
Structure of the *uvrB* gene of *Escherichia coli*. Homology with other DNA repair enzymes and characterization of the *uvrB5* mutation

C.Backendorf, H.Spank, A.P.Barbeiro and P.van de Putte

Laboratory of Molecular Genetics, University of Leiden, PO Box 9505, 2300 RA Leiden, The Netherlands

Received 17 February 1986; Accepted 11 March 1986

ABSTRACT

1. The complete nucleotide sequence of the *Escherichia coli uvrB* gene has been determined. The coding region of the *uvrB* gene consists of 2019 nucleotides which direct the synthesis of a 673 amino acid long polypeptide with a calculated molecular weight of 76.614 daltons.

2. Comparison of the UvrB protein sequence to other known DNA repair enzymes revealed that 2 domains of the UvrB protein (domain I = 6 amino acids, domain II = 14 amino acids) are also present in the protein sequence of the *uvrC* gene. We show that the structural homologies between UvrB and UvrC are as well reflected by the cross-reactivity of anti-*uvrB* and anti-*uvrC* antibodies with UvrC and UvrB protein respectively. In the N-terminal part of UvrB, domain III (17 amino acids) shows a strong homology with one part of the AtkA gene product. Adjacent to domain III, an ATP binding site consensus sequence is found in domain IV.

3. The *uvrB5* mutant gene from strain AB1885 has been cloned on plasmid pBL01. We show that the *uvrB5* mutation is due to a point deletion of a CG basepair and results in the synthesis of an 18 kD protein composed of the 113 N-terminal amino acids of the wild type *uvrB* gene and a 43 amino acid long tail coded in the -1 frame.

INTRODUCTION

In *Escherichia coli* excision repair of damaged DNA is mediated by the UvrA, B and C gene products. The three proteins, which function in concert, recognize a large number of different DNA lesions, affecting both purine and pyrimidine nucleotides (for a review see 1). Such a broad action spectrum suggests that the UvrABC excinuclease recognizes a deformation or irregularity in the DNA helix rather than the DNA adduct itself (2). Consistent with such a concept is the observation that ABC excinuclease hydrolyses the 8th phosphodiesterbond 5' and the 4th or 5th phosphodiesterbond 3' to the modified base (3-5), which suggests that the enzyme might approach the DNA helix from the side opposite to the DNA lesion (2). The individual functions of the different components of the UVR excision nuclease are only beginning to be understood. UvrA is an

ATPase (6) and binds to single-stranded and UV irradiated duplex DNA with a higher affinity than to unirradiated duplex DNA. The UvrB protein stimulates the ATPase activity of UvrA by a factor of 2.5 in the presence of UV-irradiated double-stranded DNA (2). Only a slight stimulation is observed in the presence of unirradiated duplex DNA whereas in the absence of DNA no stimulation is registered. These experiments have been interpreted to indicate that the enzyme complex is assembled on damaged DNA and does not exist as such free in solution (2). Indeed the UvrB component enhances the ability of the UvrA protein to bind to UV-irradiated duplex DNA. This enhanced binding is ATP dependent (7). After ATP hydrolysis a stable UvrAB-irradiated DNA complex resistant to chelators and high salt is formed (7). It has been suggested that the AB complex is first formed at a non-specific site in UV-irradiated DNA and is then translocated to the damaged site. This translocation would need ATP hydrolysis (8). Once the AB-damaged site complex has been preformed the only addition of the UvrC component results in DNA incision with no need of extra energy. After the incision event the excinuclease-DNA complex persists (9) and both the UvrD protein (DNA helicase II) and the PolI gene product (DNA polymerase I) are necessary to release the 12-13 basepair-long oligonucleotide harboring the DNA adduct, and promote turn-over of the ABC excinuclease (9,10). The simultaneous requirement of all three excinuclease components to promote DNA incision *in vitro* (11) suggested that the expression of UvrA, B and C might be coordinately regulated. Indeed it has been shown that both UvrA and UvrB are subjected to SOS regulation (12,13). The kinetics of UvrC transcription were different and more difficult to interpret (14).

The structural organization of the regulatory region of the uvrB gene has been described in detail (15,16). The *recA*-*lexA* regulated expression of the uvrB gene has been studied both *in vitro* (16) and *in vivo* (17). In addition it has been proposed that *uvrB* expression might be coupled to DNA replication (18).

A final understanding of the molecular mechanism of action of the different Uvr components in excision repair or in other cellular processes will depend on the analysis of the specific aminoacid to aminoacid and aminoacid to nucleotide contacts between the different Uvr proteins and DNA. Such studies will depend on the knowledge of the primary sequence of the different proteins. Furthermore the molecular characterization of mutations in the different gene products will be of great value to pinpoint

