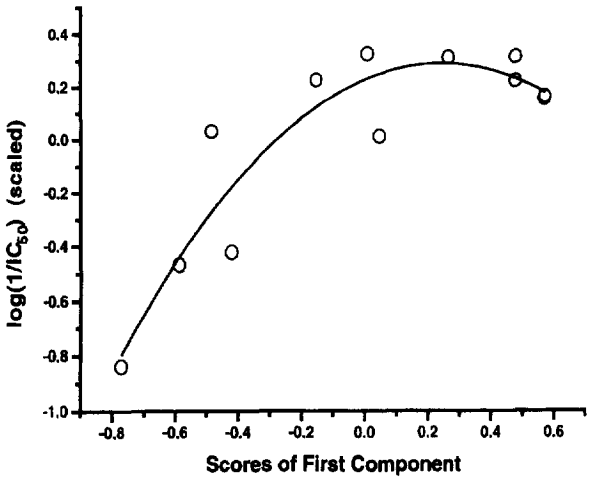


Fig 2. Relationship between inhibitory activity and the scores of the first component extracted by PLS using the first alignment and the following CoMFA settings: n = 12, field cut-off = 3 kcal, minimum sigma = 0.5 kcal (see also table III).

in the electrostatic fields at the NH group are too small for detection by CoMFA. Interestingly, a plot of the $\log(1/IC_{50})$ values for *E coli* against the scores of the first component shows a nonlinear, probably bilinear, relationship (fig 2) analogous to that given in eq 7 for the relationship between chemical shift and $\log(1/IC_{50})$ values. This result can be expected as we are approaching a final value in electron-attracting properties with NO_2 and SO_2 substituents in the *para*-position. It also explains the high value for the optimal number of components.

Principal component analysis

In addition, a principal component analysis (PCA) was performed to prove the conclusion of the multiple linear regression and CoMFA-analysis, *ie* different structural dependences of binding to DHFR from *E coli* and other sources, especially *M lufu*. Using the activity data of the nine new compounds as analyzed

Table V. Principal component analysis of four different IC_{50} values for 9 new compounds as used in eq 3.

% variance explained

PC	*Eigen value*	*Percent (%)*	*Sum (%)*
1	3.30	82.54	82.54
2	0.47	11.74	94.28
3	0.14	3.59	97.87
4	0.09	2.13	100.00

Loadings after Varimax rotation

Parameter	*PC 1*	*PC 2*
$\log(1/IC_{50})$ *M lufu*	0.3230	0.9382
$\log(1/IC_{50})$ *E coli*	0.9202	0.2801
$\log(1/IC_{50})$ *C albicans*	0.8841	0.3858
$\log(1/IC_{50})$ rat liver	0.7396	0.6194

Table VI. Principal component analysis of four different IC_{50} values for the data set of all 10 new compounds (eq 4).

% variance explained

PC	*Eigen value*	*Percent (%)*	*Sum (%)*
1	1.96	49.12	49.12
2	1.71	42.76	91.87
3	0.29	7.14	99.02
4	0.04	0.98	100.00

Loadings after Varimax rotation

Parameter	*PC 1*	*PC 2*
$\log(1/IC_{50})$ *M lufu*	–0.0071	0.9290
$\log(1/IC_{50})$ *E coli*	–0.0024	0.9289
$\log(1/IC_{50})$ *C albicans*	0.9867	0.0121
$\log(1/IC_{50})$ rat liver	0.9869	0.0214

Table VII. Principal component analysis of four different IC_{50} values for the data set used in eq 6 (n = 12, KC-1300 omitted).

% variance explained

PC	*Eigen value*	*Percent (%)*	*Sum (%)*
1	2.66	66.65	66.65
2	0.81	20.20	86.85
3	0.38	9.50	96.35
4	0.14	3.65	100.00

Loadings after Varimax rotation

Parameter	*PC 1*	*PC 2*
$\log(1/IC_{50})$ *M lufu*	0.8635	0.1351
$\log(1/IC_{50})$ *E coli*	0.8356	0.3425
$\log(1/IC_{50})$ *C albicans*	0.9448	0.1074
$\log(1/IC_{50})$ rat liver	0.1802	0.9787

in eq 3, 94% of the information content was extracted in two new orthogonal principal components (table V). The first PC was loaded by the IC_{50} values of *E coli* and *C albicans*, the second mainly by IC_{50} values of *M lufu* and the information content of the IC_{50} values from rat liver DHFR is distributed between the two PCs.

