Journal of Peptide Science
*J. Peptide Sci.* **7**: 74–81 (2001)

# Antibiotic Activity of Pentadecapeptides Modelled from Amino Acid Descriptors

**TORE LEJON*, MORTEN B. STRØM and JOHN S. SVENDSEN[1]**

Department of Chemistry, University of Tromsø, Tromsø, Norway

*Received 24 July 2000*
*Accepted 27 July 2000*

Abstract: Pentadecapeptides based on modified murine lactoferricin (LFM) sequences show varying degrees of antibacterial activity against *Escherichia coli* and *Staphylococcus aureus*. By means of projections to latent structures (PLS), a good correlation is obtained if the biological activity is modelled as a function of variables describing peptide properties, e.g. $\alpha$-helicity, hydrophobicity/hydrophilicity and charge. Using variables derived from a principal component analysis (PCA) of all naturally occurring amino acids, it is possible to describe the amino acid content of the peptides using three variables per amino acid position. The resulting descriptor matrix is then used to develop quantitative structure–activity relationships (QSAR). It is shown that the theoretically derived descriptors model the activity of the peptides better than the earlier model, and that properties of the peptides other than antibacterial activity can be predicted. 

Keywords: pentadecapeptides; antibacterial activity; amino acids; antibiotic activity

## INTRODUCTION

### Drug Design

The continuing search for new, pharmaceutically interesting compounds can, from a chemist's point of view, be divided into two main activities: (i) the search for, and characterization of, new active substances (lead compounds); and (ii) the modification of lead compounds into useful and efficient drugs. When searching for new compounds it is common to screen biological material for new substances that show activity in assays, to synthesize chemical libraries either by using combinatorial chemistry or by using a competitor's active substance as a starting point in the search for compounds that can be patented. When trying to modify lead compounds, new compounds are synthesized based on (qualified) guesses as to which of the molecule's features are important for its activity, i.e. we try to establish qualitative structure–activity relationships (QSAR). In this way, a large number of substances are produced and the possibility of finding a substance with better qualities than the lead compound increases with the number of substances synthesized. Instead of this trial and error approach it should be possible to use some strategy to design experiments that will minimize the work needed and maximize the information gained. In order to achieve this, it is imperative that the molecule can be described in such a way that its effect can be predicted prior to synthesis of a compound.

### Quantitative Structure–Activity Relationships (QSAR)

The concept of QSAR is based on the assumption that individual molecules can be described by physicochemical variables and that the model obtained can be related to the (biological) activity of

Abbreviations: LFB, bovine lactoferrricin; LFM, murine lactoferricin; MIC, minimum inhibitory concentration; PCA, principal component analysis; PLS, projections to latent structures; QSAR, quantitative structure–activity relationships.

* Correspondence to: Department of Chemistry, University of Tromsø, N-9037 Tromsø, Norway; e-mail: tore.lejon@chem.uit.no
[1] Correspondence can also be addressed to J.S. Svendsen; e-mail: johns@chem.uit.no

the molecule. Examples that are relevant to drug discovery were already being published during the 19th century by Crum-Brown and Fraser [1], who related the water solubility of alkaloids to their toxicity, and by Overton [2] and Meyer [3], who developed models in which hydrophobicity was related to narcotic effect. It should be obvious that models are merely descriptions of reality and are only valid for the substances under investigation or for very similar substances. If compounds of a very different structure are included, it is likely that other properties, e.g. bioavailability, will have to be considered. In the work by Hansch *et al.* [4], this problem was partly solved by using several variables describing hydrophobicity and electronic properties and by relating them to the effect using multiple regression. Improvements in instrumentation have led to data becoming cheaper all of the time and molecules can now be described by more and more variables; hence, data matrixes are getting larger. Regression methods are not well suited for use when the number of objects (equations) is smaller than the number of variables, and difficulties arise when the number of studied responses increases, e.g. toxicity and therapeutic activity studied simultaneously. A method such as projections to latent structures (PLS) [5,6] copes with these types of problems and has found widespread use in QSAR. There are numerous examples in the literature, e.g. the study by Hellberg *et al.* [7], in which the anaesthetic activity and toxicity of halogenated ethyl methyl ethers were examined, and later work where biologically active peptides were studied [8,9]. As descriptors for the peptides, the authors used three scales for each amino acid derived from a principal component analysis (PCA) [10] of a large number of variables describing the individual amino acids.

