A theoretical analysis of secondary structural characteristics of anticancer peptides

Sarah R Dennison[1], FrederickHarris[2], Tailap Bhatt[2], Jaipaul Singh[2], and David A Phoenix[3]

[1] School of Pharmacy and Pharmaceutical Science, University of Central Lancashire, Preston PR1 2HE, UK

[2] School of Forensic and Investigative Sciences, University of Central Lancashire, Preston PR1 2HE, UK

[3]Deputy Vice Chancellor, University of Central Lancashire, Preston PR1 2HE, UK.

Tel: +44 (0) 1772 892504

Fax: +44 (0) 1772 892936

E-mail: daphoenix@uclan.ac.uk.

Abstract

Here, cluster analysis showed that a database of 158 peptides formed 21 clusters based on net positive charge, hydrophobicity and amphiphilicity. In general these clusters showed similar median toxicities ($p = 0.176$) against eukaryotic cell lines and no single combination of these properties was found optimal for efficacy. The database contained 14 peptides, which showed selectivity for tumour cell lines only (ACP_{CT}), 123 peptides with general toxicity to eukaryotic cells (ACP_{GT}) and 21 inactive peptides (ACP_I). Hydrophobic arc size analysis showed that there was no significant difference across the datasets. Even though there was no correlation there was no correlation observed, peptides with wide hydrophobic arcs (> 270°) appeared less toxic. Extended hydrophobic moment plot analysis predicted that over 50% of ACP_{CT} and ACP_{GT} peptides would be surface active, which led to the suggestion that amphiphilicity is a key driver of the membrane interactions for these peptides but probably plays a role in their efficacy rather than their selectivity. This analysis also predicted that only 14% of ACP_{CT} peptides compared to 45% of ACP_{GT} peptides were candidates for tilted peptide formation. This implies that those peptides with non-specific activity may have a tendency towards the utilisation of membrane disruptive structures such as tilt peptides which led to the suggestion that the absence of this structure may support cancer cell selectivity. However, these analyses predicted that ACP_I peptides, which possess no anticancer activity, would also form surface active and tilted α-helices, clearly showing that other factors are involved in determining the efficacy and selectivity of ACPs.

Key words:

Abbreviations:

Acknowledgements: The authors would like to thank Dr Nadia Chuzhanova, University of Central Lancashire for her assistance with the cluster analysis.

Introduction

Defence peptides are naturally occurring antimicrobial molecules and many have been found to possess toxicity to tumour cells. However, these anticancer peptides (ACPs) differ widely in their efficacy and selectivity for cancer cells. Some ACPs exhibit general toxicity (ACP_{GT} peptides) to eukaryotic cells, killing cancer and non-cancer alike, whilst others show toxicity to cancer cells alone (ACP_{CT} peptides). Moreover, ACP_{CT} peptides can also show selectivity between different types of cancer cells. At present, the factors that determine these differences in efficacy and selectivity are poorly understood, limiting efforts to develop ACPs as therapeutically useful anticancer agents [1, 2].

Current understanding is that the efficacy and selectivity of ACPs for cancer cells is dependent upon their ability to interact with membranes of the target cells. This ability appears to depend upon the characteristics of the target cell membrane along with a range of physiochemical properties possessed by ACPs with net positive charge, amphiphilicity and hydrophobicity making major contributions [3]. These physiochemical properties are determined by amino acid composition and the secondary structure adopted by the parent molecule with the majority of studies involving peptides that adopt amphiphilic α-helical structure [3]. Many of these peptide α-helices may be classed as surface active and interact with the bilayer such that their orientation is approximately parallel to the membrane surface albeit, sometimes, as an initial step leading to further membrane interactions or cell internalisation. Adopting such orientations allows the polar face of these α-helices to interact with the bilayer head group region whilst their polar face penetrates the hydrophobic membrane core [4]. However, it is becoming increasingly clear that a number of ACPs adopt membrane interactive oblique orientated α-helical structure [5-7]. Also known as tilted peptides, these are a highly specialised class of amphiphilic α-helix that show some structural similarities to surface active α-helices but differ primarily by possessing an asymmetric distribution of hydrophobicity along the α-helical long axis. This structural feature facilitates membrane penetration by the segment at a shallow angle of between 30^{o} and 60^{o}, thereby inducing a range of membrane - related effects such as the destabilisation of lipid packing [8, 9].

To date, few large scale databases of ACPs appear to have been presented, limiting the potential for theoretical investigations into factors that influence the anticancer activity of these peptides. However, Owen *et al.*, [10] recently introduced an extensive database of naturally occurring and synthetic α-helical ACPs and in the present study, we have used a variety of theoretical techniques to identify factors that may contribute to differences in the efficacy and selectivity of these peptides.