A PCA performed with all new derivatives (eq 4) shows a different trend. From the loadings it can be seen that the first PC is mainly loaded by IC_{50} values of *E coli* and *M lufu*, the second PC by IC_{50} values of rat liver and *C albicans* (table VI).

Using the 12 derivatives analyzed in eq 6 leads again to a change in the result of the PCA (table VII). Again two PCs are extracted explaining 87% of the total variance. From the loadings it can be seen that one PC is loaded by the IC_{50} values of *E coli*, *C albicans* and *M lufu* and the second by the IC_{50} values of rats. If the naphthyl derivative KC-1300 is included in the data set (n = 13) a similar distribution of the IC_{50} values in the PCs is found as in table VI.

Two conclusions can be derived from these results. First, one extreme data point (IC_{50}) can very much influence the result of a PCA as seen by comparison of tables III–V. Secondly, the close similarity derived for the IC_{50} data of *E coli* and *M lufu* according to table VI is not supported by regression analysis and CoMFA.

If the scores of the first PC (table VI) are used in regression analysis, again no satisfying correlation with the physicochemical parameters is obtained. This is in agreement with the results from multiple linear regression and CoMFA analysis especially for *M lufu*.

Conclusion

Using multiple linear regression analysis it could be shown that despite results of recent modelling studies the *sec*-amino moiety is important for the binding affinity of the reported compounds to DHFR derived from *E coli* and to a smaller extent to DHFR derived from *M lufu* and *C albicans*. The increased activity of K-122, K-130 and K-150 against DHFR derived from *E coli* despite the negative steric effect released by the *para*- substituents in the K-series may be explained by an overcompensation of this negative steric effect by an additional binding of the extended structures of the K-series. This seems to be supported by the nonlinear fit using MR and MR^2. The different structural dependences of binding to DHFR derived from various species detected by multiple linear regression analysis could be supported by PCA and CoMFA.

This example also shows that the derivation of structure–activity relationships for small data sets non-homogeneous in structure and/or conformation can lead to erroneous results. In such cases, the application of independent methods may minimize the risk of unreliable correlations and interpretations.

Experimental protocols

General methods

All solvents and reagents for synthesis were of reagent-grade quality and used without further purification. Melting points were determined in open capillaries on a Dr Tottoli apparatus (Büchi) and are uncorrected. Elemental analyses were performed by the Mikroanalytisches Laboratorium Ilse Beetz, Kronach, Germany, and were within 0.4% of the theoretical values. High resolution mass spectra were determined by the staff of Prof Tochtermann of the Christian Albrechts University, Kiel, Germany, and were within 3 mmasses of the calculated values. All ^{1}H-NMR spectra were in full accordance with the assumed structures. The NMR spectra were recorded on a Bruker AM 360 L spectrometer, 360 MHz for protons, in DMSO-d_6 at 20°C with TMS as an external standard.

Chemical procedures

The synthesis of compounds K-122, K-130 and K-150 is described elsewhere [3]. The alkylarylbromides used were synthesized *via* alkylation of the corresponding sulfonamides as described earlier [6].

General procedure for the alkylation of HDMP
Under an N_2 atmosphere 1.1 equiv sodium trimethylsilanolate was added to a solution of HDMP (1.5–5 mmol) stirred in 20 mL DMF. After stirring at room temperature for 0.15 h, 1 equiv of alkylarylbromide and a spatula point of KI were added to the suspension. The reaction was monitored by TLC on silica gel with CH_2Cl_2/*iso*-C_3H_7OH/NH_3 (conc) (9:1: saturated) as the mobile phase. When it was considered complete, generally within one week, the whole suspension was subjected to column chromatography on 200 g alumina oxide (basic) and eluted with EtOH. Those fractions containing the desired product were pooled and concentrated to dryness at reduced pressure. The remaining oil was subjected to column chromatography on silica gel and eluted with CH_2Cl_2/*iso*-C_3H_7OH/NH_3 (conc) (9:1: saturated). Those fractions containing the desired product were pooled and the solvent removed at reduced pressure to a few millilitres. After keeping the solution overnight in a refrigerator a white product precipitated.