0002
 0003
 0004
 0005
 0006
 0007
 0008
 0009
 0010
 0011
 0012
 0013
 0014
 0015
 0016
 0017
 0018
 0019
 0020
 0021
 0022
 0023
 0024
 0025
 0026
 0027
 0028
 0029
 0030
 0031
 0032
 0033
 0034
 0035
 0036
 0037
 0038
 0039
 0040
 0041
 0042
 0043
 0044
 0045
 0046
 0047
 0048
 0049
 0050
 0051
 0052
 0053
 0054
 0055
 0056
 0057
 0058
 0059
 0060
 0061
 0062
 0063
 0064
 0065
 0066
 0067
 0068
 0069
 0070
 0071
 0072
 0073
 0074
 0075
 0076
 0077
 0078
 0079
 0080
 0081
 0082
 0083
 0084
 0085
 0086
 0087
 0088
 0089
 0090
 0091
 0092
 0093
 0094
 0095
 0096
 0097
 0098
 0099
 0100
 0101
 0102
 0103
 0104
 0105
 0106
 0107
 0108
 0109
 0110
 0111
 0112
 0113
 0114
 0115
 0116
 0117
 0118
 0119
 0120
 0121
 0122
 0123
 0124
 0125
 0126
 0127
 0128
 0129
 0130
 0131
 0132
 0133
 0134
 0135
 0136
 0137
 0138
 0139
 0140
 0141
 0142
 0143
 0144
 0145
 0146
 0147
 0148
 0149
 0150
 0151
 0152
 0153
 0154
 0155
 0156
 0157
 0158
 0159
 0160
 0161
 0162
 0163
 0164
 0165
 0166
 0167
 0168
 0169
 0170
 0171
 0172
 0173
 0174
 0175
 0176
 0177
 0178
 0179
 0180
 0181
 0182
 0183
 0184
 0185
 0186
 0187
 0188
 0189
 0190
 0191
 0192
 0193
 0194
 0195
 0196
 0197
 0198
 0199
 0200
 0201
 0202
 0203
 0204
 0205
 0206
 0207
 0208
 0209
 0210
 0211
 0212
 0213
 0214
 0215
 0216
 0217
 0218
 0219
 0220
 0221
 0222
 0223
 0224
 0225
 0226
 0227
 0228
 0229
 0230
 0231
 0232
 0233
 0234
 0235
 0236
 0237
 0238
 0239
 0240
 0241
 0242
 0243
 0244
 0245
 0246
 0247
 0248
 0249
 0250
 0251
 0252
 0253
 0254
 0255
 0256
 0257
 0258
 0259
 0260
 0261
 0262
 0263
 0264
 0265
 0266
 0267
 0268
 0269
 0270
 0271
 0272
 0273
 0274
 0275
 0276
 0277
 0278
 0279
 0280
 0281
 0282
 0283
 0284
 0285
 0286
 0287
 0288
 0289
 0290
 0291
 0292
 0293
 0294
 0295
 0296
 0297
 0298
 0299
 0300
 0301
 0302
 0303
 0304
 0305
 0306
 0307
 0308
 0309
 0310
 0311
 0312
 0313
 0314
 0315
 0316
 0317
 0318
 0319
 0320
 0321
 0322
 0323
 0324
 0325
 0326
 0327
 0328
 0329
 0330
 0331
 0332
 0333
 0334
 0335
 0336
 0337
 0338
 0339
 0340
 0341
 0342
 0343
 0344
 0345
 0346
 0347
 0348
 0349
 0350
 0351
 0352
 0353
 0354
 0355
 0356
 0357
 0358
 0359
 0360
 0361
 0362
 0363
 0364
 0365
 0366
 0367
 0368
 0369
 0370
 0371
 0372
 0373
 0374
 0375
 0376
 0377
 0378
 0379
 0380
 0381
 0382
 0383
 0384
 0385
 0386
 0387
 0388
 0389
 0390
 0391
 0392
 0393
 0394
 0395
 0396
 0397
 0398
 0399
 0400
 0401
 0402
 0403
 0404
 0405
 0406
 0407
 0408
 0409
 0410
 0411
 0412
 0413
 0414
 0415
 0416
 0417
 0418
 0419
 0420
 0421
 0422
 0423
 0424
 0425
 0426
 0427
 0428
 0429
 0430
 0431
 0432
 0433
 0434
 0435
 0436
 0437
 0438
 0439
 0440
 0441
 0442
 0443
 0444
 0445
 0446
 0447
 0448
 0449
 0450
 0451
 0452
 0453
 0454
 0455
 0456
 0457
 0458
 0459
 0460
 0461
 0462
 0463
 0464
 0465
 0466
 0467
 0468
 0469
 0470
 0471
 0472
 0473
 0474
 0475
 0476
 0477
 0478
 0479
 0480
 0481
 0482
 0483
 0484
 0485
 0486
 0487
 0488
 0489
 0490
 0491
 0492
 0493
 0494
 0495
 0496
 0497
 0498
 0499
 0500
 0501
 0502
 0503
 0504
 0505
 0506
 0507
 0508
 0509
 0510
 0511
 0512
 0513
 0514
 0515
 0516
 0517
 0518
 0519
 0520
 0521
 0522
 0523
 0524
 0525
 0526
 0527
 0528
 0529
 0530
 0531
 0532
 0533
 0534
 0535
 0536
 0537
 0538
 0539
 0540
 0541
 0542
 0543
 0544
 0545
 0546
 0547
 0548
 0549
 0550
 0551
 0552
 0553
 0554
 0555
 0556
 0557
 0558
 0559
 0560
 0561
 0562
 0563
 0564
 0565
 0566
 0567
 0568
 0569
 0570
 0571
 0572
 0573
 0574
 0575
 0576
 0577
 0578
 0579
 0580
 0581
 0582
 0583
 0584
 0585
 0586
 0587
 0588
 0589
 0590
 0591
 0592
 0593
 0594
 0595
 0596
 0597
 0598
 0599
