Pergamon

Bioorganic & Medicinal Chemistry, Vol. 5, No. 12, pp 2221–2234, 1997

PII: S0968-0896(97)00163-6
0968-0896/97 $17.00 + 0.00

Binding to δ and μ Opioid Receptors by Deltorphin I/II Analogues Modified at the Phe[3] and Asp[4]/Glu[4] Side Chains: a Report of 32 New Analogues and a QSAR Study

Stephen E. Schullery,[a] Tasneem Mohammedshah,[a] Hafida Makhlouf,[a] Eleanor L. Marks,[a] Benjamin S. Wilenkin,[a] Sharleen Escobar,[a] Carol Mousigian[b] and Deborah L. Heyl[a,*]

[a]*Department of Chemistry, Eastern Michigan University, Ypsilanti, MI 48197, U.S.A.*
[b]*College of Pharmacy, University of Michigan, Ann Arbor, MI 48109, U.S.A.*

Abstract—The synthesis and binding affinities of 32 X^3Gly^4 dual-substitution analogues of the natural opioid heptapeptides deltorphin I and II are reported. A multiple regression QSAR analysis was performed using those results along with literature data for the X^3Asp^4 and Phe^3X^4 side chain analogues. Fitting to a three-term potential well model with hydrophobic and van der Waals attraction terms and a steric repulsion term indicates that the δ and μ receptor sites for binding the residue three side chain are similar, and that the binding interaction is primarily van der Waals and secondarily hydrophobic. Further analysis indicates that both sites are more constrained with respect to side chain length than width or thickness, and the μ site appears to be somewhat larger. A binding model consistent with these findings pictures the native third residues Phe ring laying on a step notched out of the receptor surface, pointing toward the back (riser) of the step, and sandwiched between the receptor and ligand. However, the binding sites for the residue four side chains are quite different on δ and μ receptors. Binding to the δ site appears to involve both electrostatic attraction (probably to a partial positive charge) and van der Waals attraction, but not necessarily hydrogen bonding, and more constraint with respect to side chain length than width or thickness. In contrast, there is no evidence for any kind of binding attraction between the side chain of residue four and the μ site, which acts more as steric repulsion site, as though the space that is a pocket on the δ receptor is filled in on the μ receptor. A regression model based only on steric repulsion by van der Waals bulk and/or the effective bulk of a hydration layer accounts for over 80% of the residue four related variation in μ affinity. © 1997 Elsevier Science Ltd.

Introduction

The deltorphins are a family of three heptapeptides isolated from amphibian skin that display extremely high affinities and selectivities for δ opioid receptors: deltorphin I, or C, (Tyr-*D*-Ala-Phe-Asp-Val-Val-Gly-NH_2); deltorphin II, or B, (Tyr-*D*-Ala-Phe-Glu-Val-Val-Gly-NH_2); and dermenkephalin, or deltorphin A, (Tyr-*D*-Met-Phe-His-Leu-Met-Asp-NH_2).[1] It is generally accepted that the N-terminal tripeptide message domain is responsible for interaction with opioid receptors,[2] and that the first and third residues, Tyr and Phe, are important elements in the binding pharmacophore.[3–5] Because the N-terminal tetrapeptide itself is μ-selective,[6] the C-terminal residues of the native peptides are thought to play an 'addressing' role, simultaneously decreasing μ affinity and increasing δ affinity.[7–10] The molecular-level details have been the subject of much research and speculation.

The role of residue four, anionic Asp or Glu in deltorphins I or II, respectively, and partially positive His in dermenkephalin has been particularly puzzling. Substitution of residue four in deltorphin I/II with Gly results in high μ affinity while preserving the δ affinity, a phenomenon referred to as opioid infidelity.[9,11,12] Other replacements, whether hydrophobic or hydrophilic, typically result in modest increases in μ binding.[6,11–15] Although cDNA site-directed mutagenesis indicates that a negative charge may exist in the μ receptor binding pocket,[12,16] substitution of a positively charged amino acid at position four is detrimental to both μ and δ binding.[6] The surprising inference from these observations is that the importance of residue four in receptor discrimination is not related to the charge of its side chain. Further clouding the issue is the observation that the dual affinity peptides bind to the δ receptor via a two-site model, consistent with the presence of δ-1 and δ-2 sites, or subtypes, while the μ receptor binding profiles fit a one site model.[12]

Our previous work has focused mainly on modification of the Phe^3 residue of deltorphin I to vary steric, electronic and hydrophobic properties in an effort to ascertain the important features in the δ receptor binding interaction.[17–19] The Phe^3 residue was chosen because it is a crucial receptor binding element, yet, unlike Tyr^1, it is relatively tolerant to modification. However, the low μ affinity of the parent compound

Key words: deltorphin, QSAR, opioid, peptide, binding.

made it difficult to assess effects on μ binding. We therefore felt it would be informative to study third-residue modifications of the dual-affinity, Tyr-*D*-Ala-Phe-Gly-Val-Val-Gly-NH_2, which should result in easily detectable shifts in both δ and μ binding. By using many of the same amino acid substitutions employed in the earlier X^3Asp^4 study, we hoped to determine not only what features are important for improving or maintaining δ and μ binding affinities, but also which of those features are responsible for targeting δ or μ receptor sites.

The desirability of a quantitative structure–activity relationship (QSAR) study is clear. The substituent properties commonly used to interpret the behavior of substitution analogues (e.g., hydrophobicity, steric repulsion, and electronic induction) are highly interrelated. The fact that it is impossible to systematically vary one without simultaneously changing the others, hopelessly confounds the quest for understanding. The mathematical deconvolution of such simultaneous effects is possible by multiple regression analysis, but sufficiently large data sets have only recently been gathered. Specifically, three promisingly large series of deltorphin I or II analogues are available, (i) 33 [X^3Gly^4] deltorphin I/II compounds reported herein, (ii) 50 experiments involving 41 [X^3] deltorphin I analogues and (iii) 25 experiments involving 18 [X^4] deltorphin I/II compounds. Data for the latter two series are taken from a recent review.[20] The X^4 substitution series can be viewed, in part, as a study of the transition between the dual high affinity Gly^4 form and the δ selective native Asp^4/Glu^4 forms.

An important approximation was made to allow the isolation of individual residue side chains effects; QSAR behaviors were calculated for hypothetical isolated side chain moieties capped with a hydrogen at their normal α-carbon attachment site and molecular mechanics optimized. It was assumed that calculation errors due to the H-capping would be more constant through a series of analogues, and smaller than errors resulting from attempts to study attached side chains. Also, analogues that involved any covalent modification of the backbone were excluded from the study. The following properties were investigated: octanol-water partition coefficient (log *P*); polarizability; refractivity; molar mass; van der Waals and solvent-accessible surface area and volume; diameter of the equivalent sphere based on solvent-accessible and van der Waals surface areas and volumes; length, width and thickness of the equivalent rectangular solid; flatness; ovality based on solvent-accessible and van der Waals dimensions. For the side chains of the Phe^3X^4 series the following were also studied: electrical charge; absolute value of the charge; hydrogen bond donor number; hydrogen bond acceptor number; total hydrogen bonding capacity and charge-weighted hydrogen bonding capacity.

The binding data are analyzed in terms of the logarithm of enhancement factors, E_δ and E_μ, which are defined as the ratio of the native $K_{i\delta}$ or $K_{i\mu}$ to the analogue $K_{i\delta}$ or $K_{i\mu}$, respectively. For example,

$$\log E_\delta = \log (\text{native } K_{i\delta} / \text{analogue } K_{i\delta})$$

Thus, if E_δ or E_μ is larger than 1, or if $\log E_\delta$ or $\log E_\mu$ is larger than zero, then this corresponds to tighter binding than the native form. The log function has two advantages: (i) in that the QSAR variables directly reflect either energy or entropy effects, the log function is more likely to exhibit a linear dependence, in accordance with,

$$\ln K = -\Delta G^0/RT = -\Delta H^0/RT + \Delta S^0/R$$

and (ii) the log function conveniently gathers and distributes data spanning orders of magnitude. To minimize the effect of inter-laboratory variability, all enhancement factors are calculated relative to the native $K_{i\delta}$ or $K_{i\mu}$ from the same laboratory. This normalization, which also allows IC_{50} binding data to be included, reduced the noise in the K_i data from about two to one orders of magnitude.

Multiple regression analysis permits dissection of the linear effect of each member of a group of independent variables, while correcting for the linear effects of the other members. The regression equation equates the dependent *y* variable, here $\log E_\delta$ or $\log E_\mu$, to a y-intercept constant k_0, plus a term consisting of the product of a linear coefficient k_i and its variable x_i, for each independent *x* variable:

$$y = k_0 + k_1x_1 + k_2x_2 + \ldots$$

Each regression coefficient k_i is the slope (partial derivative) of the dependent *y* variable with respect to the corresponding x_i variable, controlling for the simultaneous apparent linear effects of the other *x* variables in the equation. Selection of the best set of variables to use in a regression analysis is not an exact science, but, rather, requires judgements involving trade-offs between improved fit and physical interpretability, as well as a variety of statistical issues including multicollinearity, nonlinearity, and parsimony.[21]

Our approach begins with inspection of the complete bivariate correlation matrix for the dependent and all of the independent variables. This permits any particularly important linear relationships to be identified. It also reveals combinations of variables that will be hampered by multicollinearity problems. If the bivariate correlations suggest a promising lead variable, complementary regression variables are sought by systematically testing for increased predictive power of all possible combinations with the lead variable. Also, the least useful variables are identified by starting from regression on all of the variables and systematically testing and eliminating those which, upon removal, result in the least decrease of the regression equations predictive power. The possibility of curvature is tested by inclusion of both linear and squared terms in each variable. For

example, a positive coefficient on the linear term coupled with a negative coefficient on the square term suggests a convex-upward function consistent with the existence of an optimum value for the variable. Once the field of potentially important variables has been narrowed, and we are certain that no singularly important variable has been overlooked, a physical model is sought that dictates the final set of variables that will combine physical interpretability with maximum predictive power. Finally, to maximize the information content of the regression equation, where appropriate, the independent variables are centered on the native compound by fitting to the difference, Δ, between the analogue and native forms (i.e., Δ = analogue−native). Thus, if the independent variables indeed capture all systematic sources of variability in the independent variable, the best fit should have an intercept of zero, thereby providing an additional check on the validity of the model.

