**Supporting Information**

# HUMAN DNA POLYMERASE β MUTATIONS ALLOWING EFFICIENT ABASIC SITE BYPASS

**Sonja Gieseking, Konrad Bergen, Francesca Di Pasquale, Kay Diederichs, Wolfram Welte, and Andreas Marx***

**Alignment of DNA polymerase beta genes**

1st Line: human DNA polymerase beta (*E. coli* optimized)
2nd Line: human DNA polymerase beta (*E. coli* optimized) translated to protein
3rd Line: human DNA polymerase beta (wildtype)
4th Line: human DNA polymerase beta (wildtype) translated to protein

* marks equivalent nucleotides and amino acids

```
ATGAGCAAACGTAAAGCGCCGCAGGAAACCCTGAACGGCGGCATTACCGATATGCTGACCGAACTGGCCAACTTTGAAAAAACGTGAGCCAGGCGATCC
M  S  K  R  K  A  P  Q  E  T  L  N  G  G  I  T  D  M  L  T  E  L  A  N  F  E  K  N  V  S  Q  A  I  H
ATGAGCAAACGGAAGGCGCCGCAGGAGACTCTCAACGGGGGAATCACCGACATGCTCACAGAACTCGCAAACTTTGAGAAGAACGTGAGCCAAGCTATCC
M  S  K  R  K  A  P  Q  E  T  L  N  G  G  I  T  D  M  L  T  E  L  A  N  F  E  K  N  V  S  Q  A  I  H
*********** ** *********** ** ** ***** ** ** ***** ***** ** ***** ** ******** ** ********** ** ****
*  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

ATAAATATAACGCGTATCGTAAAGCGGCGAGCGTGATTGCGAAATATCCGCACAAAATTAAAAGCGGTGCGGAAGCGAAAAAACTGCCGGGCGTGGGCAC
  K  Y  N  A  Y  R  K  A  A  S  V  I  A  K  Y  P  H  K  I  K  S  G  A  E  A  K  K  L  P  G  V  G  T
ACAAGTACAATGCTTACAGAAAAGCAGCATCTGTTATAGCAAAATACCCACACAAAATAAAGAGTGGAGCTGAAGCTAAGAAATTGCCTGGAGTAGGAAC
  K  Y  N  A  Y  R  K  A  A  S  V  I  A  K  Y  P  H  K  I  K  S  G  A  E  A  K  K  L  P  G  V  G  T
* ** ** ** ** **  * ***** **    ** ** ** ***** ** ******** ** ** ** ** ***** ** *** **** ** ** ** **
  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

CAAAATTGCGGAAAAAATCGATGAATTTCTGGCCACCGGCAAACTGCGTAAACTGGAAAAAATTCGCCAGGATGATACCAGCAGCAGCATTAACTTTCTG
 K  I  A  E  K  I  D  E  F  L  A  T  G  K  L  R  K  L  E  K  I  R  Q  D  D  T  S  S  S  I  N  F  L
AAAAATTGCTGAAAAGATTGATGAGTTTTTAGCAACTGGAAAATTACGTAAACTGGAAAAGATTCGGCAGGATGATACGAGTTCATCCATCAATTTCCTG
 K  I  A  E  K  I  D  E  F  L  A  T  G  K  L  R  K  L  E  K  I  R  Q  D  D  T  S  S  S  I  N  F  L
 ******** ***** ** ***** *** * ** ** ** *** * ************** ***** ********** **      *** ** ** ***
 *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

ACCCGTGTGAGCGGCATTGGTCCGAGCGCGGCGCGTAAATTTGTGGATGAAGGCATCAAAACCCTGGAGGATCTGCGTAAAAACGAAGATAAACTGAACC
T  R  V  S  G  I  G  P  S  A  A  R  K  F  V  D  E  G  I  K  T  L  E  D  L  R  K  N  E  D  K  L  N  H
ACTCGAGTTAGTGGCATTGGTCCATCTGCTGCAAGGAAGTTTGTAGATGAAGGAATTAAAACACTAGAAGATCTCAGAAAAAATGAAGATAAATTGAACC
T  R  V  S  G  I  G  P  S  A  A  R  K  F  V  D  E  G  I  K  T  L  E  D  L  R  K  N  E  D  K  L  N  H
** ** ** ** ***********    ** **  * ** ***** ******** ** ***** ** ** *****  * ***** ********* ******
*  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

ATCATCAGCGTATTGGCCTGAAATATTTTGGCGATTTCGAAAAACGTATTCCGCGTGAAGAAATGCTGCAGATGCAGGATATTGTGCTGAACGAAGTGAA
  H  Q  R  I  G  L  K  Y  F  G  D  F  E  K  R  I  P  R  E  E  M  L  Q  M  Q  D  I  V  L  N  E  V  K