 0600
 0601
 0602
 0603
 0604
 0605
 0606
 0607
 0608
 0609
 0610
 0611
 0612
 0613
 0614
 0615
 0616
 0617
 0618
 0619
 0620
 0621
 0622
 0623
 0624
 0625
 0626
 0627
 0628
 0629
 0630
 0631
 0632
 0633
 0634
 0635
 0636
 0637
 0638
 0639
 0640
 0641
 0642
 0643
 0644
 0645
 0646
 0647
 0648
 0649
 0650
 0651
 0652
 0653
 0654
 0655
 0656
 0657
 0658
 0659
 0660
 0661
 0662
 0663
 0664
 0665
 0666
 0667
 0668
 0669
 0670
 0671
 0672
 0673
 0674
 0675
 0676
 0677
 0678
 0679
 0680
 0681
 0682
 0683
 0684
 0685
 0686
 0687
 0688
 0689
 0690
 0691
 0692
 0693
 0694
 0695
 0696
 0697
 0698
 0699
 0700
 0701
 0702
 0703
 0704
 0705
 0706
 0707
 0708
 0709
 0710
 0711
 0712
 0713
 0714
 0715
 0716
 0717
 0718
 0719
 0720
 0721
 0722
 0723
 0724
 0725
 0726
 0727
 0728
 0729
 0730
 0731
 0732
 0733
 0734
 0735
 0736
 0737
 0738
 0739
 0740
 0741
 0742
 0743
 0744
 0745
 0746
 0747
 0748
 0749
 0750
 0751
 0752
 0753
 0754
 0755
 0756
 0757
 0758
 0759
 0760
 0761
 0762
 0763
 0764
 0765
 0766
 0767
 0768
 0769
 0770
 0771
 0772
 0773
 0774
 0775
 0776
 0777
 0778
 0779
 0780
 0781
 0782
 0783
 0784
 0785
 0786
 0787
 0788
 0789
 0790
 0791
 0792
 0793
 0794
 0795
 0796
 0797
 0798
 0799
 0800
 0801
 0802
 0803
 0804
 0805
 0806
 0807
 0808
 0809
 0810
 0811
 0812
 0813
 0814
 0815
 0816
 0817
 0818
 0819
 0820
 0821
 0822
 0823
 0824
 0825
 0826
 0827
 0828
 0829
 0830
 0831
 0832
 0833
 0834
 0835
 0836
 0837
 0838
 0839
 0840
 0841
 0842
 0843
 0844
 0845
 0846
 0847
 0848
 0849
 0850
 0851
 0852
 0853
 0854
 0855
 0856
 0857
 0858
 0859
 0860
 0861
 0862
 0863
 0864
 0865
 0866
 0867
 0868
 0869
 0870
 0871
 0872
 0873
 0874
 0875
 0876
 0877
 0878
 0879
 0880
 0881
 0882
 0883
 0884
 0885
 0886
 0887
 0888
 0889
 0890
 0891
 0892
 0893
 0894
 0895
 0896
 0897
 0898
 0899
 0900
 0901
 0902
 0903
 0904
 0905
 0906
 0907
 0908
 0909
 0910
 0911
 0912
 0913
 0914
 0915
 0916
 0917
 0918
 0919
 0920
 0921
 0922
 0923
 0924
 0925
 0926
 0927
 0928
 0929
 0930
 0931
 0932
 0933
 0934
 0935
 0936
 0937
 0938
 0939
 0940
 0941
 0942
 0943
 0944
 0945
 0946
 0947
 0948
 0949
 0950
 0951
 0952
 0953
 0954
 0955
 0956
 0957
 0958
 0959
 0960
 0961
 0962
 0963
 0964
 0965
 0966
 0967
 0968
 0969
 0970
 0971
 0972
 0973
 0974
 0975
 0976
 0977
 0978
 0979
 0980
 0981
 0982
 0983
 0984
 0985
 0986
 0987
 0988
 0989
 0990
 0991
 0992
 0993
 0994
 0995
 0996
 0997
 0998
 0999
 1000
 1001
 1002
 1003
 1004
 1005
 1006
 1007
 1008
 1009
 1010
 1011
 1012
 1013
 1014
 1015
 1016
 1017
 1018
 1019
 1020
 1021
 1022
 1023
 1024
 1025
 1026
 1027
 1028
 1029
 1030
 1031
 1032
 1033
 1034
 1035
 1036
 1037
 1038
 1039
 1040
 1041
 1042
 1043
 1044
 1045
 1046
 1047
 1048
 1049
 1050
 1051
 1052
 1053
 1054
 1055
 1056
 1057
 1058
 1059
 1060
 1061
 1062
 1063
 1064
 1065
 1066
 1067
 1068
 1069
 1070
 1071
 1072
 1073
 1074
 1075
 1076
 1077
 1078
 1079
 1080
 1081
 1082
 1083
 1084
 1085
 1086
 1087
 1088
 1089
 1090
 1091
 1092
 1093
 1094
 1095
 1096
 1097
 1098
 1099
 1100
 1101
 1102
 1103
 1104
 1105
 1106
 1107
 1108
 1109
 1110
 1111
 1112
 1113
 1114
 1115
 1116
 1117
 1118
 1119
 1120
 1121
 1122
 1123
 1124
 1125
 1126
 1127
 1128
 1129
 1130
 1131
 1132
 1133
 1134
 1135
 1136
 1137
 1138
 1139
 1140
 1141
 1142
 1143
 1144
 1145
 1146
 1147
 1148
 1149
 1150
 1151
 1152
 1153
 1154
 1155
 1156
 1157
 1158
 1159
 1160
 1161
 1162
 1163
 1164
 1165
 1166
 1167
 1168
 1169
 1170
 1171
 1172
 1173
 1174
 1175
 1176
 1177
 1178
 1179
 1180
 1181
 1182
 1183
 1184
 1185
 1186
 1187
 1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359

