**Recent Advances in Synthesis and Chemical Biology VI**
**14th December, 2007**
**Trinity College Dublin**

# IS AMINO ACIDS HYDROPHOBICITY A MATTER OF SCALE?

Sorana D. BOLBOACĂ[1] and Lorentz JÄNTSCHI[2]

[1]*„Iuliu Haţieganu" University of Medicine and Pharmacy Cluj-Napoca, 13 Emil Isac, 400023 Cluj-Napoca, Romania, http://sorana.academicdirect.ro*
[2]*Technical University of Cluj-Napoca, 15 Constantin Daicoviciu, 400020 Cluj-Napoca, Romania, http://lori.academicdirect.org*

## ABSTRACT

Hydrophobic/hydrophilic character of an amino acid, an important property in protein structure and protein-protein interactions, is one of the most studied properties of amino-acids. To date, many hydrophobicity scales have been reported, the differences between them being significant: Janin (1979) and Kyte and Doolittle (1982) classify cistein as the most hydrophobic while Wolfenden et al. (1981) or Rose et al. (1985) do not. The aim of the research was to perform the structural modelling of amino acids hydrophobicities in order to identify and characterize the relationships between the structure and the property, and to quantify these relationships on a series of twenty-four scales.

## DATA SETS AND METHODOLOGY

The sample of twenty essential amino-acids (alanine, arginine, asparagine, aspartate, cysteine, glutamine, glutamate, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine) has been investigated by using the Molecular Descriptors Family on Structure-Property Relationships approach [1]. The property of interest was represented by hydrophobic or hydrophilic character measured on twenty-four different scales [2] (see Table I). The amino acids structural information has been used in order to generate and calculate the Molecular Descriptors Family [3]. The best performing models in terms of goodness-of-fit were collected and analyzed [4]. The resulted models have been used in order to predict the hydrophobicity of a sample of eleven amino acids (seleno-L-cysteine, pyrrolysine, lanthionine, 2-aminoisobutyric acid, dehydroalanine, gamma-aminobutyric acid, ornithine, citrulline, homocysteine, hydroxyproline, and dopamine)

Table I. Amino acid hydrophobicity scales

| Scale - Set abb. | Scale - Set abb. |
|---|---|
| Black *et al.*, 1991 - Hyd_01 | Bull-Breese, 1974 - Hyd_13 |
| Kyte-Doolittle, 1982 - Hyd_02 | Roseman, 1988 - Hyd_14 |
| Wimley-White, 1996 - Hyd_03 | Welling *et al.*, 1985 - Hyd_15 |
| Hessa *et al.*, 2005 - Hyd_04 | Parker *et al.*, 1986 - Hyd_16 |
| Sereda *et al.*, 1994 - Hyd_05 | Cowan-Whittaker, 1990 - Hyd_17 |
| Hopp-Woods, 1981 - Hyd_06 | Manavalan-Ponnuswamy, 1978 - Hyd_18 |
| Cornette *et al.*, 1987 - Hyd_07 | Fauchere-Pliska, 1983 - Hyd_19 |
| Eisenberg *et al.*, 1984 - Hyd_08 | Rao-Argos, 1986 - Hyd_20 |
| Janin, 1979 - Hyd_09 | Wilson *et al.*, 1981 - Hyd_21 |
| Rose *et al.*, 1985 - Hyd_10 | Cowan-Whittaker, 1990 - Hyd_22 |
| Engelman *et al.*, 1986 - Hyd_11 | Urry, 2004 - Hyd_23 |
| Sweet-Eisenberg, 1983 - Hyd_12 | Monera *et al.*, 1995 - Hyd_24 |

