© 1990 Oxford University Press

Nucleotide sequence of rabbit *NAT1* encoding monomorphic arylamine N-acetyltransferase

Martin Blum, Markus Heim and Urs A.Meyer
Department of Pharmacology, Biocenter, University of Basel, 4056 Basel, Switzerland

Submitted July 28, 1990

EMBL accession no. X53765

A genomic library in λEMBL3 was constructed from liver DNA of an *in vivo* phenotyped rapid acetylator rabbit. It was screened with the cDNA *rnat* (1) encoding polymorphic arylamine N-acetyltransferase (NAT; EC2.3.1.5). A gene designated *NAT1* was isolated; it contained an open reading frame of 870 bp and displayed 97% sequence homology to the open reading frame of the polymorphic cDNA *rnat*. The homology at the deduced amino acid level was 96%. *NAT1* was contained on a 4.8 kb EcoRI fragment which was shown to be present both in slow and rapid acetylator rabbits (2). Upon expression of *NAT1* in COS-1 cells (2) a NAT enzyme protein was produced which differed from NAT expressed from *rnat* in its affinity for arylamine substrates (3). Thus rabbits, like humans (4) possess two highly homologous genes encoding NAT enzyme proteins.

A putative polyadenylation signal in the 3′ noncoding region of *NAT1* is underlined.