RESULTS AND DISCUSSION

A. Nucleotide sequence of the *uvrB* gene

The construction of recombinant plasmids carrying the *Escherichia coli* *uvrB* gene, including its transcriptional and translational regulatory elements has been reported previously (15). Plasmid pNP12, derived from pACYC177 (20) harbours a 4.8 kb long PstI-BamHI insert originating from the colony bank plasmid pLC25-23 (21) which carries the complete wild type *uvrB* gene. The promoter as well as the start of the *uvrB* coding sequence have been localized upstream of a unique EcoRI site (15). The DNA sequence around the EcoRI site (nucleotides 1-219 in fig. 2) has been published previously (van den Berg et al., 1981). In order to determine the complete sequence of the *uvrB* gene we started sequencing at the unique EcoRI site in pNP12 and proceeded towards the BamHI site. Sequences were determined by the dideoxy chain termination method (22) using the M13 phage system (23). Clones for sequencing were generated by introducing either selected restriction fragments from the EcoRI-BamHI region of pNP12 or fragments generated by Bal31 digestion into mp8, mp9 and mp10 phages. A complete overview of the sequencing strategy is presented in fig. 1. Most of the sequence was determined in both strands with the only exception of 280 nucleotides between position 1000 and 1280 which were sequenced only from the (-) strand but in three different clones (fig. 1). Examination of the complete nucleotide sequence (fig. 2) reveals only one long open reading frame initiating at an ATG codon at position 133-135. This initiation codon has previously been assessed as being the start of the *uvrB* gene coding region (15). The open reading frame is terminated at an TAA stop codon at position 2152-2154, leaving an open reading frame of 673 amino acids. From these results a molecular weight of 76.614 daltons can be predicted for the UvrB gene product which is in good agreement with the values of 80 K (15) and 84 K (24) determined previously from SDS-PAGE. Fig. 3 shows that the *uvrB* coding frame can be correctly predicted from theoretical considerations by using the "positional base preference" method of Staden (25). The method relies on the observation that the four bases are unevenly used at different positions in codons namely a preference of G, and to a

Figure 2. Nucleotide sequence of the *uvrB* gene. Promoters and LexA binding site (15) are indicated. Palindromic sequences downstream of the structural gene are represented by arrows. Domains I-IV (see text) are boxed. C471 deleted in UvrB5, initiation and termination codons of UvrB and UvrB5 are indicated.

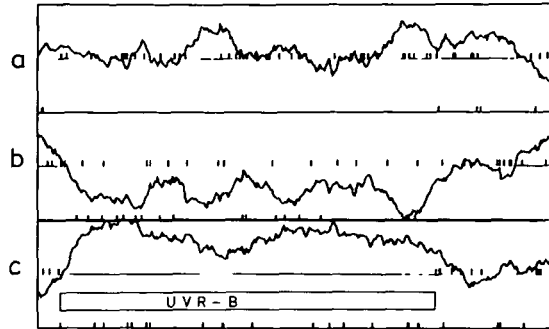


Figure 3. Prediction of the *uvrB* coding sequence using the "positional base preference" method (Staden, 1984). 3 different frames are indicated (a,b,c). The coding part of the *uvrB* gene is indicated in the coding frame (c).

lesser extent A, in position 1 and A rather than G in position 2. The slight uncertainty of the plot around codons 250-300 is probably due to a concentration of hydrophobic amino acids in this region (more T's in second position) (25). Downstream of the termination codon a region of extensive dyad symmetry covering more than 100 nucleotides is encountered (fig 2.). Parts of this region fit with the consensus established for repetitive extragenic sequences (REP sequences) (Fig. 4). These structures have been found in non-coding regions at many other locations on the *E.coli* chromosome (26-28). Their function is not yet known but it has been suggested that they might play a role in mRNA stability (27). The data presented here show that *uvrB* has two REP sequences.