## RESULTS

The interrelation between hydrophobicity of essential amino acids and their structural information was investigated. One monovariate model was obtained for each hydrophobicity scale. The model and its characteristics for each hydrophibicity scale are:

| Amino acid property | Welling et all., 1985 | Wilson et all., 1981 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = -1.23 + x \cdot 0.39$ | $\hat{y} = 6.55 + x \cdot (-27.79)$ |
| SPR Determination (%) | 44 | 66 |
| MDF Descriptor (x) | amMRLQt | immRoQg |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Bonds (topology) | Space (geometry) |
| Interaction Model | $Q \cdot d$ | $Q^{-1}$ |
| Structure on Property Scale | Inversed | Inversed |
| | $n = 20$<br>$r = 0.6649$; $F = 14^*$; $s = 1.21$<br>$r_{loo} = 0.5961$; $F_{loo} = 7^*$; $s_{loo} = 1.37$ | $n = 20$<br>$r = 0.8163$; $F = 36^*$; $s = 2.19$<br>$r_{loo} = 0.7740$ ; $F_{loo} = 27^*$; $s_{loo} = 2.41$ |

| Amino acid property | Cornette et all., 1987 | Wimley-White, 1996 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = -2.88 + x \cdot (-1.73)$ | $\hat{y} = -3.37 + x \cdot 7.35$ |
| SPR Determination (%) | 69 | 71 |
| MDF Descriptor (x) | LmDROQg | iBmrWQt |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Bonds (topology) |
| Interaction Model | Q | $Q^2/d$ |
| Structure on Property Scale | Logarithmic | Inversed |
| | $n = 20$<br>$r = 0.8309$; $F = 40^*$; $s = 1.70$<br>$r_{loo} = 0.7936$; $F_{loo} = 30^*$; $s_{loo} = 1.87$ | $n = 20$<br>$r = 0.8434$; $F = 44^*$; $s = 0.48$<br>$r_{loo} = 0.8009$ ; $F_{loo} = 32^*$; $s_{loo} = 0.54$ |

| Amino acid property | Hopp-Woods, 1981 | Cowan-Whittaker, 1990 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = -1.99 + x \cdot 10.63$ | $\hat{y} = 1.47 + x \cdot (-6.57)$ |
| SPR Determination (%) | 74 | 75 |
| MDF Descriptor (x) | iMPRoQg | AmDROQg |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Space (geometry) |
| Interaction Model | $Q^{-1}$ | Q |
| Structure on Property Scale | Inversed | Absolute |
| | $n = 20$<br>$r = 0.8608$; $F = 52^*$; $s = 1.01$<br>$r_{loo} = 0.8288$; $F_{loo} = 39^*$; $s_{loo} = 1.11$ | $n = 20$<br>$r = 0.8661$; $F = 54^*$; $s = 0.66$<br>$r_{loo} = 0.8344$; $F_{loo} = 41^*$; $s_{loo} = 0.73$ |

| Amino acid property | Manavalan-Ponnuswamy, 1978 | Fauchere-Pliska, 1983 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = 14.55 + x \cdot 23.43$ | $\hat{y} = -4.36 + x \cdot 5.94$ |
| SPR Determination (%) | 78 | 78 |
| MDF Descriptor (x) | inMrpQg | ibDRPQg |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Space (geometry) |
| Interaction Model | $Q^{-2}$ | $Q^2$ |
| Structure on Property Scale | Inversed | Inversed |
| | $n = 20$<br>$r = 0.8814$; $F = 63^*$; $s = 0.76$<br>$r_{loo} = 0.8546$; $F_{loo} = 49^*$; $s_{loo} = 0.84$ | $n = 20$<br>$r = 0.8832$; $F = 65^*$; $s = 0.50$<br>$r_{loo} = 0.8611$; $F_{loo} = 51^*$; $s_{loo} = 0.54$ |

| Amino acid property | Rao-Argos, 1986 | Janin, 1979 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = 1.43 + x \cdot (-2.73)$ | $\hat{y} = 0.86 + x \cdot 1.74$ |
| SPR Determination (%) | 79 | 81 |
| MDF Descriptor (x) | AmDROQg | inMrpQg |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Space (geometry) |
| Interaction Model | Q | $Q^{-2}$ |
| Structure on Property Scale | Proportional | Inversed |
| | $n = 20$<br>$r = 0.8901$; $F = 69^*$; $s = 0.24$<br>$r_{loo} = 0.8545$; $F_{loo} = 48^*$; $s_{loo} = 0.28$ | $n = 20$<br>$r = 0.8974$; $F = 74^*$; $s = 0.05$<br>$r_{loo} = 0.8744$; $F_{loo} = 58^*$; $s_{loo} = 0.06$ |