Methods

Database assembly

A database of anticancer peptides (ACPs) was constructed using data presented by Owen *et al.,* [10] and α-helical structure in these peptides was confirmed using the secondary structure prediction programme: Profile Network from Heidelberg [11]. Toxicity data for these ACPs were also extracted from the database of Owen *et al.*, [10] as the half lethal dose (LD_{50}) against WI38 - a normal fibroblast cell line of the lung diploid cells, MCF7 - a breast adenocarcinoma tumour cell line, SW480 - a colon adenocarcinoma tumour cell line, BMKC - a cloned melanoma cell line, H1299 - a lung large cell carcinoma tumour cell line, HeLaS3 - a cervical epithelial carcinoma cell line and PC3 – a prostate adenocarcinoma tumour cell lines. Based on these toxicity data, the dataset of ACPs was further divided into 3 subsets, which included: 14 peptides which showed selectivity for tumour cell lines only (ACP_{CT}), 123 peptides with general toxicity to eukaryotic cells (ACP_{GT}) and 21 inactive peptides (ACP_I).

Extended hydrophobic moment plot analysis

After removal of peptides < 11 residues, peptides in each of the ACP_{GT}, ACP_{CT} and ACP_I datasets were analysed using extended hydrophobic moment plot methodology [12]. According to this methodology, the hydrophobicity of successive amino acids in these sequences are treated as vectors and summed in two dimensions, assuming an amino acid side chain periodicity of 100°. The resultant of this summation, the hydrophobic moment (μH) provides a measure of α-helix amphiphilicity [13]. Our analysis used a moving window of 11 residues and for each sequence under investigation the window with the highest hydrophobic moment was identified. For these windows, the mean hydrophobic moment, $< \mu_H >$, and the corresponding mean hydrophobicity, < H >, which provides a measure of α-helix affinity for the membrane interior, were computed using the normalised consensus hydrophobicity scale of Eisenberg *et al*., [14]. For each of these datasets, these parameters were then plotted

on the extended hydrophobic moment plot diagram of Harris *et al*., [12] and the location of the data points used to identify sequences that were predicted to be surface active, globular, transmembrane or candidates to form tilted peptides

Cluster analysis of ACPs

Identification of clusters

For peptides in the ACP dataset, the net positive charge was determined along with $< \mu_H >$ and $< H >$, which were computed as described above. Three-dimensional clustering of the dataset was performed using the statistical unweighted pair group arithmetic averaging (UPGMA), methodology, which is an agglomerative hierarchical technique [15]. The UPGMA tree was reconstructed using Phylip v3.63 (http://evolution.genetics.washington.edu) with the results presented in the form of a dendrogram. As part of the pair-wise algorithmic process, Euclidean distances between the data points in multi-dimensional space were used to construct an un-rooted tree together with joining nodes forming a branch. The peptides were clustered according to net positive charge, $<H>$ and $<\mu_H>$, and were considered as part of a cluster when the ACPs, were linked to other ACPs by a maximum of two nodes.

Toxicity analysis of clusters

The cell line toxicities of ACPs within clusters were subjected to box plot analysis in order to observe any outliers. This analysis was also used to determine the range of toxicities shown by peptides in each cluster and to enable comparison of the toxicities between clusters. The efficacy of ACPs in each cluster was studied using the cell lines: W138, MCF7, SW480, BMKC, H1299, HeLaS3 and PC3. To statistically compare the cell line toxicities of ACPs between clusters, the Anderson-Darling test was applied to investigate the normal distribution of the data. If $p < 0.005$, the data were considered to be non-normally distributed, in which case the Kruskal-Wallis non-parametric test was applied to test the null hypothesis (H_0) that there was no significant difference ($p > 0.05$) between cell line toxicities for the clusters analysed.

Results

Using UPGMA cluster analysis a dendrogram of ACPs in the database was constructed, which grouped peptides with similar net positive charge, < H > and < μ_H >. This analysis produced 1 tree, which was congruent in nodes showing 21 peptide clusters that demonstrated similarities in these physiochemical properties. These clusters of ACPs were designated A to U and in order to observe if any given arrangement of net positive charge, <μ_H> and < H > maximised the efficacy of ACPs, a box plot was used to represent the toxicity values for peptides within each cluster. In the case of the fibroblast cell line, W138 (Figure 1), the non-parametric test, Kruskal-Wallis, confirmed that there was a significant difference between medians across the range of clusters (Kruskal-Wallis = 38.261; p = 0.005). Further analysis of Figure 1 showed that the medians for clusters J and U were higher than those of the other clusters but that there were no significant difference between the medians of the remaining clusters (Kruskal-Wallis = 21.074; p = 0.176).