B. Cloning and characterization of the mutant *uvrB5* gene

Chromosomal DNA from strain AB1885 (*uvrB5*) (29) was digested with *Pst*I and *Bam*HI and ligated to pACYC177 (20) digested with the same enzymes. The ligation mixture was used to transform strain HP3435 (Δ *uvrB*) (30) to Km^R . Resistant clones were plated on nitrocellulose filters at a density of 12,000 colonies per 13 cm filter, lysed and hybridized to ^{32}P -nicktranslated *Pst*I-*Bam*HI fragment from plasmid pNP12, harboring the wild type *uvrB* gene. Positive colonies were isolated and found to contain a recombinant plasmid, pBL01, undistinguishable from pNP12 by restriction analysis. However, pBL01 did not confer UV resistance to strain HP3435 (Δ *uvrB*) as pBL01 harbors the *uvrB5* mutant gene. The mutation in pBL01 responsible for the *uvrB5* phenotype was localized in the 1360 nucleotide long *Eco*RI-*Bgl*II fragment (fig. 1) representing the N-terminal part of the

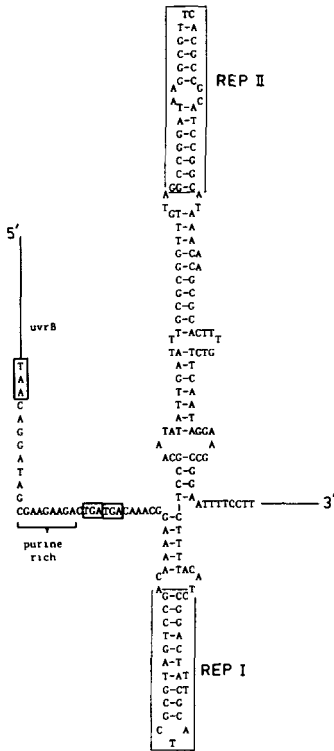


Figure 4. Palindromic sequences downstream of the *uvrB* coding region: the *uvrB* TAA stop codon, as well as two downstream TGA stop codons in the same frame, are boxed. The highly conserved REP sequences are boxed.

uvrB gene, by showing that the mutant phenotype could be transferred from pBL01 to pNP12 by exchanging the EcoRI-BglII fragment between the two plasmids. Hence the EcoRI-BglII fragment from pBL01 was sequenced completely in parallel with the corresponding fragment from the wild type *uvrB* gene. The only difference between the wild type and mutant fragment resides in the absence of a C residue at position 471 in fig. 2 in the case of the *uvrB5* mutant gene (fig. 5). This result indicates that the *uvrB5* phenotype is due to a point deletion of a C residue 338 nucleotides downstream of the *uvrB* initiation codon. The localization of the *uvrB5* mutation in this part of the *uvrB* coding sequence is further substantiated by the following experiment: pBL10 is derived from pBR329 (31) and harbors a 428 nucleotide long TaqI fragment derived from the wild type *uvrB* gene (position 203-631 in fig. 2) cloned in the Tc^R gene. pBL-10 was found to be able to rescue the *uvrB5* mutation by double crossing-over with a frequency of 2×10^{-5} whereas when the vector pBR329 was used alone no rescue was found with a frequency higher than 10^{-9} . The experiment indicates

Downloaded from https://academic.oup.com/nar/article/14/7/2877/2358723 by Medisurf user on 24 April 2024

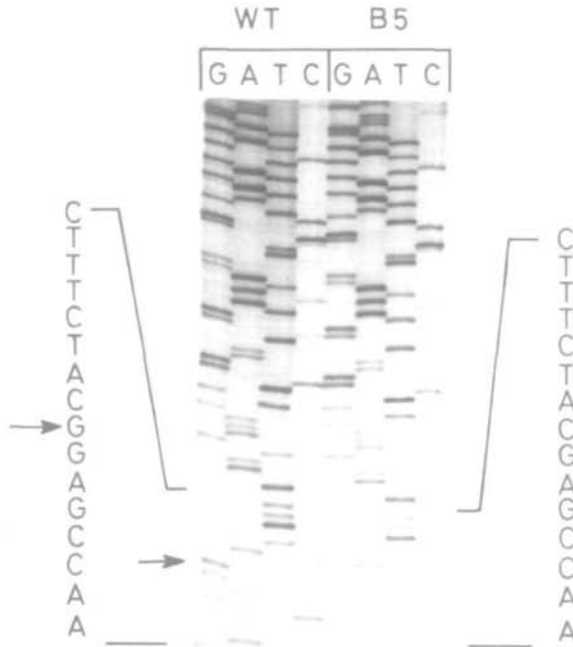


Figure 5. Direct comparison of the DNA sequence of the w.t and uvrB5 gene in the region of the uvrB5 mutation.

that the chromosomal mutation in strain AB1885 is situated in the uvrB coding sequence between position 203 and 631, that it is the only mutation in uvrB responsible for the uvrB5 phenotype and that the possibility that the mutation at position 471 in pBL01 has arisen during the cloning procedure is very unlikely.

From the sequencing data it can be predicted that the UvrB5 protein is composed of the 113 N-terminal amino acids of the wild-type uvrB gene and a 43 aminoacid long tail coded in the -1 frame and terminated at an TGA stop codon at position 602-604 (fig. 2). From these data a molecular weight of 18.047 D can be calculated for the UvrB5 protein. Fig. 6 shows that, in agreement with the theoretical predictions, plasmid pBL01 directs the synthesis of a protein of approximately 17-18 K in an in vitro transcription-translation system.