Two related indicators of predictive power and fit quality are focused on, namely the coefficient of multiple determination R^2 for the overall fit, and the probability p (or 100 minus the significance level) of the individual regression coefficient(s). R^2 equals the fraction of the observed variability in the independent variable that can be attributed to, or explained by, the combination of independent variables as presented in the proposed regression equation. A less optimistic related quantity, also considered, is the adjusted multiple coefficient of determination, adjusted R^2 which is the fraction of variability that the regression equation could be expected to account for if presented with a new (comparably scattered) set of data from the same population. This adjustment is an attempt to compensate for small sample size and/or too many parameters. The regression coefficient p-value is the probability that the apparent relationship between that variable and the independent variable could be due to chance, in other words, the probability that the regression coefficient is actually zero and that the variable is irrelevant. $p \leq 0.05$ is a common standard for a significant relationship. In addition, the F statistic and its significance level (the probability that all of the regression coefficients are zero) are given for the primary regression results.

In addition to the various approximations that are intrinsic to the multiple regression method, there are several concerns regarding the study at hand. First, there is a tremendous amount of noise in the data. Normalizing all binding data to the same laboratory value for the native compound reduces the variability from about two to one orders of magnitude, leaving signal-to-noise ratios of only about three or four. As no model can be expected to account for inter-lab variability, we infer that R^2 values of about 75% are the best that should be hoped for. Second, none of the analogue series portend easy compliance with the ten-points-per-parameter guideline. The most promising, and numerous, is the series of 50 experiments involving 41 [X^3] deltorphin I analogues. The next most numerous series is the 33 [X^3Gly^4] deltorphin I/II compounds, followed by the series of 25 experiments involving 18 [X^4] deltorphin I/II compounds. Third, the various algorithms used to calculate the QSAR properties are approximate, and those approximations are twice compounded, first before the fact by applying them to H-capped, molecular mechanics optimized structures, and second after the fact by additional derivations, such as equivalent-sphere/ellipsoid/box modeling.

Results and Discussion

Binding and physicochemical properties of the X^3Gly^4 series

All newly synthesized deltorphin analogues were assayed for binding to μ (versus [^{3}H]DAMGO], or [^{3}H][*D*-Ala^2, *N*-MePhe^4, Gly^5-ol]enkephalin), to δ (versus [^{3}H]DPDPE, or [^{3}H][*D*-Pen^2, *D*-Pen^5]enkephalin), and to κ (versus [^{3}H]U69,593, or 5a,7a,8b-(−)-*N*-[7-(1-pyrrolidinyl)-1-oxaspiro(4,5)dec-8-yl]benzeneacetamide) opioid receptors. Table 1 summarizes the binding affinity data obtained. Also included in Table 1 for comparison are the opioid binding affinities of deltorphin I and its nonselective [Gly^4]analogue. A measure of selectivity is provided as a ratio of μ to δ binding affinities. Binding to κ receptors was negligible for all compounds. Physicochemical data for all newly reported analogues are provided in Table 2. The following amino acid replacements were made at residue three of the [Gly^4] parent peptide: side chain lengthened and shortened phenylalanine analogues homophenylalanine (Hfe) and phenylglycine (Pgl); heterocyclic amino acids histidine (His), 4-thiazolylalanine (Taz), 2-thienylalanine (Thi), 3-benzothienylalanine (Bth), tryptophan (Trp), 2-, 3-, and 4-pyridylalanine (2-Pal, 3-Pal, 4-Pal); phenylalanine analogues with ring substitutents *para*-, *ortho*-, and *meta*-fluorophenylalanine (*p*-FPhe, *o*-FPhe, *m*-FPhe), pentafluorophenylalanine (F_5Phe), *meta*-trifluoromethylphenylalanine (*m*-CF_3Phe), *para*-chlorophenylalanine (*p*-ClPhe), *para*-bromophenylalanine (*p*-BrPhe), *para*-aminophenylalanine (*p*-NH_2Phe), *para*-nitrophenylalanine (*p*-NO_2Phe), tyrosine (Tyr) and tyrosine-4-methyl ether or *p*-methoxyphenylalanine (Tyr(OMe)); nonaromatic amino acids cyclohexylalanine (Cha), cyclopentylglycine (Chg), valine (Val), leucine (Leu), isoleucine (Ile), alanine (Ala), α-aminobutyric acid (Abu), norvaline (Nva), norleucine (Nle), *tert*-leucine or *tert*-butylglycine (Tle) and *tert*-butylalanine (*t*-BuAla).

QSAR properties

Table 3 presents the values calculated for the primary QSAR properties used in this study. Replacement amino acids not previously defined, (Tables 1 and 2), are as follows: 1-naphthylalanine (1-Nal); 2-naphthylalanine (2-Nal); arginine (Arg); asparagine (Asn); aspartic acid (Asp); aspartic acid-γ-benzyl ester (Asp(OBzl)); aspartic acid-γ-*tert*-butyl ester (Asp(O*t*-Bu)); bis-(benzo)cycloheptylglycine-amino-(10,11-di-

Table 1. Opioid receptor binding affinities[a] of [X^3, Gly^4]deltorphin I/II analogues

Peptide	Compound	$K_{i\delta}$ (nM) DPDPE	$K_{i\mu}$ (nM) DAMGO	$K_{i\kappa}$ (nM) (U69,593)	$K_{i\mu}/K_{i\delta}$
Deltorphin I	—	1.73±0.211	677±66.0	>10,000	391
[Gly^4]	—	2.6±0.3	4.7±1.5	4900±2500	1.81
[Pgl^3, Gly^4]	**1**	34.2±6.45	118±8.25	5.3%	3.45
[Hfe^3, Gly^4]	**2**	5.59±0.682	4.57±0.314	36.4%	0.818
[His^3, Gly^4]	**3**	162±20.5	169±15.9	11.9%	1.04
[Taz^3, Gly^4]	**4**	85.0±2.70	42.2±3.19	**0%**	0.496
[Thi^3, Gly^4]	**5**	10.9±2.78	13.1±0.776	12.9%	1.20
[Bth^3, Gly^4]	**6**	9.13±2.62	3.49±0.924	56.9%	0.382
[Trp^3, Gly^4]	**7**	483±81.2	131±10.4	5690±168	0.27
[2-Pal^3, Gly^4]	**8**	148±13.3	88.5±8.02	4.0%	0.598
[3-Pal^3, Gly^4]	**9**	729±161	1320±141	3.1%	1.82
[4-Pal^3, Gly^4]	**10**	27.3±5.32	182±17.6	7.6%	6.67
[*p*-FPhe^3, Gly^4]	**11**	18.6±0.734	27.0±1.62	30.5%	1.45
[*o*-FPhe^3, Gly^4]	**12**	10.2±0.730	12.1±0.875	18.3%	1.19
[*m*-FPhe^3, Gly^4]	**13**	10.5±0.413	18.3±2.11	25%	1.74
[F_5Phe^3, Gly^4]	**14**	191±13.5	205±8.31	22.1%	1.07
[*m*-CF_3Phe^3, Gly^4]	**15**	13.3±8.27	95.1±7.19	31.2%	7.15
[*p*-ClPhe^3, Gly^4]	**16**	106±13.5	70.0±5.15	28.2%	0.660
[*p*-BrPhe^3, Gly^4]	**17**	99.7±6.50	178±9.89	24.5%	1.79
[*p*-NH_2Phe^3, Gly^4]	**18**	1260±59.1	1600±46.5	10%	1.27
[*p*-NO_2Phe^3, Gly^4]	**19**	641±23.8	540±45.1	40.2%	0.842
[Tyr^3, Gly^4]	**20**	348±25.7	177±19.8	6.0%	0.509
[Tyr(OMe)^3, Gly^4]	**21**	431±24.6	1460±185	6.1%	3.38
[Cha^3, Gly^4]	**22**	8.22±0.419	11.7±0.993	22.8%	1.42
[Chg^3, Gly^4]	**23**	16.8±1.70	11.1±1.50	21.2%	0.661
[Val^3, Gly^4]	**24**	219±17.0	2030±183	6.6%	9.27
[Leu^3, Gly^4]	**25**	63.4±18.2	134±16.3	7.9%	2.11
[Ile^3, Gly^4]	**26**	180±25.1	325±22.7	14.0%	1.81
[Ala^3, Gly^4]	**27**	1110±6.37	4670±128	5.1%	4.21
[Abu^3, Gly^4]	**28**	490±55.3	2410±466	7.5%	4.92
[Nva^3, Gly^4]	**29**	162±31.5	498±48.8	2.9%	3.07
[Nle^3, Gly^4]	**30**	23.3±4.70	41.1±3.91	11.6%	1.76
[Tle^3, Gly^4]	**31**	44%	5840±595	2.9%	>0.584
[*t*-BuAla^3, Gly^4]	**32**	106±12.0	150±9.59	8.1%	1.42

[a]Average values were determined from two to four assays performed in triplicate, ±standard error of the mean. Where the K_i was determined to be greater than 10 μM (no substantial binding), percent displacements at the 10 μM limit are given.

hydro-5H-dibenzo{a,d}cyclohepten-5-yl) acetic acid (Bhg); β-methylphenylalanine (β-MePhe); β-hydroxyphenylalanine (β-OHPhe); biphenylalanine (Bip); benzoylphenylalanine (Bpa); cyclopentylalanine (Cpa); cysteine (Cys); diphenylalanine (Dip); glutamine (Gln); glutamic acid (Glu); glycine (Gly); 2′-iodotyrosine (*o*-ITyr); 2′,6′-iodotyrosine (*o*-I_2Tyr); lysine (Lys); pentamethylphenylalanine (Me_5Phe); methionine (Met); methionine sulfoxide (Met(O)); ornithine (Orn); penicillamine or β,β-dimethylcysteine (Pen); phenylalanine (Phe); *para*-iodophenylalanine (*p*-IPhe); *para*-methylphenylalanine (*p*-MePhe); serine (Ser); threonine (Thr); thyronine (Thy). A variety of secondary properties are derived from these data, as described in the Experimental section. The charge and hydrogen bonding related properties, which were used only for the Phe^{3}X^4 series, are presented in Table 4.