## Peptides as Drugs

Over the years, peptides have proved to be interesting in different types of therapy, e.g. insulin in the treatment of diabetes, and peptides as vaccines [11–13] and antibiotics [14–18]. A molecule like insulin is naturally difficult to modify partly because it is produced biochemically but mostly because the large number of amino acids makes it difficult to scan the activity as a function of amino acid content in a rational way. When dealing with smaller peptides, the use of machines for synthesis has made it possible to gain rapid access to a number of newly modified peptides that can be compared with the lead compound. Modified pentadecapeptides of murine lactoferricin (LFM) have been studied with respect to their antibiotic activity against *Escherichia coli* and *Staphylococcus aureus* [19]. The study revealed a clear correlation between the structure of the peptides, as described by a number of variables used to characterize the molecule, and the antibiotic activity. A drawback associated with the use of descriptors that describe the entire molecule, instead of its building blocks, is that it is virtually impossible to predict which amino acids should be replaced when attempting to improve the activity of the peptides. The peptides used in the study had been modified in only four of the 15 positions and we were interested in finding out whether it would be possible to use the descriptors for amino acids published by Hellberg *et al.* [9] to predict the activity of the individual peptides.

## Peptide Descriptors

Strøm *et al.* [19] characterized 19 peptides with 12 descriptors that were either measured, as helicity in different media and HPLC retention time, or calculated. The calculated variables were: net charge at pH 7; micelle affinity; Eisenberg, Garnier and Chou–Fasman $\alpha$-helix propensities; Kyte–Dolittle hydrophobicity; Emini surface index; mean hydrophobic moment; and mean charge moment. The minimum inhibitory concentration (MIC) against both *E. coli* and *S. aureus* was used as the dependent variable (Table 1). The descriptors used explained 82% of the variation in antibacterial activity by using 56% of the variation in the descriptors. These results were very encouraging and made an extension of the study of interest, in order to see whether descriptors for the amino acids could be used for describing the entire peptide. The theoretical description of the peptides was based on the work by Hellberg *et al.* [9], in which a property matrix of 29 descriptors for each coded amino acid had been used in a PCA. The original variables contained information about hydrophilicity/hydrophobicity, size, acidity constants, chemical shifts and HPLC retention times. The outcome of the analysis was three descriptors — $z_1$, $z_2$ and $z_3$ — describing each amino acid (Table 2). For the modified LFM peptides, only four positions had variations in amino acid substitution, resulting in a matrix of 19 objects (peptides) described by 12 new variables ($z_1$, $z_2$ and $z_3$ for each of the four amino acids). The resulting matrix was then used to model the measured/calculated descriptors for the peptides or for the activity of the peptides (as a logarithm of the MIC values).

Table 1 Descriptors for Bovine Lactoferricin (LFB), Murine Lactoferricine (LFM) and Modified LFMs

