

\* \* \* \* \*  
 302 **M A E V E T T T N V G V P G R W** 16  
 cgcacagacagatgcccaggtggagaccacacacacagtggtgtgcccggggcgtg  
 362 **F R T D D A Q W R V G G C G H T T S T** 36  
 gcgttcagatcgtgatgcccaggtgcccgtggggcgtggccgcaacaacacagctg  
 422 **L A L R P C L N G K Q V D D C V T G** 56  
 tgctggccctggcccctgctcaacggggcaagtgcacgacgctgctgacccggc  
 482 **N P S Y T C C L S G F T G R R C L D** 76  
 aaccctctcaacacgctcctgctcctcgggtctcaacggggcggaggtgcaacctggac  
 542 **N E C A S Q P C Q N G G T C T H** 96  
 gtagaagaatgctcccagcctgcaagaatggtggacctgactcaacggcatcaac  
 602 **S F C G C F A S F S G F C A R A O** 116  
 agttccgctgcaagtcggcgtggtgtgggacacacagctggagagcaacacacacac  
 662 **G C D T T F E C Q H G G Q V E N G S A** 136  
 cccgtgacacaaagagtgtaacatgtggccagtgcaggtggagaatgctctgctg  
 722 **G T C V C Q A G Y T G A A C E M D V D D C** 156  
 vtgtgtgtgcccagccgatacacgggacacctggagatgagatggagcagctg  
 782 **S P D P C L N G G S C V D L V G N Y T C** 176  
 agccctgacccctgctgaatggagctctgtgtgactagtgggaattacacctgc  
 842 **L C A E P F K G L R C E T G D H P V P D** 196  
 ttgtgtgcccagcctcaaggagctcgtgtgacagagacacacacacacacacacac  
 902 **A C L S A P C H N G G T C V D A D Q G Y** 216  
 gctgctcctcggcccctggcaacgggggacactgtgtgatgagagagagagagagag  
 962 **V C E C P E G F M G L D C R E R V P D D** 236  
 gttgtgagtgcccgaaggtctcatggtgactgcaagagagagagagagagagagagag  
 1022 **C E C R N G G R C L G A N T T L C Q C P** 256  
 tttgtgtgcccgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 1082 **L G F F G L L C E P E I T A M P C N M N** 276  
 ctgggattcttgggtctctgtgaattgaaatcaacagcactgcccgtgcaacatgaa  
 1142 **T O C P D G G Y C M E H G G S Y L C V C** 296  
 acacagtgcccagatgggggctcactgactgagcagcggggagctcactctgctctg  
 1202 **H T D H N A S H S L P S P C D S D P C F** 316  
 cacaccgac  
 1262 **N G G S C D A H D D S Y T C E C P R G F** 336  
 aacggagctcctgctgactgcccagcactcctcaacacacacacacacacacacacac  
 1322 **H G K H C E K A R P H L C S S G P C R N** 356  
 caccgcaagcactgcaagaagccggcacaacacacacacacacacacacacacacac  
 1382 **G G T C K E A G G E Y H C S C P Y R F T** 376  
 gggggcagctgcaagagcggggcagctgcaacacacacacacacacacacacacacac  
 1442 **G R H C E I G K P D S C A S G P C H N G** 396  
 gggagcactgtgagatcgggaagcagactcgtgtgctcctggcccctgcaacaaggg  
 1502 **G T C F H Y I G K Y K C D C P P G F S G** 416  
 ggcacactgctcactacatggcaatacagaatgtagctgctgcccagcactctcctgg  
 1562 **R H C E I A P S P C F R S P C V N G G T** 436  
 cggcactgagagatgcccctcctcctgctccagcagcagcagcagcagcagcagcagc  
 1622 **C E D R D T D F F C H C Q A G Y M G R R** 456  
 tggagagcgggagacaggttctctgcaactgccaagcaggtgacatgggagcggc  
 1682 **C Q A E V D C G P P E E V K H A H L R F** 476  
 tgcagcagagagtgatgagcggcccggggaggtgagcagcagcagcagcagcagcagc  
 1742 **N G T R L G A V A L Y A C D R G Y S L S** 496  