Bivariate correlation results for the X^{3}Asp^4 and X^{3}Gly^4 series

Table 5 presents the bivariate correlations between the primary QSAR variables and E_δ, log E_δ and log E_μ for the X^{3}Asp^4 series, and log E_δ and log E_μ for the X^{3}Gly^4 as well as the Phe^{3}X^4 series to be discussed later. Such bivariate correlation coefficients, r, indicate the strength of whatever linear relationship exists between pairs of variables, with $r = 0$ indicating no relationship, and $r = \pm 1$ indicating exact linearity, with corresponding positive or negative slope. The coefficient of determination, r^2, indicates the fraction of the total variability in one variable that can be attributed to, or explained by, changes in the other variable. The thresholds for 95% certainty ($p = 0.05$) of the existence of a correlation are $|r| \geq 0.29$ for the $N = 50$, X^{3}Asp^4 data set, and $|r| \geq 0.35$ for the $N = 33$, X^{3}Gly^4 data set. The μ binding correlations for the X^{3}Asp^4 series cannot be interpreted quantitatively because μ binding by 19 of these compounds was below the protocol detection limit (10,000 nM); in order to retain as much information as possible, their reported detection limits were used to calculate approximate E_μ values. This artificial stacking up of the data at the weak-binding end of the range undoubtedly contributed to the very low correlations for this series.

Table 2. Physicochemical data for [X^3, Gly^4] deltorphin I/II analogues

Peptide	Compd	HPLC[a]	Purity (%)[b]	Mol wt[c]
[Pgl^3, Gly^4]	**1**	21.24	98	697.0
[Hfe^3, Gly^4]	**2**	22.42	96	725.0
[His^3, Gly^4]	**3**	11.92	95	701.0
[Taz^3, Gly^4]	**4**	19.54	98	717.5
[Thi^3, Gly^4]	**5**	19.90	95	717.3
[Bth^3, Gly^4]	**6**	23.34	95	766.4
[Trp^3, Gly^4]	**7**	24.01	96	749.5
[2-Pal^3, Gly^4]	**8**	17.78	95	712.0
[3-Pal^3, Gly^4]	**9**	18.04	99	712.0
[4-Pal^3, Gly^4]	**10**	19.54	95	712.0
[*p*-$FPhe^3$, Gly^4]	**11**	22.76	97	728.5
[*o*-$FPhe^3$, Gly^4]	**12**	23.66	95	728.5
[*m*-$FPhe^3$, Gly^4]	**13**	23.08	96	728.5
[F_5Phe^3, Gly^4]	**14**	23.84	96	800.5
[*m*-CF_3Phe^3, Gly^4]	**15**	24.32	95	778.6
[*p*-$ClPhe^3$, Gly^4]	**16**	23.18	95	745.3
[*p*-$BrPhe^3$, Gly^4]	**17**	23.63	96	790.0
[*p*-NH_2Phe^3, Gly^4]	**18**	18.16	95	725.3
[*p*-NO_2Phe^3, Gly^4]	**19**	21.12	97	755.3
[Tyr^3, Gly^4]	**20**	17.15	96	727.0
[$Tyr(OMe)^3$, Gly^4]	**21**	21.40	95	740.5
[Cha^3, Gly^4]	**22**	23.66	97	716.8
[Chg^3, Gly^4]	**23**	22.74	99	702.6
[Val^3, Gly^4]	**24**	21.91	99	662.8
[Leu^3, Gly^4]	**25**	21.35	97	677.0
[Ile^3, Gly^4]	**26**	21.92	96	677.0
[Ala^3, Gly^4]	**27**	19.42	98	634.5
[Abu^3, Gly^4]	**28**	19.76	97	649.0
[Nva^3, Gly^4]	**29**	20.98	96	662.8
[Nle^3, Gly^4]	**30**	21.15	95	677.0
[Tle^3, Gly^4]	**31**	20.62	95	677.0
[*t*-$BuAla^3$, Gly^4]	**32**	21.92	96	690.4

[a]HPLC retention time in min on a Vydac 218TP C-18 column (0.46 cm × 25 cm); gradient of 0–75% organic component in 25 min; flow rate of 1 mL/min. Solvent system was 0.1% TFA in water, 0.1% TFA in acetonitrile. Solvent front breakthrough at 3.5 min.
[b]Purity of final product peptide as assessed by RP-HPLC peak integration at 214 or 230 nm (whichever is lower).
[c]Molecular weight obtained by electrospray mass spectrometry.

The most important point to notice is that none of the QSAR variables by itself has a dominating effect on binding. In the X^3Asp^4 series, only three variables, all geometrical, even meet the $p = 0.05$ significance level for net linear correlation with log E_δ. They are the equivalent rectangle thickness Y, and the ovalities based on the solvent-accessible and van der Waals dimensions. Of these, solvent accessible ovality is the most strongly correlated, accounting for 18% (-0.425^2) of the variation of log E_δ. Our choice of log E_δ over E_δ as the independent variable is supported, as all but one of the correlations are stronger with log $E_\delta$than with E_δ. In the X^3Gly^4 data set, none of the variables meet the $p = 0.05$ significance level with log E_δ, whereas over half correlate significantly with log E_μ, although none accounting for more than 25% of the variability. The complete correlation matrices (not shown) confirmed that all of the variables are, as expected, highly intercorrelated, warning of multicollinearlity problems to be expected during multiple regression analysis. This is also evident in the similarity of correlations for the various geometrical variables within a series.

The pattern of predominantly positive correlations for the X^3Gly^4 series, and predominantly negative correlations for the X^3Asp^4 series (and the Phe^3X^4 series to be discussed later) is striking, but only a reflection of the different average side chain sizes for these series. The side chains used in the X^3Asp^4 series were generally larger (average van der Waals volume, 116 Å^3) than those in the X^3Gly^4 (average van der Waals volume, 99 Å^3), so the dominant effect of the X^3Asp^4 series was to degrade binding by being too large, whereas the opposite was true for the X^3Gly^4 series.

Multiple regression results for the X^3Asp^4 and X^3Gly^4 series

Multiple regression analysis was carried out as described in the introduction. Hundreds of combinations of variables were tested, focusing first on the most-promising, $N = 50$, series of X^3Asp^4 analogues. During this process it was observed that the combination of a linear and squared term in virtually any of the size-related variables returns positive and negative coefficients on the linear and squared terms, respectively. The simple explanation would be that there exists an optimum side chain size, with decreased binding for either smaller or larger values. However, polarizability and log P, both of which should only increase binding strength, also exhibit this apparent-maximum effect, and have slight negative bivariate correlations. This indicates that the optimum-value inferences may be artifacts of, or seriously compromised by, the multicollinearity of the variables. Therefore, it was decided to pursue these nonlinearity phenomena in terms of a model analogous to the potential well commonly used to describe intermolecular interactions.

Both log P and polarizability represent plausible attraction 'potentials'. Log P corresponds to the free energy of transfer from water to a hydrophobic environment, and polarizability contributes linearly to both dipole-induced dipole and London dispersion type van der Waals interaction energies. On the other hand, virtually any of the size-related variables are plausible candidates for a repulsion potential. However, while exploring the effectiveness of the combination of log P and polarizability together with the equivalent-sphere diameter, raised to various powers, it was discovered that the diameter to the fourth power gave distinctly better fits than either lower or higher powers. The search for a possible physical basis of this fact led to the following heuristic analysis, and our selection of the product of the van der Waals surface area and the square of the diameter of the equivalent sphere, AD^2, as our standard repulsion variable.