| Name | No. | SDS[a] | HFIP[b] | Ch[c] | MA[d] | Rt[e] | Ea[f] | Ga[g] | C-Fa[h] | K-D H[i] | ESI[j] | M[k] | C[l] | Log *E. coli*[m] | Log *S. aureus*[n] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LFB | 0 | 4 | 6 | 5.88 | 0.67 | | 0.267 | 0.467 | 0.800 | 1.28 | 1.39 | 0.321 | 0.1749 | 1.380 | 1.681 |
| LFM | 1 | 9 | 13 | 1.88 | 0.69 | 3.13 | 0.400 | 0.667 | 0.733 | 1.23 | 1.30 | 0.422 | 0.0355 | 2.699 | 2.699 |
| LFM W8 | 2 | 12 | 17 | 1.88 | 0.71 | 11.65 | 0.467 | 0.667 | 0.733 | 1.05 | 1.00 | 0.333 | 0.0355 | 2.699 | 2.699 |
| LFM W8 Y13 | 3 | 20 | 20 | 1.88 | 1.00 | 11.99 | 0.600 | 0.600 | 0.667 | 1.34 | 1.34 | 0.327 | 0.0355 | 2.699 | 2.699 |
| LFM A1 W8 | 4 | 5 | 18 | 2.88 | 0.28 | 11.17 | 0.600 | 0.667 | 0.733 | 0.86 | 0.94 | 0.249 | 0.1021 | 2.592 | 2.699 |
| LFM A9 W8 | 5 | 15 | 19 | 2.88 | 0.79 | 15.22 | 0.400 | 0.667 | 0.733 | 0.7 | 0.74 | 0.301 | 0.0808 | 2.699 | 2.699 |
| LFM A1.9 W8 | 6 | 16 | 17 | 3.88 | 0.94 | 14.86 | 0.533 | 0.667 | 0.733 | 0.5 | 0.68 | 0.211 | 0.1313 | 2.127 | 2.699 |
| LFM R1 W8 | 7 | 7 | 11 | 3.88 | 0.64 | 10.22 | 0.533 | 0.467 | 0.733 | 1.09 | 1.02 | 0.448 | 0.1688 | 1.568 | 2.572 |
| LFM R9 W8 | 8 | 7 | 11 | 3.88 | 0.64 | 7.88 | 0.533 | 0.600 | 0.733 | 1.12 | 1.08 | 0.404 | 0.1438 | 2.000 | 2.699 |
| LFM A1 R9 W8 | 9 | 6 | 11 | 4.88 | 0.55 | 7.59 | 0.400 | 0.600 | 0.733 | 0.92 | 1.02 | 0.332 | 0.1815 | 1.491 | 2.409 |
| LFM A9 R1 W8 | 10 | 11 | 14 | 4.88 | 0.79 | 12.28 | 0.533 | 0.467 | 0.733 | 0.74 | 0.76 | 0.420 | 0.192 | 1.491 | 2.409 |
| LFM R1.9 W8 | 11 | 6 | 7 | 5.88 | 0.86 | 7.36 | 0.600 | 0 | 0.733 | 1.16 | 1.10 | 0.508 | 0.2326 | 1.000 | 1.568 |
| LFM A1 W8 Y13 | 12 | 15 | 17 | 2.88 | 0.88 | 11.81 | 0.467 | 0.600 | 0.667 | 1.14 | 1.28 | 0.250 | 0.1021 | 2.178 | 2.699 |
| LFM A9 W8 Y13 | 13 | 19 | 19 | 2.88 | 1.00 | 15.25 | 0.467 | 0.600 | 0.667 | 0.99 | 1.01 | 0.280 | 0.0808 | 2.699 | 2.699 |
| LFM A1.9 W8 Y13 | 14 | 20 | 18 | 3.88 | 1.11 | 14.6 | 0.533 | 0.600 | 0.667 | 0.79 | 0.95 | 0.193 | 0.1313 | 1.944 | 2.699 |
| LFM R1 W8 Y13 | 15 | 15 | 12 | 3.88 | 1.25 | 10.8 | 0.467 | 0.400 | 0.667 | 1.38 | 1.36 | 0.436 | 0.1688 | 1.278 | 1.863 |
| LFM R9 W8 Y13 | 16 | 14 | 13 | 3.88 | 1.08 | 8.34 | 0.533 | 0.600 | 0.667 | 1.41 | 1.45 | 0.411 | 0.1438 | 1.799 | 2.699 |
| LFM A1 R9 W8 Y13 | 17 | 11 | 11 | 4.88 | 1.00 | 8.10 | 0.467 | 0.600 | 0.667 | 1.21 | 1.39 | 0.347 | 0.1815 | 1.397 | 1.602 |
| LFM A9 R1 W8 Y13 | 18 | 13 | 12 | 4.88 | 1.08 | 13.08 | 0.467 | 0.400 | 0.667 | 1.02 | 1.03 | 0.398 | 0.192 | 1.079 | 1.397 |
| LFM R1.9 W8 Y13 | 19 | 15 | 15 | 5.88 | 1.00 | 7.27 | 0.667 | 0 | 0.667 | 1.44 | 1.47 | 0.507 | 0.2326 | 1.000 | 1.079 |