C. Sequence homologies between the UvrB and UvrC proteins

From the action mechanism of the Uvr excinuclease (simultaneous incision of the damaged DNA strand at two different positions, one at each side of

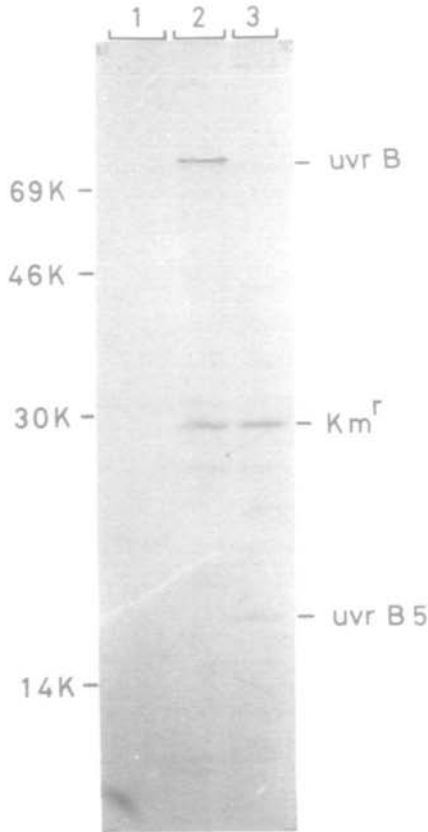


Figure 6. *In vitro* transcription-translation directed either by pNP12 (lane 2) or pBL01 (lane 3). In lane 1 no plasmid has been added to the system.

the DNA adduct) one might expect to find in the protein sequence of the 3 excinuclease subunits homologous regions which might be localized either on one Uvr protein or on several proteins. In this respect it was interesting to compare the protein sequence of the uvrB gene to the published sequence of the UvrC protein (Sancar et al., 1984). Protein sequences were compared using the computer program "DIAGON" of Staden (32). The results represented here were obtained with the "perfect matching" option. Two regions of significant homology between the UvrB and UvrC proteins have been identified (fig. 7). Domain I contains 6 consecutive identical residues (amino acids 352-357 in UvrB and 326-331 in UvrC) flanked by a tyrosine residue in

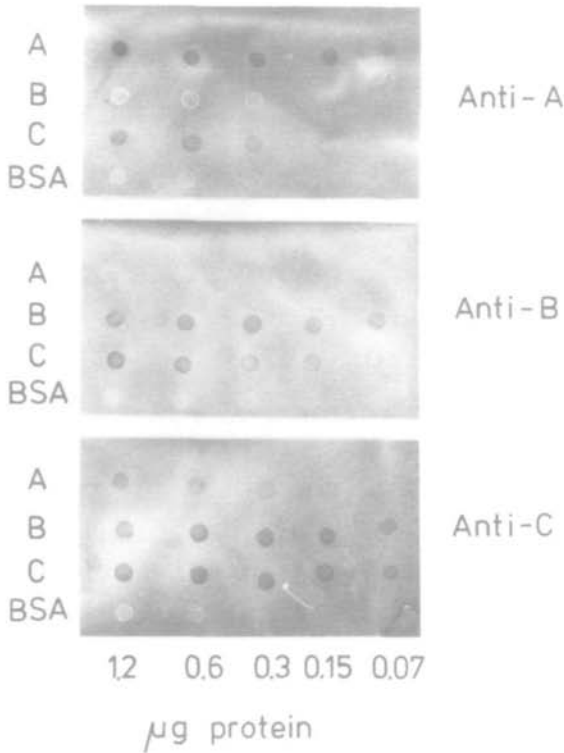


Figure 8. Dot immunobinding assay with anti-UvrA, anti-UvrB and anti-UvrC antibodies. Purified Uvr proteins (43) were spotted onto nitrocellulose membranes at the amounts indicated and assayed with the specific antibody essentially as described by Hawkes et al., 1982. Bovine serum albumin (BSA) was used as a control.

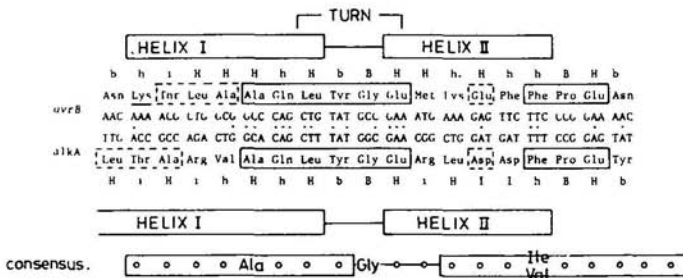


Figure 9. Homology between the UvrB and the AlkA proteins: The Chou and Fassmann helical assessments for each aminoacid are indicated (H: strong helix former; h: helix former; I: weak helix former; i: indifferent; B: strong helix breaker; b: weak helix breaker) (Chou and Fassman, 1978). The consensus proposed for bihelical regions is represented (Pabo and Sauer, 1984).