Viewing steric repulsion as a distortion of bond angles of entities comprising the binding site, and assuming the resistive force is harmonic, that force should be

Table 3. Calculated QSAR values[a] for isolated, H-capped, MM+ optimized side chains

Group	*X*	*Y*	*Z*	Normal	sa Vol	sa Area	v Vol	v Area	vA D	v Ovality	log *P*	Pol.	Refrc.
1-Nal	5.71	1.82	7.2	6.27	504	322	149	173	7.41	2.06	3.52	9.54	47.55
2-Nal	5.34	1.81	8.03	7.63	510	328	149	173	7.42	2.08	3.52	19.54	47.55
2-Pal	4.48	1.81	5.7	5.3	370	261	98	124	6.28	1.76	1.38	11.56	28.8
3-Pal	4.33	1.81	5.76	5.29	368	261	98	123	6.26	1.72	1.2	11.56	28.94
4-Pal	4.34	1.81	4.7	4.71	368	259	98	124	6.28	1.75	1.2	11.56	28.94
Abu	2.06	1.92	2.33	2.19	218	178	46	70	4.71	1.41	1.3	4.44	11
Ala	1.61	1.55	1.74	1.11	159	140	29	48	3.89	1.14	1.09	2.61	6.35
Arg	3.65	3.1	6.53	5.81	426	296	111	151	6.92	2.44	na	na	27.67
Asn	3.13	1.81	3.91	3.39	258	201	58	84	5.18	1.58	−0.86	5.88	14.59
Asp	2.35	1.81	2.6	2.49	238	189	52	77	4.94	1.49	na	4.71	9.31
Asp(OtBu)	4.37	3.66	6.28	5.66	451	308	122	161	7.16	2.47	0.7	12.51	31.22
Asp(OBzl)	4.85	1.84	8.88	8.24	526	349	146	181	7.59	2.45	1.64	16.66	42.03
Bhg	5.52	3.75	9.08	5.42	643	398	200	224	8.45	2.49	4.56	24.82	64.08
β-MePhe	4.33	2.22	7.06	5.44	426	288	117	146	6.83	2.02	2.91	14.1	35.7
β-OHPhe	4.53	2.6	5.95	5.43	401	274	109	135	6.56	1.85	1.51	12.91	32.87
Bip	4.36	1.81	10.2	9.78	590	379	177	204	8.05	2.39	4.2	21.93	56.24
Bpa	5.19	1.81	10.9	10.44	648	411	196	231	8.57	2.84	3.74	23.85	61.17
Bth	5.17	1.81	6.69	6.42	471	307	136	161	7.16	2.01	2.81	19.79	46.38
Cha	4.35	2.86	5.98	5.45	429	291	119	149	6.89	2.05	2.71	12.85	32.15
Chg	4.68	2.7	4.96	3.97	380	265	103	129	6.41	1.81	2.38	11.01	27.61
Cpa	4.3	2.75	4.93	4.61	388	271	103	133	6.5	1.93	2.31	11.01	27.55
Cys	2.31	1.82	2.43	2.4	220	181	47	70	4.73	1.37	0.32	5.61	14.39
Dip	4.35	3.83	8.83	5.43	595	378	178	209	8.16	2.56	4.13	21.93	55.8
F_5Phe	6.18	1.81	5.12	5.65	425	290	115	145	6.79	2.03	3.21	11.81	32.18
Gln	2.91	1.82	5.5	4.61	313	233	75	105	5.77	1.8	−0.23	7.72	19.21
Glu	2.47	2.33	4.01	3.18	293	221	69	98	5.58	1.75	na	6.55	13.94
Gly	1[b]	1[b]	1[b]	1[b]	85	90	10	22	2.63	1.02	−0.65	0.77	1.6
Hfe	4.76	1.81	7.35	6.31	435	293	120	149	6.89	2.02	2.91	14.1	35.7
His	4.2	1.81	4.74	4.57	327	239	82	109	5.88	1.68	−0.23	9.56	24.16
o-I_2Tyr	6.7	1.82	6.15	6.27	556	351	168	195	7.88	2.3	4.74	22.96	57.61
o-ITyr	5.6	1.82	6.08	6.24	483	315	139	168	7.31	2.15	3.49	17.93	45.2
Ile	3.1	2.85	5.2	4.17	323	237	80	111	5.95	1.89	2.09	8.11	20.21
Leu	4.36	2.71	3.97	3.51	322	236	80	110	5.92	1.84	2.02	8.11	20.15
Lys	3.18	2.96	5.26	4.45	363	260	94	128	6.38	2.09	na	na	21.97
Me_5Phe	7.12	1.82	6.69	6.44	606	371	186	217	8.32	2.61	4.85	21.44	56.31
Met	3.19	2.95	5.33	4.3	327	239	82	112	5.97	1.84	0.8	9.28	23.93
Met(O)	3.48	2.71	5.09	4.28	342	246	88	119	6.15	1.93	0.45	na	25.38
m-CF_3Phe	5.5	2.24	6.16	5.94	465	309	128	160	7.14	2.21	3.4	13.83	37.07
m-FPhe	5.06	1.81	5.49	5.42	394	271	106	132	6.48	1.8	2.65	12.18	31.32
Nle	3.1	2.85	5.2	4.17	323	237	80	111	5.95	1.89	2.09	8.11	20.21
Nva	2.18	1.8	4.37	3.51	273	210	63	90	5.34	1.6	1.69	6.28	15.61
o-FPhe	4.98	1.81	5.58	5.43	393	270	106	132	6.47	1.8	2.65	12.18	31.32
Orn	3.11	2.87	4.97	3.94	310	231	76	105	5.78	1.78	na	na	17.37
p-BrPhe	4.35	1.81	6.64	6.21	449	300	125	153	6.97	2.01	3.31	14.9	38.72
p-ClPhe	4.34	1.81	6.48	6.04	430	290	118	146	6.81	1.96	3.03	14.2	35.9
Pen	4.38	2.8	3.98	3.52	326	239	81	111	5.95	1.84	1.08	9.28	23.55
p-FPhe	4.33	1.81	6.08	5.65	394	272	106	132	6.48	1.81	2.65	12.18	31.32
Pgl	4.33	1[b]	5	3.9	332	238	87	108	5.85	1.46	2.05	10.43	26.06
Phe	4.34	1.81	5.85	5.42	386	266	104	129	6.4	1.75	2.51	12.27	31.1
p-IPhe	4.35	1.81	6.82	6.39	468	310	133	160	7.14	2.05	3.77	17.3	43.51
p-MePhe	4.32	1.81	6.72	6.37	439	292	121	149	6.89	2.02	2.98	14.1	36.14
p-NH_2Phe	4.34	1.81	6.52	6.16	419	283	114	143	6.74	1.98	1.73	13.62	35.8
p-NO_2Phe	4.36	1.81	6.78	6.44	445	299	123	154	7	2.14	2.47	14.11	38.42
Ser	1.84	1.84	2.13	1.91	183	157	35	57	4.25	1.29	−0.27	3.25	8.26
Taz	4.37	1.81	4.48	4.62	348	251	90	116	6.06	1.67	0.04	11.08	29.97
t-BuAla	3.52	3.16	3.53	3.5	367	255	97	130	6.42	2.06	2.46	9.95	24.63
Thi	4.2	1.81	4.6	4.63	344	245	89	114	6.01	1.62	1.2	11.79	31.2
Thr	2.26	1.94	3.33	2.6	240	190	53	77	4.95	1.45	0.08	5.08	13.01
Thy	4.34	3.52	11.1	19.27	648	417	192	231	8.58	2.96	3.66	23.2	59.03
Tle	3.81	2.71	4.37	3.51	322	236	80	110	5.92	1.84	2.02	8.11	20.15
Trp	5.28	1.81	6.89	6.44	465	307	132	161	7.15	2.11	2.29	17.54	42.19
Tyr	4.35	1.81	6.61	6.24	407	277	110	138	6.63	1.92	2.23	12.91	32.79
Tyr(OMe)	4.58	1.86	8.04	7.6	462	309	128	159	7.12	2.19	2.26	14.74	37.56
Val	2.18	1.8	4.37	3.51	272	211	63	90	5.34	1.6	1.69	6.28	15.61

[a]References to calculation algorithms used are given in the text. Abbreviations: sa, solvent-accessible; Vol., volume; vA D, diameter of equivalent sphere based on van der Waals surface area; v, van der Waals; log *P*, octanol/water partition coefficient; Pol., polarizability; refrc., refractivity; na, not available—incalculable due to missing parameters. All distance-related units are in terms of Å raised to the appropriate power; polarizability and refractivity are in units of $Å^3$.

[b]Estimated values of appropriate magnitude.

Table 4. Charge and hydrogen bonding related properties for the Phe^3X^4 series of analogues

Group	Charge	Absolute value of charge	H-acceptor number	H-donor number	Total H-bond capacity	Charge-weighted H-bond capacity
Ala	0	0	0	0	0	0
Arg	1	1	0	2	2	4
Asn	0	0	1	1	2	2
Asp	−1	1	2	0	2	4
Asp(OBzl)	0	0	1	0	1	1
Asp(O*t*Bu)	0	0	1	0	1	1
Cys	0	0	1	0	1	1
Gln	0	0	1	1	2	2
Glu	−1	1	2	0	2	4
Gly	0	0	0	0	0	0
His	0	0	1	0	1	1
Leu	0	0	0	0	0	0
Lys	1	1	0	2	2	4
Orn	1	1	0	2	2	4
Phe	0	0	0	0	0	0
Ser	0	0	1	1	2	2
Thr	0	0	1	1	2	2
Val	0	0	0	0	0	0

proportional to the displacement (i.e., roughly to the diameter of the intruder). The energy cost of displacement against such a force should be proportional to the square of the displacement, that is to the diameter (or some linear dimension) squared of the intruder. Also, the repulsion energy should be proportional to the total surface area, or number of contacts, over which these displacements occur (i.e., proportional to the surface area of the intruder). Thus, the total steric repulsion energy should be proportional to the product of the area and the diameter squared. The van der Waals area and the area-based diameter performed slightly better than the solvent-accessible or the volume-based quantities. Therefore, we settled on the product of the van der Waals surface area and the square of the diameter of the equivalent sphere, AD^2, as our general steric repulsion variable. The credibility of the resulting potential well model is strongly supported by the fact that it results in the best three-variable regression equation of any of the numerous alternative combinations tested with our largest data set ($N = 50$).