[a] Helicity in the presence of SDS; [b] helicity in hexafluoro isopropanol; [c] net charge at pH 7; [d] micelle affinity; [e] HPLC retention time; [f] Eisenberg $\alpha$-helix propensity; [g] Garnier $\alpha$-helix propensity; [h] Chou–Fasman $\alpha$-helix propensity; [i] Kyte–Dolittle hydrophobicity; [j] Emini surface index; [k] mean hydrophobic moment; [l] mean charge moment; [m] logarithm of MIC against *E. coli*; [n] logarithm of MIC against S. *aureus*.

Table 2 *Z* Scales for Coding of Amino Acids

| Amino acid[a] | | $z_1$ | $z_2$ | $z_3$ |
|---|---|---|---|---|
| Ala | A | 0.07 | −1.73 | 0.09 |
| Val | V | −2.69 | −2.53 | −1.29 |
| Leu | L | −4.19 | −1.03 | −0.98 |
| Ile | I | −4.44 | −1.68 | −1.03 |
| Pro | P | −1.22 | 0.88 | 2.23 |
| Phe | F | −4.92 | 1.30 | 0.45 |
| Trp | W | −4.75 | 3.65 | 0.85 |
| Met | M | −2.49 | −0.27 | −0.41 |
| Lys | K | 2.84 | 1.41 | −3.14 |
| Arg | R | 2.88 | 2.52 | −3.44 |
| His | H | 2.41 | 1.74 | 1.11 |
| Gly | G | 2.23 | −5.36 | 0.30 |
| Ser | S | 1.96 | −1.63 | 0.57 |
| Thr | T | 0.92 | −2.09 | −1.40 |
| Cys | C | 0.71 | −0.97 | 4.13 |
| Tyr | Y | −1.39 | 2.32 | 0.01 |
| Asn | N | 3.22 | 1.45 | 0.84 |
| Gln | Q | 2.18 | 0.53 | −1.14 |
| Asp | D | 3.64 | 1.13 | 2.36 |
| Glu | E | 3.08 | 0.39 | −0.07 |

[a] Symbols and abbreviations are in accordance with IUPAC recommendations.

## RESULTS AND DISCUSSION

### Analysis of Peptide Properties

Even though it is feasible to create a model from measured or calculated peptide properties, as shown by Strøm *et al.* [19], that gives a good correlation between peptide properties and antibacterial activities, it is not implied that it is straightforward to use this model when designing peptides with enhanced antibacterial activities. The problem is inherent in the way that the model is designed; the authors show that it is possible, or even easy, to tell which macroscopic properties in the peptides should be adjusted in order to achieve the preferred biological activity. The almost insurmountable task is to determine how to modify the amino acid sequence in order to give the peptide the necessary bulk properties. There are two closely related strategies to resolve this problem: one is to try to derive quantitative relations between the amino acid sequence and the macroscopic properties of the corresponding peptide; and the other is to derive directly a QSAR between the amino acid sequence and the antimicrobial activities. Both of these strategies are addressed in the present study.

If the aim is to predict the outcome of any amino acid replacement, it is imperative that the variables of the model contain information about the individual amino acids and not only macroscopic values for the entire peptide. Thus, if it were possible to use numerical values that describe each individual amino acid it would be possible to predict the properties or the activity of any modified peptide.