A box plot was used to represent the toxicity values for peptides within each cluster for each of the different cancer cell lines investigated (Figure 2). The Kruskal-Wallis test confirmed that there was a significant difference between the medians across the range of clusters for MCF7 (Kruskal-Wallis = 36.44; p = 0.009), SW480 (Kruskal-Wallis = 31.21; p = 0.027) and BMKC (Kruskal-Wallis = 36.17; p = 0.01). Further analysis of Figure 2 showed that for MCF7, SW480 and BMKC the median for clusters J, P, Q and U were higher than those of the other clusters. When these clusters are removed from the analysis there is no significant difference between the medians for the remaining clusters (MCF7 Kruskal-Wallis = 15.08; p = 0.446, SW480 Kruskal-Wallis = 16.02; p = 0.312 and BMKC Kruskal-Wallis = 14.46; p = 0.491). There was no significant difference between the medians across the range of clusters for the remaining cell lines H1229 (Kruskal-Wallis = 19.89; p = 0.401), HeLaS3 (Kruskal-Wallis = 26.28; p = 0.123) and PC3 (Kruskal-Wallis = 28.29; p = 0.007).

The box plot in Figure 2 also showed that peptides in clusters J and U are less effective across all cell lines. The peptides in these clusters had wide hydrophobic arc sizes ~ 220° implying that a wide hydrophobic arc reduced toxicity. Further statistical analysis of hydrophobic arc sizes for the complete dataset was undertaken and

represented as a boxplot in Figure 3. These data showed that the hydrophobic arc size of the ACP_I, ACP_{CT} and ACP_{GT} peptides ranged from 60° to 260° with ACP_{GT} peptides possessing the widest range (60° to 260°) and ACP_{CT} peptides possessing the narrowest range (80° to 240°) but the median arc size showed no significant difference (Kruskal Wallis = 0.01; p = 0.997). The database of ACPs was then interrogated and for each member peptide, its hydrophobic arc size along with its toxicity values (LD_{50}) against the 7 cell lines studied were extracted, collated and plotted on Cartesian planes to produce scatter plots (Figure 4). Whilst regression analysis of these data using the method of least squares showed that there was no statistical linear relationship between these two parameters ($R^2 = 0.03$), Figure 4 does indicate that in general peptides with arc sizes > 270° were less toxic with no toxicity values < 500 μM in contrast to peptides from all other arc sizes.

Peptides in each of the ACP_{CT}, ACP_{GT} and ACP_I datasets were analysed according to extended hydrophobic moment plot methodology [12]. For the ACPs of each of these datasets, plots of $< \mu_H >$ versus $< H >$ were then constructed and according to the location of their data points on the plot diagram (Figure 5, A-C), peptides were defined as potentially either: surface active, globular or tilted [5]. This analysis predicted that 7 of the ACP_I peptides were surface active with 6 potentially able to form tilted peptides. The remaining 2 ACP_I peptides were predicted to be globular (Figure 5A). This analysis also predicted that 4 ACP_{CT} peptides were surface active with 1 showing the potential to adopt tilted structure. The remaining 2 ACP_{CT} peptides were predicted to be globular (Figure 5B). ACP_{GT} peptides represented the biggest single group of ACPs in the database and 57 of these peptides were predicted to be surface active. The remaining ACP_{GT} peptides included 49 that were candidates to form tilted peptides and 4 that showed the potential to be globular (Figure 5C). These data are summarised in Table 1.

Given the apparent importance of surface activity the level of amphiphilicity was further investigated. Figure 5 showed that the $< \mu_H >$ of the ACP_I, ACP_{GT} and $ACPC_T$ datasets ranged from 0.33 to 1.05. The widest range in the $< \mu_H >$ was observed for the ACP_{GT} peptides (0.33 to 1.05) and the narrowest range in $< \mu_H >$ was observed for the ACP_{CT} peptides (0.53 to 0.78). Comparison of the medians across the dataset showed that the median $< \mu_H >$ for ACP_{GT} was 0.74, which was greater

than the median of ACP_I (median = 0.71) and ACP_{CT} (median = 0.73) peptide datasets. Since there was evidence of non-normality in the $< \mu_H >$ values of peptides in the datasets (Anderson-Darling = 6.315, $p < 0.005$), the non-parametric test, Kruskal-Wallis, was used to test the null hypothesis that the medians were the same across the 3 datasets. Here, the null hypothesis was accepted, confirming that there was no significant difference between the $< \mu_H >$ medians of the ACP_I, ACP_{CT} and ACP_{GT} peptides (Kruskal Wallis = 4.391; $p = 0.11$).

Discussion

At present, the factors that determine differences in the efficacy and selectivity of ACPs are poorly understood, inhibiting the potential to develop these peptides as therapeutically useful anticancer agents [1, 2]. However, it is generally accepted that net positive charge, hydrophobicity and amphiphilicity play major roles in the anticancer action of ACPs [1] and here, theoretical analyses are used to study the contribution of these physiochemical properties to the efficacy and selectivity of these peptides.