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-nitro-N-(4-toluenesulfonyl)anilino)propoxy]benzyl}pyrimidine ***4a***. Yield 13%, mp 92–103°C, ^{1}H-NMR δ 1.66 (m, 2H, CH_2), 2.40 (s, 3H, CH_3), 3.51 (s, 2H, CH_2), 3.67 (s, 6H, OCH_3), 3.81 (m, 4H, CH_2), 5.69 (s, 2H, NH_2), 6.08 (s, 2H, NH_2), 6.53 (s, 2H, TMP-H), 7.39–7.47 (m, 6H, Tos-H, H2'/6'), 7.51 (s, 1H, TMP-H), 8.23 (AA'XX', 2H, H3'/5').

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-dimethylamino-N-(4-toluenesulfonyl)anilino)propoxy]benzyl}pyrimidine ***4b***. Yield 27%, mp 190–191°C, ^{1}H-NMR δ 1.62 (m, 2H, CH_2), 2.40 (s, 3H, CH_3), 2.88 (s, 6H, CH_3), 3.51 (s, 2H, CH_2), 3.61 (t, 2H, CH_2), 3.67 (s, 6H, OCH_3), 3.78 (t, 2H, CH_2), 5.68 (s, 2H, NH_2), 6.08 (s, 2H, NH_2), 6.52 (s, 2H, TMP-H), 6.61 (AA'XX', 2H, H3'/5'), 6.76 (AA'XX', 2H, H2'/6'), 7.38 (AA'XX', 2H, Tos-

H3/5), 7.44 (AA'XX', 2H, Tos-H2/6), 7.50 (s, 1H, TMP-H), $C_{31}H_{38}N_6O_5S$ (C, H, N).

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(N-(4-toluenesulfonyl)-4'-trifluoromethylanilino)propoxy]benzyl}pyrimidine ***4c***. Yield 21%, mp 142–151°C, ^{1}H-NMR δ 1.63 (m, 2H, CH_2), 2.40 (s, 3H, CH_3), 3.51 (s, 2H, CH_2), 3.67 (s, 6H, OCH_3), 3.76–3.81 (m, 4H, CH_2), 5.70 (s, 2H, NH_2), 6.09 (s, 2H, NH_2), 6.53 (s, 2H, TMP-H), 7.31 (AA'XX', 2H, H2'/6'), 7.41 (AA'XX', 2H, Tos-H3/5), 7.45 (AA'XX', 2H, Tos-H2/6), 7.51 (s, 1H, TMP-H), 7.75 (AA'XX', 2H, H3'/5').

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-cyano-N-(4-toluenesulfonyl)anilino)propoxy]benzyl}pyrimidine ***4d***. Yield 53%, mp 125–143°C, ^{1}H-NMR δ 1.63 (b, 2H, CH_2), 2.39 (s, 3H, CH_3), 3.52 (s, 2H, CH_2), 3.66 (s, 6H, OCH_3), 3.79 (m, 4H, CH_2), 5.69 (s, 2H, NH_2), 6.08 (s, 2H, NH_2), 6.53 (s, 2H, TMP-H), 7.30 (AA'XX', 2H, H2'/6'), 7.40 (AA'XX', 2H, Tos-H3/5), 7.43 (AA'XX', 2H, Tos-H2/6), 7.51 (s, 1H, TMP-H), 7.86 (AA'XX', 2H, H3'/5').

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(N-(4-toluolsulfonyl)-1-naphthylamino)propoxy]benzyl}pyrimidine ***4e***. Yield 22%, mp 206–214°C, ^{1}H-NMR δ 1.49/1.68 (b, each 1H, CH_2), 2.43 (s, 3H, CH_3), 3.50 (s, 2H, CH_2), 3.58 (s, 6H, OCH_3), 3.72/4.08 (b, 3H/1H, CH_2), 5.71 (s, 2H, NH_2), 6.10 (s, 2H, NH_2), 6.48 (s, 2H, TMP-H), 6.8 (d, 1H, Ar-H), 7.41–7.59 (m, 8H, Ar-H), 7.96 (m, 2H, Ar-H), 8.12 (m, 1H, Ar-H).