Table 7 presents the results of multiple regression analyses based on the potential well model for δ binding

Table 5. Bivariate correlations between binding enhancement factors and the primary QSAR properties

QSAR Property	X^3Asp^4			X^3Gly^4		Phe^3X^4	
	E_δ	log E_δ	log E_μ	log E_δ	log E_μ	log E_δ	log E_μ
log E_δ	0.727	1	0.678	1	0.906	1	0.486
X	0.053	0.002	0.171	0.347	0.519	−0.240	−0.389
Y	−0.148	−0.322	−0.228	0.021	0.054	−0.446	−0.712
Z	−0.055	−0.175	0.007	0.204	0.352	−0.457	−0.551
Normal	−0.065	−0.263	−0.036	0.214	0.368	−0.421	−0.487
sa Volume	−0.082	−0.240	−0.013	0.295	0.462	−0.416	−0.571
saV Diameter	−0.057	−0.196	0.025	0.314	0.483	−0.354	−0.600
sa Area	−0.096	−0.261	−0.031	0.290	0.455	−0.407	−0.593
sa A Diameter	−0.082	−0.235	−0.008	0.300	0.467	−0.370	−0.607
sa Ovality	−0.215	−0.425	−0.201	0.146	0.274	−0.446	−0.648
vdW Volume	−0.076	−0.231	−0.008	0.307	0.477	−0.420	−0.545
vV Diameter	−0.043	−0.174	0.041	0.330	0.502	−0.339	−0.592
vdW Area	−0.097	−0.260	−0.022	0.276	0.445	−0.433	−0.592
vA Diameter	−0.078	−0.225	0.009	0.294	0.464	−0.380	−0.615
vdW Ovality	−0.194	−0.382	−0.100	0.110	0.252	−0.514	−0.693
β Crowding	−0.136	−0.142	−0.222	−0.105	−0.053	−0.181	−0.521
X/Y Flatness	0.108	0.203	0.227	0.256	0.350	0.053	0.073
Molar mass	−0.074	−0.192	−0.033	0.235	0.369	−0.341	−0.554
Refractivity	−0.014	−0.17	0.030	0.304	0.493	−0.346	−0.395
log P	−0.012	−0.076	0.069	0.325	0.331	na	na
Polarizability	−0.006	−0.136	0.065	0.311	0.510	na	na

Table 6. Bivariate correlations between binding enhancement factors, and charge and hydrogen bonding related properties of the Phe^3X^4 series of analogues. (N = 25)

Variable	Correlation Coefficient log E_μ	log E_δ
Charge	−0.252	−0.659
Absolute value of charge	−0.679	−0.415
Number of H-bond acceptor sites	−0.256	0.276
Number of H-bond donor sites	−0.513	−0.402
Total H-bond capacity	−0.760	−0.268
Charge-weighted H-bond capacity	−0.822	−0.388

by the X^3Asp^4 series, and δ and μ binding by the X^3 Gly^4 series. As must be the case with such highly scattered data, there is much unexplained variability; R^2 ranges from 25% to 45%. Nonetheless, all of the coefficients meet the p = 0.05 significance test, leaving little doubt about the existence of their corresponding regression relationships. Also, as expected, the coefficients are positive on the attraction variables, log P and polarizability, and negative on the repulsion variable. However, the most important observation is that the three sets of coefficients are statistically equivalent. That implies that, when starting from base compounds having comparable δ and μ binding affinities, changes in third position side chain hydrophobicity, van der Waals attractiveness, and overall steric repulsiveness have the same effects on δ and μ binding affinity. The simplest interpretation is that the local nature of the binding site for the X^3 side chain is identical on the δ and μ receptors.

The relative importance of the two attraction variables can be estimated in two ways. First, comparison of the beta weights (Table 7) shows that the log P variable is responsible for approximately one-third as much of the observed variation in binding compared to the polarizability variable. Perhaps more directly, taking the regression equation at face value permits calculation of each variable's thermodynamic contribution to the binding interaction, as follows. Changes in binding free energy are equal to changes in log E_δ multiplied by 2.30 RT, based on $\Delta G^\circ = RT \ln K$, and recalling that E_δ is a ratio of equilibrium constants. Thus, the maximum contribution of a variable to the observed range of log E_δ is calculated as the product of the regression coefficient multiplied by the total range of the variable and by −2.30 RT. Using the X^3Asp^4 data set, the following free energy contributions are calculated: log P, −9.85 kJ/mol; polarizability, −31.4 kJ/mol and steric repulsion (AD^2), 46.9 kJ/mol. Comparison of the log P and polarizability values indicates that the source of the X^3 side chains binding free energy is about 25% hydrophobic and 75% van der Waals, consistent with the beta weight analysis. For further perspective, these calculated total free energy changes over the entire analogue data set can be compared with, for example, the 5.7 kJ/mol corresponding to one power of ten change in K_i, or the 10–20 kJ/mol for a typical hydrogen bond.

The non-zero intercept also is informative. Attempts to force the regression equation through the 0,0 origin failed completely (R = 0.0), consistent with the low p-values on the intercept. The fact of a non-zero intercept with the variables centered on the native compound proves that some essential property or properties of the X^3 side chain is, or are, not being captured by the three variables used, resulting in systematic error in addition to the sources of scatter in the data. Specifically, the intercept of −1 (Table 7) says that a binding penalty averaging about one power of ten in K_i, or 5.7 kJ/mol in free energy, is to be expected for any change to the X^3 side chain, even if hydrophobicity, polarizability and overall steric repulsion are somehow held constant at the values of the native parent compound. We assume

Table 7. Multiple regression data for log E_δ or log E_μ dependence on the three variables Δ log P, Δ polarizability, and Δ repulsion, with all Δ values referenced to the native Phe^3 side chain[a]

		Overall fit statistics						Coefficients and related statistics			
Series	Receptor	N	R^2	Adj R^2	SE[b]	F	Sig F	Intercept	Δ log P	Δ Pol	Δ AD^2
X^3Asp^4	δ	50	0.45	0.42	0.65	13	4×10^{-6}	−1.01	0.34	0.27	-5.3×10^{-4}
								SE, 0.12	SE, 0.14	SE, 6×10^{-3}	SE, 9×10^{-5}
								p, 2×10^{-10}	p, 0.016	p, 3×10^{-5}	p, 3×10^{-7}
									β, 0.45	β, 1.5	β, −2.2
X^3Gly^4	δ	33	0.25	0.17	0.75	3.2	0.04	−1.04	0.51	0.27	-5.0×10^{-4}
								SE, 0.22	SE, 0.22	SE, 0.12	SE, 2×10^{-4}
								p, 5×10^{-5}	p, 0.031	p, 0.025	p, 0.045
									β, 0.53	β, 1.2	β, -5×10^{-4}
X^3Gly^4	μ	33	0.39	0.33	0.73	6.3	0.002	−0.89	0.44	0.37	-5.7×10^{-4}
								SE, 0.21	SE, 0.22	SE, 0.11	SE, 2×10^{-4}
								p, 2×10^{-4}	p, 0.051	p, 2.6×10^{-3}	p, 0.021
									β, 0.42	β, 1.5	β, -6×10^{-4}

[a]Abbreviations used: Pol, polarizability; SE, standard error; p, probability that coefficient is zero; β, beta weight; AD^2, repulsion variable calculated as product of van der Waals surface area and square of equivalent sphere diameter; Sig F, probability related to the significance level of the F statistic for the overall fit.

[b]Standard error of y-values estimated by the regression equation.

this reflects the role of very specific, local interactions. Nonetheless, whatever missing variable(s) the intercept represents, as with the other coefficients, there is no apparent difference for δ versus μ binding.

Additional support for the hypothesis that the X^3 side chain binding sites are very similar in δ and μ receptors is presented in Figure 1. A plot of log E_μ versus log E_δ for the X^3Gly^4 analogues has a slope of 0.98, intercept of 0.056, and correlation coefficient of 0.91, all within experimental error of the values 1, 0, and 1, respectively, expected if X^3 side chain modifications have identical effects on δ and μ binding. The coefficient of determination, r^2, corresponds to 82% of the variation observed in μ binding being attributable to variation in δ binding. Although the zero intercept of Figure 1 suggests that there is no major systematic difference between the δ and α binding pockets for the X^3 side chain, the possibility of significant minor differences with respect to the other QSAR variables was investigated by multiple regression analysis.

We sought another independent variable which, when combined with log E_δ (now being treated as an independent variable), produced a multiple R^2 against log E_μ substantially greater than the 82% bivariate R^2 between log E_μ and log E_δ alone. There is no rule, but a 5% increase in R^2 is often used as a minimum to justify addition of a new variable to a regression equation. In the present case, inclusion of almost any of the size-related variables results in a few percent increase in R^2. Their coefficients all are positive, implying that μ binding favors a somewhat larger X^3 side chain. Polarizability is, by a small margin, the most effective new variable, just meeting the 5% improvement

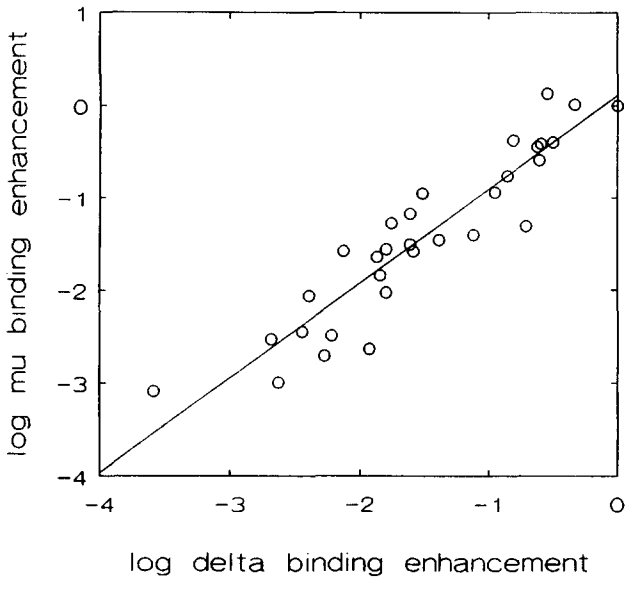

Figure 1. Plot of log E_μ versus log E_δ for the X^3Gly^4 series of deltorphin I/II analogues. The line is the least-squares best fit to the data points shown as open circles: *N*, 33; slope, 0.978; intercept, 0.056; *r*, 0.906.

criterion (R^2 = 0.879, K_{pol} = 0.0635±0.017 SE, p = 0.0007). This is consistent with a van der Waals type interaction between the side chain and the binding pocket driving the size preference. Inclusion of the log *P* variable is unproductive (p = 0.29), implying no significant difference in hydrophobicity between the δ and μ pockets.

