**Supplementary Materials**

**Exploring the structural aspects of ureido amino acid-based APN inhibitors: A validated comparative multi-QSAR modeling study**

Suvankar Banerjee, Sk. Abdul Amin, Sandip Kumar Baidya, Nilanjan Adhikari, Tarun Jha\*

Natural Science Laboratory, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, P. O. Box 17020, Jadavpur University, Kolkata 700032, West Bengal, India

\*Correspondence should be addressed to Tarun Jha, Natural Science Laboratory, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, Jadavpur University, Kolkata – 700032, India. Email: tjupharm@yahoo.com

**Keywords**: Aminopeptidase-N; APN inhibitors, 2D-QSAR; LDA; Bayesian classification model.

**Supplementary Table S1.** Compound table of hydroxamate-based ureido amino acid derivatives along with their HDAC8 IC50 and *pIC50* values.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | | | | |
| **Cpd** | **R1** | **R2** | **R3** | **IC50(nM)** | ***pIC50*** |
| 1*a* | CONHOH | (CH2)4NHCOOBnz | Bnz | 16,400 | 4.785 |
| 2 | CONHOH | (CH2)4NHCOOBnz | CH2CONHOH | 20,400 | 4.690 |
| 3 | CONHOH | (CH2)4NHCOOBnz | CH(CH3)CONHOH | 19,900 | 4.701 |
| 4 | CONHOH | (CH2)4NHCOOBnz | CH(CH2CH(CH3)2)CONHOH | 4,510 | 5.346 |
| 5 | CONHOH | (CH2)4NHCOOBnz | CH(CH2Ph)CONHOH | 14,500 | 4.839 |
| 6 | CONHOH | (CH2)4NHCOOBnz | CH(CH2CH2SCH3)CONHOH | 14,400 | 4.842 |
| 7 | H | CONHOH | Bnz | 32,000 | 4.495 |
| 8*a* | *(S)-*CH3 | CONHOH | Bnz | 21,000 | 4.678 |
| 9 | *(S)-*CH(CH3)2 | CONHOH | Bnz | 17,000 | 4.770 |
| 10 | *(S)-*H(CH3)CH2CH3 | CONHOH | Bnz | 4,100 | 5.387 |
| 11*a* | *(S)-*Bnz | CONHOH | Bnz | 5,600 | 5.252 |
| 12*a* | *(S)-* (CH2)2SCH3 | CONHOH | Bnz | 3,800 | 5.420 |
| 13 | H | CONHOH | (CH2)2Ph | 52,000 | 4.284 |
| 14 | *(S)-CH3* | CONHOH | (CH2)2Ph | 21,000 | 4.678 |
| 15*a* | *(S)-*CH2CH(CH3)2 | CONHOH | (CH2)2Ph | 1,100 | 5.959 |
| 16 | *(S)-*Bnz | CONHOH | (CH2)2Ph | 6,500 | 5.187 |
| 17 | *(S)*- CH2CH(CH3)2 | CONHOH | (CH2)3Ph | 3,400 | 5.469 |
| 18 | *(S)*-Ph | CONHOH | 4-F-Bnz | 5,000 | 5.301 |
| 19 | *(S)*- (CH2)2SCH3 | CONHOH | (CH2)3Ph | 7,000 | 5.155 |
| 20 | *(S)*-Ph | CONHOH | (CH2)3Ph | 2,200 | 5.658 |
| 21 | *(S)*-Bnz | CONHOH | (CH2)3Ph | 2,800 | 5.553 |
| 22 | *(S)*-Ph | CONHOH | Bnz | 980 | 6.009 |
| 23 | *(S)*-Ph | CONHOH | (CH2)2Ph | 1,900 | 5.721 |
| 24 | *(S)*- CH2CH(CH3)2 | CONHOH | Ph | 2,700 | 5.569 |
| 25 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-Furylmethyl | 2,300 | 5.638 |
| 26 | *(S)*- CH2CH(CH3)2 | CONHOH | Napthalenyl | 99 | 7.004 |
| 27 | *(S)*- CH2CH(CH3)2 | CONHOH | 1,1'-biphenyl | 2,100 | 5.678 |
| 28 | *(S)*- CH2CH(CH3)2 | CONHOH | 4-OMePh | 570 | 6.244 |
| 29*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 4-FPh | 2,500 | 5.602 |
| 30 | *(S)*- CH2CH(CH3)2 | CONHOH | 4-MePh | 330 | 6.481 |
| 31 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-CH2Thiophene | 6,60 | 6.180 |
| 32 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-OMeBnz | 570 | 6.244 |
| 33*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 3-OMeBnz | 290 | 6.538 |
| 34 | *(S)*- CH2CH(CH3)2 | CONHOH | 3-F-Bnz | 690 | 6.161 |
| 35*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 4-OMeBnz | 830 | 6.081 |
| 36 | *(S)*- CH2CH(CH3)2 | CONHOH | 4-MeBnz | 410 | 6.387 |
| 37 | *(S)*- CH2CH(CH3)2 | CONHOH | 4-F-Bnz | 730 | 6.137 |