In order to investigate whether the theoretical descriptors developed by Hellberg *et al.* [9] contain the same information as the variables describing the peptides in Strøm *et al.*'s [19] study, a PLS analysis was performed. The theoretical variables were used as independent variables (*X* matrix) and the variations in the measured/calculated variables describing the peptides were used as dependent variables (*Y* matrix). The analysis yielded five significant variables explaining 80% of the variation in *Y* using 98% of the variation in *X*. The result implies that even though the variation in peptide descriptors can be explained by the theoretically derived descriptors, the peptide descriptors used by Strøm *et al.* [19] contained information that is not present in the descriptors of the amino acids, *vide infra*. On the other hand, the graphical presentation of the results of the PLS analysis presented in Figure 1 (observed *vs.* predicted helicity of peptides in the presence of SDS) and in Figure 2 (observed *vs.* predicted retention time on HPLC) shows that there is a surprisingly good correlation between the observed properties of the peptides and the calculated properties based on Hellbergs *z* descriptors.

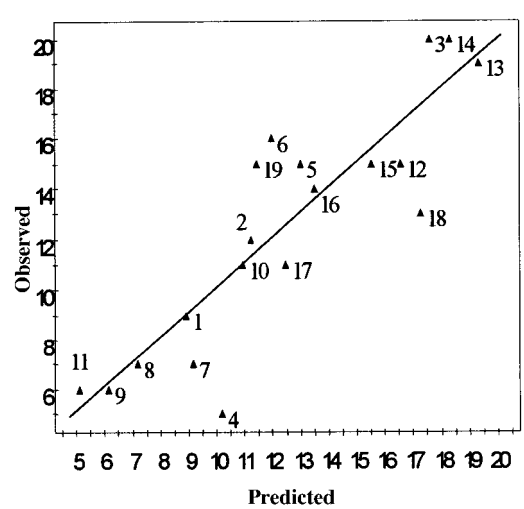


Figure 1 Observed *vs.* predicted peptide helicity in the presence of SDS.

A few points should be made here about the information contained in the theoretically derived *z* scales. HPLC data were included in the original matrix of Hellberg *et al.*, from which the descriptors were derived; however, the solvent systems than were used for the HPLC analysis of the peptides in Strøm *et al.* [19] were completely different. Still, retention times are generally considered to correlate fairly well with hydrophilicity/hydrophobicity because molecular interaction with solvent and stationary phase, i.e. solubility, is what governs the retention time. The obvious conclusion is that information about solubility is additive and that, in some sense, the solubility of the peptide is a sum of the solubilities of the amino acids included in the peptide.

As mentioned above, there is more information contained in the peptide matrix than in the matrix consisting of the *z* scales. One way of explaining this is to look at how multivariate analytical methods, e.g. PLS, treat data. The way to describe an object in such an analysis is to consider it as a point in an *n*-dimensional space, where *n* is equal to the number of variables. The position of each point is defined by the measured values for each of the variables. The result of this is that peptides containing the same amino acids will be positioned at the same point regardless of the sequence of the amino acids because no sequence information is present. Against this background, it is even more surprising that such a good correlation exists between the observed and the predicted $\alpha$-helicity. Whether this is a coincidence due to the small variation in helicity of the peptides studied (5–20%) or whether it implies that secondary structure is mainly governed by which amino acids are present and not by their position in the peptide needs to be further investigated.

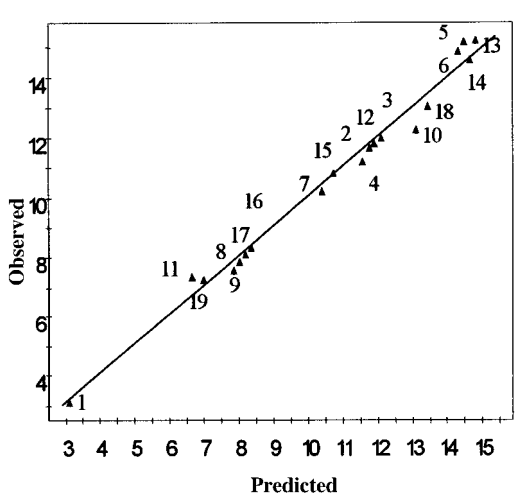


Figure 2 Observed *vs.* predicted HPLC retention times.