Owen *et al.*, [10] recently introduced an extensive database of α-helical ACPs, which included sequence information and toxicity data for a variety of cancer cell lines. Here, three-dimensional clustering techniques [15], were used to group ACPs of this database with similar net positive charge, $< H >$ and $< \mu_H >$, which produced 21 clusters, A to U. Box plots of these clusters were constructed for each cell line studied and in all cases, toxicities to fibroblast and cancer cell lines were in the low micromolar range (Figure 1 and 2). These box plot analyses also showed that in the vast majority of cases, clusters showed similar median cell line toxicities ($p = 0.176$) and examination of Figures 1 and 2 showed that no single combination of these 3 properties seemed to increase the overall efficacy of the ACPs analysed.

The hydrophobic arc sizes of the peptides in the ACP database ranged between 60° and 260°, which is comparable to those observed by Dennison *et al.*, [3] for ACPs (20° to 240°). Figure 4 indicates that in general peptides with arc sizes > 270° were

less toxic (MIC values > 500 μM) in contrast to peptides from all other arc sizes. However, whilst less active clusters appeared to have wider arc size, statistical analysis could find no direct correlation between arc size and toxicity. Nonetheless, it is generally accepted that hydrophobicity is a key driver of the ability of ACP peptides to penetrate membranes and hence their anticancer action [3]. Thus, the fact that ACP_I and ACP_{GT} peptides have similar hydrophobic arc sizes to ACP_{CT} peptides clearly suggests that although these physiochemical properties are important to the anticancer action of these latter peptides, other factors must be involved.

In response, we have used $< H >$ and $< \mu_H >$ along with extended hydrophobic moment plot methodology to consider the impact of the overall molecular architecture on the membrane interactive potential of peptides in the ACP_{GT}, ACP_{CT} and ACP_I datasets. Use of this methodology predicted that 56% of these ACP_{CT} peptides and 52% of these ACP_{GT} peptides would be surface active (Figure 5, B and C), which is consistent with previous work where it was predicted that surface activity may play an important role in the membrane interactions of ACPs [1]. Moreover, these results clearly suggest that amphiphilicity is a key driver of the membrane interactions of over half the peptides in the dataset and given its importance to both ACP_{CT} and ACP_{GT} peptides, it probably plays a role in the efficacy of the peptides against target cells rather than their selectivity. Extended hydrophobic moment plot analysis further showed that 14% of the ACP_{CT} peptides and 45% of the ACP_{GT} peptides studied were candidates for oblique orientated α-helix formation (Figure 5, C). Although the ACP_{CT} dataset is small, these results would seem to indicate that in contrast to ACP_{CT} peptides, membrane interaction of ACP_{GT} peptides may more commonly utilise tilted peptide structure. Thus, it may be that use of oblique orientated α-helical structure by ACP_{GT} peptides is associated with their broader spectrum of target specificity as compared to ACP_{CT} peptides. This structure has been associated with relatively non-specific means of cell lysis [1], which would fit this suggestion.

In the case of amphiphilicity, comparisons across the ACP_I, ACP_{AO} and ACP_T datasets suggested that there may be an optimal range of values between 0.33 and 1.05, which is in close in agreement with that obtained by Dennison *et al.*, [3], indicating their high potential as surface active compounds This architecture is required for their ability to interact with the hydrophobic section of a cancer cell

membrane, whilst the charged residues remain in contact with the anionic phospholipids on the outer surface, thus ensuring their aggregation onto the surface consequently leading to membrane permeabilization [16, 17]. Statistical analysis showed that there were significant differences between the $< \mu_H >$ values of peptides across the $ACP_{I,}$ ACP_{CT} and ACP_{GT} datasets although those peptides showing cancer cell specificity fell within a narrow range of amphiphilicity (0.53-0.78)

In summary, amphiphilicity appears to be a key driver in the efficacy of most ACP_{CT} peptides and *circa* half the ACP_{GT} peptides studied here against cell lines. The adoption of oblique orientated α-helical structure by many of the remaining ACP_{GT} peptides may be important to their broad spectrum activity. However, it can be seen from Figure 3A and Table 1 that peptides of the ACP_I dataset, which possess no anticancer activity, would also be predicted to form surface active and tilted α-helices. This is an important result in that it clearly shows that other factors are involved in determining the efficacy and selectivity of ACP_{CT} and ACP_{GT} peptides and strongly supports our previous work where it was suggested that the anticancer activity of ACPs was determined by the interplay of a range of physiochemical characteristics rather than any single over riding factor [3].

