Caractérisation de l'interactome protéique des lésions de l'ADN : application aux lésions d'oxydation

Ewa Barbier

▶ **To cite this version:**

Ewa Barbier. Caractérisation de l'interactome protéique des lésions de l'ADN : application aux lésions d'oxydation. Biologie moléculaire. Université de Grenoble, 2013. Français. NNT : 2013GRENV032 . tel-01680308v2

HAL Id: tel-01680308

https://theses.hal.science/tel-01680308v2

Submitted on 11 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Chimie Biologie**

Arrêté ministériel : 7 août 2006

Présentée par

Ewa BARBIER née PIETRAS

Thèse dirigée par **Didier GASPARUTTO** et
codirigée par **Jean BRETON**

préparée au sein du **Laboratoire Lésions des Acides Nucléiques (SCIB – UMR-E 3 CEA-UJF**)
dans **l'École Doctorale Chimie et Sciences du Vivant**

Caractérisation de l'interactome protéique des lésions de l'ADN : application aux lésions d'oxydation.

Thèse soutenue publiquement le **2 octobre 2013**
devant le jury composé de :

Pr Eric PEYRIN
Professeur, Université de Grenoble, Président
Dr Wilfrid BOIREAU
Chargé de recherche, CNRS – Besançon, Rapporteur
Dr Dimitar ANGUELOV
Directeur de recherche, CNRS – Lyon, Rapporteur
Pr Fabrice CONFALONIERI
Professeur, Université Paris-Sud, Examinateur
Dr Jean BRETON
Maître de conférences, Université de Grenoble, Examinateur
Dr Didier GASPARUTTO
Ingénieur CEA, Grenoble, Examinateur

Université Joseph Fourier / Université Pierre Mendès France / Université Stendhal / Université de Savoie / Grenoble INP

Sommaire

Liste des figures 5
Liste des tableaux 9
Abréviations 11
Introduction 15
I Analyse bibliographique 17
1. Stress oxydant et son effet sur les cellules 17
1.1. Formation des radicaux libres 18
1.1.1. Espèces réactives de l'oxygène 18
1.1.2. Espèces réactives de l'azote 21
1.2. Sources du stress oxydant 21
1.2.1. Stress oxydant endogène 22
1.2.1.1. Respiration cellulaire 22
1.2.1.2. Inflammation 22
1.2.2. Stress oxydant exogène 23
1.2.2.1. Substances chimiques 23
1.2.2.2. Rayonnement ionisant 25
1.2.2.3. Photosensibilisation 26
1.2.2.4. Agents infectieux 27
1.3. Systèmes de défense antioxydants des cellules 28
1.4. Effet biologique du stress oxydant 29
1.4.1. Influence du stress oxydant sur les cellules 29
1.4.1.1. Apoptose 29
1.4.1.2. Nécrose 31
1.4.1.3. Régulation du cycle cellulaire et sénescence 31
1.4.1.4. Vieillissement 32
1.4.2. Interactions avec les composants cellulaires 33
1.4.2.1. Protéines 33
1.4.2.2. Lipides 35
1.4.2.3. Acides nucléiques 37
2. Lésions de l'ADN générées par le stress oxydant 39
2.1. Oxydation du désoxyribose 39
2.2. Oxydation des bases : lésions simples 40
2.2.1. Types de lésions et leur formation 40
2.2.1.1. Oxydation des purines 40
2.2.1.2. Oxydation de la thymine 42
2.2.1.3. Oxydation de la cytosine 45
2.2.1.4. Autres lésions simples 45
2.2.2. Prise en charge par le système de réparation BER 46
2.2.2.1. BER chez les eucaryotes 47
2.2.2.2. BER chez les procaryotes 51
2.2.3. Conséquences de la présence des lésions simples non réparées dans l'ADN 52
2.3. Lésions complexes 54
2.3.1. Formation et types 54
2.3.2. Réparation 55
2.3.3. Conséquences 57
2.3.4. Cas particulier des cassures de l'ADN 58
3. Cyclonucléosides et leurs conséquences pour les cellules 60
3.1. Caractéristiques chimiques des cyclonucléosides 60

3.1.1. Structure et mécanisme de formation 60
3.1.2. Formation des cyclonucléosides dans les cellules 62
3.1.3. Effet sur la structure de l'ADN 63
3.2. Conséquences biologiques 64
3.2.1. Influence sur la synthèse de l'ADN 64
3.2.2. Influence sur la transcription de l'ADN 65
3.2.3. Influence sur l'activité d'autres enzymes 66
3.3. Réparation 66
3.3.1. Prise en charge par le système de réparation NER 67
3.3.2. Conséquences de l'insuffisance du NER 72
3.3.3. NER et cyclonucléosides : tests de réparation *in vitro* et dans des modèles cellulaires 74
3.3.4. NER et cyclonucléosides : tests de réparation sur les modèles animaux 75
3.4. Implication dans les pathologies humaines 76
3.4.1. Maladies neurodégénératives (XP , CS) 76
3.4.2. Cancers 78

II Stratégie et objectifs 81
1. Notion d'interactome et objectif de la thèse 81
2. Stratégie globale 82
2.1. Screening des protéines 82
2.2. Travail sur les protéines isolées 84
3. Modèles utilisés 85
3.1. Lignée humaine HeLa 85
3.2. Enzyme d'*Escherichia coli* : la Fpg 86
3.3. Enzymes de *Thermococcus gammatolerans* 88
3.3.1. Genre *Thermococcus* 88
3.3.2. Description de l'organisme et de ses conditions de vie 88
3.3.3. Génome et protéome de *T. gammatolerans* 89
3.3.4. Systèmes de réparation de l'ADN chez les Archées thermophiles 90
3.3.5. Radiorésistance et événements dans l'organisme après l'irradiation 91
4. Structure du manuscrit 92

III Matériel et méthodes 93
1. Screening des protéines : construction du piège à protéines 93
1.1. Principe de la technique du ligand fishing 93
1.2. Etapes d'optimisation du piège 95
1.3. Protocoles détaillés 95
1.3.1. Synthèse des oligonucléotides 95
1.3.2. Purification des oligonucléotides 98
1.3.3. Phosphorylation des oligonucléotides 98
1.3.4. Formation des concatémères sur les billes 99
1.3.5. Culture d'*Escherichia coli*, extraction des protéines bactériennes et leur dosage 101
1.3.6. Exposition du piège aux extraits et récupération des protéines 102
1.3.7. Récupération des protéines en utilisant le bras photoclivable 102
1.3.8. Coloration au nitrate d'argent 103
1.3.9. Analyse protéomique 103
2. Techniques de validation des interactions 104
2.1. Western blot 105
2.2. EMSA 106
2.2.1. Principe 106
2.2.2. Protocole détaillé 107

2.2.3. Analyse des résultats et calculs de K_d 109
3. Tests d'activité enzymatique 110
3.1. Test de coupure sur gel 110
3.1.1. Principe 110
3.1.2. Protocole détaillé 111
3.2. Analyse de masse par MALDI-TOF 112
3.2.1. Principe 112
3.2.2. Protocole détaillé 113
4. Dosage des lésions de l'ADN 114
4.1. Principe général 114
4.2. Protocoles détaillés 115
4.2.1. Culture de *Thermococcus gammatolerans* 115
4.2.2. Extraction et digestion de l'ADN de *Thermococcus gammatolerans* 116
4.2.3. Dosage des lésions 118
4.2.3.1. Dosage de la 8-oxo-désoxyguanosine 118
4.2.3.2. Dosage des cyclonucléosides 119
4.2.4. Tests statistiques 120

IV Résultats 121
1. Screening des extraits protéiques dans le but d'identifier l'interactome des cyclonucléosides 121
1.1. Premier format de piège testé : courtes sondes oligonucléotidiques 121
1.1.1. Mise au point du bras photoclivable 121
1.1.2. Essais de piégeage de HMGB1 sur les adduits du cisplatine 123
1.1.3. Analyse des protéines piégées sur les cyclonucléosides 124
1.2. Deuxième format de piège testé : concatémères 126
1.2.1. Mise au point des conditions de la phosphorylation des oligonucléotides 127
1.2.2. Mise au point des conditions de la ligation des oligonucléotides 133
1.2.3. Formation des concatémères sur les billes 136
1.2.4. Comparaison de différentes conditions de récupération des protéines 139
1.3. Exposition des pièges aux extraits protéiques 140
1.4. Identification par analyse protéomique des protéines piégées 142
1.4.1. Aspects quantitatifs des trois analyses protéomiques 142
1.4.1.1. Première analyse et choix des conditions d'incubation 142
1.4.1.2. Deuxième analyse : sondes 5'S-cyclo-dG et 5'R-cyclo-dA 144
1.4.1.3. Troisième analyse : utilisation du panel complet des cyclonucléosides 145
1.4.2. Analyse qualitative des protéines identifiées 146
1.4.3. Choix des candidats pertinents 150
1.4.3.1. NONO 150
1.4.3.2. SFPQ 150
1.4.3.3. PARP1 151
1.4.3.4. DREF 152
1.5. Validation des candidats identifiés 152
1.5.1. Western Blot 153
1.5.2. EMSA 157
1.6. Identification par analyse protéomique des protéines bactériennes piégées sur les cyclonucléosides 159
1.6.1. Premier format de piège : sondes courtes 160
1.6.2. Deuxième format de piège : concatémères 161
2. Travail sur les protéines isolées 163
2.1. Etude biochimique de la protéine Fpg 163
2.1.1. Activité de coupure 163

2.1.1.1. Test d'activité de la Fpg sur les lésions simples 163
2.1.1.2. Test d'activité de la Fpg sur les lésions complexes d'oxydation 166
2.1.2. Analyse des interactions entre les cyclonucléosides et la Fpg 166
2.1.2.1. Optimisation des conditions de l'EMSA 168
2.1.2.2. Formation des complexes entre la Fpg et les sondes 171
2.1.3. Influence de la longueur de la sonde sur les interactions non-spécifiques avec la Fpg 180
2.2. Etude biochimique des protéines isolées chez *Thermococcus gammatolerans* 184
2.2.1. Etude d'un modèle original : *Thermococcus gammatolerans* 184
2.2.1.1. Mise au point du protocole d'extraction de l'ADN de *Thermococcus gammatolerans* 184
2.2.1.2. Dosage de la 8-oxo-désoxyguanosine 186
2.2.1.3. Dosage des cyclonucléosides 188
2.2.2. Activité de coupure 190
2.2.2.1. Test d'activité sur gel 190
2.2.2.2. Vérification du mécanisme de coupure par MALDI-TOF/MS 192
2.2.2.2.1. Coupure de la 8-oxo-dG 193
2.2.2.2.2. Coupure de la 5-OH-5-Me-dHyd 197
2.2.2.3. Activité de coupure en fonction de la base opposée 202
2.2.2.4. Test d'activité des protéines de *T.gammatolerans* sur les lésions complexes d'oxydation 205
2.2.3. Analyse des interactions entre les cyclonucléosides et les protéines de *Thermococcus gammatolerans* 206

V Discussion 211
1. Screening des extraits protéiques 211
1.1. Considérations techniques 211
1.1.1. Utilité du bras photoclivable 211
1.1.2. Exhaustivité des protéines piégées 212
1.1.3. Application à différents types d'extraits protéiques 213
1.1.4. Utilisation des pièges de type concatémères 214
1.2. Pertinence et analyse des protéines sélectionnées 216
1.2.1. PARP1 217
1.2.2. DREF 220
2. Travail sur les protéines isolées 222
2.1. Modèle bactérien : glycosylase Fpg d'*Escherichia coli* 222
2.1.1. Interactions de la Fpg avec l'ADN non-lésé 223
2.1.2. Interactions de la Fpg avec les cyclonucléosides 224
2.2. *Thermococcus gammatolerans* 228
2.2.1. Dosage des lésions d'oxydation 229
2.2.1.1. Dosage de la 8-oxo-dG 229
2.2.1.2. Dosage des cyclonucléosides 231
2.2.2. Analyse des protéines purifiées 232
2.2.2.1. Activité de coupure 232
2.2.2.2. Interactions avec les cyclonucléosides 234

VI Conclusions et perspectives 237

Bibliographie 243

Annexes 263

Liste des figures

Figure 1 : Formation des espèces réactives de l'oxygène dans les mitochondries. 19
Figure 2 : Formation des espèces réactives de l'oxygène et de l'azote. 20
Figure 3 : Formation photochimique des espèces réactives de l'oxygène. 27
Figure 4 : Induction de l'apoptose par la voie extrinsèque ou intrinsèque. 30
Figure 5 : Exemples des produits d'oxydation des acides aminés. 34
Figure 6 : Peroxydation des lipides. 36
Figure 7 : Exemples des produits d'oxydation du désoxyribose. 39
Figure 8 : Schéma de l'oxydation de la guanine. 41
Figure 9 : Schéma de l'oxydation de la thymine. 43
Figure 10 : Schéma de l'oxydation de la cytosine. 44
Figure 11 : Représentation schématique du système BER. 47
Figure 12 : Mécanisme de β- et β,δ-élimination. 49
Figure 13 : Types de lésions complexes. 55
Figure 14 : Destin de différents types des clusters générés par le rayonnement ionisant. 57
Figure 15 : Formation des cyclonucléosides. 61
Figure 16 : Configuration adoptée par le sucre dans la 5'S-cyclo-dG. 63
Figure 17 : Exemples d'adduits volumineux étant les substrats du NER. 68
Figure 18 : Représentation schématique du système NER 69
Figure 19 : Stratégie globale des travaux effectués. 82
Figure 20 : Détermination de l'interactome des lésions complexes de l'ADN. 83
Figure 21 : Stabilisation de la base de Schiff à l'aide du $NaBH_4$. 87
Figure 22 : Représentation schématique du piège à protéines. 94
Figure 23 : Synthèse d'ADN par la méthode phosphoramidite. 96
Figure 24 : Piège à protéines (« ligand fishing »), basé sur une approche concatémères. 99
Figure 25 : Structure d'une biotine TEG et d'une biotine couplée à un bras photoclivable. 100
Figure 26 : Quantification de gel radioactif. 110
Figure 27 : Fragmentation de l'ion parent de la 8-oxo-dG. 119
Figure 28 : Fragmentation des ions parents de la cyclo-dG et de la cyclo-dA. 120
Figure 29 : Structure d'un bras photoclivable. 122
Figure 30 : Décrochage des sondes contenant un bras photoclivable à l'aide des UV. 123
Figure 31 : Piégeage de HMGB1 sur une lésion générée par le cisplatine. 124
Figure 32 : Analyse quantitative des protéines récupérées par irradiation UVA. 125
Figure 33 : Phospohorylation des oligonucléotides fluorescents. 128
Figure 34 : Spectre MALDI-TOF de l'oligonucléotide fluorescent. 129
Figure 35 : Spectre MALDI-TOF de l'oligonucléotide fluorescent, phosphorylé avec 0.5 U de T4 kinase pour 100 pmol d'oligonucléotide. 130
Figure 36 : Spectre MALDI-TOF de l'oligonucléotide fluorescent, phosphorylé avec 2 U de T4 kinase pour 100 pmol d'oligonucléotide. 130
Figure 37 : Spectre MALDI-TOF de l'oligonucléotide fluorescent, phosphorylé avec 10 U de T4 kinase pour 100 pmol d'oligonucléotide. 130
Figure 38 : Spectre MALDI-TOF de l'oligonucléotide 5'S-cyclo-dG. 132
Figure 39 : Spectre MALDI-TOF de l'oligonucléotide 5'S-cyclo-dG phosphorylé. 132
Figure 40 : Optimisation de la formation des concatémères en solution : conditions d'hybridation et température de ligation. 134
Figure 41 : Optimisation de la formation des concatémères en solution : quantité de la T4 ligase et quantité relative des séquences « stop ». 135
Figure 42 : Vérification de la dégradation des concatémères par les irradiations. 137
Figure 43 : Détermination de la taille des concatémères formés directement sur les billes. 138