 aacggcagcggctggggcggggcggcggcggcggcggcggcggcggcggcggcggcggc  
 1802 **A P S R I R V C Q P H G V W S E P P Q C** 516  
 gcccacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 1862 **L E T D E C R S Q P C L H G G S C Q D R** 536  
 cttgaaatcagatgagtcggcgtcagcggcggcggcggcggcggcggcggcggcggcggc  
 1922 **V A G Y L C A S T G Y E G A H C E L E** 556  
 gttgtggtacctgctctgcaacagcagctatgagggcggcagcagcagcagcagcagc  
 1982 **R D E C R A H P C R N G G S C R N L E E** 576  
 agggatgagtcggagctcaccctgcaagaatggaggtctcctgaggaacctccaggg  
 2042 **L Y Y C P A H C C A C A H D** 596  
 gctatgctcggcgtgcccagcctgctgtgagctccactgtagagcagagtgtagc  
 2102 **A C D S E F C Q H G R C E S G G G A** 616  
 gctgagcactccagcccctgcaagcagtgagggctgtagagcggcggcggcggcggcggc  
 2162 **L C P E S F F G Y H C T E T V S D P C** 636  
 ctgtgctgctgcccagagagctctctgctgacacactgagagcagtgagtgaccctg  
 2222 **F S S P C G G R G Y C L A S N G S H S C** 656  
 tctccagcccctggtggggcggctgctatgctgcaacgcaacagcctcccaagctg  
 2282 **T C K V G Y T G E D C A K E L F P P T A** 676  
 accctgcaagtggtgctacagggcagagcagcagcagcagcagcagcagcagcagcagc  
 2342 **L K M E R V E E S G V S I S W N P N G** 696  
 ctaagatggagagagtgaggagagtggtgctctctctcctggaaacccgccaatggt  
 2402 **P A A R O M L D G Y A V T Y V S D G S** 716  
 ccagccggcagcagatgctgtagctgagcagcagcagcagcagcagcagcagcagcagc  
 2462 **F Y R R T D F V D R T R S S H L Q A L A** 736  
 taccggcagcagacttggagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 2522 **A G R A Y N I S V F V K R N S N N K N** 756  
 gcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag  
 2582 **D I S R P A V L L A R T R P R P V E G F** 776  
 gacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 2642 **G A G A G A G A G A G A G A G A G A G A** 796  
 gaggtcaccatgtgacggtagcacaactcagtgagtgggccctgcaacagatccg  
 2702 **G A G A G A G A G A G A G A G A G A G** 816  
 catgcccagcagtggtgggtcgtgtgctcactcggcaccctgagcagcagcagcagcagc  
 2762 **T D V D R S R D R F T C L A L L P S** 836  
 gccaccagtgagcagagtgtagcaggtcactcttagggccctcgtgctgggaag  
 2822 **S Y L L D D S G L R S E E R P P E** 856  
 aggtacacatccagctgaccacccctcagtggtcaggggagagagcaccacacagag  
 2882 **B L A T A F T H Y N T R P L P P A N** 876  
 agcctggcaccggccgagcagcagctgtggaccggcccctgctccagcaaacctgacc  
 2942 **A K A V W A A S A H V V D A P T P S** 896  
 gcccccagtgctgcccactctgcccagtggtggtgagtgcccagcactccagggagc  
 3002 **L E A N T V I N V T T S C S I T K S D Y** 916  
 ttgctggagcttatgcatcaatgtagcaccacagcagcagcagcagcagcagcagcagc  
 3062 **N R L A S C V R D E L P R R R Y** 936  
 cccaagggagctggcgtcctcaacgctgagcagcagcagcagcagcagcagcagcagcag  
 3122 **S Y L L D D S G L R S E E R P P E** 956  
 ctctctgtagcagtgcaagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 976 **V I T S P R D G A D R R W H G E E**