References

1. Hoskin, D. W. & Ramamoorthy, A. (2008) Studies on anticancer activities of antimicrobial peptides, *Biochim Biophys Acta. 1778*, 357-75.
2. Leuschner, C. & Hansel, W. (2004) Membrane disrupting lytic peptides for cancer treatments, *Curr Pharm Des. 10*, 2299-310.
3. Dennison, S. R., Whittaker, M., Harris, F. & Phoenix, D. A. (2006) Anticancer alpha-helical peptides and structure/function relationships underpinning their interactions with tumour cell membranes, *Curr Protein Pept Sci. 7*, 487-99.
4. Papo, N. & Shai, Y. (2005) Host defense peptides as new weapons in cancer treatment, *Cell Mol Life Sci. 62*, 784-90.
5. Dennison, S. R., Harris, F. & Phoenix, D. A. (2005) Are oblique orientated alpha-helices used by antimicrobial peptides for membrane invasion?, *Protein Pept Lett. 12*, 27-9.
6. Dennison, S. R., Harris, F. & Phoenix, D. A. (2007) The interactions of aurein 1.2 with cancer cell membranes, *Biophys Chem. 127*, 78-83.
7. Marcotte, I., Wegener, K. L., Lam, Y. H., Chia, B. C., de Planque, M. R., Bowie, J. H., Auger, M. & Separovic, F. (2003) Interaction of antimicrobial peptides from Australian amphibians with lipid membranes, *Chem Phys Lipids. 122*, 107-20.
8. Lins, L. & Brasseur, R. (2008) Tilted peptides: a structural motif involved in protein membrane insertion?, *J Pept Sci. 14*, 416-22.
9. Lins, L., Decaffmeyer, M., Thomas, A. & Brasseur, R. (2008) Relationships between the orientation and the structural properties of peptides and their membrane interactions, *Biochim Biophys Acta. 1778*, 1537-44.
10. Owen, D. R. (2005) Short bioactive peptides in, Helix BioMedix. Inc., USA.
11. Rost, B. & Sander, C. (1993) Prediction of protein secondary structure at better than 70% accuracy, *J Mol Biol. 232*, 584-99.
12. Harris, F., Wallace, J. & Phoenix, D. A. (2000) Use of hydrophobic moment plot methodology to aid the identification of oblique orientated alpha-helices, *Mol Membr Biol. 17*, 201-7.
13. Eisenberg, D., Weiss, R. M. & Terwilliger, T. C. (1982) The helical hydrophobic moment: a measure of the amphiphilicity of a helix, *Nature. 299*, 371-4.
14. Eisenberg, D., Weiss, R. M., Terwillinger, T. C. & Wilcox, W. (1982) Hydrophobic moment and protein structure, *Farad Symp Chem Soc 17*, 109-120.

15. Kaufman, K. & Rousseeuw, P. J. (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley and Sons, Brisbane.
16. Bechinger, B., Zasloff, M. & Opella, S. J. (1993) Structure and orientation of the antibiotic peptide magainin in membranes by solid-state nuclear magnetic resonance spectroscopy, *Protein Sci. 2*, 2077-84.
17. Tossi, A., Tarantino, C. & Romeo, D. (1997) Design of synthetic antimicrobial peptides based on sequence analogy and amphipathicity, *Eur J Biochem. 250*, 549-58.