Figure 44 : Comparaison de la quantité de protéines récupérée après le piégeage, en effectuant une élution dénaturante (Laemmli) ou une irradiation d'un bras photoclivable. 139
Figure 45 : Comparaison de la quantité des protéines piégées dans différentes conditions d'incubation. 141
Figure 46 : Aspects comparatifs de la première analyse protéomique 143
Figure 47 : Aspects comparatifs de la deuxième analyse protéomique. 145
Figure 48 : Aspects comparatifs de la troisième analyse protéomique. 146
Figure 49 : Identification par Western blot de NONO et SFPQ 154
Figure 50 : Identification par Western blot de PARP1 et DREF. 155
Figure 51 : Cinétique du piégeage de PARP1 156
Figure 52 : Interactions entre DREF et les sondes : non-lésée, 5'S-cyclo-dG ou scramble 158
Figure 53 : Interactions entre PARP1 et les sondes : non-lésée ou 5'S-cyclo-dG. 159
Figure 54 : Analyse quantitative des protéines d'*E.coli*, identifiées sur les sondes 5'S-cyclo-dG ou non-lésées, récupérées par irradiation UVA. 160
Figure 55 : Comparaison de la quantité de protéines d'*E.coli* piégées dans différentes conditions d'incubation 161
Figure 56 : Structures des lésions simples utilisées lors des tests d'activité enzymatique. 164
Figure 57 : Activité de coupure de Fpg et d'Endo III sur les lésions simples d'oxydation. 165
Figure 58 : Activité de coupure de Fpg et d'Endo III sur les cyclonucléosides 166
Figure 59 : Structure d'un site abasique « naturel » et de son homologue THF. 167
Figure 60 : Formation des complexes spécifiques et non-spécifiques entre la Fpg et les sondes lésées ou non. 169
Figure 61 : Formation des complexes spécifiques et non-spécifiques entre la Fpg et les sondes lésées (5'S-cyclo-dA) ou non, après la modification des conditions de l'incubation. 170
Figure 62 : Interactions entre un oligonucléotide non-lésé et la Fpg. 171
Figure 63 : Interactions entre une 5'S-cyclo-dA et la Fpg. 172
Figure 64 : Interactions entre une 5'R-cyclo-dA et la Fpg. 172
Figure 65 : Interactions entre une 5'S-cyclo-dG et la Fpg 173
Figure 66 : Interactions entre une 5'S-cyclo-dHT et la Fpg. 173
Figure 67 : Interactions entre un THF et la Fpg 174
Figure 68 : Différents modèles mathématiques utilisés pour déterminer la K_d 177
Figure 69 : Interations entre les sondes lésées ou non, et la protéine Fpg. 179
Figure 70 : Interaction entre la Fpg et les sondes non-lésées de différentes longueurs. 181
Figure 71 : Optimisation du protocole d'extraction de l'ADN de *Thermococcus gammatolerans*. 185
Figure 72 : Dosage de la 8-oxo-dG, exemple des chromatogrammes. 186
Figure 73 : Dosage de la 8-oxo-dG dans l'ADN de *Thermococcus gammatolerans*. 187
Figure 74 : Dosage des cyclo-dA dans l'ADN de *Thermococcus gammatolerans*. 189
Figure 75 : Activité de coupure des enzymes de *Thermococcus gammatolerans* sur les lésions simples d'oxydation. 191
Figure 76 : Mécanisme de coupure d'un site abasique par β- et δ-élimination 192
Figure 77 : Spectres MALDI-TOF de l'oligonucléotide 8-oxo-dG, digéré par la Fpg ou par une des deux enzymes isolée de *T.gammatolerans*. 194
Figure 78 : Oligonucléotide 8-oxo-dG et de ses fragments obtenus après la coupure 197
Figure 79 : Spectres MALDI-TOF de l'oligonucléotide 5-OH-5-Me-dHyd, digéré par la Fpg, l'Endo III ou tg1277. 199
Figure 80 : Spectre MALDI-TOF de l'oligonucléotide 5-OH-5-Me-dHyd, digéré par tg1653. 200
Figure 81 : Oligonucléotide 5-OH-5-Me-dHyd et de ses fragments obtenus après la coupure. 202
Figure 82 : Hybridation des oligonucléotides 8-oxo-dA et 8-oxo-dG avec leurs brins complémentaires. 203
Figure 83 : Coupure de la 8-oxo-dA par les enzymes de *T.gammatolerans* en fonction de la base complémentaire. 203

Figure 84 : Coupure de la 8-oxo-dG per les enzymes de *T.gammatolerans* en fonction de la base complémentaire. 204
Figure 85 : Activité de coupure des enzymes de *T.gammatolerans* sur les cyclonucléosides. 206
Figure 86 : Interactions entre un THF et les enzymes : Fpg, tg1653 et tg1277, dans les conditions optimisées pour la Fpg. 207
Figure 87 : Interactions entre les cyclonucléosides et tg1653. 208
Figure 88 : Interactions entre les cyclonucléosides et tg1277. 209
Figure 89 : Schéma présentant l'hypothèse de la reconnaissance des sondes par PARP1 et son départ consécutif dans le temps. 220

Liste des tableaux

Tableau 1 : Lésions de l'ADN formées de manière endogène dans les cellules humaines.....46
Tableau 2 : Glycosylases humaines et leurs substrats. 48
Tableau 3 : ADN glycosylases identifiées chez les *E.coli* et chez les hommes.....52
Tableau 4 : Exemples des niveaux de le 5'S-cyclo-dA mesurée par LC/MS dans différents tissus humains et animaux non-irradiés. 62
Tableau 5 : Modification des angles de torsion dans le sucre suite à la formation d'une 5'S-cyclo-dG. 63
Tableau 6 : Principales protéines du système NER.....72
Tableau 7 : Séquences et modifications des oligonucléotides utilisés au cours de ce travail.....97
Tableau 8 : Concentrations des oligonucléotides utilisés pour préparer les pièges. 98
Tableau 9 : Tampons utilisés au cours de l'optimisation du piégeage des protéines. 102
Tableau 10 : Anticorps utilisés pour les expériences de Western Blot. 106
Tableau 11 : Protéines recombinantes utilisées lors des expériences de retard sur gel. 108
Tableau 12 : Enzymes utilisées pour effectuer les tests de coupure. 111
Tableau 13 : Identification protéomique des protéines : NONO, SFPQ, PARP1 et DREF. 149
Tableau 14 : Protéines bactériennes identifiées dans différentes conditions de l'incubation.....162
Tableau 15 : R^2, K_d et n obtenus avec le modèle de Hill, pour les interactions entre la Fpg et les sondes lésées ou non-lésées. 179
Tableau 16 : R^2, K_d et n obtenus avec le modèle de Hill, pour les interactions entre la Fpg et les sondes non-lésées de différentes longueurs. 183
Tableau 17 : Exemples des sondes utilisées pour le piégeage des protéines reconnaissant différentes modifications de l'ADN, et les protéines identifiées avec ces approches. 215
Tableau 18 : K_d de la Fpg pour différentes lésions retrouvées dans la littérature.....227
Tableau 19 : Dosage de la 8-oxo-dG dans différents modèles avant ou après traitement avec des rayonnements γ. 230

Abréviations

3-HPA	acide 3-hydroxypicolinique
5-OH-dHyd	5-hydroxy-désoxyhydantoïne
5-Me-5-OH-dHyd	5-méthyl-5-hydroxy-désoxyhydantoïne
5'R-cyclo-dA	(5'R)-5'-8-cyclo-2'-désoxyadénosine
5'S-cyclo-dA	(5'S)-5'-8-cyclo-2'-désoxyadénosine
5'R-cyclo-dG	(5'R)-5'-8-cyclo-2'-désoxyguanosine
5'S-cyclo-dG	(5'S)-5'-8-cyclo-2'-désoxyguanosine
5'S-cyclo-dHT	(5'S)-5'-6-cyclo-2'-désoxydihydrothymidine
6-4 PP	6-4 photoproduit
8-Br-dG	8-bromo-2'-désoxyguanosine
8-Me-dG	8-méthyl-2'-désoxyguanosine
8-oxo-dA	8-oxo-2'-désoxyadénosine
8-oxo-dG	8-oxo-2'-désoxyguanosine
8-oMe-dG	8-méthoxy-2'-désoxyguanosine
8-oxo-dI	8-oxo-2'-désoxyinosine
8-oxo-dN	8-oxo-2'-désoxygébularine
αKGDH	α-kétoglutarate déshydrogénase
ACN	acétonitrile
AMP	adénosine monophosphate
APE1	apurinic endonuclease 1
ATM	kinase−transducteur : ataxia telangiectasia mutated
ATR	kinase−transducteur : ataxia telangiectasia mutated Rad3-related
ATP	adénosine triphosphate
ATT	6-aza-2-thiothymine
BER	Base Excision Repair ; réparation par excision de bases
BRCA1	breast cancer 1
BRCT	BRCA1 C Terminus (BRCT) domain
BSA	bovine serum albumin; albumine de serum bovin
CAK	cdk-activating kinase
Cb5R	cytochrome b5 réductase
Cdc25A	Cell division cycle 25 homolog A
CDK	cyclin-dependent kinase
CHK1	Serine/threonine-protein kinase Chk1
CHK2	Serine/threonine-protein kinase Chk2
CPD	cyclobutane pyrimidine dimer
CS	syndrome de Cockayne
cyclo-dA	cyclodésoxyadénosine
cyclo-dG	cyclodésoxyguanosine
D-loop	boucle de déplacement ; displacement-loop
DDB1	DNA-damage binding protein 1
DDB2	DNA-damage binding protein 2 (=XPE)
DMTr	diméthoxytrityle
DNA-PK_{cs}	protéine kinase ADN-dépendante
DREF	zinc finger BED domain-containing protein 1 (ZBED1)
EF-Tu 2	facteur d'élongation
EMSA	Electrophoretic Mobility Shift Assay
Fapy-dA	Fapy-désoxyadénosine ; 4,6-diamino-5-formamidopyrimidine
Fapy-dG	Fapy-désoxyguanosine ; 2,6-diamino-4-hydroxy-5-formamidopyrimidine

Fpg	formamidopyrimidine DNA glycosylase (également MutM)
GC-MS	chromatographie gaseuse couplée à la spectométrie de masse
GG-NER	global genome nucleotide excision repair
GPDH	glycérol-3-phosphate déshydrogénase
Gpx	Glutathione peroxidase
GSH	L-γ-glutamyl-L-cysteinylglycine ; glutathion
H2AX	H2A histone family, member X
H_2O_2	peroxyde d'hydrogène
HDL	lipoprotéines à haute densité
HMGB	hifh mobility group box
$HO^{\bullet}$	radical hydroxyle
HPLC−MS/MS	high performance liquid chromatography – tandem mass spectrometry; chromatographie liquide à haute performance couplée à la spectrométrie de masse en tandem
hUBF	human upstream binding factor
HR	Homologous Recombination ; Recombinaison Homologue
LC−MS/MS	liquid chromatography – tandem mass spectrometry; chromatographie liquide couplée à la spectrométrie de masse en tandem
LDL	lipoprotéines à basse densité
$[M-H]^-$	ion moléculaire monochargé
$[M-2H]^-/2$	ion moléculaire dichargé
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation – Time of Flight
MAO	monoamine oxydase
MDA	malondialdéhyde
MMR	réparation des mésappariements
MnSOD	superoxyde dismutase
MRN	complexe MRN composé de trois protéines : MRE11, RAD50 et NBS1
mtHsp70	mitochondrial heat shock protein 70
MutL	DNA mismatch repair protein MutL
MutS	DNA mismatch repair protein MutS
NAD/NADH	nicotinamide adénine dinucléotide
NEIL1	Endonuclease VIII-like 1
NER	Nucleotide Excision Repair ; réparation par excision de nucléotides
NHEJ	Non-Homologous End-Joining
N(*i*Pr$_2$)	groupement diisopropylamine
NMR	résonance magnétique nucléaire
$NO^{\bullet}$	oxyde nitrique
$NO_2^{\bullet}$	dioxyde de l'azote
NO_2^-	nitrite
NO_3^-	nitrate
NONO	Non-POU domain-containing octamer-binding protein
$O_2^{\bullet -}$	anion superoxyde
OGG1	8-oxoguanine glycosylase
PAR	poly(ADP-ribose)
PARP1	poly(ADP-ribose) polymerase 1
PARylation	poly(ADP-ribos)ylation
PCR	polymerase chain reaction ; réaction de polymérase en chaîne
PDH	déshydrogénase
PNKP	polynucleotide kinase/phosphatase
PNUTS	phosphatase 1 nuclear targeting subunit
Polβ	polymérase β

PUA	3'-aldéhyde-phosphor-α,β-insaturé
redAP	site abasique réduit
ROS	reactive oxygen species ; espèces réactives de l'oxygène
RNS	reactive nitrogen species ; espèces réactives de l'azote
RPA	Replicative Protein A ; protéine de réplication A
SDS-PAGE	sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SFPQ	splicing factor, proline- and glutamine-rich
SOD	superoxyde dismutase
SPRi	Surface Plasmon Resonance imaging
RRM	RNA recognition motif
TBP	TATA Binding Protein
TC-NER	transcription coupled nucleotide excision repair
TDG	thymine ADN glycosylase
TEAA	acétate de triéthylamine
TFIIH	Transcription factor II Human
THF	tétrahydrofurane
TOX4	TOX HMG box family member 4
THAP	2,4,6-trihydroxyacétophénone
XP	*Xeroderma pigmentosum*
XPC/hHR23	*Xeroderma pigmentosum group C-complementation protein*
XRCC1	X-ray repair cross-complementing protein 1
UNG	uracile ADN glycosylase
WDR82	WD repeat-containing protein 82
ZBTB2	zinc finger and BTB domain-containing protein 2

Introduction

L'ADN est une molécule dont l'intégrité est cruciale pour le bon fonctionnement de tous les organismes vivants. Il est cependant constamment soumis à différents stress, qui peuvent l'endommager. La nature et les conséquences de ces dommages sont variées et dépendent du type de stress. Toutes ces lésions, quelle que soit leur nature, vont devoir être prises en charge par les systèmes de réparation cellulaires, afin d'éviter l'apparition des mutations menant à la cancérogénèse ou à la mort cellulaire.

Les lésions d'oxydation sont un vaste groupe de dommages, dont les sources sont multiples : les processus endogènes contribuent à leur formation quotidienne, mais un stress exogène supplémentaire, par exemple les rayonnements ionisants, peut augmenter considérablement leur niveau dans les cellules. Parmi ces lésions, des dommages simples ou complexes sont retrouvés, leurs conséquences pour les cellules étant différentes.

La majorité des dommages formés au niveau de l'ADN sont pris en charge par des systèmes de réparation spécialisés, composés de dizaines de protéines interagissant les unes à la suite des autres pour restaurer l'intégrité de l'ADN. D'autres protéines peuvent également reconnaître les lésions de l'ADN : des facteurs de transcription ou des protéines de remodelage chromatinien peuvent ainsi interagir avec des dommages de l'ADN et constituer ainsi son interactome.

Les travaux décrits dans le présent manuscrit s'intéressent à l'interactome des lésions d'oxydation de l'ADN, à la fois simples et complexes. Ceci peut être étudié en utilisant diverses techniques, deux approches distinctes ont été appliquées ici. La recherche des membres de l'interactome de lésions complexes a été effectuée par une technique de screening dite « ligand fishing », dont le but est d'identifier les nouvelles protéines possédant une affinité pour le dommage étudié. Une deuxième approche consiste en l'analyse des interactions entre une protéine isolée et sa cible, ce qui permet d'approfondir les connaissances concernant ce membre particulier de l'interactome étudié.

Ce manuscrit commence par une analyse bibliographique qui permet de placer la problématique de cette thèse dans son contexte. Ainsi, les facteurs et les conséquences cellulaires

du stress oxydant seront abordés en premier lieu, suivis de la description des lésions simples et, enfin, des cyclonucléosides. Ces derniers sont des représentants intéressants des lésions complexes, et ils sont peu connus par rapport à de nombreux autres dommages de l'ADN.

Dans les parties suivantes de ce manuscrit, l'objectif de cette thèse va être détaillé, ainsi que la stratégie employée. La description des méthodes utilisées précédera ensuite la présentation des résultats obtenus. Ces derniers sont divisés en deux parties, qui correspondent aux deux approches mentionnées ci-dessus : recherche des nouveaux membres de l'interactome des lésions complexes d'oxydation, et l'étude des protéines isolées.

Les résultats obtenus seront ensuite analysés dans la partie « Discussion », avant de terminer ce manuscrit avec les conclusions générales et les perspectives pour de futurs travaux.

I Analyse bibliographique

1. Stress oxydant et son effet sur les cellules

Toutes les cellules sont constamment exposées à différents stress qui peuvent avoir un effet délétère sur leur fonctionnalité et leur intégrité : les agents chimiques comme pesticides ou polluants atmosphériques, et les agents physiques tels que le rayonnement UV. Certains de ces agents sont capables de générer un stress oxydant. Parmi les facteurs à l'origine de ce stress, on va retrouver à la fois des sources internes, comme la respiration mitochondriale, et externes, tels que les rayonnements ou des xénobiotiques. Dans cette analyse bibliographique nous allons décrire les conséquences du stress oxydant sur les modèles eucaryotes, qui sont le plus couramment étudiés. Cependant, d'autres organismes tels que les procaryotes ou les archées subissent également les effets du stress oxydant.

Le stress oxydant est impliqué dans de nombreuses pathologies, telles que le cancer, le diabète de type II ou des maladies neurodégénératives, mais aussi dans le vieillissement de l'organisme [1]. Quelle que soit sa source, il est exercé via les espèces réactives de l'oxygène (reactive oxygen species ; ROS) – ces composés très réactifs envers toutes les biomolécules sont générés lors de l'activité « normale » d'une cellule. Ainsi dans toutes les cellules il existe un équilibre entre les ROS et les molécules dites antioxydantes, qui servent à diminuer le caractère nocif des premières citées. Afin de garder cet équilibre, la concentration cellulaire de l'oxygène est contrôlée et gardée entre 0.5 et 5%, ce qui est inférieur à sa concentration dans l'air (autour de 20%) – elle dépend du type cellulaire et de la fonction de la cellule, mais également de son compartiment [2]. Au-dessous de ces limites, les cellules souffrent d'hypoxie, tandis que quand ces valeurs sont dépassées, le stress oxydant augmente. Les ROS peuvent être également générées suite à un bouleversement extérieur à la cellule, comme par exemple certaines substances chimiques ou des rayonnements, ce qui va perturber cet équilibre. Ainsi, le stress oxydant peut être défini comme les conditions dans lesquelles la quantité des ROS dépasse les capacités défensives de l'organisme, qui n'arrive pas à fournir suffisamment d'antioxydants pour contrebalancer l'effet des espèces réactives [3].

1.1. Formation des radicaux libres

Les radicaux libres sont des particules extrêmement réactives, formées quand un électron est arraché d'une couche externe d'un atome ou d'une molécule. Ces radicaux libres peuvent être créés à partir d'une molécule d'ADN, des peptides, des protéines, ou d'autres molécules biologiques, mais la majorité provient de la radiolyse de l'eau, ces radicaux étant aussi les plus toxiques.

1.1.1. Espèces réactives de l'oxygène

Les espèces réactives de l'oxygène sont des molécules générées à partir de l'oxygène moléculaire. On sait depuis 40 ans que de nombreuses ROS sont formées majoritairement dans les mitochondries, en tant que produits secondaires du métabolisme de l'oxygène [4–7]. La première étape qui mène à leur formation consiste en la réduction de l'oxygène moléculaire par un électron – cette réaction donne l'anion superoxyde ($O_2^{\bullet -}$) qui sert de précurseur dans la formation de toutes les autres espèces réactives.