| Amino acid property | Roseman, 1988 | Rose et all., 1985 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = -3.36 + x \cdot 3.76$ | $\hat{y} = 0.48 + x \cdot (-137.72)$ |
| SPR Determination (%) | 81 | 81 |
| MDF Descriptor (x) | iBDdwQg | IHPrFQt |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Bonds (topology) |
| Interaction Model | $Q^2/d$ | $Q^2/d^2$ |
| Structure on Property Scale | Inversed | Logarithmic |
| | $n = 20$<br>$r = 0.8986$; $F = 75^*$; $s = 0.45$<br>$r_{loo} = 0.8812$; $F_{loo} = 62^*$; $s_{loo} = 0.48$ | $n = 20$<br>$r = 0.8997$; $F = 76^*$; $s = 0.32$<br>$r_{loo} = 0.8599$; $F_{loo} = 56^*$; $s_{loo} = 0.36$ |

| Amino acid property | Urry, 2004 | Engelman et all., 1986 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = -29.73 + x \cdot (-11.96)$ | $\hat{y} = 1.85 + x \cdot (-753.09)$ |
| SPR Determination (%) | 82 | 83 |
| MDF Descriptor (x) | iBDMkEt | INPrWQg |
| Dominant Atomic Property | Electronegativity (E) | Charge (Q) |
| Interaction via | Bonds (topology) | Space (geometry) |
| Interaction Model | $Q^2 \cdot d^{-1}$ | $Q^2/d$ |
| Structure on Property Scale | Inversed | Logarithmic |
| | $n = 20$<br>$r = 0.9047$; $F = 81^*$; $s = 1.07$<br>$r_{loo} = 0.8819$; $F_{loo} = 63^*$; $s_{loo} = 1.18$ | $n = 20$<br>$r = 0.9116$; $F = 89^*$; $s = 2.07$<br>$r_{loo} = 0.8731$; $F_{loo} = 51^*$; $s_{loo} = 2.56$ |

| Amino acid property | Eisenberg et all., 1984 | Cowan-Whittaker, 1990 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = 1.68 + x \cdot (-0.92)$ | $\hat{y} = 4.64 + x \cdot (-2.16)$ |
| SPR Determination (%) | 83 | 84 |
| MDF Descriptor (x) | IAMdKQg | IbmdKQg |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Space (geometry) |
| Interaction Model | $Q^2 \cdot d$ | $Q^2 \cdot d$ |
| Structure on Property Scale | Logarithmic | Logarithmic |
| | $n = 20$<br>$r = 0.9128$; $F = 90^*$; $s = 0.42$<br>$r_{loo} = 0.8935$; $F_{loo} = 70^*$; $s_{loo} = 0.46$ | $n = 20$<br>$r = 0.9182$; $F = 97^*$; $s = 0.52$<br>$r_{loo} = 0.8984$; $F_{loo} = 75^*$; $s_{loo} = 0.58$ |

| Amino acid property | Roseman, 1988 | Sereda et all., 1994 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = 5.30 + x \cdot (-3.78)$ | $\hat{y} = 81.72 + x \cdot 817.95$ |
| SPR Determination (%) | 85 | 85 |
| MDF Descriptor (x) | IAmrLQg | inMrpQg |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Space (geometry) |
| Interaction Model | $Q \cdot d$ | $Q^{-2}$ |
| Structure on Property Scale | Logarithmic | Inversed |
| | $n = 20$<br>$r = 0.9208$; $F = 100^*$; $s = 0.80$<br>$r_{loo} = 0.9073$; $F_{loo} = 84^*$; $s_{loo} = 0.86$ | $n = 20$<br>$r = 0.9232$; $F = 104^*$; $s = 20.73$<br>$r_{loo} = 0.9082$; $F_{loo} = 85^*$; $s_{loo} = 22.58$ |