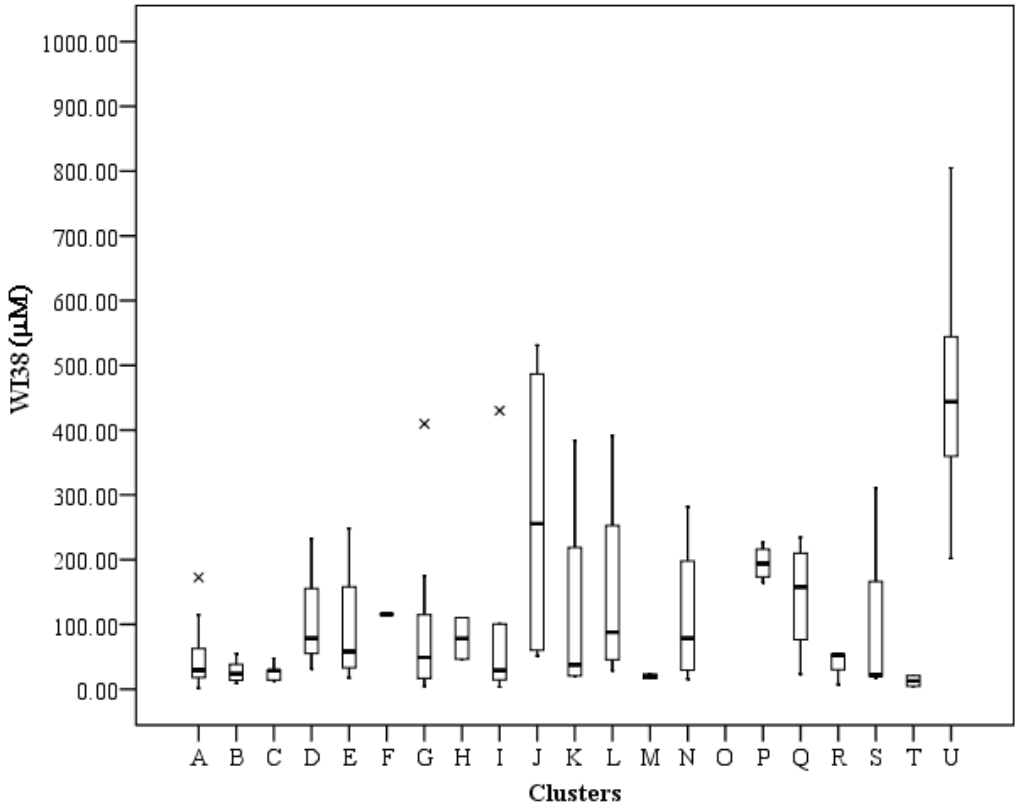

Figure 1. Peptides were grouped by net positive charge, $< \mu_H >$ and $< H >$, as described in the methods to form clusters A to U. For each cluster box plot analysis of the toxicity of ACPs when directed against the W138 cell line is shown. The dark band within each plot indicates median toxicity of the ACPs analysed and positive outliers are annotated (×) are Shive10 AC in cluster A, Modelin-5-COOH in cluster G and FLAK50T8 in cluster I.

Figure 2

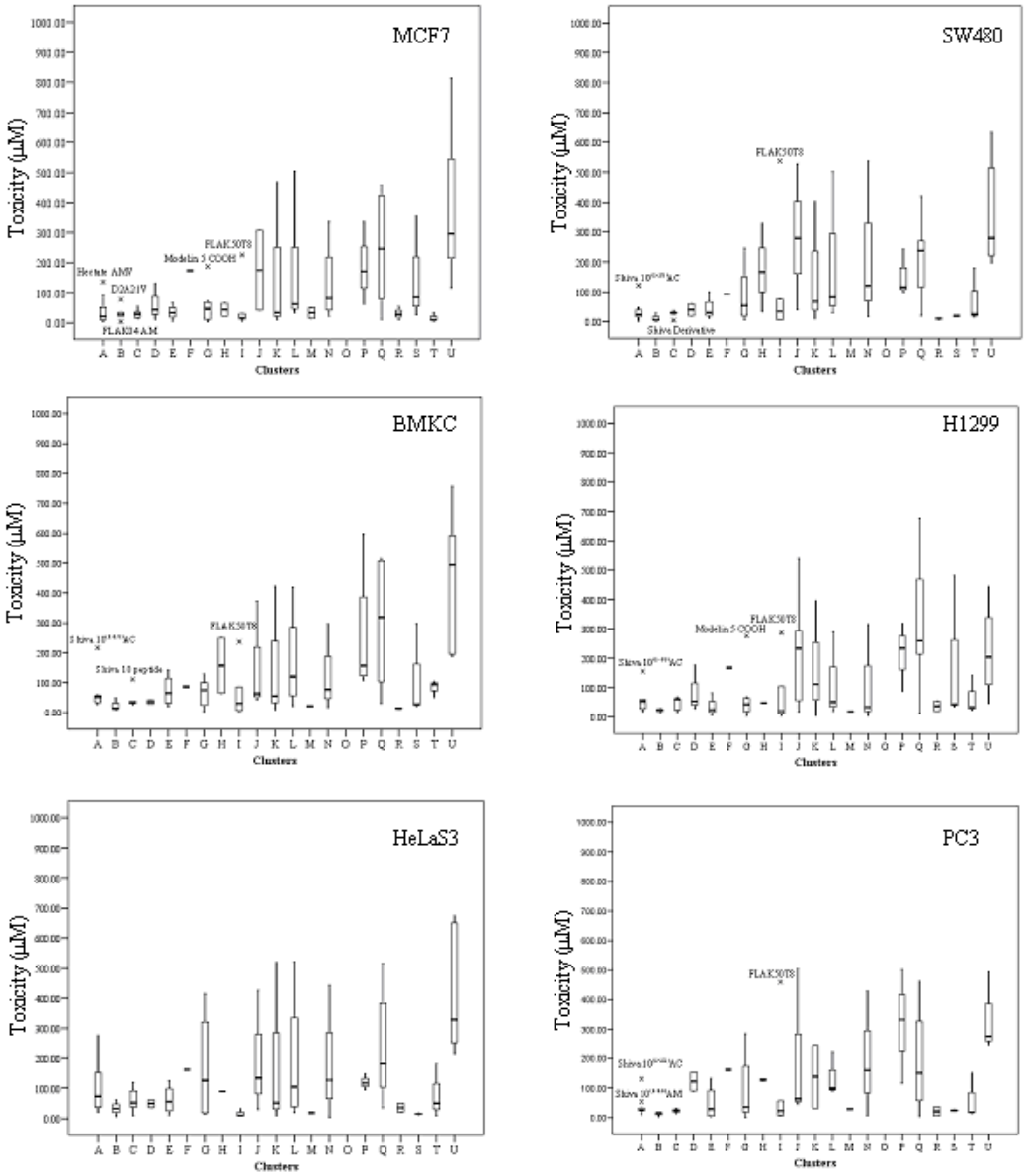

Figure 2. Peptides were grouped by net positive charge, $< \mu_H >$ and $< H >$, as described in the methods to form clusters A to U. For each cluster box plot analysis of the toxicity of ACPs when directed against the different cell line is shown. The dark band within each plot indicates median toxicity of the ACPs analysed and positive outliers are annotated (×).