Le potentiel redox de la transformation de l'oxygène moléculaire en anion superoxyde n'étant pas très élevé (-0.160V), cette réaction peut être accomplie par de nombreuses enzymes mitochondriales à différentes étapes de la chaîne respiratoire. Il existe ainsi plusieurs sources d'$O_2^{\bullet -}$ au sein des mitochondries, parmi lesquelles deux semblent être les plus importantes : le complexe I et, dans une moindre mesure, le complexe III. Leur rôle varie en fonction des tissus : le complexe I est la source majeure des anions superoxyde dans le cerveau [8,9] et il est également impliqué dans de nombreuses pathologies, tandis que le complexe III serait plutôt responsable de la génération des ROS dans les mitochondries des tissus cardiaque et pulmonaire [10,11].

Le complexe I, appelé également NADH-ubiquinone oxydoréductase, est situé dans la membrane interne de la mitochondrie. Il oxyde NADH en utilisant le coenzyme Q comme accepteur d'électrons, et génère en même temps une quantité importante d'$O_2^{\bullet -}$. Le complexe III (ubiquinone : cytochrome c réductase), situé également dans la membrane interne, oxyde à son tour le coenzyme Q, avec le cytochrome c en tant qu'accepteur d'électrons. D'autres enzymes peuvent contribuer aussi à la formation de l'anion superoxyde, telles que : monoamine oxydase (MAO), cytochrome b5 réductase (Cb5R), glycérol-3-phosphate déshydrogénase (GPDH), les cytochromes P450, ainsi que des enzymes de la matrice : aconitase, pyruvate déshydrogénase (PDH) et α-kétoglutarate déshydrogénase (αKGDH). Dans la majorité des cas, ces réactions

impliquent un électron et donnent naissance à un anion superoxyde. Il peut néanmoins arriver qu'une réaction à deux électrons survienne, ce qui donne lieu à une transformation directe de l'oxygène en peroxyde d'hydrogène. Le fonctionnement de la chaîne respiratoire avec les enzymes impliquées dans la formation des ROS, est présenté sur la Figure 1 [12,13].

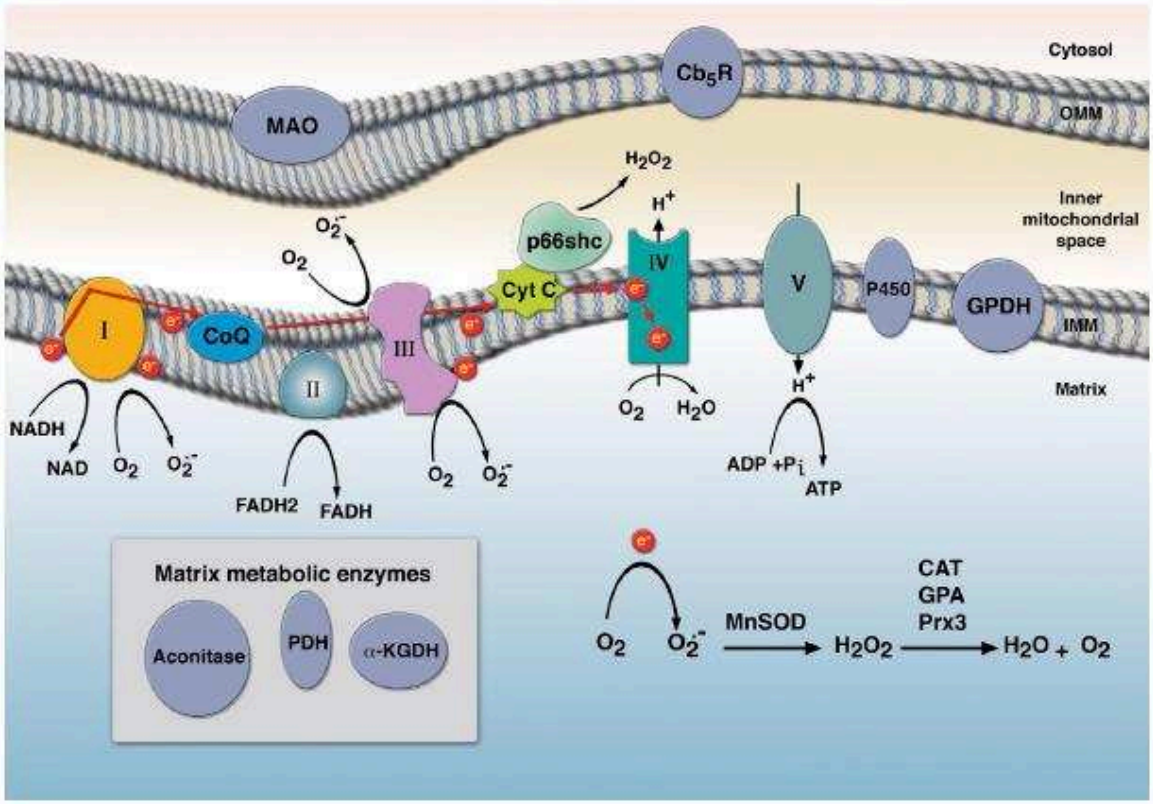


Figure 1 : Formation des espèces réactives de l'oxygène dans les mitochondries. Source : [12].

L'anion superoxyde peut également être créé par une voie non-enzymatique, quand un électron est transféré sur l'oxygène depuis une coenzyme ou un xénobiotique réduits. L'anion superoxyde est relativement peu réactif, cependant il peut être transformé en peroxyde d'hydrogène (H_2O_2) à l'aide de la superoxyde dismutase (SOD). Les niveaux de ces deux ROS : $O_2^{\bullet-}$ et H_2O_2, sont maintenus stables par les cellules, et ils sont estimés à 10^{-10} M pour $O_2^{\bullet-}$ et 5×10^{-9} M pour H_2O_2 (dosages effectués dans les mitochondries du cœur et du foie de rat) [14].

Le peroxyde d'hydrogène peut, à son tour, être réduit en H_2O par la catalase ou la glutathion peroxydase, ou bien en radical hydroxyle ($HO^{\bullet}$) en présence des métaux de transition, dans la réaction de Fenton :

$$H_2O_2 + Fe^{2+} \rightarrow {}^{\bullet}OH + OH^- + Fe^{3+}$$

HO$^{\bullet}$ peut également être formé quand les deux espèces réactives d'oxygène citées auparavant réagissent entre elles, dans la réaction de Haber-Weiss [15] :

$$O_2^{\bullet -} + H_2O_2 \rightarrow O_2 + OH^{\bullet} + OH^-$$

Le radical hydroxyle est extrêmement réactif et il interagit avec toute molécule biologique se trouvant dans sa proximité immédiate, sa distance de diffusion dans une cellule étant de 2 nm [16]. Il peut être également réduit en H_2O, terminant ainsi la chaîne des réactions radicalaires (Figure 2).

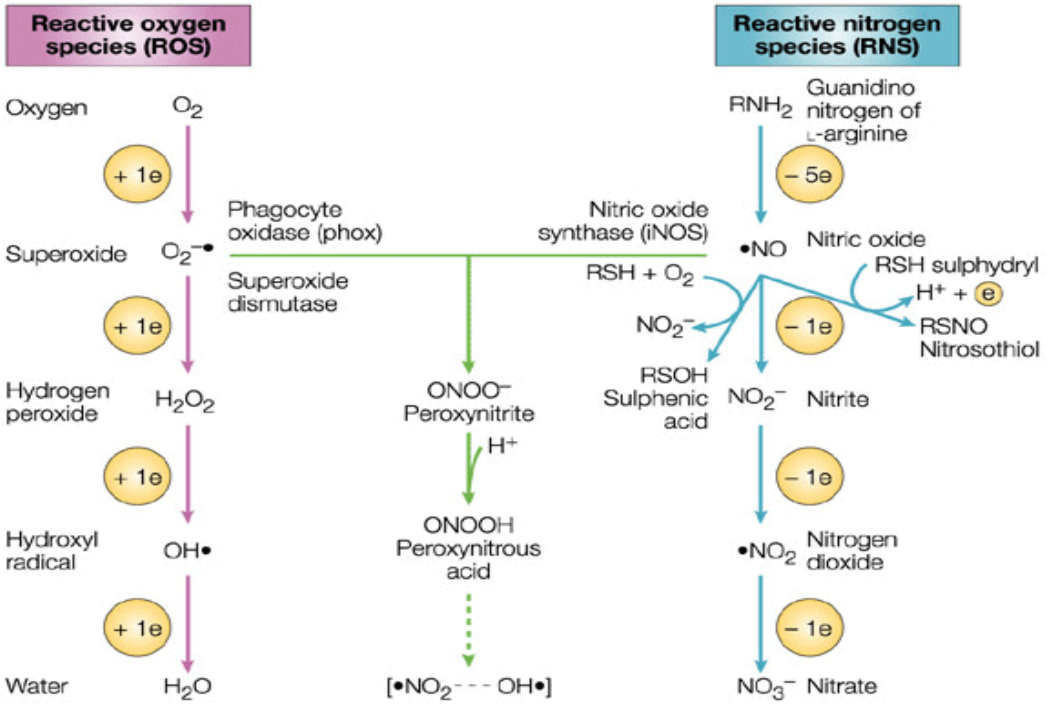


Figure 2 : Formation des espèces réactives de l'oxygène et de l'azote. Source : [17].

Le peroxyde d'hydrogène n'a pas qu'un rôle destructeur dans les cellules. En dehors de son rôle de précurseur dans la formation du radical hydroxyle, il devient, grâce à sa stabilité et à sa neutralité qui lui permettent de traverser les membranes (à l'aide des aquaporines), une importante molécule signalisatrice dans de nombreux processus tels que le contrôle du cycle cellulaire ou la réponse au stress [18].

1.1.2. Espèces réactives de l'azote

Bien que les espèces réactives de l'oxygène soient les plus connues et les plus étudiées, il existe également d'autres molécules nuisibles aux cellules, notamment les espèces réactives de l'azote (reactive nitrogen species ; RNS). L'oxyde nitrique $NO^{\bullet}$, généré par la $NO^{\bullet}$ synthase, est le précurseur de toutes les autres RNS, dont les caractéristiques sont très hétérogènes [19,20]. $NO^{\bullet}$ est un gaz assez instable, impliqué dans le transfert des signaux cellulaires en tant que neurotransmetteur, dans la régulation de la vasodilatation et dans l'agrégation des plaquettes [19,21–24]. Néanmoins, quand il est présent en quantité élevée, il joue un rôle dans les systèmes de défense contre les microorganismes et les cellules tumorales [25,26].

L'oxyde nitrique peut réagir directement avec de nombreuses molécules. Parmi ses actions directes on retrouve :

- réaction avec les groupements thiols des cystéines, pour donner lieu à la formation des ponts disulfure, à la S-nitrosylation, ou bien à l'oxydation vers un acide sulfénique (RSOH), les trois réactions étant réversibles [17] ;
- formation des complexes avec des enzymes fer-soufre [27] ;
- désamination et cross-linking de l'ADN [28] ;
- inactivation des enzymes antioxydantes (glutathion peroxydase, SOD) [29,30] ;
- induction de l'apoptose via l'activation de la caspase-3 et la modification de la perméabilité membranaire [31].

Alternativement, $NO^{\bullet}$ peut réagir avec l'oxygène et être transformé en nitrite (NO_2^-) ou en dioxyde d'azote ($NO_2^{\bullet}$), qui peut initier l'oxydation des lipides, et enfin en nitrate (NO_3^-) [17,20]. Une autre voie possible implique la réaction de l'oxyde nitrique avec le superoxyde pour donner le peroxynitrite ($ONOO^-$) [29,32] – celui-ci étant instable, il réagit rapidement avec toutes les molécules biologiques : molécules avec des cycles aromatiques, thiols, lipides ou protéines à hème [20] – et l'acide peroxynitrique (ONOOH), qui va se décomposer en radicaux : dioxyde d'azote et radical hydroxyle [17] (Figure 2).

1.2. Sources du stress oxydant

Le stress oxydant et ses acteurs décrits dans les paragraphes précédents peuvent provenir de plusieurs sources, que l'on peut généralement diviser en deux groupes : le stress généré par la

cellule elle-même, dit endogène, et le stress provoqué par les facteurs extérieurs à la cellule et même à tout l'organisme, dit exogène.

1.2.1. Stress oxydant endogène

1.2.1.1. Respiration cellulaire

Bien que le rôle des sources exogènes du stress oxydant soit important, la majeure partie des ROS est générée par les cellules elles-mêmes au cours de la respiration mitochondriale. Cette formation des ROS par les mitochondries a été constatée pour la première fois au début des années 70' [33]. Elle est due à ce qui a été nommé par Kelvin Davies « le paradoxe de la vie en aérobie » : toutes les cellules eucaryotes ont besoin de l'oxygène afin de fabriquer l'énergie pour leur fonctionnement, mais en même temps cet oxygène présente un réel danger pour elles, puisqu'il est la source des espèces réactives [34].

Les mitochondries fournissent à toutes les cellules aérobies de l'énergie sous forme d'ATP, qu'elles fabriquent en présence de l'oxygène. Ce sont alors les organelles où la concentration de l'oxygène cellulaire est la plus élevée et qui, par conséquent, sont les plus exposées au risque des dommages d'oxydation. Lors de la respiration mitochondriale, environ 1-2 % de l'oxygène que la cellule consomme, est transformé en radicaux, qui sont considérés comme des produits secondaires du fonctionnement normal d'une cellule. L'oxygène moléculaire donne ainsi l'anion superoxyde qui est un intermédiaire stable dans la formation d'autres ROS, et sa création peut avoir lieu à différentes étapes du transport d'électrons dans les mitochondries [35].

1.2.1.2. Inflammation

L'inflammation est une autre source de stress oxydant endogène. Les causes de l'inflammation peuvent être nombreuses : infections virales ou bactériennes, exposition aux allergènes, obésité, maladies auto-immunes, consommation d'alcool ou régime calorique [36–38]. Les ROS jouent un rôle important dans les voies de communication cellulaire, notamment dans le système immunitaire, et c'est pourquoi ils sont produits en grande quantité pendant les réponses inflammatoires [3].

L'inflammation se manifeste souvent en deux étapes : le premier, l'inflammation aiguë, est bénéfique pour l'organisme car elle permet de combattre l'infection, mais si elle se prolonge, l'inflammation devient chronique et dangereuse pour le corps [39]. Les cellules inflammatoires secrètent des molécules telles que les cytokines, qui favorisent la formation des ROS et des RNS. Ainsi, le recrutement des leucocytes vers l'endroit endommagé va engendrer leur accumulation. Les cytokines attirent également de nouvelles cellules inflammatoires, ce qui provoque un cercle vicieux de génération du stress oxydant. Ceci peut avoir pour conséquence l'endommagement des cellules voisines, et mener à long terme à la cancérogénèse, qui peut avoir un lien direct avec l'inflammation, si l'ADN est endommagé, ou indirect, lorsque les ROS endommagent les protéines et les lipides, qui, à leur tour, forment les adduits de l'ADN [36,40–42]

1.2.2. Stress oxydant exogène

En plus du stress endogène généré dans le corps humain, de nombreuses sources extérieures sont capables de générer un stress oxydant supplémentaire, appelé exogène. Tout facteur extérieur au corps humain, qu'il soit de nature physique ou chimique, capable de produire les ROS, peut être ainsi nocif pour les cellules.

1.2.2.1. Substances chimiques

Parmi les substances chimiques, appelées aussi xénobiotiques, pouvant être à l'origine du stress oxydant exogène, on retrouve par exemple les médicaments tels que le cisplatine ou la doxorubicine, les métaux ou les polluants atmosphériques. Puisque chacun de ces agents agit de manière différente sur les mécanismes cellulaires, quelques exemples de ces interactions vont être présentés ci-dessous.

Un premier exemple est la **fumée de cigarettes**, qui est un mélange complexe de molécules telles que les hydrocarbures aromatiques polycycliques, les nitrosamines, les amines aromatiques et autres. Elle contient également des radicaux libres qui peuvent générer au niveau de l'ADN des lésions d'oxydation, comme par exemple la 8-oxo-désoxyguanosine (8-oxo-dG) [43,44]. L'effet de la fumée de cigarettes est exercé également par la réaction inflammatoire qui peut avoir pour effet l'oxydation du glutathion, la diminution de la concentration sanguine des antioxydants et l'augmentation du niveau d'$O_2^{\bullet -}$ [43,45]. De plus, les substances cancérigènes

présentes dans la fumée de cigarettes sont transformées par le cytochrome P450 en métabolites réactifs envers l'ADN [43,46].

L'**alcool** a été classé cancérigène et il est, après les cigarettes, le deuxième facteur de risque pour les cancers ORL (au niveau de la tête et du cou), car il favorise la formation des ROS et la libération des cytokines. De plus, il diminue le niveau des antioxydants tels que le GSH (L-γ-glutamyl-L-cysteinylglycine ; glutathion) [43,47–50]. Bien que le mécanisme par lequel l'alcool modifie le statut redox de la cellule ne soit pas encore connu, il serait possible que l'oxydation de l'éthanol en acétaldéhyde réduise le NAD en NADH ce qui va ensuite favoriser la formation des ROS [51].

La **doxorubicine** est un des médicaments bien documentés, qui a à la fois un effet anticancéreux et cardiotoxique, ces deux rôles semblent pourtant être indépendants. La doxorubicine est capable d'interagir avec la bicouche lipidique, inhibant ainsi de nombreuses enzymes dont l'activité dépend de la présence des lipides, telles que la NADH déshydrogénase, la cytochrome c oxydase ou la cytochrome c réductase. Ceci peut avoir pour effet l'augmentation du taux de formation des ROS, mais la principale voie de formation des ROS par la doxorubicine passe par sa réduction en radical libre de semiquinone, qui, à son tour, va réduire l'oxygène moléculaire en anion peroxyde [52].