Downloaded from https://academic.oup.com/nar/article/14/7/2877/2358723 by Medisurf user on 24 April 2024

Protein	Residues	Sequences
ATPase β , ox	150-171	G G K I G <u>L</u> <u>F</u> <u>G</u> G A G V <u>C K T</u> V F I M E L <u>T</u>
ATPase β	143-164	G G <u>K</u> V C <u>L</u> <u>F</u> <u>G</u> G A C V C K T V N M M E L <u>I</u>
ATPase α	162-183	G Q <u>R</u> E L I I G D R Q T C K T A L A I D A I <u>V</u>
AMP kinase, pig	8- 29	S K I I F V V G G P G S G K G T Q C E K I <u>V</u>
Myosin, nematode	162-183	N Q S M L I T G E S G A G K T E N T K K V I <u>I</u>
Myosin, rabbit	171-192	N Q S I L I T G E S G A C K T V N T K R V I <u>I</u>
RecA protein	59- 80	G R I V E I Y G P E S S C K T T L T L Q V I <u>I</u>
DnaB protein	223-244	S <u>D</u> L I I V A A R P S M G K T T F A M N L V <u>I</u>
Rho protein	172-192	G Q R G L I V A P P K A G K T M L L Q N I A <u>I</u>
TdR kinase, HSF	47- 69	L L <u>E</u> V Y I D G P H G M C K T T T T Q L L V <u>I</u>
UvrD protein	22-43	R S <u>N</u> L L V <u>L</u> <u>A</u> G A G S <u>C K T</u> R V L V H R <u>I</u>
UvrB protein	32-53	L A <u>H</u> Q T <u>L</u> <u>L</u> C V T G S <u>C K T</u> F T I A N V <u>I</u>

Figure 10. Alignment of homologous sequences in the UvrB protein and other adenine nucleotide binding proteins (34). Similar sequences are boxed. Conserved basic residues are underlined. The organism was *E.coli* unless specified.

flanked by chemically related amino acids. Using the Chou and Fassman rules it is possible to predict a helix-turn-helix motive for this region. Helix I fits the consensus proposed by Pabo and Sauer (36), whereas helix II is shorter in the case of UvrB and AlkA and the conserved Ile/Val residue is replaced by another hydrophobic amino acid namely phenylalanine. Such bihelical domains have been found to be implicated in protein-DNA recognition in a number of different DNA binding proteins (for a review see 36). It is interesting, in this respect, to mention that the AlkA enzyme, like the Uvr excinuclease does not recognize one specific DNA adduct, but a large number of different (alkylation-) products, both on purine and pyrimidine nucleotides (37). As with the Uvr excinuclease, it has been reasoned that AlkA recognizes in a first instance a deformation or irregularity in the DNA helix rather than the DNA adduct itself (37). Hence, Uvr excinuclease and AlkA might harbour homologous domains, involved in helix irregularity recognition. That such a domain would be localized on the UvrB excinuclease subunit, is in agreement with experimental findings (see introduction), which show that UvrB targets the excinuclease specifically to damaged DNA. Obviously, such a hypothesis has to be considered with care. However, we believe that the results reported here might provide a valuable tool to study the damage recognition process in both systems.

Comparison of the UvrB and UvrD sequences (34) revealed the presence of an identical pentapeptide, Gly-Ser-Gly-Lys-Thr, localized at position 42-46 in UvrB and 32-36 in UvrD. Interestingly, this pentapeptide, in UvrD, is part of a consensus sequence found in a large number of ATP binding proteins (34,38). The sequence found in UvrB completely fits into this

proposed consensus (fig. 10; domain IV in fig. 2). What might be the reason for such an ATP binding site in UvrB? UvrB is not an ATPase by itself, but stimulates efficiently the UvrA ATPase activity, when the AB complex is assembled on damaged DNA (see Introduction). These results, taken together with the structural data reported here, are in favor of a model where ATP binding site (situated on UvrB) and ATPase site (situated on UvrA) have to be correctly juxtaposed (a situation obtained when the complex is assembled at a damaged site in DNA), in order to be able to efficiently hydrolyse ATP. Physical measurements have indicated that the extent of the deformation induced in DNA by different agents can vary greatly (Pearlman et al., 1985). Nevertheless it has been shown that DNA damaged by these agents is incised by the Uvr excinuclease in a very similar way (see Introduction). We tentatively propose that ATP hydrolysis is needed to bring damaged DNA into a common, incisable, conformation. Interestingly, an ATP binding site consensus has as well been identified in the *rad3* gene (40) as well as in several other *rad* genes (E.C. Friedberg, personal communication), all involved in excision repair in *Saccharomyces cerevisiae*, suggesting that the model proposed here might not only be confined to *Escherichia coli*.

ACKNOWLEDGEMENTS

The valuable contributions of G.F.D. Moolenaar, T. Kartasova and P.A. van Rijn are acknowledged. A. Yasui and C.A. van Sluis were helping with computer analysis. Thanks to E.A. van den Berg, J.A. Brandsma and J. Brouwer for reading the manuscript and N. van Hoek for the typographic work. This work was supported by EURATOM (contract No. BIO-E;408-NL) and FUNGO (# 13-23-35).