Another piece of information that needs to be mentioned is that peptide 1, native LFM, is the peptide that exhibits the lowest correlation between observed and predicted values. This is explained by the fact that LFM is the only peptide that does not contain a Trp residue in position 8. (In a paper by Strøm *et al.* [20], it was shown that including Trp in this position leads to more active peptides and that this substitution has been made in all modified molecules.) Still, the correlation is at such a level that it should be possible to predict the activities of peptides with a completely different amino acid content than the ones included in this study.

### Analysis of Peptide Activity

To be able to predict as many properties as possible for a peptide is of course of interest, but the main objective is to make good predictions for biological response. If amino acid descriptors contain information about, for example, the antibiotic activity of the peptides, models can be developed that predict which amino acids should be included in the peptide before its synthesis. To do this, a new PLS analysis was performed using the amino acid descriptors for modelling the activity of the peptides, using the logarithm of the activity of the peptides as the response. The analysis resulted in three significant components explaining 99% of the variation in antibacterial activity, using 81% of the variation in the *X* matrix. As for the previous analysis, the fact that not all of the variation in the independent variables was explained is partly due to the lack of sequence information in the variables. As is seen from the results, the theoretically derived variables gave an equally good, or better, model for the activities compared with the earlier study [19,21] and, hence, the same information is present in both matrices.

The question to be answered is whether the model is able to predict the activity of the peptides. In Figure 3, predicted *vs.* observed antibacterial activity against *E. coli* is plotted, while the same plot for *S. aureus* is presented in Figure 4. The fact that the peptides exhibiting the lowest response all have the same observed value while predicted values differ is a result of the way in which the biological testing was performed. For compounds with low activity, the actual MIC value is not measured because the

 

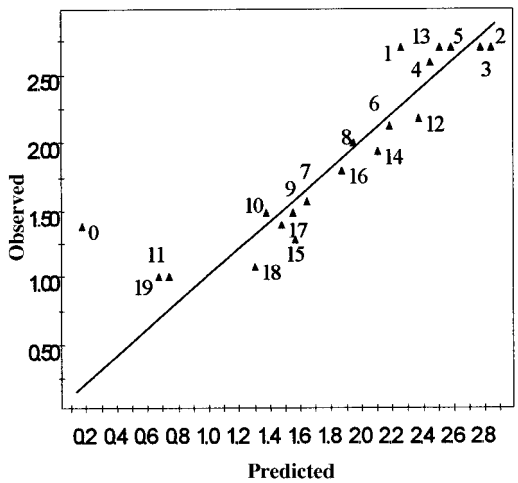


Figure 3 Observed *vs.* predicted activity against *E. coli*.

concentrations would be too high. Instead, the value is reported as higher than the threshold value of 500 μM. One way of presenting results that seemingly would be better is to exclude those peptides that do not fit the model; however, their inclusion shows the stability of the model.

Another way of checking the validity of the model is to include objects (peptides) that were not part of the calculations in order to see how well they are described by the model. In this case, a pentadecapeptide from bovine lactoferrricin (LFB) is included as object 0. The plot reveals that this peptide is well described by the model, even though the predicted value is somewhat lower than the observed one. This is still a good result considering that the sequence homology between LFB and LFM is only 53%.

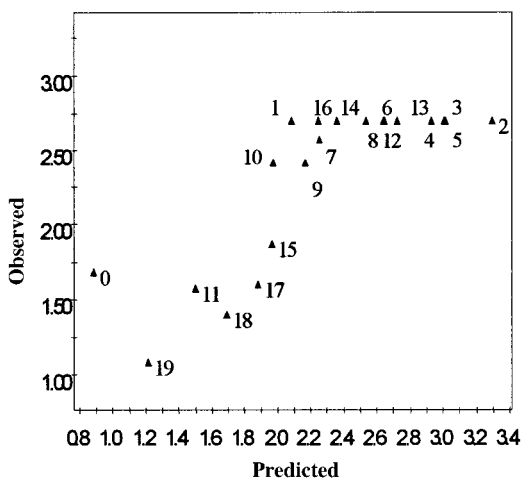


Figure 4 Observed *vs.* predicted activity against *S. aureus*.