Il existe plusieurs études qui prouvent également l'implication des **métaux** dans la formation des radicaux libres. Parmi les métaux les plus étudiés, on peut énumérer : le fer, le cuivre, le chrome, le cobalt ou encore le cadmium, qui sont tous capables d'augmenter le niveau du stress oxydant dans les cellules [53]. Le fer joue un rôle extrêmement important dans les cellules, puisqu'il participe au transport de l'oxygène vers tous les tissus. Cependant, s'il est présent en quantité dépassant les besoins de la cellule, ce transport excessif va engendrer la formation des ROS qui dépassent les défenses antioxydantes. Ainsi, un des mécanismes de développement du cancer résultant de l'exposition à l'amiante, serait dû à la formation des radicaux libres par le fer, qui constitue 30% de la masse de l'amiante [53,54]. Le fer, comme le cuivre, le chrome, ou le cobalt, est aussi catalyseur dans la réaction de Fenton, qui transforme le peroxyde d'hydrogène en radical hydroxyle. Le cadmium, quant à lui, n'a pas de capacités redox, et ne peut donc pas participer directement à la formation des ROS. Son rôle dans la génération du stress oxydant a pourtant été démontré et le mécanisme par lequel il procède a été élucidé : le cadmium peut remplacer le fer dans les protéines cytoplasmiques ou membranaires (par exemple la ferritine), ce qui va avoir pour effet l'augmentation de la concentration de Fe^{2+} libre [53,55,56].

1.2.2.2. Rayonnement ionisant

L'effet néfaste des irradiations sur le matériel génétique, pouvant avoir pour conséquence des mutations chez les descendants, a été découvert dans les années 1920 chez *Drosophila melanogaster*. Cependant, cette problématique chez les hommes n'a suscité l'intérêt dans le milieu scientifique qu'après les explosions des bombes atomiques à Hiroshima et Nagasaki en 1945. Depuis, de nombreux programmes sont menés dans le but de déterminer les conséquences génétiques et les effets héréditaires chez les enfants des personnes ayant survécu à ces événements [57]. La majorité du rayonnement que l'on reçoit provient des sources naturelles : de l'espace, des radionucléides présents naturellement dans la couche terrestre, tels que l'uranium, le thorium et le potassium, ou bien du radon, qui est présent sous forme de gaz généré pendant la désintégration de l'uranium. Le reste est dû à l'exposition artificielle, notamment aux examens médicaux ou radiothérapie, et aussi à l'industrie nucléaire, cette dernière ne constituant que 1% de la dose totale de radioactivité reçue [58].

Les rayonnements sont une forme d'énergie qui se propage en tant qu'ondes ou particules, et qui est libérée au contact de la matière. Il en existe de différents types, parmi lesquels :

- les rayons X : ils traversent la peau, les os et le métal, ils sont utilisés en médecine dans un but diagnostique ;
- les rayons gamma : ils sont plus pénétrants que les rayons X, on les utilise en radiothérapie ;
- les rayons cosmiques : présents sous forme de particules notamment au-dessus de l'atmosphère ;
- les particules alpha : émises par certains éléments naturels (uranium, radium), à très faible potentiel de pénétration, mais très dangereuses dans le cas de l'exposition interne ;
- les particules bêta : des électrons à relativement faible capacité de pénétration ;
- les neutrons : des particules libérées lors de la fission nucléaire, à très forte capacité de pénétration [59].

L'effet exercé par les rayonnements va également dépendre du type d'exposition : l'exposition externe, qui correspond à la durée de l'exposition à la source de la radioactivité, est ainsi moins dangereuse que l'exposition interne. Celle-ci a lieu quand une particule radioactive pénètre à l'intérieur du corps et continue à exercer son effet pendant une période plus longue, dont la durée dépend du type de la molécule [58].

Le rayonnement ionisant peut agir sur les molécules biologiques via une action directe ou indirecte. On parle de l'effet direct quand le rayonnement est directement absorbé par une

particule, ce qui donne lieu à son endommagement ou dégradation. Si l'on considère le rayonnement à faible transfert linéaire d'énergie (TEL), un tiers des dommages de l'ADN est dû à l'effet direct de l'ionisation. Les deux tiers de ces dommages sont générés par l'effet indirect, qui implique la radiolyse de l'eau comme première étape. Quand une molécule d'eau absorbe des irradiations, ceci a pour l'effet son excitation et, par la suite, la formation des radicaux libres. D'après Platzman [60], les événements qui suivent peuvent être classés en quatre étapes :

1) étape « physique » – le dépôt de l'énergie venant des irradiations génère la formation des électrons secondaires et des espèces instables, par exemple $H_2O^{\bullet+}$
2) étape « physico-chimique » – les espèces formées dans la première étape de radiolyse sont rapidement transformées en d'autres radicaux, plus stables, tandis que les électrons secondaires sont soit capturés soit hydratés ; les produits générés à cette étape sont organisés d'une manière non-homogène le long d'une « piste »
3) étape de « chimie non-homogène » – les espèces réactives diffusent et réagissent entre elles
4) étape « biologique » – cette étape comprend la réaction de la cellule au dommage subi.

Dans une cellule aérobie à pH physiologique, les espèces qui prévalent suite à la radiolyse d'eau sont : $O_2^{\bullet-}$, $^{\bullet}OH$ et H_2O_2.

L'abstraction d'hydrogène par $^{\bullet}OH$ mène aussi à la formation des radicaux organiques, qui sont rapidement transformés en radicaux peroxydes qui, quant à eux, peuvent ensuite réagir avec des lipides [61].

1.2.2.3. Photosensibilisation

Le rayonnement ultraviolet (UV), bien que connu le plus souvent pour la formation des dommages directs au niveau des composants cellulaires, joue également un rôle important dans le stress oxydant. La formation des ROS suite à l'action des UV est présentée sur la Figure 3 : la première étape consiste en l'excitation d'un photosensibilisateur, qui est ensuite réduit, et cette forme réduite est capable de réagir avec l'oxygène moléculaire, pour donner l'anion superoxyde. Ceci régénère également le photosensibilisateur qui peut entrer à nouveau dans le même cycle et contribuer à nouveau à la formation des ROS [62].

Il a été démontré que la cataracte, une maladie oculaire répandue dans les pays à haut indice UV, est due au mécanisme radicalaire initié par les rayons solaires [62]. La riboflavine,

présente dans la cornée et connue pour ses propriétés de photosensibilisant, génère les ROS qui oxydent le tryptophane dans le cristallin [62,63].

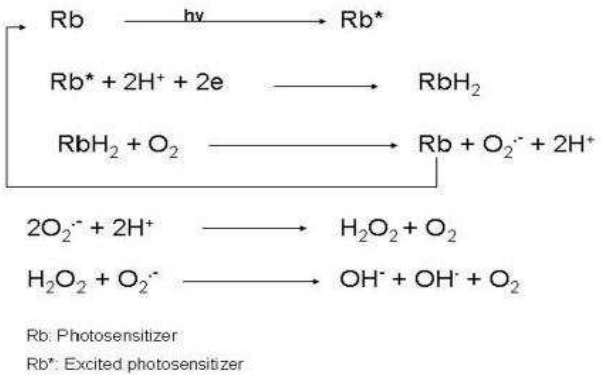


Figure 3 : Formation photochimique des espèces réactives de l'oxygène. Rb : photosensibilisateur (p.ex. la riboflavine), Rb* : photosensibilisateur excité. Source : [62].

1.2.2.4. Agents infectieux

Les infections constituent également un facteur d'induction du stress oxydant. Les pathogènes, y compris virus et bactéries, peuvent être ainsi associés de manière directe ou indirecte avec des processus de cancérogénèse. La voie impliquant le stress oxydant est la voie indirecte, et elle implique dans la majorité des cas une inflammation chronique générée par l'infection. Les mécanismes de défense antioxydants qui servent dans un premier temps à éliminer le pathogène, peuvent ainsi se retourner contre l'hôte. Néanmoins, la relation entre une infection virale et l'apparition de cancer n'est pas toujours facile à identifier, car une dizaine d'années au minimum passe entre le moment de l'infection et la détection du cancer [64,65].

Le lien entre le stress oxydant et les infections chroniques a cependant été établi pour de nombreux virus, dont les mieux documentés sont l'hépatite B et C, et le virus HIV. L'action du virus de l'hépatite B est complexe et son influence sur le stress oxydant intervient à plusieurs niveaux : formation des espèces réactives de l'oxygène et de l'azote, déstabilisation du cycle redox de GSH ou la péroxydation des lipides [66–69]. Des niveaux élevés de lésions d'oxydation de l'ADN, essentiellement la 8-oxo-désoxyguanosine, ont été mesurés à plusieurs reprises chez les patients atteints d'hépatite B [70,71]. Des quantités élevées de dommages d'oxydation ont également été observées chez les porteurs du virus d'hépatite C [72]. Les patients atteints du SIDA présentent, quant à eux, un niveau important d'apoptose et de formation des ROS, ainsi qu'un système de réparation BER moins efficace et un niveau plus élevé de 8-oxo-désoxyguanosine [73,74].

1.3. Systèmes de défense antioxydants des cellules

Afin de ne pas subir de dommages trop importants, les cellules sont dotées d'un système de défense antioxydant, qui leur permet de neutraliser les ROS. Ce système est constitué des molécules qui entrent en compétition avec d'autres particules facilement oxydables (telles que les acides nucléiques, les protéines et les lipides) en les protégeant ainsi contre les dommages. La molécule la plus importante dans ce système de protection est le GSH – bien que les mitochondries ne soient pas capables de le synthétiser, elles l'importent depuis le cytoplasme. Le stock de GSH mitochondrial est ainsi estimé à 10 à 12% du GSH total, mais sa concentration y est plus élevée que dans le cytoplasme. Les réserves de GSH mitochondrial sont nécessaires pour la survie cellulaire, et elles sont aussi plus résistantes à la réduction de leur niveau que les réserves cytosoliques [35,75–77].

De nombreuses enzymes « de détoxification » utilisent le GSH en tant que donneur d'électrons. Parmi ces enzymes on retrouve entre autres les glutathion peroxydases, qui catalysent la réduction de H_2O_2 vers H_2O. Les membres les plus importants de cette famille d'enzymes sont Gpx 1 et Gpx 4. Le premier est localisé surtout dans le cytosol, mais on le retrouve également dans la matrice mitochondriale, tandis que le deuxième est associé aux membranes mitochondriales. Gpx 4 est considérée comme l'enzyme la plus importante dans le système de défense antioxydante, car elle réduit les groupements hydroperoxides localisés sur les phospholipides, les lipoprotéines et les esters de cholestérol, ce qui protège la structure et la fonction des membranes lipidiques [35,78].

Une autre enzyme détoxifiante est la superoxyde dismutase MnSOD – à la différence des glutathion peroxydases, elle n'a pas besoin de cofacteurs tels que le GSH, son activité dépend ainsi uniquement de sa quantité [76,79].

Des substances non-enzymatiques sont également qualifiées d'antioxydantes et jouent un rôle important dans la protection des cellules contre le stress oxydant. L'α-tocophérol (vitamine E) est présent dans les membranes mitochondriales et il est capable de réduire les radicaux lipidiques, pour être ensuite régénéré par la forme réduite du coenzyme Q. Le β-carotène ou l'ascorbate (vitamine C) sont également connus pour leurs qualités d'antioxydants [1,76]. Le cytochrome c est appelé « l'antioxydant idéal », car sa régénération par une cytochrome c oxidase génère l'énergie qui peut être utilisée par les mitochondries pour produire l'ATP [76,80,81].

1.4. Effet biologique du stress oxydant

L'impact du stress oxydant sur le fonctionnement des cellules est complexe, puisque les espèces réactives de l'oxygène peuvent interagir avec presque tous les composants cellulaires, que ce soit des lipides, des protéines ou des acides nucléiques. C'est pour cela que l'effet exercé sur la cellule va constituer la somme de toutes ces perturbations.

1.4.1. Influence du stress oxydant sur les cellules

1.4.1.1. Apoptose

L'apoptose, nommée également « mort cellulaire programmée », est un processus complexe. Dans les organismes multicellulaires elle garantit le maintien de l'homéostasie des tissus, et son dérèglement peut mener à des déformations au niveau fœtal, à des cancers, ou à des maladies neurodégénératives. L'apoptose peut être engendrée de deux manières, par la voie intrinsèque ou extrinsèque :

- La voie intrinsèque peut être déclenchée par un stimulus apoptotique, en réponse aux produits toxiques ou rayonnements, et elle implique l'activation des mitochondries. Deux familles de protéines rentrent en compétition afin de réguler les signaux pro- et antiapoptotiques : suite à un signal proapoptotique, les protéines Bak et Bax stimulent la libération du cytochrome c des mitochondries, processus qui est normalement bloqué par Bcl-2. Le cytochrome c active ensuite la caspase-9 en présence d'APAF-1, et la caspase-9 agit directement sur la caspase-3 qui déclenche la cascade des événements menant à la mort de la cellule.
- La voie extrinsèque est activée quand des récepteurs de mort (par exemple TNF-R) se fixent à des ligands, ce qui provoque l'activation directe de la caspase 8 et de la caspase 3, sans passer par l'endommagement des mitochondries. Néanmoins, la protéine proapoptotique Bid peut faire le lien entre les voies intrinsèque et extrinsèque – activée par la caspase-8, elle provoque la libération du cytochrome c des mitochondries via son action sur Bak et Bax [82,83].

Le résumé du mécanisme de l'apoptose est présenté sur la Figure 4.

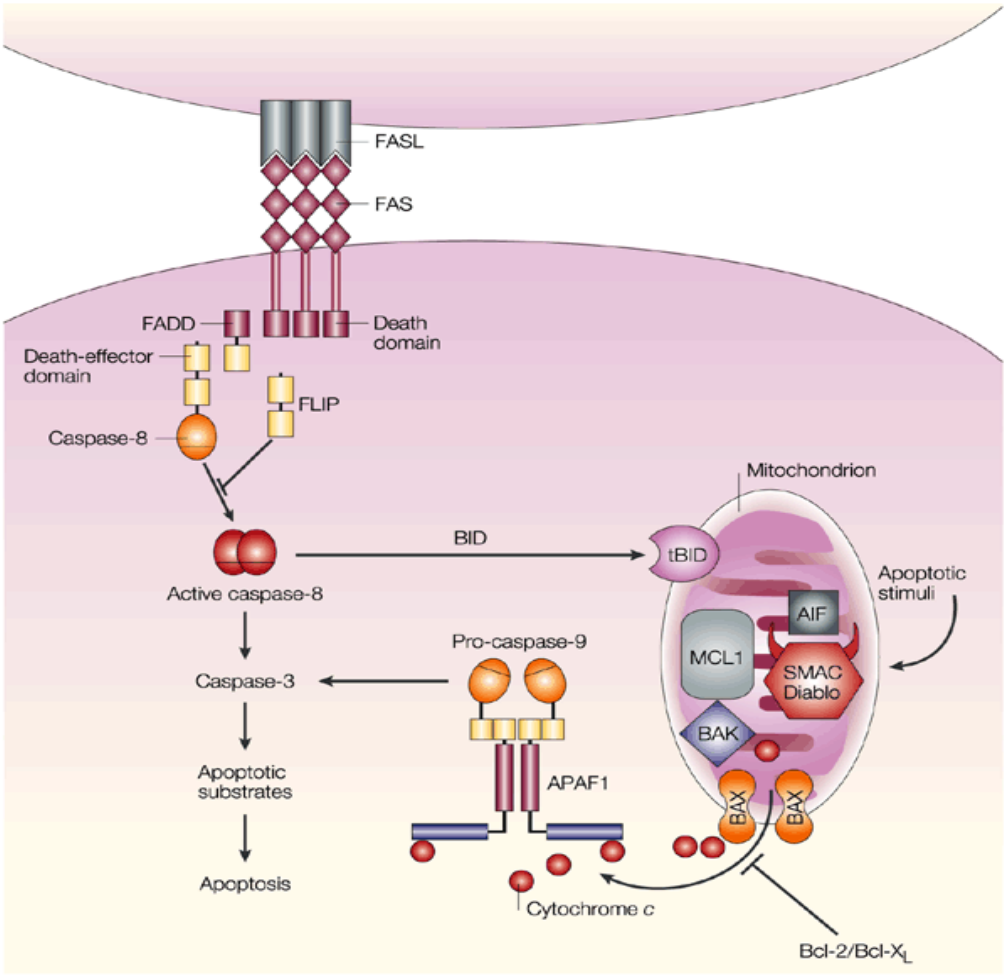


Figure 4 : Induction de l'apoptose par la voie extrinsèque ou intrinsèque. La voie extrinsèque passe par la fixation des ligands sur les récepteurs de mort et l'activation de la caspase-8. La voie intrinsèque implique la libération du cytochrome c des mitochondries et l'activation de la caspase-9. Les caspases 8 et 9 activent ensuite la caspase-3 qui déclenche la cascade des événements menant à l'apoptose. Source : [82].

Les ROS peuvent déclencher l'apoptose aussi bien par la voie intrinsèque qu'extrinsèque. Quelle que soit la voie d'activation, il est maintenant admis qu'il existe une corrélation entre l'augmentation du niveau des ROS dans les cellules et le déclenchement de l'apoptose, ainsi que le fait que les antioxydants sont des agents antiapoptotiques [84,85].

Il existe de nombreux mécanismes par lesquels le stress oxydant peut déclencher l'apoptose. Les protéines du pore de transition de perméabilité mitochondriale sont cibles des ROS : leur ouverture induit la dépolarisation des membranes et la libération consécutive du cytochrome c. Les espèces réactives de l'oxygène peuvent également endommager les lysosomes, le réticulum endoplasmique ou les protéines, toutes ces altérations étant capables de déclencher

l'apoptose par la voie intrinsèque : le potentiel de la membrane mitochondriale est modifié, le cytochrome c est libéré dans le cytoplasme et la cascade des caspases est lancée [86].