REFERENCES

1. Friedberg, E.C. (1985) DNA Repair, Freeman Co., New York.
2. Thomas, D.C., Levy, M. and Sancar, A. (1985) *J. Biol. Chem.* **260**, 9875-9883.
3. Sancar, A. and Rupp, W.D. (1983) *Cell* **33**, 249-260.
4. Beck, D.J., Popoff, S., Sancar, A. and Rupp, W.D. (1985) *Nucl. Acids Res.* **13**, 7395-7412.
5. Sancar, A., Franklin, K.A. and Sancar, G. (1985) *J. Mol. Biol.* **184**, 725-734.
6. Seeberg, E. and Steinum, A-L. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 988-992.
7. Yeung, A.T., Mattes, W.B., OH, E.Y. and Grossman, L. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6157-6161.
8. Seeberg, E. and Steinum, A-L. (1983) in: Cellular Responses to DNA Damage, E.C. Friedberg and B.A. Bridges eds., Alan R. Liss Inc., New York, pp 39-49.
9. Caron, P.R., Kushner, S.R. and Grossman, L. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4925-4929.

10. Kumura, K., Sekiguchi, M., Steinum, A-L. and Seeberg, E. (1985) *Nucl. Acids Res.* 13, 1483-1492.
11. Seeberg, E., Nissen-Meyer, J. and Strike, P. (1976) *Nature* 263, 524-526.
12. Kenyon, C.J. and Walker, G.C. (1981) *Nature* 289, 808-810.
13. Fogliano, M. and Schendel, P. (1981) *Nature* 289, 196-198.
14. Van Sluis, C.A., Moolenaar, G.F. and Backendorf, C. (1983) *EMBO J.* 2, 2313-2318.
15. Van den Berg, E.A., Zwetsloot, J., Noordermeer, I.A., Pannekoek, H., Dekker, R.B., Dijkema, R. and van Ormondt, H. (1981) *Nucl. Acids Res.* 9, 5623-5643.
16. Sancar, G.B., Sancar, A., Little, J.W. and Rupp, W.D. (1982) *Cell* 28, 523-530.
17. Van den Berg, E.A., Geerse, R.H., Pannekoek, H. and van de Putte, P. (1983) *Nucl. Acids Res.* 13, 4355-4363.
18. Van den Berg, E.A., Geerse, R.H., Memelink, J., Bovenberg, R.A.L., Magnee, E.A. and van de Putte, P. (1985) *Nucl. Acids Res.* 13, 1829-1840.
19. Sancar, G.B., Sancar, A. and Rupp, W.D. (1984) *Nucl. Acids Res.* 12, 4593-4608.
20. Chang, A.C.J. and Cohen, S.N. (1978) *J. Bacteriol.* 134, 1141-1156.
21. Clarke, L. and Carbon, J. (1976) *Cell* 9, 91-99.
22. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
23. Messing, J. and Vierra, J. (1982) *Gene* 19, 269-279.
24. Sancar, A., Clarke, N.D., Griswold, J., Kennedy, W.J. and Rupp, W.D. (1981) *J. Mol. Biol.* 148, 63-76.
25. Staden, R. (1984) *Nucl. Acids Res.* 12, 521-538.
26. Higgins, C.F., Ferro-Luzzi Ames, G., Barnes, W.M., Clement, J-M. and Hofnung, M. (1982) *Nature* 298, 760-762.
27. Stern, M.J., Ferro-Luzzi Ames, G., Smith, N.H., Robinson, E.C. and Higgins, C.F. (1984) *Cell* 37, 1015-1026.
28. Gilson, E., Clement, J-M., Brüttag, D. and Hofnung, M. (1984) 3, 1417-1421.
29. Howard-Flanders, P., Boyce, R.P. and Theriot, L. (1966) *Genetics* 53, 1119--1136.
30. Pannekoek, H., Noordermeer, I.A. and van de Putte, P. (1979) *J. Bacteriol.* 139, 48-53.
31. Covarrubias, L. and Bolivar, F. (1982) *Gene* 17, 79-89.
32. Staden, R. (1982) *Nucl. Acids Res.* 0, 2951-2961.
33. Chou, P.J. and Fassman, G.D. (1978) *Adv. enzymol.* 47, 45-148.
34. Finch, P.W. and Emmerson, P.T. (1984) *Nucl. Acids Res.* 12, 5789-5799.
35. Nakabeppu, J., Miyata, T., Kondo, H., Iwanaga, S. and Sekiguchi, M. (1984) *J. Biol. Chem.* 259, 13730-13736.
36. Pabo, C.O. and Sauer, R.T. (1984) *Ann. Rev. Biochem.* 53, 293-321.
37. McCarthy, T.V., Karran, P. and Lindhal, T. (1984) *EMBO J.* 3, 545-550.
38. Walker, J.E., Sarasie, M., Runswick, M.J. and Gay, N.J. (1982) *EMBO J.* 1, 945-951.
39. Pearlman, D.A., Holbrook, S.R., Pirkle, D.H. and Kim, S-H (1985) *Science* 227, 1304-1308.
40. Reynolds, P., Higgins, D.R., Prakash, L. and Prakash, S. (1985) *Nucl. Acids Res.* 13, 2373-2386.
41. Hawkes, R., Niday, E. and Gordon, J. (1982) *Anal. Biochem.* 119, 142-147.
42. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual*, CSH.
43. Zwetsloot, J.M.C., Barbeiro, A.P., Vermeulen, W., Arthur, H.M., Hoeijmakers, J.H.J. and Backendorf, C. (1986) *Mutation Res.*, accepted.