Details about the apparent size difference between the δ and μ binding pockets for the X^3 side chain were sought by testing for binding maxima as a function of *X*, *Y*, *Z*, and normal dimensional variables. Log E_δ and log E_μ were regressed on the linear and square term pair of variables for each of the dimensions. The combination of a positive coefficient on the linear term and a negative coefficient on the square term indicates the existence of a maximum value of the binding enhancement. The corresponding optimum value of the variable can be extracted by setting equal to zero the derivative of log E_δ or log E_μ with respect to the variable. Neither *X* nor *Y* variables returned significant regression coefficients under this analysis, suggesting that binding was not systematically dependent on these variables at least not over their range in the present data set. However, some significant coefficients were returned for the *Z* and Normal variables, presented in Table 8. Two points are noteworthy; first, the apparent optima are, not surprisingly, close to the values of the native Phe^3 side chain: *Z*, 5.85Å; Normal, 5.42Å. All of the optima derived from $p < 0.05$ data differ from the native by less than 1 Å. Second, the anticipated pattern of μ site preference for larger side chains is observed—albeit for less than a 1 Å difference. These analyses were performed also in combination with the attraction-repulsion variables used previously (e.g., Table 7), from which significant coefficients were returned only for the *Z* variable. However, instability of the attraction-repulsion coefficients (due to their multicollinearity with the *Z* variable) suggested that Z/Z^2 coefficients from such combinations are also unreliable. Nonetheless, in all cases, the apparent optimum *Z* value for the μ site is larger than for the δ site. This apparent preference of the μ site for a longer X^3 side chain is consistent with its slightly larger Δ polarizability regression coefficient, seen in Table 7.

The fact that only the length of the X^3 side chain appears to have a distinct optimum value, whether for δ or μ binding, invites further speculation about the nature of the binding site. In particular, the unimportance of the *X* and *Y* dimensions argues against a simple binding pocket (with defined walls) perpendicular to the ligand backbone. The only model we can imagine that is compatible with these features, is a binding site similar to a step notched out of the receptor surface, with the Phe ring laying on and directed toward the back (riser) of the step. The distance to the back of the step would define the optimum side chain length (*Z*) beyond which a high steric price must be paid, perhaps due to holding the ligand's other pharmacophore elements out of their proper binding ranges. By contrast, excess side chain width (*X*) might merely hang

off the edges of the step, and excess side chain thickness (Y) merely pile deeper on the step, both apparently with less steric consequence. Further, based on the positive regression coefficient on log P, it follows that the bound ring is probably sandwiched between the receptor and the ligand, rather than exposed to water.

Bivariate correlation results for the Phe^{3}X^4 series

The role of the Asp^4/Glu^4 side chain with respect to δ selectivity has been the subject of much speculation. Early suggestions that the negative carboxylate is responsible for weak μ binding by deltorphins I and II due to direct repulsion by an alleged negative group on the μ receptor, or due to intramolecular electrostatic or hydrogen bonding induced stabilization of a conformation incompatible with μ binding, were confounded by similarly weak μ binding by the positive charged Lys^4, Orn^4, and Arg^4 analogues. In order to explore such interactions, a variety of charge and hydrogen bonding variables were tested. Unfortunately, insufficient parameterization prevented calculation of log P and polarizability for several of the most important (charged) analogues, so those variables are excluded from the analysis of the Phe^{3}X^4 series. However, refractivity correlates very highly with polarizability (r = 0.995), and so provides an alternative measure of van der Waals force susceptibility.

The Phe^{3}X^4 series bivariate correlations of log E_δ or log E_μ with the primary QSAR variables are shown in Table 5, and with the charge and hydrogen bonding related variables in Table 6. The critical value for p = 0.05 confidence in the existence of a correlation in these data (N = 25) is 0.40. The first point to notice is, in contrast to the X^{3}Gly^4 series, there is only weak correlation between δ and μ binding behaviors of the X^4 side chain analogues (Table 5). This is expected, in as much as the μ binding affinity is essentially maximized by removing the side chain (Gly^4). The second point to notice is the generally negative correlation—stronger for μ than δ between side chain size and binding affinity (Table 5). This suggests that most of the analogue side chains are too large, and that this oversize is more of a problem for μ binding. However, the most telling correlations are in Table 6, where μ binding is seen to correlate most strongly with the charge-weighted hydrogen bonding capacity (r = −0.82), followed by total hydrogen bonding capacity (r = −0.76) and the absolute value of the charge (r = −0.68), and not at all with signed charge. This behavior suggests an interaction with the amphiprotic solvent water, rather than with a receptor site (of presumably defined charge and hydrogen bonding preference). In contrast, δ binding correlates strongly only with signed charge (r = −0.66), suggesting an interaction with a positive-charged receptor site. These data provided obvious lead variables for further, multiple regression analysis.

Multiple regression results for μ binding by the Phe^{3}X^4 series

The role of the X^4 side chain in μ binding was explored by multiple regression analysis to determine what variables might supplement the substantial μ binding predictive power (r^2 = 68%) of the charge-weighted hydrogen bonding capacity. Almost any of the size-related variables correlate (negatively) at the p = 0.05 level, and increase R^2 by a few percent over r^2. Of these, the van der Waals based ovality is best (R^2 = 0.826, k_{vOv} = −0.92±0.21 SD, p = 0.00024), followed closely by the steric repulsion variable AD^2 (R^2 = 0.813, $k_{AD}{}^2$ = 0.00013±3 × 10^{-5}, p = 0.00055). Regression of the Δ form of either of these variables (calculated with respect

Table 8. Data related to apparent optimum values of the Z and Normal variables

Series	Receptor	R^2	Coefficients, probabilities, apparent optimum values[a]				
			Z				
			k_Z	p_Z	$k_Z{}^2$	$p_Z{}^2$	Optimum value (Å)
X^{3}Gly^4	μ	0.24	1.36	0.017	−0.114	0.040	5.99
X^{3}Gly^4	δ	0.14	(1.066)	(0.054)	(−0.095)	(0.081)	(5.63)
X^{3}Asp^4	μ	0.09	0.541	0.038	−0.0408	0.036	6.64
X^{3}Asp^4	δ	0.15	0.791	0.031	−0.0667	0.015	5.93
			Normal				
			k_{Normal}	p_{Normal}	$k_{Normal}{}^2$	$p_{Normal}{}^2$	Optimum value (Å)
X^{3}Gly^4	μ	0.30	1.47	0.0046	−0.137	0.013	5.38
X^{3}Gly^4	δ	0.18	1.17	0.021	−0.116	0.034	5.04
X^{3}Asp^4	μ	0.06	(0.178)	(0.144)	(−0.0092)	(0.108)	(9.73)
X^{3}Asp^4	δ	0.12	(0.159)	(0.348)	(−0.0124)	(0.119)	(6.41)

[a]Data corresponding to $p > 0.05$ are shown in parentheses.

to the native Asp^4) together with the Δ form of the charge-weighted hydrogen bonding capacity yields a regression intercept that is not significantly different than zero (ovality k_0, $p = 0.91$; AD^2 k_0, $p = 0.64$), indicating that such a combination can account for nearly all of the change in μ binding related to the X^4 side chain.

The simplest interpretation of these results is that the role of the X^4 side chain, with respect to μ binding, is one of simply being in the way. Both the innate van der Waals volume and the bulk of a hydration layer are effective in blocking μ binding. Possible mechanisms are (i) a direct steric clash of the X^4 side chain with the μ receptor and (ii) an unfavorable ligand conformation driven by nonspecific steric repulsion from the X^4 side chain. Arguing against the latter is the relative insignificance of variables that would seem most likely to induce such backbone strain (e.g., *X*, *Y*, and β-crowding). Similarly, the fact that there is no evidence of a charge-specific or a hydrogen bond polarity-specific interaction argues against conformational stabilization by some specific interaction of the X^4 side chain. Further, the inability to justify a positive, attraction-type variable in the regression equation argues against even the existence of a μ receptor binding site per se for the X^4 side chain; regression of log E_μ onto refractivity and charge-weighted hydrogen bonding capacity, either alone or also with the AD^2 repulsion variable, returns values of K_{ref} that are insignificant or negative, respectively. Thus, that region of the μ receptor closest to the X^4 side chain appears to be more akin to an anti-binding site.

Multiple regression results for δ binding by the Phe^3X^4 series

The role of the X^4 side chain in δ binding was explored by seeking variables that would supplement the predictive power of the side chain charge variable. Because neither polarizability nor log *P* values were available for all of the analogues, the potential well model was approximated by using refractivity in place of polarizability (and disregarding log *P*). Good results were obtained ($R^2 = 0.63$) with the three variables charge ($k_{chg} = -0.94$, $p = 0.001$), refractivity ($k_{ref} = 0.084$, $p = 0.028$), and AD^2 steric repulsion ($k_{AD2} = -0.00044$, $p = 0.0067$). This suggests that the binding interaction is due to a combination of van der Waals attraction and electrical attraction to an apparently positive-charged receptor site. The rather surprising absence of a correlation with the side chains' hydrogen bond acceptor ability, argues against a fully-charged protonated receptor moiety (e.g., Lys) and in favor of an only partially-positive polar moiety. The high δ affinity of the natural His^4 ligand (dermenkephalin), which is predicted to have approximately +0.1 charge at neutral pH, also argues against a fully-charged receptor site. A plausible receptor group would appear to be another His, which could enjoy interacting both in its charged state with a ligand's Glu or Asp, and in its uncharged state with a ligand's His, either by π bonding if both His residues are uncharged or by hydrogen bonding if one of them is charged.