Les dommages de l'ADN peuvent aussi déclencher l'apoptose. Les lésions les plus cytotoxiques sont les cassures double-brin, 1 Gy de radiations ionisantes en induit environ 35 par cellule [87]. Des cassures simple-brin et des lésions simples peuvent constituer un obstacle à la réplication, et être transformées en cassures double-brin lors de leur prise en charge. Ceci peut également arriver quand les modifications de bases sont réparées par le BER. Dans tous ces cas, la formation des cassures double-brin de manière directe ou indirecte va être un facteur de déclenchement de l'apoptose [88].

1.4.1.2. Nécrose

Outre l'apoptose, le stress oxydant peut également déclencher la mort cellulaire par nécrose. Il existe deux voies par lesquelles ce processus peut être provoqué : soit par l'oxydation et, de ce fait, l'inactivation des caspases, soit par la perturbation de la chaîne respiratoire entrainant une baisse de concentration de l'ATP [19,89–91]. Les caspases possèdent dans leur site actif une cystéine, qui est très susceptible à l'oxydation et à la S-nitrosylation, les deux modifications étant capables d'inhiber son activité et de changer le destin de la cellule en la menant sur la voie de la nécrose [19].

1.4.1.3. Régulation du cycle cellulaire et sénescence

La sénescence est un état particulier de la cellule, caractérisé par un blocage définitif de la prolifération, contrairement à l'arrêt du cycle cellulaire qui est réversible [92]. Ce phénomène a été observé pour la première fois dans les années '60 dans les cellules primaires, qui arrêtaient leur prolifération après avoir atteint un certain nombre de passages, et c'est cette observation qui a permis de définir le rôle de la sénescence : empêcher les cellules trop vieilles (et ayant accumulé trop de dommages) de se diviser et de transférer leurs mutations [93,94]. Le stimulus principal qui déclenche la sénescence est le raccourcissement des télomères, qui a lieu à chaque division cellulaire. Il a néanmoins été observé que les télomères perdent plus de paires de bases par division, quand la cellule est exposée au stress oxydant, ce qui indique l'implication de celui-ci dans la sénescence. Les médiateurs du stress oxydant sont les ROS, qui peuvent également activer

la protéine oncogène Ras qui, à son tour, va activer p53. La protéine p53 possède un double rôle, puisqu'en fonction de son expression elle peut déclencher soit la sénescence, quand son niveau augmente légèrement, soit l'apoptose, quand elle est largement surexprimée [86].

Le processus exact par lequel les ROS entraînent la sénescence n'est pas clair, deux voies semblent néanmoins possibles. La première met en avant les dommages aléatoires au niveau de tous les composants cellulaires, qui auraient un rôle de médiateurs non-spécifiques de la sénescence. La deuxième hypothèse se base sur l'activation des voies spécifiques au niveau de la cellule par les ROS [95].

1.4.1.4. Vieillissement

Le vieillissement est un processus lié à l'âge qui consiste en déclin progressif des fonctions vitales. Chez les organismes multicellulaires, il comprend les changements qualitatifs dans le fonctionnement d'un organisme, tels que la dégénération progressive de ses fonctions biologiques, l'augmentation de la susceptibilité aux maladies ainsi que la probabilité accrue de mourir. D'après la théorie du vieillissement lié aux radicaux libres, cette dégénérescence des fonctions vitales d'un organisme serait due au stress oxydant et à ses effets, bien qu'il ne soit pas clair si c'est l'accumulation des dommages d'oxydation ou l'augmentation de la production des ROS qui est âge-dépendante [1,96].

Le génome mitochondrial semble être le plus susceptible au stress oxydant lié à l'âge [1]. Afin d'éviter l'endommagement des mitochondries par les ROS, celles-ci possèdent les systèmes de défense et de protection : on y inclue les enzymes de réparation des dommages subis par l'ADN, mais aussi des antioxydants, que ce soient des enzymes ou des petites molécules non-enzymatiques. Malgré ces systèmes, une accumulation des protéines oxydées peut être observée dans les mitochondries, ce qui est en rélation avec la durée de vie [97]. Cette accumulation serait due à l'affaiblissement des capacités de défense antioxydante dans les cellules, ce qui constitue un des signes de vieillissement. Ainsi, l'augmentation du niveau des ROS avec l'âge a été constatée dans de nombreux organes, comme le foie, le cœur, le cerveau et les muscles squelettiques [86,98–102]. De nombreuses modifications sont aussi observées dans les cellules âgées, comme par exemple l'oxydation des lipides ou des protéines, ainsi que les modifications du potentiel redox [1].

Une autre cible du stress oxydant qui peut impacter le vieillissement, est l'ADN. L'ADN mitochondrial, contrairement à l'ADN nucléaire, est dépourvu d'histones, ce qui le rend plus

fragile vis-à-vis des attaques du stress oxydant, qui, de plus, est plus élevé dans les mitochondries que dans d'autres compartiments cellulaires. Ainsi, McCord et Fridovich ont trouvé que le taux de la 8-hydroxyguanine était 10 à 20 fois plus élevé dans l'ADN mitochondrial que dans l'ADN nucléaire [103]. Un tel dommage dans l'ADN qui code pour les porteurs d'électrons peut engendrer l'arrêt de la chaîne respiratoire et la formation des ROS [13].

1.4.2. Interactions avec les composants cellulaires

Des systèmes de défense antioxydants existent dans les cellules pour neutraliser les espèces réactives d'oxygène. Néanmoins, les ROS qui n'ont pas été prises en charge peuvent attaquer tout type de macromolécule : protéines, lipides ou acides nucléiques. Ceci va influencer le bon fonctionnement ou même la survie de la cellule concernée.

1.4.2.1. Protéines

L'oxydation des protéines est un phénomène qui touche directement de nombreuses fonctions cellulaires telles que la transmission de signal, la structure ou le métabolisme, puisqu'une protéine oxydée est très souvent incapable d'exercer sa fonction.

Les ROS peuvent modifier les acides aminés des chaînes polypeptidiques – les résidus les plus sensibles à l'oxydation sont ceux qui contiennent un cycle aromatique ou bien un groupement sulfhydrile, les modifications au niveau des résidus sulfurés étant réversibles [104,105]. Les exemples des produits de leur oxydation sont présentés sur la Figure 5.

Quel que soit le mécanisme qui mène à l'oxydation des protéines, le produit obtenu le plus souvent est une protéine carbonylée, qui peut interagir par la suite avec les résidus de lysine afin de former des pontages intra- ou intermoléculaires, qui vont favoriser la formation d'agrégats protéiques. Ces agrégats ne pouvant pas être éliminés via le processus habituel de protéolyse, ils vont avoir tendance à s'accumuler et à altérer le fonctionnement de la cellule [104].

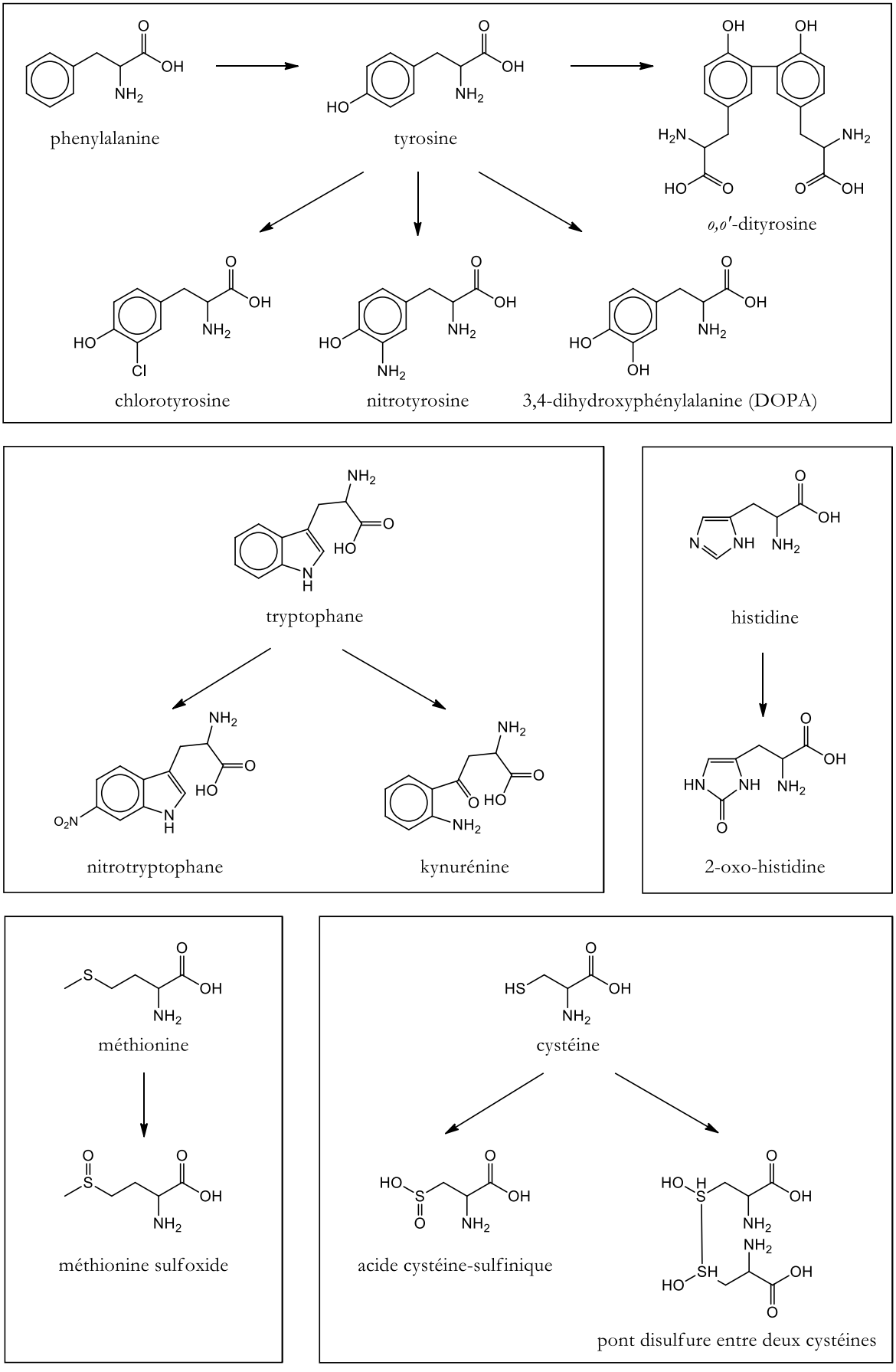


Figure 5 : Exemples des produits d'oxydation des acides aminés. Adapté de [105].

Un autre mécanisme par lequel les protéines peuvent être oxydées implique la participation des métaux : les ions Fe^{2+} ou Cu^{+} sont présents dans les sites de fixation des métaux au sein de la protéine, où ils catalysent la formation de $^{\bullet}OH$ à partir de H_2O_2, $^{\bullet}OH$ réagissant ensuite avec les acides aminés situés à sa proximité [104–107]. Alternativement, $O_2^{\bullet-}$ peut oxyder les protéines fer-soufre, telles que par exemple l'aconitase mitochondriale ou la NADH déshydrogénase, ce qui va avoir pour effet leur inactivation. L'inactivation de l'aconitase impacte immédiatement le cycle de Krebs et ainsi la production de l'énergie par la cellule. Le deuxième effet d'une telle oxydation est la libération de Fe^{2+} et H_2O_2, qui vont générer encore plus de dégâts, notamment via la formation de $^{\bullet}OH$ [35]. Ce radical hydroxyle est ensuite capable de casser des liaisons peptidiques, et générer par la suite un radical alkyle.

1.4.2.2. Lipides

Les lipides constituent une autre cible des attaques radicalaires dans les cellules. Parmi de nombreux produits de leur oxydation, certains sont électrophiles, capables de réagir avec des protéines ou avec l'ADN, ce qui leur donne un caractère mutagène. Dans les cellules, ce sont notamment les acides gras insaturés qui vont être endommagés par les agents d'oxydation [108].

Les acides gras insaturés sont de longues chaînes carboxyliques aliphatiques contenant au moins une double liaison qui détermine la conformation *cis* ou *trans* de la molécule, en fonction de la position des groupes méthylène par rapport à celle-ci. Si plusieurs doubles liaisons sont présentes dans la chaîne, la conformation de chacune d'elle peut être déterminée. Les stéréoisomères présents naturellement dans les cellules sont majoritairement de conformation *cis*, c'est-à-dire que les deux groupes méthylène sont localisés du même côté de la double liaison. Ces groupes présentent une réactivité élevée vis-à-vis des molécules à capacité d'oxydation, ce qui fait que les acides gras insaturés sont facilement transformés en carboradicaux (radicaux alkyles) par abstraction d'un atome d'hydrogène. Ces radicaux sont instables et réagissent immédiatement avec l'oxygène pour donner les radicaux peroxydes. La position du radical dans la chaîne carboxylique détermine la suite des réactions :

- s'il est localisé au bout de la chaîne, il va être réduit en hydroperoxyde qui, en présence des métaux, peut être transformé en divers produits, par exemple époxyde ou aldéhyde ; la réduction du radical peroxyde en hydroperoxyde est catalysée soit par une molécule de vitamine E, soit par une autre molécule d'acide gras, ce qui va donner naissance à un nouveau radical – on appelle ceci la propagation de l'oxydation des acides gras

- si le radical peroxyde est localisé à l'intérieur de la chaîne, il peut être soit réduit en hydroperoxyde, comme décrit ci-dessus, soit il peut réagir avec la double liaison adjacente dans la réaction de cyclisation ; le radical cyclique formé ainsi peut être réduit ou encore il peut cycliser une deuxième fois pour donner un peroxyde bicyclique, précurseur du malondialdéhyde (MDA) [108] (Figure 6).

Le malondialdéhyde est considéré comme le produit le plus mutagène de la peroxydation des lipides, puisqu'il peut former des adduits avec dG, dA ou dC, ceci ayant pour conséquences des substitutions de bases (transversions ou transitions) ou des mutations avec décalage du cadre de lecture [108].

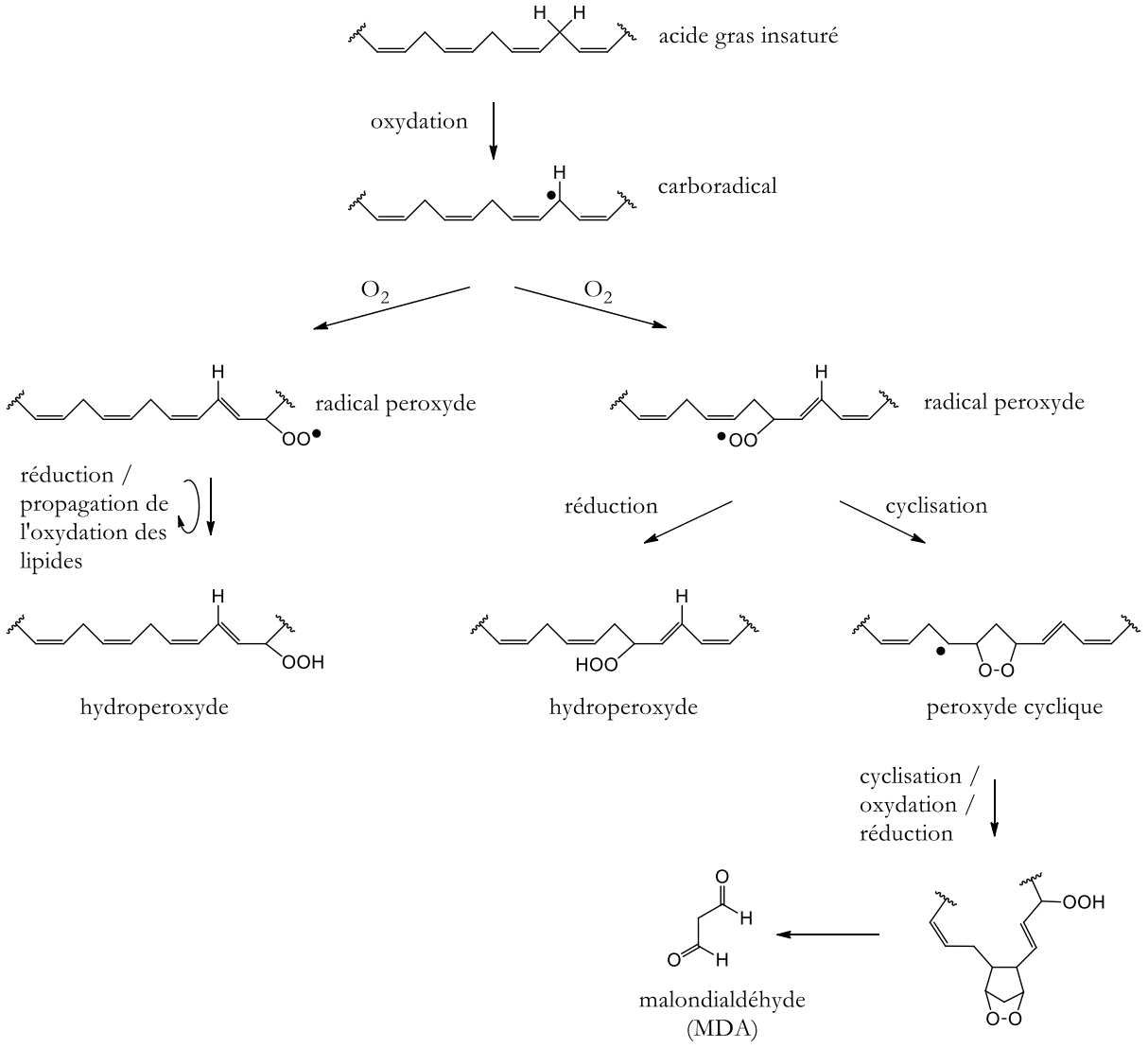


Figure 6 : Peroxydation des lipides. Adapté de [108].