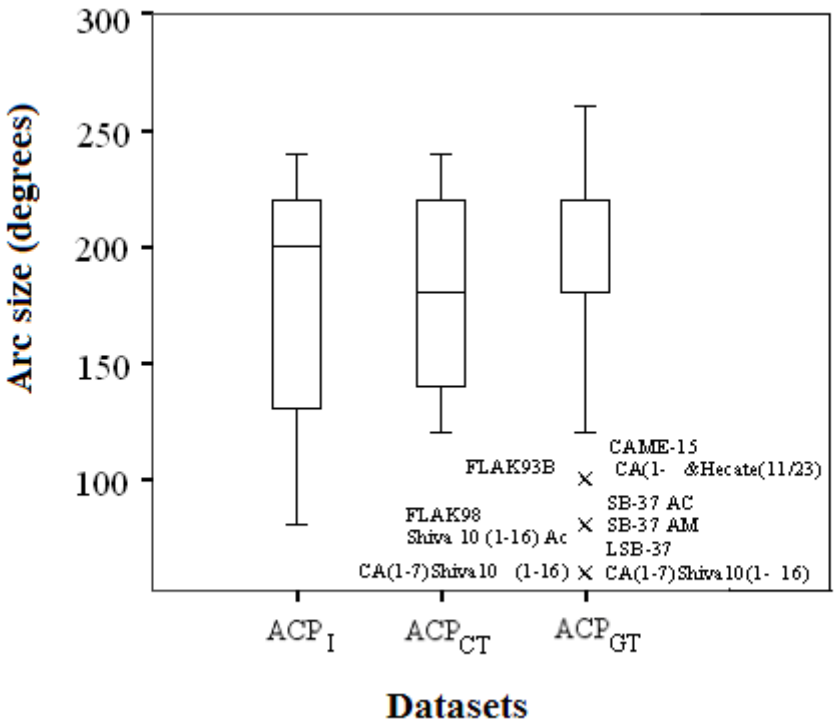

Figure 3: The hydrophobic arc size of peptides within the ACP_I, ACP_{CT} and ACP_{GT} datasets were subjected to box plot analysis. The plot shows the median hydrophobic arc size (dark band) along with the minimum and maximum hydrophobic arc size. The box represents the lower (Q1 = 25%) and upper (Q3 = 75%) quartile range of hydrophobic arc size.

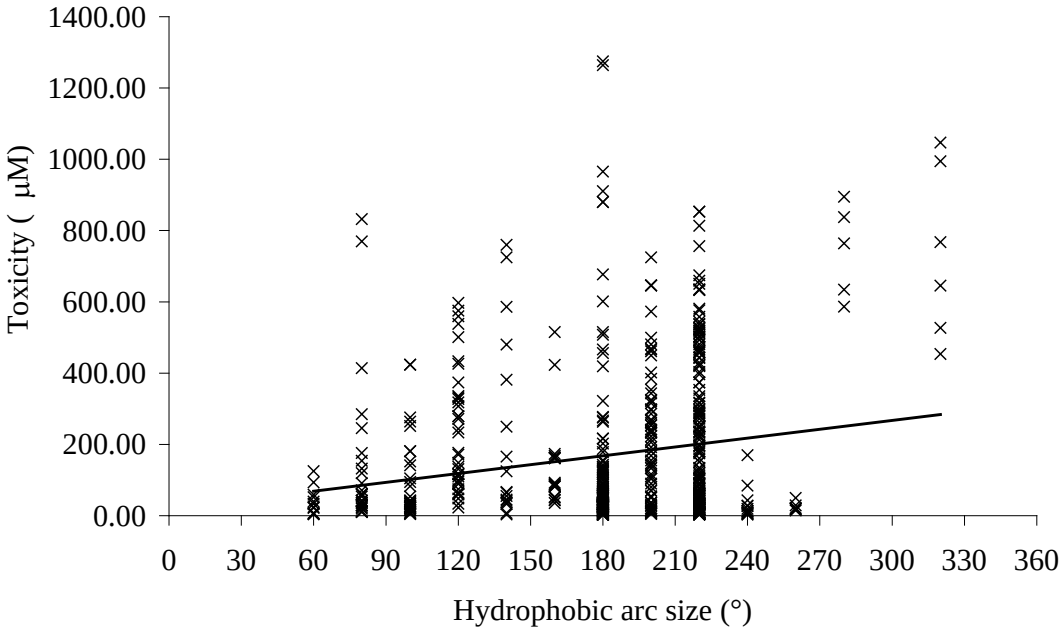

Figure 4 Shown above is a plot of hydrophobic arc size versus the toxicity (LD_{50}) of ACPs in the database for the combined number of cell lines investigated, which were: WI38, MCF7, SW480, BMKC, H1299, HeLaS3 and PC3