2,4-Diamino-5-{3,5-dimethoxy-4-[3-{4'-(4-toluenesulfonyl)-1-naphthylamino)propoxy]benzyl}pyrimidine KC-1300. *N*-(3-Bromopropyl)-4-(4'-toluenesulfonyl)-1-naphthylamine prepared from *N*-(3-bromopropyl)-*N*-(4'-toluenesulfonyl)-1-naphthylamine *via* acyl migration in H_2SO_4 (conc) was reacted with HDMP to give KC-1300 in 21% yield, mp 101°C (decomp), ^{1}H-NMR δ 1.80 (b, 2H, CH_2), 2.30 (s, 3H, CH_3), 3.33 (b, 2H, CH_2), 3.52 (s, 2H, CH_2), 3.60 (s, 6H, OCH_3), 3.80 (t, 2H, CH_2), 5.81 (s, 2H, NH_2), 5.92 (t, 1H, NH), 6.23 (s, 2H, NH_2), 6.54 (s, 2H, TMP-H), 7.34 (AA'XX', 2H, H3"/5"), 7.51 (s, 1H, TMP-H), 7.51–7.56 (m, 2H, H2'/6'), 7.64 (t, 1H, H7'), 7.81 (AA'XX', 2H, H2"/6"), 7.86 (d, 1H, H3'), 7.91 (d, 1H, H8'), 8.21 (d, 1H, H5'), $C_{33}H_{35}N_5O_5S$ (C, H, N).

2,4-Diamino-5-{3,5-dimethoxy-4-(4-(4"-dimethylaminophenyl-4'-sulfonylphenyl)butoxy]benzyl}pyrimidine KC-146. 4-(4-Bromobutyl)-4'-dimethylaminodiphenylsulfone prepared from 4-bromobutylsulfonylchloride and *N,N*-dimethylaniline *via* Friedel–Crafts acylation was reacted with HDMP to give KC-146 in 3% yield, mp 183–196°C (decomp), ^{1}H-NMR δ 1.58 (m, 2H, CH_2), 1.68 (m, 2H, CH_2), 2.66 (t, 2H, CH_2), 2.97 (s, 6H, NCH_3), 3.51 (s, 2H, CH_2), 3.66 (s, 6H, OCH_3), 3.78 (t, 2H, CH_2), 5.71 (s, 2H, NH_2), 6.11 (s, 2H, NH_2), 6.52 (s, 2H, TMP-H), 6.76 (AA'XX', 2H, H3"/5"), 7.38 (AA'XX', 2H, H2'/6'), 7.50 (s, 1H, TMP-H), 7.66 (AA'XX', 2H, H2"/6"), 7.75 (AA'XX', 2H, H3'/5'), $C_{31}H_{37}N_5O_5S$.

General procedures for removal of the protecting group. Method A

A solution of compounds **4a–d** (0.2–1.1 mmol) in H_2SO_4 (conc) (3.5 mL/mmol of compound) was stirred at room temperature and monitored by TLC. When the reaction was considered complete the solution was poured into 100 mL of ice water. The solid was filtered off and dried. Further purification was carried out by column chromatography on silica gel as mentioned above.

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-nitroanilino)propoxy]-benzyl}pyrimidine KC-1303. **4a** (1.1 mmol) was reacted to give KC-1303 in a 55% yield, mp 193–195°C, ^{1}H-NMR δ 1.87 (m, 2H, CH_2), 3.36 (m, 2H, CH_2), 3.53 (s, 2H, CH_2), 3.70 (s, 6H, OCH_3), 3.92 (t, 2H, CH_2), 5.71 (s, 2H, NH_2), 6.11 (s, 2H, NH_2), 6.57 (s, 2H, TMP-H), 6.65 (AA'XX', 2H, H2'/6'), 7.30 (t, 1H, NH), 7.52 (s, 1H, TMP-H), 8.00 (AA'XX', 2H, H3'/5'), $C_{22}H_{26}N_6O_5$ (C, H, N).

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-dimethylaminoanilino)-propoxy]benzyl}pyrimidine KC-1308. **4b** (0.3 mmol) was reacted to give KC-1308 in a 43% yield, mp 111–133°C, ^{1}H-NMR δ 1.83 (b, 2H, CH_2), 2.70 (s, 6H, CH_3), 3.11 (b, 2H, CH_2), 3.52 (s, 2H, CH_2), 3.71 (s, 6H, OCH_3), 3.90 (t, 2H, CH_2), 4.89 (b, 1H, NH), 5.69 (s, 2H, NH_2), 6.09 (s, 2H, NH_2), 6.51 (AA'XX', 2H, H3'/5'), 6.56 (s, 2H, TMP-H), 6.63 (AA'XX', 2H, H2'/6'), 7.51 (s, 1H, TMP-H), $C_{24}H_{32}N_6O_3$ (C (calcd: 63.7, found: 63.2), H, N), mass spectrum.