Toutes les membranes cellulaires sont composées de lipides, tels que les phospholipides, le cholestérol ou les acides gras, qui assurent leur structure et leur fonction. Ainsi, leur oxydation influence l'organisation des membranes en modifiant leur fluidité et leur perméabilité [109,110]. Ces lipides sont aussi souvent associés aux protéines pour former des macromolécules appelées lipoprotéines, parmi lesquelles on retrouve par exemple les lipoprotéines de basse densité (LDL) et les lipoprotéines de haute densité (HDL). Les premières sont beaucoup plus sensibles au dommage oxydant que les deuxièmes, et leur oxydation peut provoquer une réponse inflammatoire et mener à la formation des lésions arthérosclérotiques [111–114].

La peroxydation des lipides n'est pas non plus anodine pour les mitochondries, puisqu'elle peut influencer la respiration et la phosphorylation oxydative, mais aussi altérer les propriétés des membranes mitochondriales. Les membranes mitochondriales sont les plus touchées par le phénomène de l'oxydation des lipides, puisqu'elles sont riches en acides gras insaturés et proches du lieu de formation des ROS. La membrane interne est constituée à 18% de cardiolipine, un lipide présent uniquement dans les mitochondries et indispensable pour leur fonctionnement. Son oxydation modifie la fluidité de la membrane et son potentiel, ce qui a un impact direct sur sa perméabilité et la régulation de la concentration en Ca^{2+} [35,52,115–117]. La cardiolipine est aussi extrêmement importante pour la régulation de l'apoptose, parce qu'elle est complexée avec le cytochrome c. Une fois oxydée, elle ne fixe plus le cytochrome c avec autant d'efficacité, ce qui permet de le libérer de la membrane interne de la mitochondrie et de déclencher l'apoptose [35,118].

1.4.2.3. Acides nucléiques

Les interactions des ROS avec l'ADN sont beaucoup plus néfastes pour les cellules que les dommages au niveau des autres biomolécules. Si les protéines ou les lipides oxydés peuvent être éliminés grâce au renouvellement naturel de ces molécules, ceci n'est pas le cas de l'ADN. Les modifications au niveau de sa séquence, si elles ne sont pas réparées à temps, vont être transférées d'une génération de cellules à l'autre en provoquant les mutations, voire la perte de l'information génétique [119].

L'ADN présent dans les cellules est constamment exposé à des agents toxiques qui altèrent sa structure et, par la suite, sa fonction. Ces agents peuvent être d'origine interne, comme certains métabolites, ou externes : rayons UV ou ionisants, produits chimiques et autres. En

fonction de la nature de l'agent génotoxique, différents types de lésions au sein de l'ADN peuvent être générés :

- des sites abasiques
- des modifications d'une base, tels que : oxydation (8-oxoguanine, 5-hydroxycytosine), alkylation, méthylation/déméthylation, hydrolyse
- des photoproduits, qui touchent surtout les pyrimidines, par exemple les dimères de thymine ou les 6 - 4 photoproduits
- des adduits volumineux, comme par exemple les adduits du cisplatine
- des pontages intra- ou interbrin, pontages ADN-protéines
- des cassures simple-brin
- des cassures double-brin

Des mésappariements, délétions ou insertions de bases peuvent être aussi générés au cours de la synthèse de l'ADN, résultats d'une copie infidèle par les polymérases.

Quel que soit le type des lésions de l'ADN, si elles ne sont pas réparées, elles peuvent avoir des conséquences néfastes pour la cellule, allant de la mort par apoptose, jusqu'aux mutations, qui peuvent engendrer le processus de cancérogénèse.

Bien que les dommages à l'ADN soient les plus étudiés, l'ARN est aussi la cible du stress oxydant. Il n'y a pas de consensus en ce qui concerne la sensibilité de l'ARN au dommage oxydant : certains chercheurs évoquent une vulnérabilité importante due à sa structure simple brin, à son manque de protection protéique et à sa localisation plus proche de la source de formation des ROS [120]. D'autres suggèrent que le groupe –OH en position 2' lui conférerait une certaine résistance à l'oxydation par la polarité que ce groupement donne et par la charge positive qu'il fait accumuler dans l'état de transition de la réaction de l'oxydation [121]. Néanmoins, les dommages, quand ils ont lieu, sont du même type que ceux observés dans l'ADN, avec la formation des mêmes bases oxydées et de cassures simple brin (celles-ci constituent 40-90% de dommage oxydant de l'ARN) [120].

Parmi tous les effets du stress oxydant au niveau des cellules, ce sont les lésions de l'ADN qui sont pour nous d'intérêt particulier, car elles constituent le cœur du sujet de cette thèse. C'est pourquoi, différents types de dommages qui peuvent être provoqués par des agents d'oxydation seront développés dans les paragraphes suivants. Les processus de leur prise en charge par les cellules vont également être expliqués, ainsi que les conséquences de l'échec éventuel de leur réparation.

2. Lésions de l'ADN générées par le stress oxydant

L'ADN est une des cibles du stress oxydant dans les cellules, mais ses lésions peuvent avoir des conséquences beaucoup plus graves pour les cellules, que les dommages survenus au niveau des protéines ou des lipides. Diverses altérations qui peuvent être générées par des agents oxydants vont être présentées dans ce chapitre. Ces dommages peuvent toucher soit le désoxyribose, soit, plus couramment, la base.

2.1. Oxydation du désoxyribose

Le radical hydroxyle est capable d'arracher un proton au niveau de chacun des cinq atomes de carbone, ce qui donne naissance à de nombreux produits d'oxydation. Le sucre endommagé peut rester accroché à la base située dans la chaîne de l'ADN ou à son extrémité, ou bien il peut être libéré de la chaîne nucléique [122]. Les exemples de ce type de dommage au niveau du désoxyribose sont montrés sur la Figure 7.

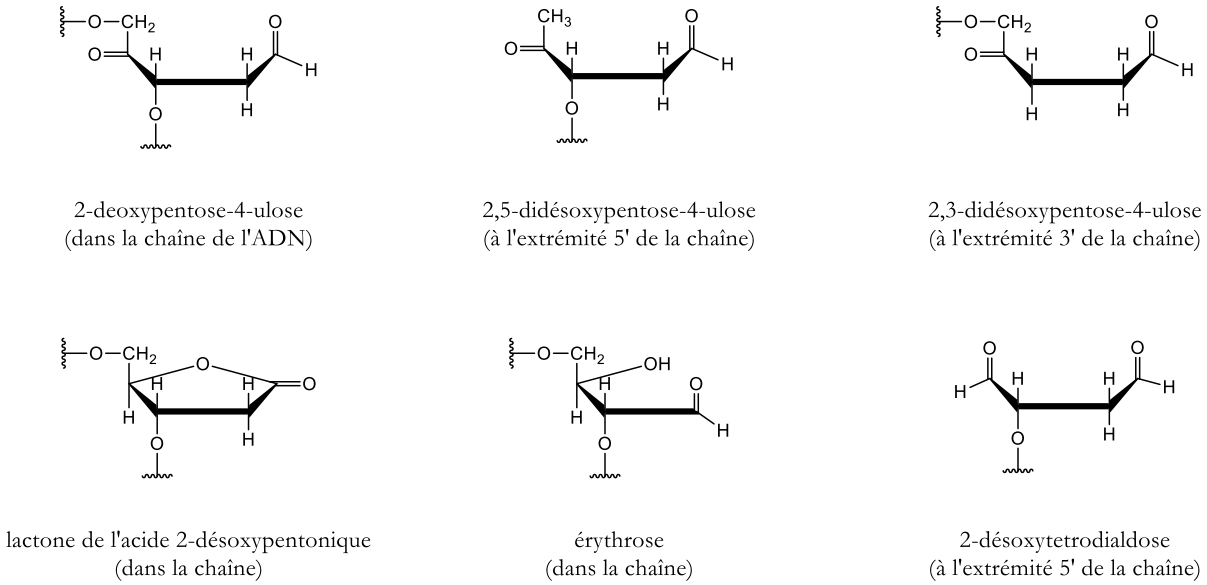


Figure 7 : Exemples des produits d'oxydation du désoxyribose. Adapté de [122].

2.2. Oxydation des bases : lésions simples

Les deux principales voies par lesquelles le stress oxydant peut endommager les bases sont : l'addition du radical hydroxyle à la double liaison d'une purine ou d'une pyrimidine, ou bien l'abstraction d'un atome d'hydrogène au niveau du groupement méthyle de la thymine ou de la liaison C−H du désoxyribose [122].

2.2.1. Types de lésions et leur formation

Les quatre bases de l'ADN : l'adénine, la guanine, la thymine et la cytosine ne sont pas toutes affectées de la même manière par le stress oxydant. La guanine possède le potentiel ionisant le plus bas, et de ce fait elle est la cible principale des agents oxydants, que ce soit pour leur effet direct ou indirect [123]. Aussi, les radicaux formés au niveau des autres bases de l'ADN peuvent être transférés vers la guanine. Tout cela fait que le niveau de la 8-oxo-désoxyguanosine, la principale lésion d'oxydation, est environ six fois plus élevé que la somme de toutes les lésions d'oxydation de la thymine [16].

2.2.1.1. Oxydation des purines

Comme indiqué dans le paragraphe précédent, la guanine est la base de l'ADN la plus touchée par les processus oxydants. Parmi les principales lésions d'oxydation de purines on retrouve : la 8-oxo-désoxyguanosine (8-oxo-dG), la 8-oxodésoxyadénosine (8-oxo-dA), mais aussi la Fapy-désoxyguanosine (Fapy-dG) et la Fapy-désoxyadénosine (Fapy-dA) (le mécanisme de leur formation sera exposé plus loin dans ce paragraphe). Parmi ces lésions, la 8-oxo-dG est la plus étudiée. Elle est mutagène, car capable de se mésapparier avec une adénine, ce qui résulte en transversion G→T, la deuxième mutation la plus fréquente dans les cellules cancéreuses [122,124–127]. Son niveau basal dans les tissus sains est estimé à 10^3 /cellule /jour, tandis que dans les tissus cancéreux il peut atteindre 10^5 /cellule /jour [128,129]. Le niveau de la 8-oxo-dA est environ dix fois inférieur à celui de la 8-oxo-dG [123], et pendant de nombreuses années cette lésion était considérée comme non-mutagène, puisque les expériences *in vitro* ont démontré que les ADN polymérases introduisaient la thymine en face de cette adénine modifiée, et n'engendraient donc pas de modifications de la séquence. Néanmoins, des expériences *in vivo* ne

confirment pas ces informations, et en fonction de la séquence dans laquelle est située la 8-oxo-dA, des mutations A→C ou A→G ont été observées [130,131]. Les deux autres lésions mentionnées : Fapy-dG et Fapy-dA, sont présentes dans le génome au même niveau que la 8-oxodésoxyguanine, et leur quantité est estimée à 1500-2500 /cellule /jour pour Fapy-dG et à 1000-2000 /cellule /jour pour Fapy-dA [128].

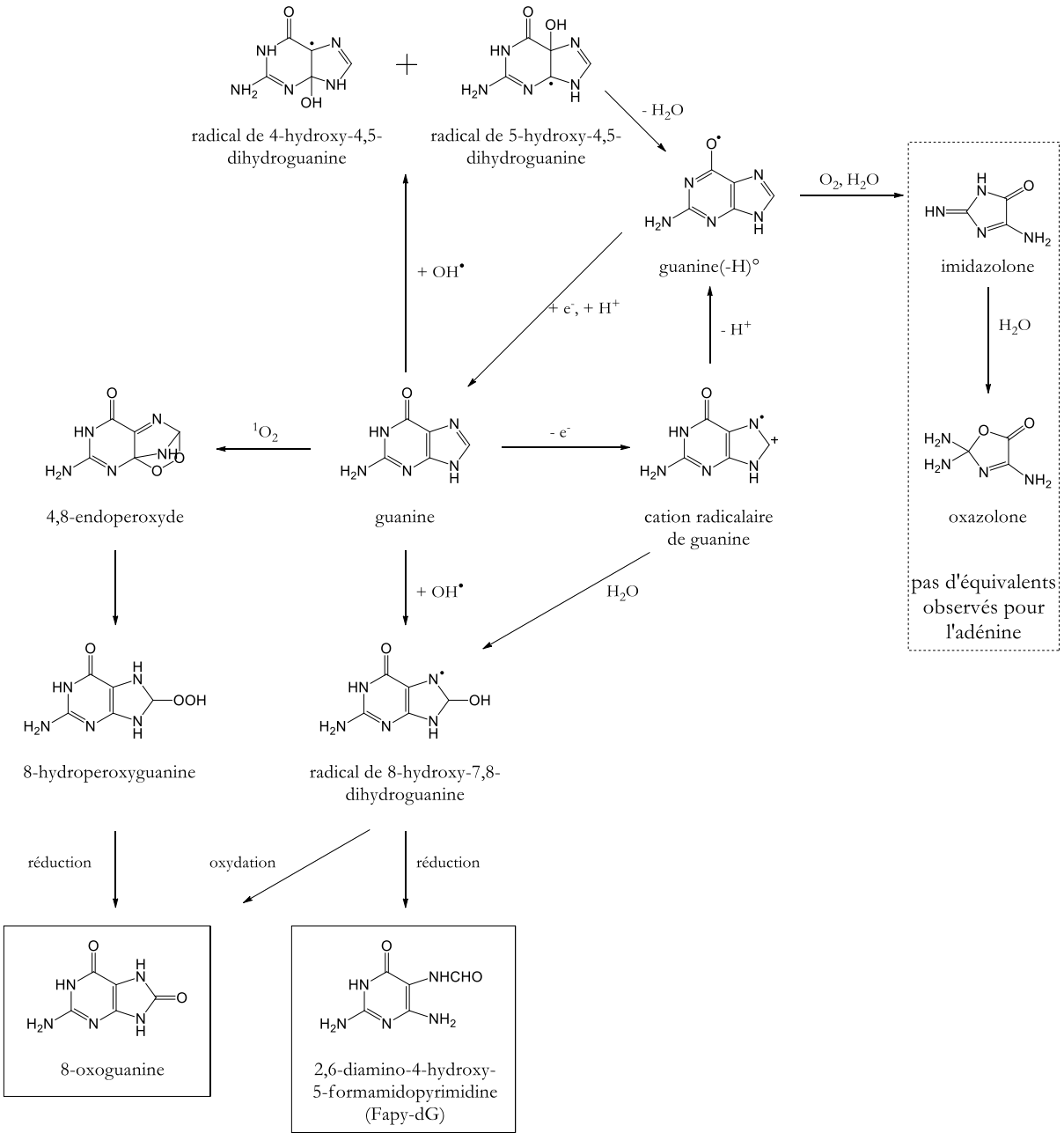


Figure 8 : Schéma de l'oxydation de la guanine. Inspiré de [119].

Les mécanismes d'oxydation de la guanine et de l'adénine sont analogues, un seul, celui de la guanine, sera décrit ici (Figure 8).

L'attaque du radical hydroxyle (˙OH) sur la guanine conduit à l'addition de celui-ci au niveau d'une des doubles liaisons du cycle. Ceci génère trois produits radicalaires, en fonction de la liaison qui va être affectée et de l'atome de carbone sur lequel va s'additionner le groupement –OH : radical de 4-hydroxy-4,5-dihydroguanine (C4−OH), radical de 5-hydroxy-4,5-dihydroguanine (C5−OH ; son équivalent dans l'adénine est formé en très faible quantité) et radical de 8-hydroxy-7,8-dihydroguanine (C8−OH). Les radicaux C4−OH et C5−OH peuvent être reconstitués en guanine, ou convertis en **oxazolone**, par l'intermédiaire d'un radical neutre de guanine(-H) (pas d'équivalents observés pour l'adénine). Quand un électron est arraché à la guanine, un cation radicalaire peut être formé – celui-ci peut être ensuite transformé en radical neutre mentionné ci-dessus ou en radical C8−OH. Le radical C8−OH peut être, quant à lui, soit oxydé soit réduit : la réduction va donner la **2,6-diamino-4-hydroxy-5-formamidopyrimidine (Fapy-dG)** comme produit final, tandis que l'oxydation va mener à la **8-oxoguanine**. Une autre réaction donne également la formation de la 8-oxoguanine – celle de l'attaque de l'oxygène singulet. La guanine est la seule nucléobase réagissant avec $^1O_2$, et cette réaction donne une 4,8-endopéroxyde, qui est ensuite transformé en 8-hydropéroxyguanine, qui est finalement réduite en 8-oxoguanine. La prévalence de la 8-oxoguanine ou de la Fapy-dG est influencée par les conditions redox : l'environnement oxydant favorise la formation de la 8-oxoguanine, et réducteur – de la Fapy-dG [119,122,123].

2.2.1.2. Oxydation de la thymine

L'attaque d'un radical hydroxyle sur la thymine à pH neutre peut s'effectuer à deux endroits : soit au niveau de la double liaison C5=C6, soit au niveau du groupement méthyle, dont il arrache un atome d'hydrogène en formant un radical méthylèneuracile ((CH$_2$)OH). L'addition de ˙OH au niveau de la double liaison C5=C6 est favorisée et elle peut donner lieu à deux types de radicaux : C5−OH ou C6−OH. Le destin des trois radicaux ainsi formés : C5−OH, C6−OH et (CH$_2$)OH, dépend ensuite de l'environnement de l'ADN, puisque la présence ou l'absence d'oxygène va influencer leurs transformations. En absence d'oxygène, le radical C5−OH, qui est le produit majoritaire de l'attaque du radical hydroxyle, est oxydé et par la suite hydraté, afin de donner le glycol de thymine, tandis que le radical C6−OH est, quant à lui, réduit et protoné, pour se transformer en 6-hydroxy-5,6-dihydrothymine. Le radical méthylèneuracile est oxydé et hydraté, comme le radical C5−OH, pour conduire à la 5-hydroxyméthyluracile.