| Amino acid property | Hessa et all., 2005 | Bull-Breese, 1974 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = -0.41 + x \cdot 7.18$ | $\hat{y} = 1.36 + x \cdot (-0.20)$ |
| SPR Determination (%) | 85 | 85 |
| MDF Descriptor (x) | AmDROQg | iIPmLQt |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Bonds (topology) |
| Interaction Model | Q | $Q \cdot d$ |
| Structure on Property Scale | Proportional | Inversed |
| | $n = 20$<br>$r = 0.9238$; $F = 105^*$; $s = 0.32$<br>$r_{loo} = 0.9018$; $F_{loo} = 78^*$; $s_{loo} = 0.58$ | $n = 20$<br>$r = 0.9252$; $F = 107^*$; $s = 0.36$<br>$r_{loo} = 0.9003$; $F_{loo} = 75^*$; $s_{loo} = 0.42$ |

| Amino acid property | Parker et all., 1986 | Kyte-Doolittle, 1982 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = 11.05 + x \cdot 1.85$ | $\hat{y} = -7.60 + x \cdot 19.17$ |
| SPR Determination (%) | 86 | 87 |
| MDF Descriptor (x) | IfPROQg | iGPdLQg |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Space (geometry) |
| Interaction Model | Q | $d \cdot \sqrt{Q}$ |
| Structure on Property Scale | Logarithmic | Inversed |
| | $n = 20$<br>$r = 0.9259$; $F = 108^*$; $s = 2.46$<br>$r_{loo} = 0.8935$; $F_{loo} = 69^*$; $s_{loo} = 2.97$ | $n = 20$<br>$r = 0.9327$; $F = 120^*$; $s = 1.11$<br>$r_{loo} = 0.9226$; $F_{loo} = 103^*$; $s_{loo} = 1.18$ |

| Amino acid property | Black et all., 1991 | Monera et all., 1995 |
|---|---|---|
| MDF SPR Equation | $\hat{y} = 0.86 + x \cdot (-0.96)$ | $\hat{y} = 86.05 + x \cdot 843.88$ |
| SPR Determination (%) | 88 | 90 |
| MDF Descriptor (x) | IAmrLQg | inMrpQg |
| Dominant Atomic Property | Charge (Q) | Charge (Q) |
| Interaction via | Space (geometry) | Space (geometry) |
| Interaction Model | $d \cdot \sqrt{Q}$ | $Q^{-2}$ |
| Structure on Property Scale | Proportional | Inversed |
| | $n = 20$<br>$r = 0.9376$; $F = 131^*$; $s = 0.12$<br>$r_{loo} = 0.9263$; $F_{loo} = 109^*$; $s_{loo} = 0.13$ | $n = 19$ (-Prolina)<br>$r = 0.9504$; $F = 159^*$; $s = 16.49$<br>$r_{loo} = 0.9382$; $F_{loo} = 125^*$; $s_{loo} = 18.37$ |

The summary of characters in molecular descriptors used by the above-presented models is presented in Table II. Based on the obtained models, the hydrophobicity of 11 non-standard amino acids was obtained and is presented in Table III.

*Table II. Character distribution in descriptors name*

| 1st letter | | 2nd letter | | 3rd letter | | 4th letter | | 5th letter | | 6th letter | | 7th letter | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cha. | $f_a$ | Cha. | $f_a$ | Cha. | $f_a$ | Cha. | $f_a$ | Cha. | $f_a$ | Cha. | $f_a$ | Cha. | $f_a$ |
| A | 4 | A | 3 | D | 7 | d | 4 | F | 1 | E | 1 | g | 19 |
| i | 16 | B | 5 | m | 11 | m | 2 | K | 3 | Q | 23 | t | 5 |
| l | 4 | f | 1 | P | 6 | r | 18 | L | 5 | | | | |
| | | G | 1 | | | | | O | 7 | | | | |
| | | H | 1 | | | | | p | 5 | | | | |
| | | I | 1 | | | | | W | 3 | | | | |
| | | m | 7 | | | | | | | | | | |
| | | n | 5 | | | | | | | | | | |