| 38 | *(S)*-Ph | CONHOH | 4-MeBnz | 1,640 | 5.785 |
| 39 | *(S)*-Ph | CONHOH | 4-OMeBnz | 1,030 | 5.987 |
| 40*a* | *(S)*-Ph | CONHOH | 4-FPh | 9,900 | 5.004 |
| 41 | *(S)*-Ph | CONHOH | Ph | 5,900 | 5.229 |
| 42*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 1-Napthylmethyl | 50 | 7.301 |
| 43*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 2-FPh | 520 | 6.284 |
| 44 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-ClPh | 93 | 7.032 |
| 45 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-BrPh | 60 | 7.222 |
| 46 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-MePh | 180 | 6.745 |
| 47*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 2-OMePh | 130 | 6.886 |
| 48 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-CF3Ph | 790 | 6.102 |
| 49 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-OHPh | 11,900 | 4.924 |
| 50 | *(S)*- CH2CH(CH3)2 | CONHOH | 3-FPh | 1,900 | 5.721 |
| 51*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 3-ClPh | 1,660 | 5.780 |
| 52*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 3-BrPh | 580 | 6.237 |
| 53 | *(S)*- CH2CH(CH3)2 | CONHOH | 3-MePh | 240 | 6.620 |
| 54 | *(S)*- CH2CH(CH3)2 | CONHOH | 3-OMePh | 1,360 | 5.866 |
| 55*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 3-CF3Ph | 1,970 | 5.706 |
| 56 | *(S)*- CH2CH(CH3)2 | CONHOH | 3-OHPh | 1,700 | 5.770 |
| 57 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-FBnz | 1,380 | 5.860 |
| 58*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 2-ClBnz | 220 | 6.658 |
| 59 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-BrBnz | 130 | 6.886 |
| 60 | *(S)*- CH2CH(CH3)2 | CONHOH | 2-MeBnz | 900 | 6.046 |
| 61*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 3-ClBnz | 350 | 6.456 |
| 62*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 3-MeBnz | 440 | 6.357 |
| 63 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,6-diMePh | 20 | 7.699 |
| 64 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,6-diEtPh | 30 | 7.523 |
| 65 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,6-diPrPh | 50 | 7.301 |
| 66 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,6-diFPh | 230 | 6.638 |
| 67*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 2,4,6-triMePh | 80 | 7.097 |
| 68 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,4-diMePh | 180 | 6.745 |
| 69 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,4-diFPh | 340 | 6.469 |
| 70 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,4-diClPh | 560 | 6.252 |
| 71 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,4-diOMePh | 460 | 6.337 |
| 72 | *(S)*- CH2CH(CH3)2 | CONHOH | 3-F-4-OMePh | 420 | 6.377 |
| 73 | *(S)*- CH2CH(CH3)2 | CONHOH | 3-Cl-4-MePh | 90 | 7.046 |
| 74*a* | *(S)*- CH2CH(CH3)2 | CONHOH | 2,6-diFBnz | 690 | 6.161 |
| 75 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,6-diClBnz | 580 | 6.237 |
| 76 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,4-diFBnz | 330 | 6.481 |
| 77 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,4-diOMeBnz | 460 | 6.337 |
| 78 | *(S)*- CH2CH(CH3)2 | CONHOH | 3,4-diFBnz | 420 | 6.377 |
| 79 | *(S)*- CH2CH(CH3)2 | CONHOH | 3,4-diMeBnz | 440 | 6.357 |
| 80 | *(S)*- CH2CH(CH3)2 | CONHOH | 2,3-diMeBnz | 70 | 7.155 |

N.B.: *a***:** Test set compounds.