NIDO  
 EGF CA  
 CCP  
 FN3

3182 ctctacatcatcacctccccaggatggcgtgacagacgctggcaccagggggagcaac  
 3242 caccctcgggtgctcaagaagcagacggccccggcggcctgcccggagctgcccctgctc 996  
 3302 aatgaccacagcggccccggagaccccccagccaggttctggagcttgggac 1016  
 3362 c R R G R V S A A F G G S S P S K A / T / R  
 3422 ggcagagaaagatgagcggcaggtctggtggtcaccagcaaaagcagccacccgtgaga  
 3482 C D P T A S A Q L D N M E A P I R V S  
 3542 tcaacaacccagcctcggcggcagctgagaacaatgggagaaagccccaaaggctcagc  
 3602 L A L Q L P E H G S K D I G N V P G N C  
 3662 ctggcctccagctcccctgaacagcagcaagcagcagcaacagcagcagcagcagcagc  
 3722 S E N P C Q N G G T C V P G G A D A H S C  
 3782 tcaaaaaacccctgcaagaagcagcagcagcagcagcagcagcagcagcagcagcagc  
 3842 D C G P G F K G R R C E L A C I K V S R  
 gactgcccggcaggggttcaaaagcagcagcagcagcagcagcagcagcagcagcagcagc  
 P C T R L F S E T K A F P V W E G G V C  
 cctgcaacaagcctgttcccgagcaaaagccttccagctgggagggagcagcagcagcagc  
 H H V Y K R V Y Y R V H Q D I C F K E S C  
 caccagctgataaaagatctcagcagctcaccagacacagcagcagcagcagcagcagc  
 E A S T S L K K T P N R K Q S K S Q T L E  
 gaagcacaagcctcaagagcccccaacagcaagaagaagtagcagcagcagcagcagcagc  
 K S \*  
 aaactttagagtttagagcttctgtttacactcccaacactcagagtttctaacac 3901

### Annotation Key

>...129 : NIDO : Extracellular domain of unknown function in nidogen (entactin) and hypothetical proteins  
 >148..177, 180..216, 256..292, 300..333, 415..446, 454..485, 490..524, 533..562, 622..658, 662..696, 698..734, 1181..1212: EGF CA : Calcium-binding EGF-like domain, present in a large number of membrane-bound and extracellular (mostly animal) proteins. Many of these proteins require calcium for their biological function.  
 \*\*\*Not annotated in sequence above, look at figure below.  
 567..620 CCF : Complement control protein (CCP) modules (aka short consensus repeats SCRs or SUSHI repeats) have been identified in several proteins of the complement system  
 377..386, 374..381, 373..1053 FN3 : Fibronectin type 3 domain; One of three types of internal repeats found in the plasma protein fibronectin. Its tenth fibronectin type III repeat contains an RGD cell recognition sequence in a flexible loop between 2 strands.  
 RED TEXT: Alpha helix domains in secondary structure.  
 PURPLE TEXT: Beta sheet domains in secondary structure  
 S-T: Kinase Dependent Phosphorylation sites score > 0.8  
 S-97 PKC 0.84  
 T-473 PKC 0.82  
 S-716 PKC 0.82  
 S-747 PKC 0.93  
 S-801 PKC 0.83  
 T-828 PKC 0.81  
 T-839 PRA 0.80  
 S-910 PKC 0.83  
 S-938 PKC 0.86  
 T-1034 PKC 0.90  
 S-1056 PRA 0.86  
 S-1170 PRA 0.83  
 Highest Score: 0.93 PKC at position 747  
 \*\*\*Acetylation prediction, 1 sequence (most likely a degradation signal)  
 Sequence # Context Score Acetylation  
 -----  
 Sequence 2 A --MAEVE 0.479 yes