Cha. = character; $f_a$ = absolute frequency

*Table III. 11 non-standard amino acids: predicted hydrophobicity*

| aa | Hyd_01 | Hyd_02 | Hyd_03 | Hyd_04 | Hyd_05 | Hyd_06 | Hyd_07 | Hyd_08 | Hyd_09 | Hyd_10 | Hyd_11 | Hyd_12 | Hyd_13 | Hyd_14 | Hyd_15 | Hyd_16 | Hyd_17 | Hyd_18 | Hyd_19 | Hyd_20 | Hyd_21 | Hyd_22 | Hyd_23 | Hyd_24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Aib | 0.71 | -9.91 | -0.22 | 0.14 | 143.07 | -1.2 | 1.56 | 0.68 | 0.99 | 0.45 | 1.69 | 0.46 | -0.27 | 0.9 | 3.58 | -1.77 | 0.94 | 16.31 | 0.63 | 1.22 | -2.82 | 0.96 | -1.87 | 149.34 |
| Cit | -0.07 | -19.45 | -0.81 | 2.95 | 105.51 | 2.98 | -1.57 | -2.5 | 0.91 | -0.83 | -8.25 | -0.73 | 0.56 | -4.59 | -1.02 | 6.16 | -2.38 | 15.23 | -0.74 | 0.15 | 0.38 | -1.61 | 0.88 | 110.59 |
| Dhd | 0.51 | -11.39 | -0.16 | 0.78 | 232.95 | -0.23 | 0.22 | 0.72 | 1.18 | -0.07 | -0.3 | 0.21 | 0.64 | -0.13 | -0.8 | 4.04 | 0.73 | 18.89 | 0.38 | 0.98 | -2.76 | 0.38 | -1.33 | 242.07 |
| Dop | 0.79 | -14.31 | 5.04 | 1.09 | 98.01 | -0.99 | -0.18 | 0.15 | 0.9 | 0.17 | 1.38 | 3.94 | 0.07 | 1.25 | -0.97 | 1.36 | 1.19 | 15.02 | 2.23 | 0.86 | 5.66 | 0.09 | -5.41 | 102.85 |
| Gab | 0.5 | -10.67 | 0.14 | 0.22 | 93.32 | -1.07 | 1.33 | 0.16 | 0.89 | 0.46 | 1.67 | 0.85 | 0.41 | -0.18 | 3.27 | 4.53 | 0.33 | 14.89 | 0.46 | 1.19 | 0.14 | 0.89 | 0 | 98.02 |
| Hcy | 0.54 | -11.47 | 0.06 | -0.29 | 96.35 | -1.83 | 4.3 | 0.39 | 0.89 | 0.49 | 1.85 | 0.84 | 0.22 | 0.03 | -1.2 | 1.12 | 0.44 | 14.97 | 1.1 | 1.39 | 5.78 | 1.36 | -1.48 | 101.15 |
| Hy | 0.33 | -11.88 | -0.42 | 0.2 | 159.64 | -1.66 | 1.38 | -0.07 | 1.03 | 0.49 | 1.84 | -0.09 | 0.7 | -1.23 | -1.1 | 1.63 | -0.37 | 16.79 | 1.63 | 1.2 | 0.53 | 0.91 | -3.27 | 166.44 |
| Lth | -0.21 | -23.87 | -1.04 | 3.19 | 353.92 | 3.33 | -1.69 | -3.54 | 1.44 | -1.32 | -15.63 | -1.31 | 0.6 | -6.18 | -1.14 | 7.63 | -2.89 | 22.35 | -0.57 | 0.06 | 4.56 | -1.83 | -0.17 | 366.88 |
| Oth | 0.38 | -13.92 | 0.07 | 2.35 | 98.86 | 2.09 | -1.23 | -0.61 | 0.9 | -1.33 | -5.84 | 0.31 | 0.4 | -0.92 | -0.53 | 5.57 | -0.35 | 15.04 | 0.34 | 0.38 | 0.11 | -1.06 | -0.96 | 103.74 |
| Pyl | 0.25 | -33.21 | -0.78 | 2.98 | 119.29 | 0.63 | -1.59 | -4.71 | 0.94 | -0.72 | -5.7 | -2.42 | 0.53 | -1.78 | -1.17 | 17.39 | -2.05 | 15.63 | -3.58 | 0.14 | 1.25 | -1.64 | -2.66 | 124.81 |
| Sec | 0.18 | -12.87 | -1.14 | 10.97 | 145.69 | 5.19 | -3.68 | -2.03 | 1 | -13.33 | -67.94 | -1.32 | 0.58 | -2.37 | -0.59 | 10.79 | -0.87 | 16.39 | -2.86 | -2.9 | 5.18 | -8.96 | -0.39 | 152.04 |
| Pro | | | | | | | | | | | | | | | | | | | | | | | | 96.57 |