**Supplementary Table S2:** Observed and predicted activities of MLR model-I and Model-II.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Comp. No.*** | ***Set*** | ***Observed activity (pIC50)*** | ***Predicted activity (Model-I)*** | ***Predicted activity (Model-II)*** |
| 1 | Test | 4.785 | 5.123 | 4.937 |
| 2 | Training | 4.690 | 4.694 | 4.578 |
| 3 | Training | 4.701 | 4.494 | 4.685 |
| 4 | Training | 5.346 | 5.474 | 5.514 |
| 5 | Training | 4.839 | 4.942 | 5.386 |
| 6 | Training | 4.842 | 4.837 | 5.078 |
| 7 | Training | 4.495 | 4.617 | 4.345 |
| 8 | Test | 4.678 | 4.465 | 4.930 |
| 9 | Training | 4.770 | 5.445 | 5.427 |
| 10 | Training | 5.387 | 5.621 | 6.083 |
| 11 | Test | 5.252 | 5.618 | 5.517 |
| 12 | Test | 5.420 | 4.981 | 5.255 |
| 13 | Training | 4.284 | 4.229 | 4.086 |
| 14 | Training | 4.678 | 4.220 | 4.668 |
| 15 | Test | 5.959 | 6.091 | 5.621 |
| 16 | Training | 5.187 | 5.238 | 5.305 |
| 17 | Training | 5.469 | 6.056 | 5.745 |
| 18 | Training | 5.301 | 5.659 | 5.683 |
| 19 | Training | 5.155 | 4.926 | 5.268 |
| 20 | Training | 5.658 | 5.590 | 5.544 |
| 21 | Training | 5.553 | 5.421 | 5.460 |
| 22 | Training | 6.009 | 5.713 | 5.615 |
| 23 | Training | 5.721 | 5.656 | 5.329 |
| 24 | Training | 5.569 | 6.054 | 5.949 |
| 25 | Training | 5.638 | 5.663 | 5.432 |
| 26 | Training | 7.004 | 6.862 | 6.519 |
| 27 | Training | 5.678 | 6.100 | 5.943 |
| 28 | Training | 6.244 | 6.009 | 6.078 |
| 29 | Test | 5.602 | 5.787 | 5.780 |
| 30 | Training | 6.481 | 6.217 | 6.405 |
| 31 | Training | 6.180 | 5.951 | 5.632 |
| 32 | Training | 6.244 | 6.296 | 5.976 |
| 33 | Test | 6.538 | 5.859 | 6.001 |
| 34 | Training | 6.161 | 6.027 | 6.084 |
| 35 | Test | 6.081 | 6.157 | 6.033 |
| 36 | Training | 6.387 | 6.427 | 6.297 |
| 37 | Training | 6.137 | 6.257 | 6.028 |
| 38 | Training | 5.785 | 5.733 | 5.937 |
| 39 | Training | 5.987 | 6.017 | 5.640 |
| 40 | Test | 5.004 | 5.458 | 5.478 |
| 41 | Training | 5.229 | 5.503 | 5.395 |
| 42 | Test | 7.301 | 6.572 | 6.148 |
| 43 | Test | 6.284 | 6.277 | 6.197 |
| 44 | Training | 7.032 | 6.558 | 6.579 |
| 45 | Training | 7.222 | 6.699 | 6.935 |
| 46 | Training | 6.745 | 6.530 | 6.658 |
| 47 | Test | 6.886 | 6.552 | 6.244 |
| 48 | Training | 6.102 | 6.275 | 6.127 |
| 49 | Training | 4.924 | 5.401 | 5.822 |
| 50 | Training | 5.721 | 5.892 | 5.974 |
| 51 | Test | 5.780 | 6.134 | 6.154 |
| 52 | Test | 6.237 | 6.341 | 5.915 |
| 53 | Training | 6.620 | 6.669 | 6.470 |
| 54 | Training | 5.866 | 5.905 | 6.025 |
| 55 | Test | 5.706 | 5.965 | 5.999 |
| 56 | Training | 5.770 | 5.293 | 5.606 |
| 57 | Training | 5.860 | 6.145 | 5.969 |
| 58 | Test | 6.658 | 6.329 | 6.323 |
| 59 | Training | 6.886 | 6.494 | 6.631 |
| 60 | Training | 6.046 | 6.589 | 6.440 |
| 61 | Test | 6.456 | 6.500 | 6.350 |
| 62 | Test | 6.357 | 6.949 | 6.494 |
| 63 | Training | 7.699 | 7.359 | 7.124 |
| 64 | Training | 7.523 | 7.338 | 7.468 |
| 65 | Training | 7.301 | 7.299 | 7.553 |
| 66 | Training | 6.638 | 6.724 | 6.612 |
| 67 | Test | 7.097 | 8.282 | 7.545 |
| 68 | Training | 6.745 | 7.198 | 7.080 |
| 69 | Training | 6.469 | 6.304 | 6.414 |
| 70 | Training | 6.252 | 6.599 | 6.990 |
| 71 | Training | 6.337 | 6.450 | 6.465 |
| 72 | Training | 6.377 | 6.119 | 5.953 |
| 73 | Training | 7.046 | 6.363 | 6.426 |
| 74 | Test | 6.161 | 6.143 | 6.109 |
| 75 | Training | 6.237 | 6.401 | 6.566 |
| 76 | Training | 6.481 | 6.180 | 6.127 |
| 77 | Training | 6.337 | 6.272 | 6.121 |
| 78 | Training | 6.377 | 6.266 | 6.130 |
| 79 | Training | 6.357 | 6.804 | 6.873 |
| 80 | Training | 7.155 | 6.858 | 6.681 |