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-trifluoromethylanilino)-propoxy]benzyl}pyrimidine KC-1311. **4c** (0.2 mmol) was reacted to give KC-1311 in a 84% yield, mp 114–142°C, ^{1}H-NMR δ 1.85 (m, 2H, CH_2), 3.25 (m, 2H, CH_2), 3.52 (s, 2H, CH_2), 3.70 (s, 6H, OCH_3), 3.92 (t, 2H, CH_2), 5.69 (s, 2H, NH_2), 6.09 (s, 2H, NH_2), 6.35 (t, 1H, NH), 6.56 (s, 2H, TMP-H), 6.66 (AA'XX', 2H, H2'/6'), 7.37 (AA'XX', 2H, H3'/5'), 7.52 (s, 1H, TMP-H), $C_{23}H_{26}F_3N_5O_3$, mass spectrum.

KC-1315, KC-1316 and KC-1317. **4d** (0.96 mmol) was reacted to give a 4% yield of KC-1315, a 29% yield of KC-1316 and a 31% yield of KC-1317.

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-cyanoanilino)propoxy]-benzyl}pyrimidine KC-1315. Mp 196–197°C, ^{1}H-NMR δ 1.84 (m, 2H, CH_2), 3.26 (m, 2H, CH_2), 3.52 (s, 2H, CH_2), 3.70 (s, 6H, OCH_3), 3.91 (t, 2H, CH_2), 5.69 (s, 2H, NH_2), 6.09 (s, 2H, NH_2), 6.56 (s, 2H, TMP-H), 6.63 (AA'XX', 2H, H2'/6'), 6.68 (t, 1H, NH), 7.45 (AA'XX', 2H, H3'/5'), 7.51 (s, 1H, TMP-H), $C_{23}H_{26}N_6O_3$, mass spectrum.

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-isopropylamidoanilino)-propoxy]benzyl}pyrimidine KC-1316. Mp 124–135°C, ^{1}H-NMR δ 1.13 (d, 6H, CH_3), 1.85 (m, 2H, CH_2), 3.25 (m, 2H, CH_2), 3.53 (s, 2H, CH_2), 3.71 (s, 6H, OCH_3), 3.92 (t, 2H, CH_2), 4.06 (m, 1H, CH), 5.70 (s, 2H, NH_2), 6.09 (b, 3H, NH/NH_2), 6.54 (AA'XX', 2H, H2'/6'), 6.56 (s, 2H, TMP-H), 7.52 (s, 1H, TMP-H), 7.63 (AA'XX', 2H, H3'/5'), 7.73 (d, 1H, AMID-H), $C_{26}H_{34}N_6O_4$, mass spectrum.

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-amidoanilino)propoxy]-benzyl}pyrimidine KC-1317. Mp 212–217°C, ^{1}H-NMR δ 1.85 (m, 2H, CH_2), 3.25 (m, 2H, CH_2), 3.52 (s, 2H, CH_2), 3.71 (s, 6H, OCH_3), 3.92 (t, 2H, CH_2), 5.69 (s, 2H, NH_2), 6.09 (s, 2H, NH_2), 6.14 (t, 1H, NH), 6.54 (AA'XX', 2H, H2'/6'), 6.56 (s, 2H, TMP-H), 6.84 (s, 1H, AMID-H), 7.52 (s, 2H, AMID-H/TMP-H), 7.64 (AA'XX', 2H, H3'/5'), $C_{23}H_{28}N_6O_4$ (C, H, N).

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-carboxyanilino)propoxy]-benzyl}pyrimidine KC-1318. A solution of 0.22 mmol KC-1317 in 8.5 mL H_2SO_4 (50% m/m) was stirred at 90°C for 3 h. The solution was poured into 100 mL ice water and the pH adjusted to 4. The fine precipitate was separated and purified as mentioned above, mp 181–186°C, ^{1}H-NMR δ 1.85 (m, 2H, CH_2), 3.27 (m, 2H, CH_2), 3.52 (s, 2H, CH_2), 3.71 (s, 6H, OCH_3), 3.92 (t, 2H, CH_2), 5.74 (s, 2H, NH_2), 6.12 (s, 2H, NH_2),

6.41 (t, 1H, NH), 6.56 (s, 2H, TMP-H), 6.57 (AA'XX', 2H, H2'/6'), 7.51 (s, 1H, TMP-H), 7.67 (AA'XX', 2H, H3'/5'), 11.96 (b, 1H, COOH), $C_{23}H_{27}N_5O_5$, mass spectrum within 3.6 mmasses.