The dimensional constraints of the δ binding site for the X^4 side chain were probed by regression of log E_δ onto linear and square terms of each of the dimensional variables, *X*, *Y*, *Z*, and Normal. No relationship could be found between δ binding and either the X/X^2 or Y/Y^2 variable pairs; their coefficient *p*-values exceeded 0.6, whether the pair was used alone or in combination with the charge or the charge/refractivity variables (all strategies that invite a variable to reveal its repulsion role). By contrast, under the same treatment the square terms of the *Z* and Normal variables returned significant negative coefficients ($p \leq 0.05$), consistent with a steric repulsion role. However, we failed to find the positive-linear and negative-square coefficient pattern characteristic of the optimum-value scenario. We suspect that failure may be due to an inadequate data set, as only four of the eighteen X^4 side chains studied are smaller than the native Asp^4. In other words, because almost all of the side chains are too large (in the *Z* and Normal dimensions), only their steric repulsion effect is observable. Indeed, either Z^2 or $Normal^2$ are more effective steric repulsion variables, returning higher R^2 and lower *p*-values, than AD^2 when combined with charge and refractivity. The importance of the Normal variable, implying that perpendicular distance out from the ligand backbone matters, and of the refractivity variable, implying that van der Waals bonding (to something) occurs, are the two pieces of evidence for the existence of an X^4 side chain binding pocket on the δ receptor. (However, the possibility of electrostatic intramolecular stabilization of a δ selective ligand conformation cannot be excluded.) The apparent van der Waals attraction to the δ receptor contrasts with the μ receptor, for which the refractivity variable would adopt only a repulsion role, indicated by a negative *k* coefficient. However, the unimportance of the *X* and *Y* variables, and the fact that the equivalent sphere-based AD^2 variable is less important than *Z*, also imply that the lateral dimensions of a X^4 δ receptor pocket must be very loose. Some degree of openness to the aqueous environment is consistent with nature's use of the hydrophilic Asp, Glu and His (dermenkephalin) residues at that position.

Conclusions

The following picture is consistent with the QSAR analysis based on a potential well ligand-receptor binding model.

The binding site for the X^3 side chain is very similar in the δ and μ receptors, although there is evidence that the μ receptor pocket is somewhat larger. The binding pocket is highly sterically constrained in the longest dimension of the X^3 side chain, but not in the other dimensions; X^3 side chains longer than the native Phe quickly degrade the binding. The δ or μ receptor-bound

Phe[3] ring is visualized to be laying on a step notched in the receptor surface. Larger side chains of X^3 analogues are blocked by the back (riser) of the step, but are able to hang over the edges. The binding interaction appears to be primarily van der Waals in nature, and secondarily hydrophobic, which, together with the geometrical constraints, suggests that the ring is sandwiched between the ligand and receptor. Finally, there is at least one piece of the puzzle yet missing; some unknown interaction not captured by the gross QSAR variables used in this study is responsible for the final power-often in the high affinity binding constant of the native structure.

The δ and μ receptor binding sites for the X^4 side chain are quite different. The δ receptor site appears to be a pocket large enough to accomodate an Asp or Glu side chain, and probably contains at least a partial positive charge. Both electrostatic and van der Waals binding appear to be supported, but not by necessarily hydrogen bonding. The depth of the pocket is much more constrained than are its other dimensions. By contrast, the X^4 side chain binding site that is a pocket on the δ receptor appears to be filled-in on the μ receptor. In general, X^4 side chains larger than Gly degrade the binding. Both the van der Waals size and the bulk associated with a hydration layer are effective—the larger or more hydrated the X^4 group, the more severe the steric problem. The mechanism whereby the steric interference from the X^4 side chain weakens binding to the μ receptor might involve simple displacement of the Phe[3] side chain so that it is out of reach of its binding pocket. This could be due either to direct ligand-receptor repulsion, or, less likely, to an unfavorable ligand conformation induced by the X^4 side chain bulk, but not by a specific charge-charge or hydrogen bond interaction.

Experimental

Peptide synthesis

Most of the protected amino acids, coupling agents and resins were purchased from Bachem (including novel amino acids) unless otherwise noted. Boc-3-(4-thiazolyl)alanine, Boc-pentafluorophenylalanine, Boc-*meta*-trifluoromethylphenylalanine, Boc-3-(2-pyridyl)alanine, Boc-3-(3-pyridyl)alanine and Boc-3-(4-pyridyl)alanine were purchased from Synthetech, Inc. Solvents and deprotecting agents were obtained from Fisher Scientific and Aldrich Chemical Co. Radioligands were purchased from New England Nuclear, Multiple Peptide Systems and Amersham, and frozen guinea pig brains were obtained from Rockland, Inc. The peptides were prepared on a St. John's Associates manual shaker using standard solid phase techniques for *N*-α-*t*-butyloxycarbonyl (Boc) protected amino acids on *p*-methylbenzhydrylamine (MBHA) resin (0.6 to 1.1 mmol/g). The side chain of Tyr was protected as the 2,6-dichlorocarbobenzyloxy derivative. The deprotection solution for the *N*-terminal amine was 30% trifluoroacetic acid (TFA) in dichloromethane (DCM). Dicyclohexylcarbodiimide (DCC) or diisopropylcarbodiimide (DIC) and hydroxybenzotriazole (HOBt) were used as coupling agents. The protocol for peptide synthesis in each cycle was as follows: (1) addition of Boc-amino acid in DCM (3 equiv), (2) addition of HOBt (2.4 equiv), (3) addition of DCC or DIC (2.4 equiv), (4) mixing/shaking for 4 h, (5) washing with DCM (3 × 2 min), (6) checking for completion of reaction with bromophenol blue test and recoupling if necessary, (7) Boc deprotection with 30% TFA in DCM (30 min), (8) washing with DCM (3 × 2 min), (9) neutralization with diisopropylethylamine (DIEA) in DCM (10 min), (10) washing with DCM (3 × 2 min). Simultaneous deprotection and cleavage from the resin were accomplished by treatment with 90% anhydrous HF (Immunodynamics apparatus) and 10% anisole scavenger (10 mL of HF and 1 mL of anisole per gram of resin) at 0 °C for 1 h. After evaporation of the HF, the peptide resin was washed with diethyl ether and the peptide was extracted with 70% acetonitrile/30% water (with 0.1% TFA), concentrated under reduced pressure, diluted with water, and lyophilized. Crude peptides were purified to homogeneity by preparative reversed-phase high performance liquid chromatography (RP-HPLC) on a Waters instrument with a Vydac C18 column (2.2 × 25.0 cm, 10 ml/min). A linear gradient of water (0.1% TFA) to 50% acetonitrile (0.1% TFA)/water (0.1% TFA) was employed, followed by lyophilization.

Peptide analysis

Peptide purity was assessed by analytical RP-HPLC. Peaks were monitored at 214, 230, and 280 nm. All compounds were at least 95% pure as analysed by peak integration. Proton nuclear magnetic resonance (^{1}H NMR) spectra were obtained on a Bruker spectrometer at 250 MHz. Samples (approx. 1 mg) were dissolved in DMSO. Diagnostic resonances and peak patterns confirmed the presence of all indicated residues. Electrospray mass spectroscopy confirmed the appropriate molecular weights.

Opioid receptor binding assays

Using a published protocol,[22] receptor binding assays measured displacement by the test compounds of radiolabelled receptor-selective ligands from guinea pig brain homogenates, using 1.2 nM [^{3}H]DAMGO for the μ receptor, 2.5 nM [^{3}H]DPDPE for the δ receptor and 1.0 nM [^{3}H]U69,593 for the κ receptor. IC_{50} values of the test compounds against the labelled receptor-specific compounds were obtained by linear regression from plots relating inhibition of specific binding to the log of 12 different ligand concentrations, using the LIGAND computer software program (Biosoft Software)[23]. For binding to κ receptors, which was expected to be weak, the protocol was altered to include only five ligand concentrations and was performed in duplicate. K_i values for each test compound were

calculated using values for K_D of each radiolabelled marker ligand. Saturation binding experiments determined the K_D range of each ligand as follows: [^{3}H]DAMGO = 1.31–3.60 nM; [^{3}H]DPDPE = 1.60–1.72 nM; [^{3}H]U69,593 = 1.13–1.40 nM. K_i values reported represent the mean of 2–4 determinations, each performed in triplicate

Calculation of QSAR properties

The structures of the analogue side chains were capped with an H atom at their normal attachment site to the α-carbon, and energy-optimized by the HyperChem ver. 4 (HyperCube, Inc.) MM+ implementation of Allinger's MM2 force field.[24] We assumed that errors flowing from use of the H-capped side chains would be more nearly constant, and possibly smaller, than would the errors resulting from attempting to calculate properties of attached side chains. The MM+ force field was used in preference to AMBER, OPLS, or BIO+ because the latter, even though customized for aqueous biomolecules, are insufficiently parameterized for the many unusual atom types in the side chain analogues.