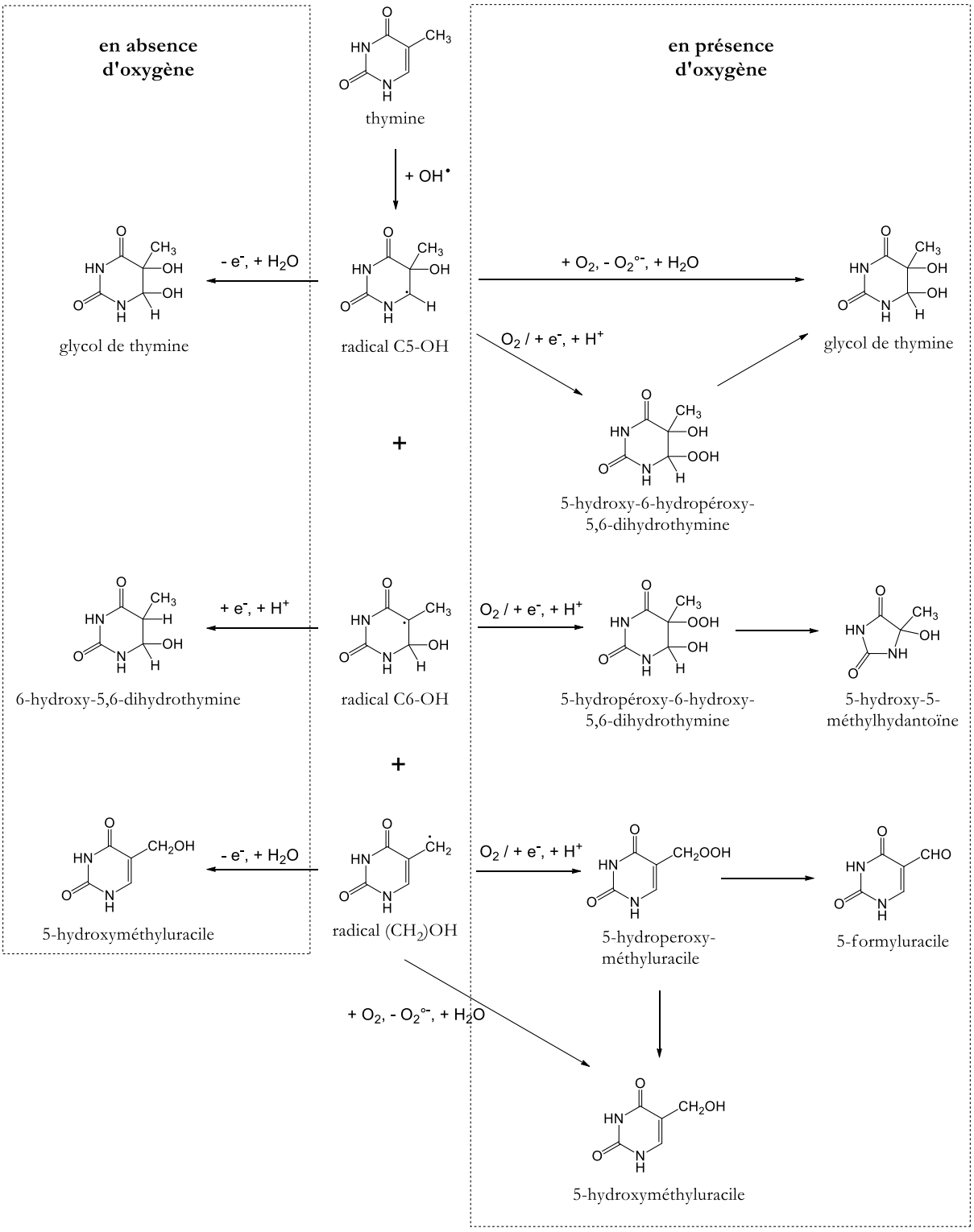


Figure 9 : Schéma de l'oxydation de la thymine. Inspiré de [119,132].

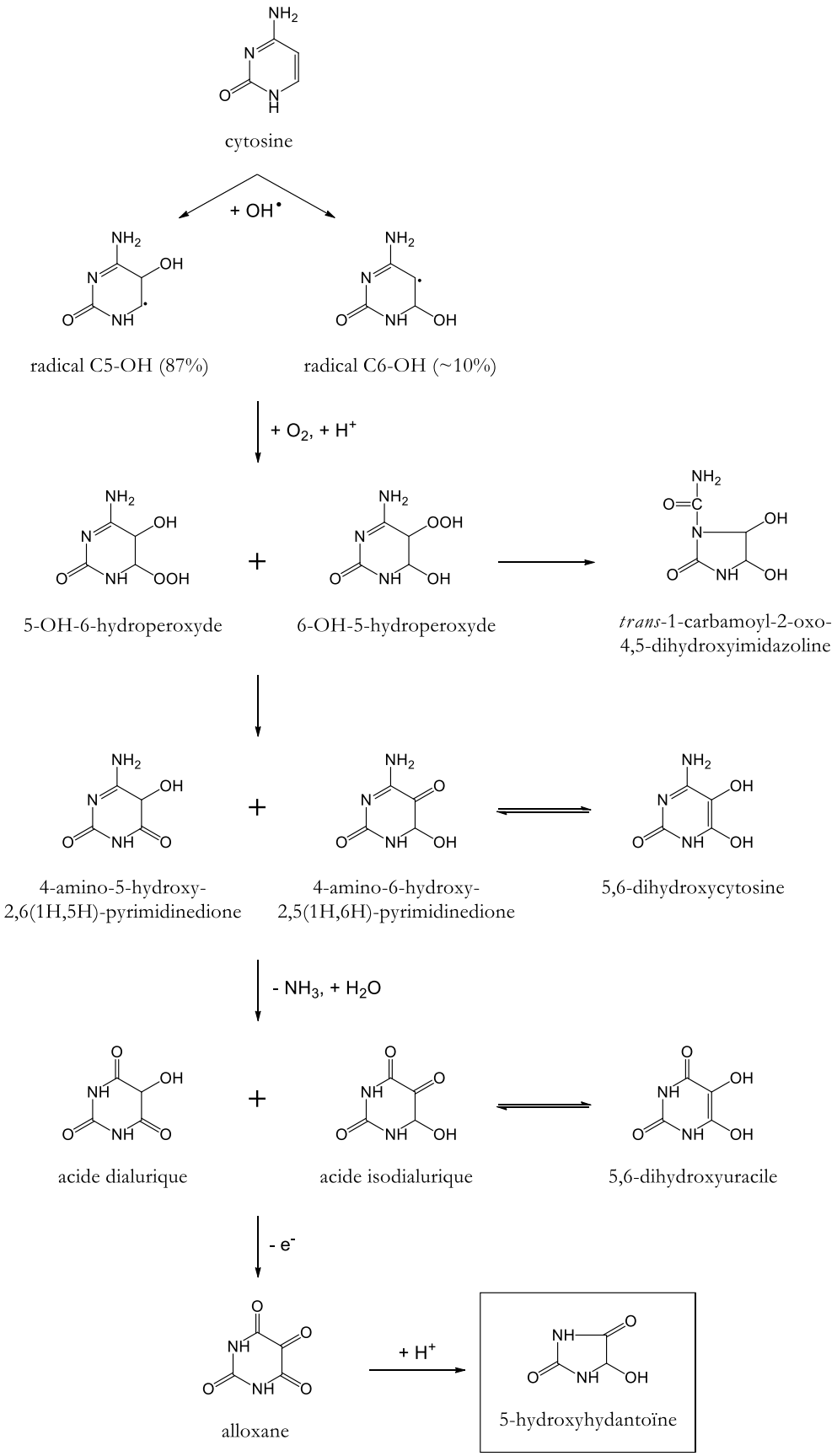


Figure 10 : Schéma de l'oxydation de la cytosine avec le radical hydroxyle, en présence d'oxygène. Adapté de [119].

En présence d'oxygène, les trois radicaux subissent un autre sort, même si certains produits finaux, tels que le **glycol de thymine** ou la **5-hydroxyméthyluracile**, peuvent être les mêmes que dans le cas d'un environnement anaérobie. La formation de ces deux dérivés passe par le transfert direct d'un électron vers l'oxygène, par l'addition de l'oxygène suivie par l'élimination de $O_2^{\bullet -}$ et l'hydrolyse, ou par la dégradation des hydropéroxydes intermédiaires instables. La **5-formyluracile** et la **5-hydroxy-5-méthylhydantoïne** sont ainsi retrouvées parmi les produits de l'oxydation de la thymine [132]. Le schéma réactionnel est présenté sur la Figure 9.

2.2.1.3. Oxydation de la cytosine

L'attaque du radical hydroxyle sur une molécule de cytosine génère deux radicaux majoritaires, résultant de l'addition de $^{\bullet}OH$ à la double liaison C5=C6 : C5−OH et C6−OH. Les deux vont ensuite réagir de manière analogue, en donnant à chaque fois deux produits, dont la seule différence va porter sur les modifications au niveau du carbone C5 ou C6, réciproquement. Dans une première étape, le radical va être oxygéné et protoné pour donner un hydropéroxyde instable, qui va ensuite se transformer en 4-amino-5-hydroxy-2,6-pyrimidinedione (ou 4-amino-6-hydroxy-2,5-pyrimidinedione). La désamination de celui-ci va donner ensuite l'acide dialurique ou isodialurique, oxydés facilement en **alloxane**. La protonation de l'alloxane donne comme produit final **5-hydroxyhydantoïne**. De nombreux produits dérivés des intermédiaires de la réaction décrite peuvent également être formés, comme par exemple la **5,6-dihydroxycytosine** ou la **5,6-dihydroxyuracile** [119,132] (Figure 10).

2.2.1.4. Autres lésions simples

Outre les lésions simples d'oxydation, il existe d'autres types de dommages non-volumineux de l'ADN, qui vont être reconnus et pris en charge par les cellules de manière similaire. Une molécule d'ADN peut ainsi subir la désamination, qui va toucher toutes les bases, sauf la thymine qui ne contient pas de groupement $–NH_2$. L'adénine, la guanine et la cytosine vont donner après leur désamination : l'hypoxanthine, la xanthine et l'uracile, réciproquement. La 5-méthylcytosine, une forme de cytosine modifiée, va se désaminer pour donner la thymine mésappariée avec une guanine. La méthylation, quant à elle, peut toucher tous les nucléosides, et de nombreux dérivés sont ainsi formés, en fonction de l'atome de carbone auquel va être fixé le

groupement $-CH_3$. Enfin, des sites abasiques sont formés quand une purine ou une pyrimidine sont enlevées en laissant le résidu de désoxyribose dans la chaîne nucléique.

Toutes ces lésions ne sont pas générées dans l'ADN au même degré ni réparées avec la même efficacité, et c'est pour cela que leur niveau basal dans les cellules est caractéristique de chaque lésion. Par exemple, environ 30 molécules de thymine mésappariées avec la guanine sont générées suite à la désamination de la 5-méthylcytosine dans le génome par jour, et leur niveau basal reste entre 10 ou 20 par génome. Au contraire, parmi les 9000 sites abasiques formés par jour dans un génome, seul 5 persistent au final (Tableau 1).

Tableau 1 : Lésions de l'ADN formées de manière endogène dans les cellules humaines [133].

Lésion	**Mécanisme de formation**	**Nombre de résidus formé par jour dans le génome humain**	**Nombre de résidus dans le génome des cellules possédant un système de réparation fonctionnel**
uracile	Désamination de la cytosine	400	~1
thymine (: guanine)	Désamination de la 5-méthylcytosine	30	10 – 20
hypoxanthine	Désamination de l'adénine	10	~1
8-oxo-dG	Oxydation de la guanine	~1000	~1
Fapy-dG	Oxydation de la guanine	~200	~5
glycol de thymine et produits similaires	Oxydation des pyrimidines	~500	~5
site abasique	Dépurination hydrolytique	9000	~5

2.2.2. Prise en charge par le système de réparation BER

Les lésions de l'ADN dites simples, c'est-à-dire des modifications de bases qui ne génèrent pas la distorsion de l'hélice, sont généralement prises en charge par le système de réparation par excision de bases (Base Excision Repair ; BER). La base endommagée ou incorrectement insérée dans la séquence est ainsi enlevée et remplacée par une base correcte. Ce système comporte de nombreuses enzymes, dont certaines sont redondantes et capables de réparer les mêmes lésions.

Le BER est présent chez tous les organismes, mais des différences peuvent être observées au niveau des enzymes et des mécanismes précis qui y sont impliquées. Nous allons présenter son fonctionnement en prenant l'exemple du BER chez les eucaryotes.

2.2.2.1. BER chez les eucaryotes

Typiquement, la prise en charge par le BER d'une base altérée se déroule en cinq étapes :

1) reconnaissance et excision la base endommagée par une ADN glycosylase, avec formation d'un site abasique
2) incision du site abasique par une endonucléase ou une lyase
3) enlèvement du résidu de sucre par une phosphodiestérase ou une lyase
4) insertion d'une nouvelle base avec une ADN polymérase
5) ligation avec une ADN ligase [134] (Figure 11)

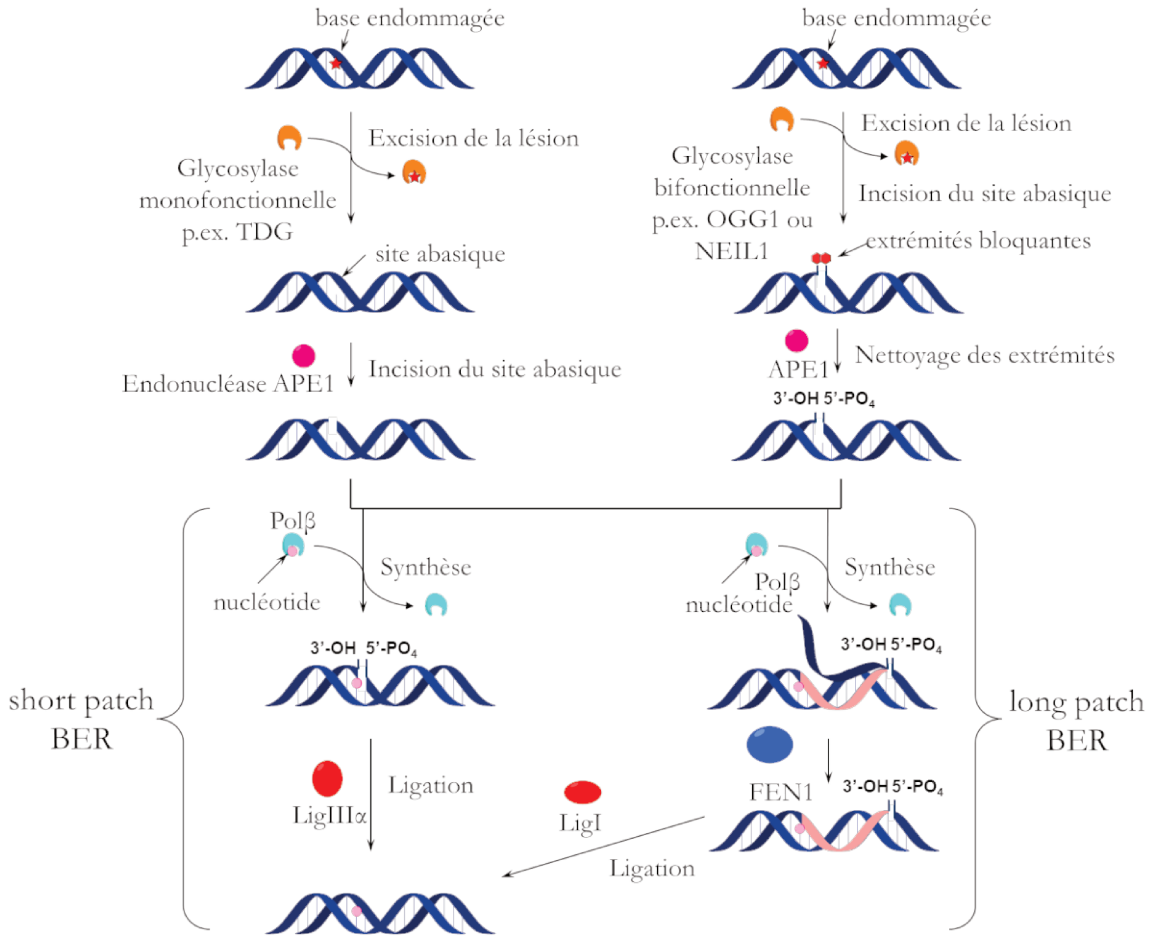


Figure 11 : Représentation schématique de la réparation de l'ADN par le système BER. Inspiré de [133,135].

1) Reconnaissance et excision de la base endommagée

La première étape de la réparation des lésions non-volumineuses consiste en l'identification et la reconnaissance de la base endommagée. Ceci est effectué par une des ADN glycosylases, qui « scannent » la séquence à la recherche des dommages. Une fois le dommage identifié, cette enzyme déforme l'ADN afin de faciliter l'entrée de la base modifiée dans le site de fixation de la glycosylase [133]. Il existe de nombreuses glycosylases, qui diffèrent par leur spécificité : il n'est pas rare qu'une glycosylase soit capable de reconnaître plusieurs lésions, ce qui fait qu'une lésion peut être réparée malgré la déficience d'une des protéines du BER. Certaines de ces glycosylases sont présentées, avec leurs substrats, dans le Tableau 2.

Tableau 2 : Glycosylases humaines et leurs substrats. Source [133].