REFERENCES

1. Blum,M., Grant,D.M., Demierre,A. and Meyer,U.A. (1989) *Nucleic Acids Res.* **17**, 3589.
2. Blum,M., Grant,D.M., Demierre,A. and Meyer,U.A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9554–9559.
3. Blum,M., Demierre,A., Beer,M., D'Orazio,D. and Meyer,U.A. (1990) *Experientia* **46**, A60.
4. Blum,M., Grant,D.M., McBride,W., Heim,M. and Meyer,U.A. (1990) *DNA and Cell Biol.* **9**, 193–203.

```
CACTACCTTTGTATGTAAATCTAAGTCCCTCTAAAAGTAAGTATTTCTGCTGGCAAGCAAGTGCTAAATCAGCAGCTGATACACAATTTTGATTCAGGAAACTTCTTATTGAAACAATCA  120
AATCAATCATTAAGTGATTGAATAAAGTTTTTGAGTGGAAAACATTATGTAAGCTTCAGCTAAAAGAATTGAATGGGTCAGGTAATATGAATACCATATACTCTATCGTATAATCTCCTG  240
TTTATATCAGGAACTATTTATAAGCCCATTTTAATTAATGAGTTGGAAAATCACATATAAGATGTTTATCATTTGACTCATCATTTAACATGGATTTCCAATTTTTTTTTAATTTTTAAA  360
TTTATTTATGTCCAATTTCTTTGAAAATCAGAGCAACAGAGTCATACAGAGAGAAAGAGTTCTTCCATCTGCAAACGCCTGTAATTCCCAAGTCAGGAGCCCGAAACTCCATCCAGGTCT  480
CCCACATTAGTAGCCAGAACCCATGTACATGAACCTTCATTGGTGCCTCCCAGGGTGCACAGTAGCAAGAAGCTGGATCAAATGCAGGCACCTTGATATTAGATGCTGGTGCCCTAAGCA  600
GCAACTTAACTGCTATATCAAAGGCATGACCCATATTTTTAATTTTTTTTGTGTTTTTTGCTTTAGAGAGTGAGATGTGCATACAAAATGAGAAAGAGAGAGAGAGAAACTCCCATCTAC  720
TGAGTTCACTTGGGATTCACTTTCCAAATACCTGAACAGCCAGAGCTGGGCCAAGATTGGGCTGGGGCTGGGTCATGCAGAAATCAGGAGTTGGGAGCTCAATCCAGGTCTCCCATGTGG  840
GTGGCAAGAACAAATTAGTTGAGCCATCACCACTGCCTCCCAGGTCTGTAACCAGGCACTCCTGTGTGGGACATATACATCTTACCAGAAGGCCAAACATCTGCCCCATATACATATCTA  960

                                                                                                    M  D  I  E  A  Y  Y  Q  R
AATATGTTACAGCTATCATTAGCCTACCCTAGGATATCCACCTTCAAAAGTGATACAAGTTGGTTTTGTTCTCTTTTCCTTGTCTTAGGGTATCATGGATATTGAAGCATATTATCAAAG 1080

 I  G  Y  K  N  P  R  N  K  L  D  L  E  S  L  T  D  I  F  Q  H  Q  I  R  T  V  P  Y  E  N  L  S  I  H  C  G  E  S  M  E
AATTGGTTACAAAAACCCAAGGAACAAATTGGACTTGGAATCACTAACAGACATTTTTCAGCACCAGATCCGAACTGTTCCCTATGAGAACCTTAGCATTCACTGTGGGGAGTCCATGGA 1200

 L  D  L  E  A  I  F  D  Q  I  V  R  R  N  R  G  G  W  C  L  Q  V  N  Y  L  L  Y  W  A  L  T  T  T  G  F  E  T  T  M  L
GTTGGACTTAGAGGCCATTTTTGATCAAATTGTGAGGAGGAATAGAGGTGGCTGGTGTCTCCAGGTCAACTATCTTTTGTACTGGGCTTTGACTACCACTGGTTTTGAGACCACAATGTT 1320

 G  G  F  V  C  G  S  H  T  D  K  Y  S  T  G  M  I  H  L  I  V  Q  V  T  I  N  G  R  N  Y  I  V  D  A  G  F  G  R  S  Y
AGGAGGGTTTGTTTGCGGCAGTCATACTGACAAATACAGCACTGGCATGATTCACCTCATAGTACAGGTGACCATCAATGGCAGGAACTACATTGTCGATGCTGGGTTTGGACGCTCCTA 1440

 Q  M  W  Q  P  V  E  L  I  S  G  K  D  Q  P  Q  V  P  S  I  F  R  L  R  E  E  G  E  T  W  Y  L  D  Q  I  R  R  Q  Q  H
CCAGATGTGGCAGCCTGTGGAGTTAATTTCTGGGAAAGATCAGCCTCAAGTGCCTTCCATATTCCGCTTGAGAGAAGAGGGAGAAACCTGGTACCTGGACCAAATCAGAAGACAGCAACA 1560

 V  P  D  Q  E  F  L  N  S  E  L  L  E  R  K  T  H  R  K  L  Y  C  F  T  L  Q  P  R  T  I  E  E  F  E  S  A  N  T  Y  L
TGTTCCAGACCAAGAATTTCTTAATTCTGAACTTCTGGAAAGGAAAACACACCGAAAACTCTATTGCTTTACTCTTCAACCTCGAACAATTGAAGAATTTGAGTCTGCAAACACATATCT 1680

 Q  I  S  P  S  S  P  F  L  D  K  S  I  C  S  L  Q  T  P  E  G  V  H  C  L  V  G  L  I  L  T  F  R  T  Y  N  Y  K  E  N
TCAAATATCTCCATCATCTCCATTTTTAGACAAATCAATTTGTTCCTTGCAGACCCCAGAAGGTGTTCACTGTTTGGTGGGTCTCATCCTCACCTTCAGGACCTACAATTACAAGGAAAA 1800

 T  D  L  V  E  F  K  V  L  T  E  E  E  V  E  E  V  L  K  T  I  F  N  I  S  L  G  K  K  L  V  S  K  N  G  N  L  F  F  T
TACAGATCTGGTAGAGTTTAAAGTTCTGACTGAGGAAGAAGTGGAGGAAGTGCTGAAAACTATATTTAATATTTCCTTGGGAAAAAAGCTTGTGTCCAAAAATGGTAATTTGTTTTTCAC 1920

 I
TATTTAGTGTATAGACGAAAAATCATCTTCAATCTTAGTACAAGTATAAGCTTTTTATTATAAAAATGATCAAACATGAAAAATTACAGAGAATAGGATAGTAAATTCATATGTATTTAT 2040
AACCAGATCACCAATTATTCACTGCTGACAGATAATATTCTCCCTACTCTCTTATTGTTTTGATGAAAATCATAAACATCAAATTATTCACCTATAAAGATGTCTACATATATAAGTTGA 2160
AACTTTGTAAAGAATTATAGCAACACTCCTTTACAAAATTAATAAAAAGGCATTTTAATAATGGTG 2226
```