### Analysis of Peptide Composition

In order to examine which amino acids are most important for modelling the activity, the loadings, $w_k$, for the 12 variables have to be considered. The numerical values are listed in Table 3, while the first loading vectors are plotted in Figure 5. The way to interpret loadings is to look at their absolute values, i.e. examine how far from the origin they are situated. From this it is seen that the amino acid residue in position 8 is the most important for explaining activity (keeping in mind that the PLS

Table 3 Loadings ($w*c$) for the Variables in the PLS Model

| Variable name | $w*c$ [1] | $w*c$ [2] | $w*c$ [3] |
|---|---|---|---|
| 1a | 0.302 | 0.116 | 0.558 |
| 1b | 0.026 | −0.460 | −0.077 |
| 1c | −0.116 | 0.521 | 0.311 |
| 8a | −0.623 | 0.111 | 0.459 |
| 8b | 0.517 | 0.225 | 0.201 |
| 8c | 0.125 | 0.092 | 0.117 |
| 9a | 0.301 | 0.095 | 0.364 |
| 9b | 0.035 | −0.381 | −0.137 |
| 9c | −0.130 | 0.383 | 0.208 |
| 13a | −0.313 | −0.277 | −0.135 |
| 13b | −0.059 | −0.216 | 0.549 |
| 13c | −0.106 | −0.124 | 0.061 |
| Log *E. coli* | 0.257 | 0.344 | 0.265 |
| Log *S. aureus* | 0.320 | 0.328 | 0.192 |

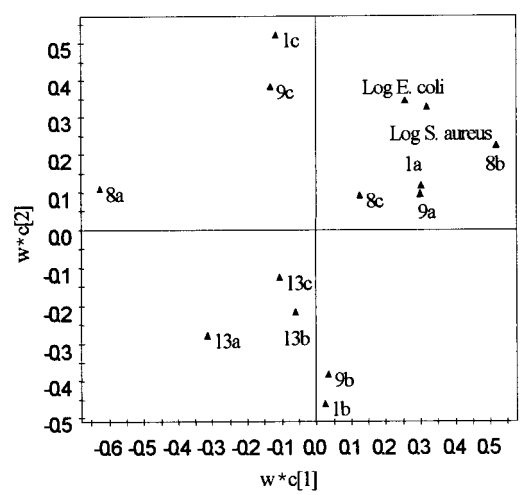


Figure 5 First to loading vectors from PLS analysis.

analysis undertaken has no sequence information). A closer examination of position 8 reveals that it should contain an amino acid that exhibits a large negative $z_1$ and a large positive $z_2$, while $z_3$ is of less importance. (The properties contained in $z_1$ are mainly related to hydrophilicity and $z_2$ contains information about size, but the simplest way to use this information is to look for amino acids that have the principal properties desired instead of looking at the molecular descriptors.) From Table 2 it seems as though Trp is in fact the best choice, which Strøm *et al.* [20] also concluded. For position 1, $z_1$ is of less importance while $z_3$ (which contains information about charge) should be large and $z_2$ should be negative. The exact same pattern is also noticeable for position 9. A substitution that would fulfil the requirements would be the incorporation of Cys. Position 13 has all three components on the negative side and, numerically, is of the least importance. A smaller effect is obtained if substitutions are made here, but any amino acid that is to be incorporated should have negative $z_1$, $z_2$ and $z_3$ values, e.g. Val, Leu, Ile or Met. Of these four possible amino acids, Ile and Met have the largest negative values for $z_2$ and should be the first choice in modified peptides.