ATCATCAGCGAATTGGGCTGAAATATTTTGGGGACTTTGAAAAAAGAATTCCTCGTGAAGAGATGTTACAAATGCAAGATATTGTACTAAATGAAGTTAA
  H  Q  R  I  G  L  K  Y  F  G  D  F  E  K  R  I  P  R  E  E  M  L  Q  M  Q  D  I  V  L  N  E  V  K
********** ***** ************** ** ** ****** * ***** ******** *** * ** ***** ******** ** ** ***** **
  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

AAAAGTGGATAGCGAATATATTGCGACCGTGTGCGGCAGCTTTCGTCGTGGCGCGGAAAGCAGCGGCGATATGGATGTGCTGCTGACCCATCCGAGCTTT
 K  V  D  S  E  Y  I  A  T  V  C  G  S  F  R  R  G  A  E  S  S  G  D  M  D  V  L  L  T  H  P  S  F
AAAAGTGGATTCTGAATACATTGCTACAGTCTGTGGCAGTTTCAGAAGAGGTGCAGAGTCCAGTGGTGACATGGATGTTCTCCTGACCCATCCCAGCTTC
 K  V  D  S  E  Y  I  A  T  V  C  G  S  F  R  R  G  A  E  S  S  G  D  M  D  V  L  L  T  H  P  S  F
**********   ***** ***** ** ** ** ***** **  *  * ** ** **   *** ** ** ******** ** ********** *****
 *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

ACCAGCGAAAGCACCAAACAGCCGAAACTGCTGCATCAGGTGGTGGAACAGCTGCAGAAAGTGCATTTTATTACCGATACCCTGAGCAAAGGCGAAACCA
T  S  E  S  T  K  Q  P  K  L  L  H  Q  V  V  E  Q  L  Q  K  V  H  F  I  T  D  T  L  S  K  G  E  T  K
ACTTCAGAATCAACCAAACAGCCAAAACTGTTACATCAGGTTGTGGAGCAGTTACAAAAGGTTCATTTTATCACAGATACCCTGTCAAAGGGTGAGACAA
T  S  E  S  T  K  Q  P  K  L  L  H  Q  V  V  E  Q  L  Q  K  V  H  F  I  T  D  T  L  S  K  G  E  T  K
**    ***   *********** ****** * ******** ***** *** * ** ** ** ******** ** *********   ** ** ** ** *
*  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

AATTTATGGGCGTGTGCCAGCTGCCGAGCAAAAACGATGAAAAAGAATATCCGCATCGCCGTATTGATATTCGTCTGATCCCGAAAGATCAGTATTATTG
  F  M  G  V  C  Q  L  P  S  K  N  D  E  K  E  Y  P  H  R  R  I  D  I  R  L  I  P  K  D  Q  Y  Y  C
AGTTCATGGGTGTTTGCCAGCTTCCCAGTAAAAATGATGAAAAAGAATATCCACACAGAAGAATTGATATCAGGTTGATACCCAAAGATCAGTATTACTG
  F  M  G  V  C  Q  L  P  S  K  N  D  E  K  E  Y  P  H  R  R  I  D  I  R  L  I  P  K  D  Q  Y  Y  C
* ** ***** ** ******** ** ** ***** ****************** **  *  * ********  *  **** ** *************** **
  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

CGGCGTGCTGTATTTTACCGGCAGCGATATCTTCAACAAAAACATGCGTGCGCATGCGCTGGAAAAAGGCTTTACCATCAACGAATACACCATTCGTCCG
 G  V  L  Y  F  T  G  S  D  I  F  N  K  N  M  R  A  H  A  L  E  K  G  F  T  I  N  E  Y  T  I  R  P
TGGTGTTCTCTATTTCACTGGGAGTGATATTTTCAATAAGAATATGAGGGCTCATGCCCTAGAAAAGGGTTTCACAATCAATGAGTACACCATCCGTCCC
 G  V  L  Y  F  T  G  S  D  I  F  N  K  N  M  R  A  H  A  L  E  K  G  F  T  I  N  E  Y  T  I  R  P
 ** ** ** ***** ** ** ** ***** ***** ** ** *** * ** ***** ** ***** ** ** ** ***** ** ******** *****
 *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

CTGGGCGTGACCGGTGTTGCGGGTGAACCGCTGCCGGTGGATAGCGAAAAAGATATCTTCGATTACATCCAGTGGAAATATCGTGAACCGAAAGATCGTA
L  G  V  T  G  V  A  G  E  P  L  P  V  D  S  E  K  D  I  F  D  Y  I  Q  W  K  Y  R  E  P  K  D  R  S
```


TTGGGAGTCACTGGAGTTGCAGGAGAACCCCTGCCAGTGGATAGTGAAAAAGACATCTTTGATTACATCCAGTGGAAATACCGGGAACCCAAGGACCGGA
L G V T G V A G E P L P V D S E K D I F D Y I Q W K Y R E P K D R S
GCGAATAA
E *
GCGAATGA
E *