Figure 5

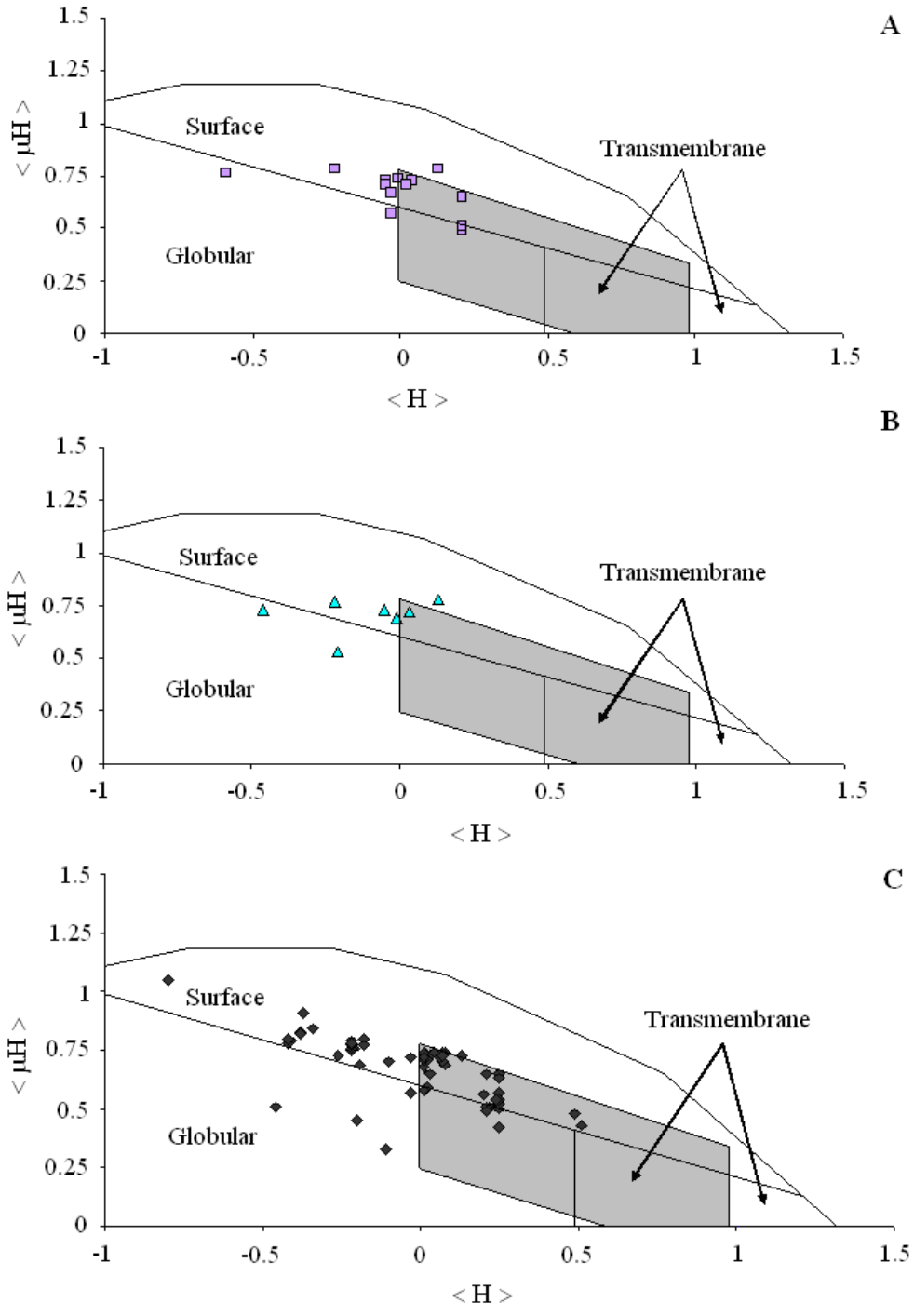

Figure 5. Shown above is extended hydrophobic moment plot analysis of peptides in the ACP_I (A), ACP_{CT} (B) and ACP_{GT} (C) datasets. It can be seen that these data points are mainly distributed over the area predicting surface activity and the shaded area, which identifies candidates for oblique orientated α-helix formation.

Predicted protein type	**ACP_I**	**ACP_{CT}**	**ACP_{GT}**
Surface active	7	4	57
Globular	2	2	4
Tilted	6	1	49

Table 1. Summarised above are the classification of peptides > 11 residues in the ACP_I, ACP_{CT} and ACP_{GT} datasets (Figure 7, A-C) when analysed according to extended hydrophobic moment plot methodology, all as described above.