Algorithms for calculation of three non-geometrical properties were available: The log of the octanol–water partition coefficient log P, and the refractivity were calculated using HyperChem's implementation of the method of Viswanadhan et al.[25] The log P algorithm is not appropriate for the ionic side chains. Polarizability was calculated by HyperChem's implementation of the method of Miller,[26] although insufficient parameterization prevented calculation for a few side chains.

Surface area and volume were calculated using HyperChem's implementation of the grid method of Bodor et al.[27] based on the atomic radii of Gavezotti,[28] and in terms of both the van der Waals dimensions and the solvent-accessible dimensions assuming a solvent probe radius of 1.4 Å (for water). The solvent-accessible values correspond to the surface painted by the nucleus of the solvent probe rolling over the van der Waals structure, and so are substantially larger than the van der Waals properties. Based on these values, four different equivalent-sphere diameters were calculated.

The equivalent-box (rectangular solid) dimensions length Z, width X, and thickness Y (the shortest dimension) were determined using HyperChem's periodic box facility, which fits a structure to the smallest possible rectangular box. The Normal, defined as the maximum distance the side chain can extend perpendicular to the backbone, was measured as the internuclear distance from the β-carbon to the most distant atom of the side chain in the MM+ optimized structure. Depending on a side chain's point of attachment, its length might differ considerably from its maximum extension from the backbone.

The flatness was calculated in two ways based on the equivalent rectangular dimensions ($Z \geq X \geq Y$): (i) the quantity, $XZ/Y^2 - YZ/X^2$, which equals zero for a cube and increases with both X and Z, and (ii) the ratio X/Y, which equals unity for a cube but is independent of length.

The ovality represents another attempt, like flatness, to isolate the effect of shape from size. Bodor et al.[27] applied the term ovality to the ratio of a molecule's surface area to the surface area of a sphere having the same volume. An equivalent view is to recognize that the ratio $(\text{surface area})^3/(36\ \pi\ (\text{volume})^2)$ equals unity for spheres of any size, and, similarly, equals some constant value greater than unity for ellipsoidsal solids of defined eccentricity but any size. By that definition, ovality increases from unity for either prolate (stretching) or oblate (flattening) distensions from spherical. For semi-axis ratios from 1.0 to about 2.0, the ovalities of prolates and oblates differ by only a few percent, while at higher ratios, oblate ovalities increase much more rapidly. For the structures used in this study, ovality correlates highly with surface area, volume and length, and so, in spite of our intentions, it behaves as a length-weighted measure of size. Ovality values were calculated based both on the van der Waals and on the solvent-accessible dimensions.

A β-crowding variable was defined to test for possible particular importance of steric crowding close to the backbone. The following simple scale was used: the sum of 1 for each non-H group attached to the β carbon, plus 1/2 for each non-H group attached to each sp^3 γ carbon. For example, Ala, 0; Asn, Cys, and Phe, 1; Lys and Gln, 1.5; Val and Thr, 2.

The Phe^{3}X^4 series involves many ionic or polar side chains, which requires that appropriate additions be made to the QSAR properties. In addition to the electrical charge and its absolute value, the following related variables were defined and tested for that series: The number of hydrogen bond H-acceptor sites and the number of hydrogen bond H-donor sites were assigned with the aid of the solute hydrogen bonding scale of Abraham.[29] It was assumed that the ammonium group on Lys and Orn could serve as H-donor in two hydrogen bonds.[30] The total hydrogen bonding capacity is the sum of the H-acceptor and H-donor sites. The charge-weighted hydrogen bonding capacity is the total hydrogen bonding capacity multiplied by a factor of two if the group is charged. We envision this variable as a measure of the extent of a group's hydrogen-bonding stabilized hydration layer, assuming that a charged group, whether positive or negative, will hydrogen bond to water more strongly than will an uncharged group.

Correlation and regression calculations were performed using the Microsoft Excel v. 5.0 spreadsheet program.

Acknowledgements

Opioid receptor binding assays were graciously performed under the direction of Dr Henry I. Mosberg at

the University of Michigan. Electrospray mass spectral analyses were provided by Dr Phil Andrews and colleagues at The University of Michigan Protein and Carbohydrate Structure Facility. Funding for this research has been provided by the Eastern Michigan University Chemistry Department, an EMU Sabbatical Leave Award (SES), an EMU Faculty Research Fellowship (DLH), an EMU Graduate School Support Award (DLH), a Josephine Nevins Keal Professional Development Award for Women Faculty (DLH) and primarily by Research Corporation (DLH).

References

1. Erspamer, V. *Int. J. Dev. Neuro.* **1992**, *10*, 3.

2. Misicka, A.; Lipkowski, A. W.; Horvath, R.; Davis, P.; Kramer, T. H. *Life Sci.* **1992**, *51*, 1025.

3. Schiller, P. W.; Weltrowska, G.; Nguyen, T. M.-D.; Wilkes, B. C.; Chung, N. N.; Lemieux, C. *J. Med. Chem.* **1992**, *35*, 3956.

4. Geza, T.; Russell, K. C.; Landis, G.; Kramer, T. H.; Fanh, L.; Knapp, R.; Davis, P.; Burks, T. F. *J. Med. Chem.* **1992**, *35*, 2384.

5. Hruby, V. J.; Gehrig, C. A. *Med. Res. Rev.* **1989**, *9*, 343.

6. Sagan, S.; Amiche, M.; Delfour, A.; Camus, A.; Mor, A; Nicolas, P. *Biochem. Biophys. Res. Commun.* **1989**, *163*, 726.

7. Salvadori, S.; Marastoni, M.; Balboni, G. *J. Med. Chem.* **1991**, *34*, 1656.

8. Lazarus, L. H.; Salvadori, S.; Attila, M.; Grieco, P.; Bundy, D. M.; Wilson, W. E.; Tomatis, R. *Peptides* **1993**, *14*, 21.

9. Sagan, S.; Charpentier, S.; Delfour, A.; Amiche, M.; Nicolas, P. *Biochem. Biophys. Res. Commun.* **1992**, *187*, 1203.

10. Mosberg, H. I.; Kroona, H. B.; Omnaas, J. R.; Sobczyk-Kojiro, K.; Bush, P.; Mousigian, C. In *Peptides: Chemistry, Structure and Biology, Proc. Am. Pept. Symp., 13th meeting date*; Hodges, R. S.; Smith, J. A., Eds.; ESCOM: Leiden, 1993; pp 514–516.

11. Melchiorri, P.; Negri, L.; Falconieri-Erspamer, G.; Severini, C.; Corsi, R.; Soaje, M.; Erspamer, V.; Barra, D. *Eur. J. Pharmacol.* **1991**, *195*, 201.

12. Lazarus, L. H.; Bryant, S. D.; Salvadori, S.; Attila, M.; Jones, L. S. *Trends Neurosci.* **1996**, *19*, 31.

13. Breveglieri, A.; Guerrini, R.; Salvadori, S.; Bianchi, C.; Bryant, D.; Attila, M.; Lazarus, L. *J. Med. Chem.* **1996**, *39*, 773.

14. Misicka, A.; Lipkowski, A. W.; Nikiforovich, G. V.; Kazmierski, W. M.; Knapp, R. J.; Yamamura, H. I.; Hruby, V. J. In *Peptides: Chemistry and Biology: Proc. Am. Pept. Symp., 12th meeting date*; Smith, J.; Rivier, J., Eds.; ESCOM: Leiden: Netherlands, 1991; pp 140–141.

15. Misicka, A.; Lipkowski, A. W.; Fang, L.; Knapp, R. J.; Davis, P.; Kramer, T.; Burks, T. F.; Yamarura, H. I.; Carr, D. B.; Hruby, V. J. *Biochem. Biophys. Res. Commun.* **1991**, *180*, 1287.

16. Surratt, C. K.; Johnson, P. S.; Moriwaki, A.; Seidleck, B. K.; Blaschak, C. J.; Wang, J. B.; Uhl, G. R. *J. Biol. Chem.* **1994**, *269*, 20548.

17. Heyl, D. L.; Schmitter, S. J.; Bouzit, H.; Johnson, T. W.; Hepp, A. M.; Kurtz, K. R.; Mousigian, C. *Int. J. Peptide Protein Res.* **1994**, *44*, 420.

18. Heyl, D. L.; Dandabathula, M.; Kurtz, K. R.; Mousigian, C. *J. Med. Chem.* **1995**, *38*, 1242.

19. Heyl, D. L.; Bouzit, H.; Mousigian, C. *Lett. Pept. Sci.* **1996**, *2*, 277.

20. Heyl, D. L.; Schullery, S. E. *Current Med. Chem.* **1997**, *4*, 117.

21. (a) Norusis, M. J. *SPSS for Windows: Base System Users Guide, Release 6.0;* SPSS Inc.: Chicago, 1993. (b) Pfaffenberger, R. C.; Patterson, J. H. *Statistical Methods*, 3rd ed.; Irwin: Homewood, IL, 1987.

22. Mosberg, H. I.; Heyl, D. L.; Omnaas, J. R.; Haaseth, R. C.; Medzihradsky, F.; Smith, C. B. *Mol. Pharmacol.* **1990**, *38*, 924.

23. Munson, P. J.; Rodbard, D. *Anal. Biochem.* **1980**, *107*, 220.

24. Allinger, N.L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

25. Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. N.; Robins, R. K. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163.

26. Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533.

27. Bodor, N.; Gabanyi, A.; Wong, C. *J. Am. Chem. Soc.* **1989**, *111*, 3783.

28. Gavezotti, A. *J. Am. Chem. Soc.* **1983**, *105*, 5220.

29. Abraham, M. A. *Chemical Society Reviews* **1993**, 73.

30. Guo, S.; Scott, R. M. *J. Mol. Struct.* **1990**, *237*, 307.

(*Received in U.S.A.* 20 *March* 1997; *accepted* 25 *July* 1997)