General procedure for removal of the protecting group. Method B

To a solution of compounds **4c** and **4e** (0.4 mmol and 0.33 mmol, respectively) in 100 mL liquid ammonia a sufficient amount of lithium was added. After completion of the reaction the ammonia was allowed to evaporate to dryness. The solid was subjected to column chromatography on silica gel as mentioned above.

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(4'-methylanilino)propoxy]-benzyl}pyrimidine KC-1310. **4c** (0.4 mmol) was reacted to give KC-1310 in a 19% yield, mp 139–143°C, ^{1}H-NMR δ 1.85 (b, 2H, CH_2), 2.50 (s, 3H, CH_3), 3.15 (b, 2H, CH_2), 3.52 (s, 2H, CH_2), 3.71 (s, 6H, OCH_3), 3.90 (t, 2H, CH_2), 5.29 (t, 1H, NH), 5.67 (s, 2H, NH_2), 6.09 (s, 2H, NH_2), 6.47 (AA'XX', 2H, H2'/6'), 6.55 (s, 2H, TMP-H), 6.88 (AA'XX', 2H, H3'/5'), 7.51 (s, 1H, TMP-H), $C_{23}H_{29}N_5O_3$, mass spectrum.

2,4-Diamino-5-{3,5-dimethoxy-4-[3-(5,8-dihydronaphthyl-1-amino)propoxy]benzyl}pyrimidine KC-1307. **4e** (0.33 mmol) was reacted to give KC-1307 in a 21% yield, mp 119–126°C, ^{1}H-NMR δ 1.89 (m, 2H, CH_2), 2.98 (b, 2H, CH_2), 3.23–3.30 (m, 4H, CH_2), 3.52 (s, 2H, CH_2), 3.70 (s, 6H, OCH_3), 3.91 (t, 2H, CH_2), 4.79 (t, 1H, NH), 5.69 (s, 2H, NH_2), 5.85 (s, 2H, H5'/8'), 6.10 (s, 2H, NH_2), 6.36 (d, 1H, H2'), 6.40 (d, 1H, H4'), 6.56 (s, 2H, TMP-H), 6.96 (t, 1H, H3'), 7.52 (s, 1H, TMP-H), $C_{26}H_{33}N_5O_3$.

Biological assays

The test strains used were *M lufu* L 209 [11], *E coli* ATCC 11775 and *C albicans* ATCC 11651. The derived enzyme DHFR was partially purified as described previously [12]. The DHFR activity was assayed in 0.1 M Tris buffer, pH 7.24, by monitoring the decrease in absorbance photometrically at 340 nm as a function of time. After incubation of the DHFR with 0.1 mM NADPH and various amounts of inhibitor for 5 min at 25°C the reaction was started by adding 0.03 mM dihydrofolate. IC_{50} values were calculated as the concentration of free inhibitor required for a 50% decrease in velocity of the enzyme reaction [12].

Acknowledgments

We would like to thank HP Cordes and B Kunz for their skillful technical assistance.

References

1 Kuyper LF, Roth B, Baccanari DP *et al* (1985) *J Med Chem* 28, 303–311
2 Czaplinski K-H, Kansy M, Seydel JK (1987) *Quant Struct–Act Relat* 6, 70–72
3 Kansy M, Seydel JK, Wiese M, Haller R (1992) *Eur J Med Chem* 27, 237–244
4 Czaplinski K-H (1993) PhD Thesis, Christian-Albrechts-Universität, Kiel, Germany
5 Brossi A, Grunberg E, Hoffer M, Teitel S (1971) *J Med Chem* 14, 58–59
6 Baker BR, Santi DV, Shapiro HS (1964) *J Pharm Sci* 53, 1317–1325
7 Dewar MJS, Thiel W (1977) *J Am Chem Soc* 99, 4899
8 SYBYL version 6.0, Tripos Associates, Saint Louis, MO 63144, USA
9 Baker DJ, Bedell CR, Champness JN *et al* (1981) *FEBS Lett* 126, 49–52
10 Mathews DA, Bolin JT, Burridge JM *et al* (1985) *J Biol Chem* 260, 381–391
11 Portaels F (1980) *Ann Soc Belge Med Trop* 60, 381–386
12 Wiese M, Kansy M, Kunz B, Seydel JK (1990) *In: Proceedings in Chemistry and Biology of Pteridines*, Walter de Gruyter & Co Berlin, Germany, 1145–1149
13 Seydel JK, Schaper KJ (1979) *Chemische Struktur und biologische Aktivität von Wirkstoffen*, Verlag Chemie, Weinheim, Germany