## CONCLUSIONS

Using PLS for QSAR it is obvious that theoretically derived descriptors for individual amino acids contain information that is necessary for explaining peptide activity. From earlier studies it is known that charge is important for modelling activity [19], but the present study shows that variables describing lipophilicity are equally important. It is surprising that, using the theoretical descriptors, it is possible to model experimentally derived values for peptides, as well as calculated descriptors. In order to obtain even higher biological activities, new modified peptides can be designed based on the results of the present study. Another extension of this study would be the inclusion of other modified peptides and the prediction of the activity of the new peptides. The new peptides should span a larger variation in amino acid composition in order to better understand which factors govern the activity.

An earlier study [8] claimed that, because the descriptors for the individual amino acids are so well correlated with the activity, the secondary structure of the peptides is of less importance. This is probably the case where short peptides are employed. We are not convinced that this is the case with the pentadecapeptides studied here, because descriptors such as $\alpha$-helicity show a good correlation between observed and predicted values. However, further investigations of longer peptides need to be undertaken in order to answer this question.

### Experimental

The program package Simca-P 8.0 (Umetrics, Umeå, Sweden) was used for all calculations. The theoretically derived $z$ scales were used without scaling, because the variables used for the original analysis had been scaled, and data were centred when using experimental variables as dependent variables. In these analyses, the experimentally derived variables were scaled to unit variance and centred. In the PLS analysis of peptide activity, the activity data were used as the logarithm but neither data set was centred.

## Acknowledgements

Alpharma AS is gratefully acknowledged for their financial support to Morten B. Strøm and John S. Svendsen.

## REFERENCES

1. Crum-Brown A, Frazer TR. *Trans. R. Soc. Edinburgh* 1868; **25**: 151–103.
2. Overton E. *Physic. Chem.* 1897; **22**: 189–209.
3. Meyer H. *Arch. Exper. Pathol. Phamacol.* 1899; **42**: 109–118.
4. Hansch C, Maloney PP, Fujita T, Muir RM. *Nature* 1962; **194**: 178–180.
5. Kvalheim OM. *Latent-variable modelling of multivariate data*. University of Bergen, 1990.
6. Carlson R. *Design and Optimisation in Organic Synthesis*. Elsevier: Amsterdam, 1992.
7. Hellberg S, Wold S, Dunn WJ, Gasteiger J, Hutchings MG. *Quant. Struct.-Act. Relat.* 1985; **4**: 1–11.
8. Hellberg S, Sjöström M, Wold S. *Acta Chem. Scand. B* 1986; **40**: 135–140.
9. Hellberg S, Sjöström M, Skagerberg B, Wold S. *J. Med. Chem.* 1987; **30**: 1126–1135.
10. Mardia KV, Kent JT, Bibby JM. *Multivariate Analysis*. Academic Press: London, 1982.
11. Vermeij P, Blok D. *Pharm. World Sci.* 1996; **18**: 87–93.
12. Tam JP. *J. Immun. Meth.* 1996; **196**: 17–32.
13. Freland S, Chambers BJ, Ljunggren HG. *Canc. J.* 1996; **9**: 304–307.

 

14. Boman HG. *Scand. J. Immunol.* 1996; **43**: 475–482.
15. Bevins CL, Zasloff M. *Annu. Rev. Biochem.* 1990; **59**: 395–414.
16. Nicolas P, Mor A. *Annu. Rev. Microbiol.* 1995; **49**: 277–304.
17. Lehrer RI, Ganz T. *Ann. NY Acad. Sci.* 1996; **797**: 228–239.
18. Nissen-Meyer J, Nes IF. *Arch. Microbiol.* 1997; **167**: 67–77.
19. Strøm MB, Rekdal Ø, Stensen W, Svendsen JS. *J. Pept. Res.* 2000; **56**: 265–274.
20. Strøm MB, Rekdal Ø, Svendsen JS. *J. Pept. Res.* (in press).
21. Rekdal Ø, Andersen J, Vorland LH, Svendsen JS. *J. Pept. Sci.* 1999; **5**: 32–45.