## CONCLUSION

The amino acids hydrophobicity is a property linear related with compounds structure, being on strong relationship with atomic charge through geometry interaction, this fact being shown by the large majority of the scales.

## QUESTIONS THAT ARISES

First compelling question that arises from the investigation of amino acids hydrophobicity scales is as follows: "How reliable the predicted values are?" To answer to this question, future research will require. A study must be conduct in order to measure the hydrophobicity of the sample of non-standard amino acids. Once the measurements are done, the experimental values could be comparing with predicted values.
Second compelling question that arises from this investigation is: "Is it possible to rescale the hydrophobicity scales?" This task could be made by taking into consideration the scale which in most of the cases give the minimum values for hydrophobicity (Sereda *et al.*, 1994), the scale that give the maximum values in most of the cases (Manavalan-Ponnuswamy, 1978), a middle scale, and considering the QSPR models obtained in the present research and the confidence intervals for intercepts and slopes. This will require future research.

## REFERENCES

1. Jäntschi L, Bolboacă S.D. Modeling the octanol-water partition coefficient of substituted phenols by the use of structure information. Int. J. Quant. Chem. 2007;107:1736-1744.
2. Bolboacă SD, Jäntschi L. Modeling Analysis of Amino Acids Hydrophobicity. ChemMod - Chemical Graph Theory and Molecular Modeling Workshop, October 23-26, Cluj-Napoca, Romania, 2007, p. 6.
3. Bolboacă SD, Jäntschi L. Modelling the Property of Compounds from Structure: Statistical Methods for Models ValidationEnv. Chem. Lett. 2007, DOI 10.1007/s10311-007-0119-9.
4. Bolboacă SD, Jäntschi L. From Structural Chemistry to Structural Biology through Information Science Study on Collagen. Chem.Biol. & Drug Design. In press, 2008.

**The research was partly supported by UEFISCSU Romania through grants.**

# IS AMINO ACIDS HYDROPHOBICITY A MATTER OF SCALE?

Sorana D. BOLBOACĂ and Lorentz JÄNTSCHI
*"Iuliu Haţieganu" Medicine and Pharmacy University and Technical University, Cluj-Napoca, 400349 and 400641, Romania*

*email: sbolboaca@umfcluj.ro, lori@academicdirect.org*

Amino acids, the building blocks of proteins, molecule that contains amine and carboxyl functional groups, play important roles in biology such as: synthesis of proteins, intermediates of metabolic pathways, neurotransmitters, antibiotics, etc. Even if there are considered the essential amino acids, the amino acids found in biological systems, abiotically synthesized, or those engineered by scientists, the quantitative investigations of the structure-property relationships of amino acids are important for biological researches.

Hydrophobic or hydrophilic character of an amino acid, the property important in protein structure and protein-protein interactions, is one of the most studied properties of amino-acids. To date, many hydrophobicity scales have been reported. Moreover, the differences between scales are significant: Janin (1979) and Kyte and Doolittle (1982) classify cistein as the most hydrophobic while Wolfenden et al. (1981) or Rose et al. (1985) do not. These differences could be explained by the fundamentally different methods used for constructing the scale.