**Supplementary Table S3:** The *t-*value, *p-*value and *VIF* values of the descriptors MLR Model-I and Model-II.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Model*** | ***Parameter*** | ***t-Value*** | ***p-Value*** | ***VIF*** |
| ***Model-I (Equation 3)*** | Intercept | 6.45 | 0.000 | -- |
| *JGI4* | 8.30 | 0.000 | 1.62 |
| *GATS7e* | 6.27 | 0.000 | 1.07 |
| *MLOGP* | 3.84 | 0.000 | 1.60 |
| *MATS5v* | -3.85 | 0.000 | 1.53 |
| *nOHPh* | -2.86 | 0.006 | 1.08 |
| *RDF045v* | -2.83 | 0.007 | 1.41 |
| ***Model-II (Equation 4)*** | Constant | 10.21 | 0.000 | -- |
| *JGI3* | 2.06 | 0.044 | 2.64 |
| *JGI4* | 5.68 | 0.000 | 2.64 |
| *MATS1e* | 5.53 | 0.000 | 1.01 |
| *R5e+* | -2.75 | 0.008 | 1.02 |
| *R4p+* | 3.06 | 0.003 | 1.12 |

**Supplementary Table S4:** Correlation matrix of the variables used in Model-I *(Equation 3)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | *JGI4* | *GATS7e* | *MLOGP* | *MATS5v* | *nOHPh* | *RDF045v* | *pIC50* |
| *JGI4* | 1.00 | 0.19 | 0.39 | -0.55 | 0.06 | 0.21 | 0.79 |
| *GATS7e* | 0.19 | 1.00 | 0.19 | -0.15 | 0.03 | 0.18 | 0.48 |
| *MLOGP* | 0.39 | 0.19 | 1.00 | -0.31 | -0.22 | 0.50 | 0.55 |
| *MATS5v* | -0.55 | -0.15 | -0.31 | 1.00 | 0.01 | -0.32 | -0.62 |
| *nOHPh* | 0.06 | 0.03 | -0.22 | 0.01 | 1.00 | -0.11 | -0.15 |
| *RDF045v* | 0.21 | 0.18 | 0.50 | -0.32 | -0.11 | 1.00 | 0.22 |
| *pIC50* | 0.79 | 0.48 | 0.55 | -0.62 | -0.15 | 0.22 | 1.00 |

**Supplementary Table S5:** Correlation matrix of the variables used in Model-II *(Equation 4)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *JGI3* | *JGI4* | *MATS1e* | *R5e+* | *R4p+* | *pIC50* |
| *JGI3* | 1.00 | 0.78 | 0.03 | -0.11 | 0.31 | 0.71 |
| *JGI4* | 0.78 | 1.00 | 0.06 | -0.11 | 0.31 | 0.79 |
| *MATS1e* | 0.03 | 0.06 | 1.00 | -0.05 | 0.02 | 0.38 |
| *R5e+* | -0.11 | -0.11 | -0.05 | 1.00 | -0.03 | -0.27 |
| *R4p+* | 0.31 | 0.31 | 0.02 | -0.03 | 1.00 | 0.43 |
| *pIC50* | 0.71 | 0.79 | 0.38 | -0.27 | 0.43 | 1.00 |

C:\Users\Tarun Jha\Desktop\001.tiff

**Supplementary Figure S1**. Good molecular fingerprints obtained from the Bayesian classification model.

**C:\Users\Tarun Jha\Desktop\SFigure S2.tiff**

**Supplementary Figure S2**. Bad molecular fingerprints obtained from the Bayesian classification model.