The aim of the research was to perform a structural modelling analysis of amino acids hydrophobicity in order to identify and characterize if there exists a relationship between the structure and the property, and to quantify these relationship. A sample of twenty essential amino-acids (alanine, arginine, asparagine, aspartate, cysteine, glutamine, glutamate, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine) has been investigated by using the Molecular Descriptors Family on Structure-Property Relationships approach. [1] The property of interest was represented by hydrophobic or hydrophilic character measured on twenty-four different scales. [2] The amino acids structural information has been used in order to generate and calculate the Molecular Descriptors Family. [3] The best performing models in terms of goodness-of-fit were collected and analyzed. [4] The identified monovariate models were significant statistically ($p < 0.0001$). An internal validation approach was applied for analyzing the validity of the obtained models. The correlation coefficient of the measured and estimated hydrophobicity varied according with the method from 0.6649 (hydrophobicity reported by Welling et al. 1985) to 0.9504 (hydrophobicity reported by Monera et al., 1995). The resulted models have been used in order to predict the hydrophobicity of a sample of eleven amino acids (seleno-L-cysteine, pyrrolysine, lanthionine, 2-aminoisobutyric acid, dehydroalanine, gamma-aminobutyric acid, ornithine, citrulline, homocysteine, hydroxyproline, and dopamine). The obtained results shown that the amino acids hydrophobicity is a property linear related with compounds structure, being on strong relationship with atomic charge through geometry interaction.

**References:**

[1] L. Jäntschi, S. D. Bolboacă, *Int. J. Quant. Chem.* 2007, *107*, 1736.
[2] S. D. Bolboacă, L. Jäntschi, in *ChemMod Workshop*, Cluj-Napoca, 2007, p. 6.
[3] S. D. Bolboacă, L. Jäntschi, *Env. Chem. Lett.* 2007, DOI 10.1007/s10311-007-0119-9.
[4] S. D. Bolboacă, L. Jäntschi, *Chem.Biol. & Drug Design* In press.

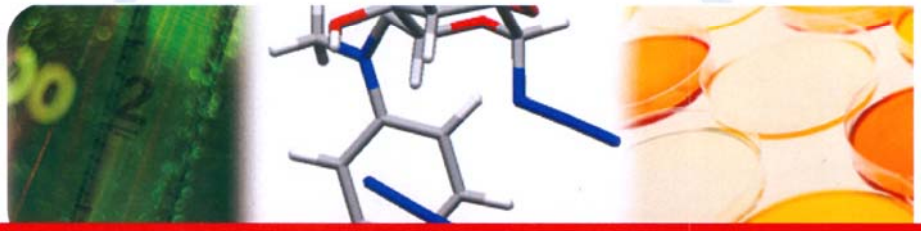
CSCB

Centre for Synthesis
& Chemical Biology

# Recent Advances in Synthesis & Chemical Biology VI

Symposium
14th December 2007

MacNeill Lecture Theatre
University of Dublin
Trinity College
Dublin 2



CSCB
Centre for Synthesis & Chemical Biology

RCSI Labs

TCD Labs

UCD Labs

RCSI

UCD DUBLIN

Funding for the establishment of the Centre for Synthesis and Chemical Biology (CSCB) was approved by the Higher Education Authority's Programme for Research in Third Level Instututions in December 2001. The CSCB assembles researchers in the chemical sciences from University College Dublin (UCD), Trinity College Dublin (TCD) and the Royal College of Surgeons in Ireland (RCSI).

Centre for Synthesis and Chemical Biology,
University College Dublin, Belfield, Dublin 4, Ireland
Tel: +353-1-716 2302 Fax: +353-1-716 2501 Email: cscb@ucd.ie
Web: www.ucd.ie/cscb

NDP NATIONAL DEVELOPMENT PLAN

HEA Higher Education Authority An tÚdarás um Ard-Oideachas

European Union Structural Funds

Funded under the Programme for Research in Third-Level Institutions (PRTLI), Administered by the HEA