Glycosylase		**Activité AP lyase**	**Substrats**
Uracile ADN glycosylase	UNG	non	U, U :dG, U :dA, 5-FU
Uracile ADN glycosylase	SMUG1	non	U
Thymine ADN glycosylase	TDG	non	U :dG, dT :dG
Alkyleadénine / méthylpurine ADN glycosylase	AAG/ MPG	non	3-Me-dA, 7-Me-dA, 3-Me-dG, 7-Me-dG
MutY homologue	MYH	β-élimination	dA :dG, dA :8oxo-dG
8-oxo-guanine glycosylase 1	OGG1	β-élimination	8-oxo-dG :dC, Fapy-dG
Endonucléase III homologue	NTH1	β-élimination	Tg, Cg, Fapy-dG, DHU, 5-OH-U, 5-OH-dC
Nei endonucléase VIII-like 1	NEIL1	β,δ-élimination	Tg, 5-OH-U, 5-OH-dC, Fapy-dA/dG, 8-oxo-dG
Nei endonucléase VIII-like 2	NEIL2	β,δ-élimination	son spectre se chevauche avec ceux de NTH1 et NEIL1
Nei endonucléase VIII-like 3	NEIL3	β,δ-élimination	Purines oxydées, Fapy-dG, Fapy-dA

En fonction de leur mode d'action, les glycosylases peuvent être définies comme mono- ou bi-fonctionnelles :

- les glycosylases monofonctionnelles, comme l'uracile ADN glycosylase (UNG) ou la thymine ADN glycosylase (TDG), sont seulement capables d'exciser la base endommagée, en coupant la liaison N-glycosidique entre la base et le 2'-désoxyribose ; elles transforment

ainsi la lésion en site abasique (les sites abasiques pouvant provenir également d'une dépurination ou dépyrimidination spontanée) ; leur action est ensuite suivie par celle d'une AP-endonucléase ;

- les glycosylases bi-fonctionnelles, telles que OGG1 ou NEIL1, en plus de leur activité glycosylase, possèdent une activité AP lyase, qui leur permet d'inciser la liaison phosphodiester située du côté 3' ou 5' du site abasique nouvellement formé ; cette incision peut être faite en une étape de β-élimination, ou bien en deux étapes de β,δ-élimination, suivant le type de coupure qui a lieu entre le sucre et les phosphates (Figure 12) ; une cassure simple brin avec un 3'-phosphate ou 3'-aldéhyde-phosphor-α,β-insaturé (PUA) est obtenue comme produit de l'action des glycosylases bi-fonctionnelles – ces extrémités étant bloquantes, elles doivent être enlevées avant que l'ADN polymérase puisse insérer une nouvelle base dans la séquence [133,136,137].

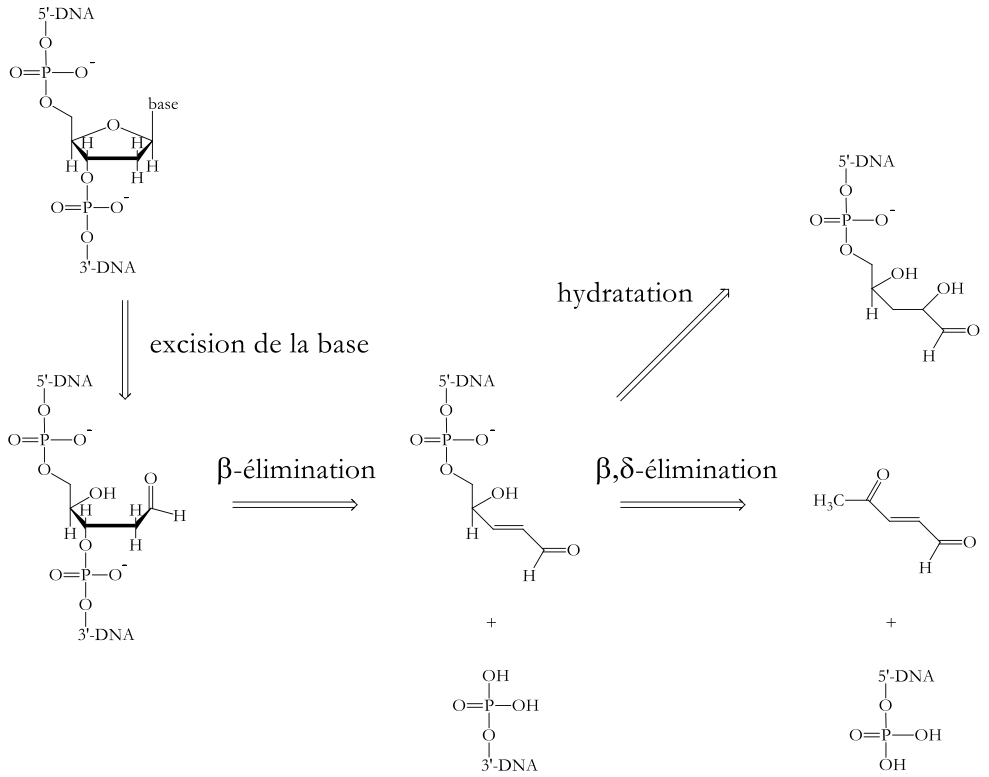


Figure 12 : Coupure d'un site abasique par les ADN glycosylases bifonctionnelles, suivant un mécanisme de β- ou β,δ-élimination. Inspiré de [138].

2) Incision du site abasique

Une fois que l'ADN glycosylase a enlevé la base endommagée de la séquence, le site abasique généré est pris en charge par une AP endonucléase. L'AP-endonucléase la plus importante chez les mammifères est APE1, et elle est responsable de plus de 95% de l'activité AP-endonucléase dans les cellules. Elle coupe la liaison située du côté 5' du site abasique et elle génère une cassure simple brin avec des extrémités 3'-OH et 5'-désoxyribose phosphate (dRP).

En plus de son activité AP endonucléase, APE1 possède également d'autres activités : elle catalyse l'enlèvement des extrémités bloquantes en 3' générées par les glycosylases bi-fonctionnelles, et elle est aussi dotée d'une activité 3' → 5' exonucléase, qui lui permet de corriger et d'enlever les nucléosides mal incorporés lors de la re-synthèse pendant le BER [133,139–142].

3) Elimination du résidu de sucre

Suite à l'action des glycosylases bifonctionnelles ou des AP-endonucléases, des cassures simple-brin avec une extrémité bloquante en 3' ou en 5' sont formées. Les groupements qui bloquent l'action des ADN polymérases sont : 5'-désoxyribose phosphate, 5'-OH, 3'-PO_4, 3'-phosphoglycolate ou 3'-aldéhyde-phosphor-α,β-insaturé. L'élimination du 5'-désoxyribose phosphate est effectuée essentiellement par la polymérase β (Polβ) qui possède, en plus de son activité polymérase, une activité lyase. Les groupes bloquants en 3', surtout 3'-PO_4, sont principalement enlevés par l'enzyme PNKP (polynucleotide kinase/phosphatase) – cette enzyme bi-fonctionnelle possède une activité 3'-phosphatase qui hydrolyse l'extrémité 3'-PO_4 en 3'-OH, et une activité 5'-kinase qui phosphoryle l'extrémité 5'-OH en 5'-PO_4, afin que les deux soient compatibles avec l'action de la polymérase. Cette étape de la réparation est également appelée « le nettoyage des extrémités » [133,143–145].

4) Synthèse

L'introduction d'une nouvelle base à la place de la base endommagée enlevée lors de la première étape de la réparation, est effectuée par les polymérases, dont la plus importante chez l'homme est la polymérase β. Elle nécessite une extrémité 3'-OH dans le brin endommagé et le brin complémentaire intact. Après avoir incorporé un désoxynucléotide, Polβ peut exercer son

activité 3' → 5' exonucléase – ceci lui permet d'enlever un nucléotide incorrect, avant qu'un autre soit incorporé à sa suite.

Le processus décrit ci-dessus correspond à ce qui est appelé short patch BER. Alternativement, la synthèse d'un fragment d'ADN de 2 à 12 bases peut avoir lieu, ce qui va être dénommé long patch BER. Dans ce cas, les polymérases δ et ε synthétisent un fragment du brin qui remplace progressivement le brin qui contenait la base endommagée, jusqu'à ce que celui-ci soit enlevé par la flap endonucléase 1.

Le facteur principal qui détermine le choix entre short patch ou long patch BER est la quantité d'ATP disponible : à forte concentration d'ATP, la ligation directe, donc short patch BER, va être favorisée, tandis qu'une faible concentration d'ATP va avantager le long patch BER. Il existe également des lésions résistantes à l'activité lyase de la polymérase β, telles que la 2-désoxyribonolactone, qui doivent être, par défaut, réparées par la voie longue [133,146–151].

5) Ligation

La dernière étape de la réparation de l'ADN consiste en la ligation des extrémités 3'-OH et 5'-PO_4 qui restent dans la séquence après l'incorporation d'un nouveau nucléotide, par formation d'une liaison phosphodiester. Ceci est effectué par une des ADN ligases – les deux les plus importantes sont : la ligase I, qui participe principalement dans la voie longue du BER, et la ligase IIIα, impliquée plutôt dans la voie courte [133,152–154].

2.2.2.2. BER chez les procaryotes

Le système BER est très conservé chez tous les organismes, depuis les bactéries jusqu'aux hommes. Ainsi, les étapes principales de son action sont fondamentalement les mêmes que celles présentées pour le BER eucaryote, les différences concernant surtout les enzymes qui y participent, et notamment celles qui reconnaissent initialement le dommage [155].

L'enzyme la plus connue du BER bactérien est la Fpg – formamidopyrimidine DNA glycosylase – appelée également MutM. C'est une glycosylase bifonctionnelle, qui reconnaît les lésions de purines, telles que Fapy-dA, Fapy-dG et 8-oxo-dG. Elle est active sur les mêmes substrats que l'OGG1 humaine, et ces deux protéines sont souvent considérées comme homologues [155]. Néanmoins, il n'est pas toujours évident de déterminer les homologues

bactériens des protéines humaines, car souvent les enzymes qui possèdent des identités au niveau de la séquence, ne reconnaissent pas les mêmes substrats, ou inversement. Aussi, de nombreuses enzymes sont redondantes dans un organisme mais pas dans l'autre. Ainsi, le Tableau 3 présente des glycosylases majeures humaines et leurs prototypes bactériens [156].

Tableau 3 : ADN glycosylases identifiées chez les *E.coli* et chez les hommes. Source : [156].

enzyme	*E.coli*		homme	
	nom	substrats	nom	substrats
uracile ADN glycosylase	Ung	U	UNG	U
			SMUG1	
			TDG	U, dT :dG
alkylbase ADN glycosylase	Tag, AlkA	3-Me-A	MPG	3-Me-dA, hypoxanthine
adénine-spécifique mismatch ADN glycosylase	MutY	dA dans 8-oxo-dG :dA	MYH	dA dans 8-oxo-dG :dA
ADN glycosylases spécifiques pour les bases oxydées	Endo III (Nth)	5-OH-dC, 5-OH-U, Tg	NTH1	Tg, 5-formyl-U
			OGG1	8-oxo-dG, Fapy-dG
	Endo VIII (Nei)	Tg, 5-OH-U	NEIL1	Fapy-dA, Tg, 5-OH-U
	Fpg (MutM)	8-oxo-dG, Fapy-dG	NEIL2	hydantoïne, 5-OH-U

2.2.3. Conséquences de la présence des lésions simples non réparées dans l'ADN

Déficience du BER et maladies

Il a été démontré que l'accumulation des dommages d'oxydation dans l'ADN contribue à de nombreuses maladies, telles que le cancer ou les maladies neurodégénératives, mais également joue un rôle dans le vieillissement ou la mort cellulaire [135]. Il en résulte que la déficience du système de réparation BER aurait une influence sur ces processus. Comme il a été indiqué dans les paragraphes précédents, la première étape du BER consiste en l'élimination de la base endommagée par l'une des glycosylases, dont les spécificités se chevauchent. Par conséquent, le déficit de l'une d'entre elles conduit rarement à un phénotype de maladie [135]. Les souris déficientes en une glycosylase sont ainsi indifférenciables des souris normales, et ce n'est que le manque de plusieurs glycosylases qui leur confère un phénotype particulier : la délétion simultanée de MYH et OGG1 augmente le risque de cancer du poumon et des ovaires, ainsi que des lymphomes [157], tandis que les souris déficientes en NTH et NEIL1 développent des tumeurs pulmonaires et hépatocellulaires [158]. Le seul cas chez l'homme, où le manque d'une

glycosylase serait à l'origine d'une maladie, est la déficience de MYH chez les patients atteints de la polypose colorectale MYH-associée [159].

Il existe des mutations des enzymes du BER qui, si elles sont homozygotes, sont létales à l'état embryonnaire – ce sont des mutations qui touchent la polymérase Polβ [160], l'endonucléase APE1 [161], la protéine XRCC1 (elle rentre en complexe avec la ligase III) [162] et les ligases I et III [163,164]. Ceci montre le caractère indispensable du BER pour le développement embryonnaire. Néanmoins, si les mutations ne touchent qu'un allèle, elles ne sont pas mortelles, mais elles sont à l'origine de nombreuses aberrations. Les souris $APE1^{+/-}$ ont une sensibilité accrue au stress oxydant et une tendance à développer des lymphomes et des adénocarcinomes spontanés [161,165]. Les souris $Polβ^{+/-}$ ont un niveau basal d'aberrations chromosomiques et de cassures simple brin plus élevé que les souris wild-type et développent également des lymphomes et des adénocarcinomes spontanés [166,167]. Quant aux souris $XRCC1^{+/-}$, elles présentent de nombreuses lésions précancéreuses dans le colon suite à l'exposition à l'azoxyméthane [168].

La déficience du BER pourrait également être l'un des facteurs à l'origine des maladies neurodégénératives, celles-ci étant liées à l'accumulation des lésions d'oxydation. La réparation réduite de ces lésions aurait ainsi un impact sur la survie cellulaire et il a été prouvé que les déficiences au niveau de certaines enzymes, telles qu'UNG ou APE1 conféraient des prédispositions à l'apoptose ainsi qu'une sensibilité accrue au stress génotoxique [169,170]. Aussi, l'activité du BER étant probablement modifiée avec l'âge, la baisse de son efficacité pourrait être liée au déclin de l'activité cérébrale chez des personnes âgées [135].

BER et réponse aux traitements antitumoraux

De nombreuses études ont démontré que la résistance de cellules cancéreuses aux traitements antitumoraux serait dans certains cas due à la surexpression et à l'activité augmentée des systèmes de réparation de l'ADN. Certaines protéines du système BER pourraient ainsi jouer un rôle dans l'efficacité de ces thérapies anticancéreuses, et les enzymes impliquées dans toutes les étapes de la réparation seraient alors concernées. Tout d'abord, il a été démontré que la diminution de l'activité des ADN glysosylases conférait une sensibilité accrue aux irradiations γ (désactivation de NEIL1 dans les cellules souches embryonnaires des souris) ou aux agents alkylants tels que le temozolomide ou la carmustine (inhibition des AAG / MPG dans les cellules HeLa). De plus, la surexpression de OGG1 confère une résistance à un agent anticancéreux thio-TEPA dans les cellules hématopoïétiques [133].

2.3. Lésions complexes

Les lésions complexes englobent un vaste groupe de dommages de l'ADN générés principalement par le rayonnement ionisant, et qui sont réfractaires au système BER. De ce fait, ces lésions sont plus mutagènes et leurs conséquences pour les cellules sont souvent plus graves que celles des lésions simples. Généralement, les lésions complexes sont définies comme deux ou plusieurs lésions simples localisées à proximité (elles sont souvent incluses dans un ou deux pas de l'hélice) qui peuvent être formées sur un même brin de l'ADN ou bien sur les brins opposés. Parmi les lésions complexes on retrouve par exemple les lésions en cluster, les lésions en tandem, ou encore les cyclonucléosides.

2.3.1. Formation et types

Les lésions complexes sont dans la majorité des cas le résultat de l'action indirecte du rayonnement ionisant, même si certains chercheurs ont démontré une faible quantité de ce type de dommages dans les cellules non-stressées [171,172]. Leur fréquence et leur complexité dépendent dans ce cas de la dose d'irradiation reçue mais également de l'énergie portée par la molécule qui traverse la matière et qui est appelée transfert d'énergie linéaire (TEL). Cette valeur dépend notamment du type de la source du rayonnement. Ainsi, d'après les simulations théoriques, les irradiations à faible TEL pourraient générer des clusters composés de 10 lésions, tandis que les irradiations à fort TEL seraient capables de former des clusters encore plus complexes, allant jusqu'à 25 lésions [173]. En parallèle, il a été démontré qu'avec l'augmentation du TEL, le nombre de lésions isolées formées dans l'ADN diminuait au profit des lésions en cluster [174]. Ces clusters sont composés des lésions simples d'oxydation, telles que les bases oxydées, les sites abasiques, les lésions des sucres, les pontages ou les cassures simple-brin. En fonction de leur composition, on distingue les clusters « avec les cassures double-brin » ou « sans cassures double-brin » et les lésions en tandem, et il a été estimé que ces premiers constituaient jusqu'à 80% de toutes les lésions complexes générées par le rayonnement ionisant [175]. Les cassures double-brin sont extrêmement toxiques et vont être décrites en détails dans le chapitre 2.3.4.

Le type le plus simple de lésions complexes est une lésion tandem, composée de deux lésions simples localisées l'une à côté de l'autre dans la séquence d'ADN. Cette lésion tandem

peut être composée de deux bases endommagées, une base endommagée et un site abasique, ou elle peut contenir également une cassure simple-brin (Figure 13).

Un autre type particulier des lésions complexes sont les cyclonucléosides. Ces dommages sont uniques, puisqu'ils touchent dans un même nucléoside la base et le sucre – le mécanisme de leur formation et leurs conséquences pour le fonctionnement de la cellule vont être détaillés dans le chapitre 3.

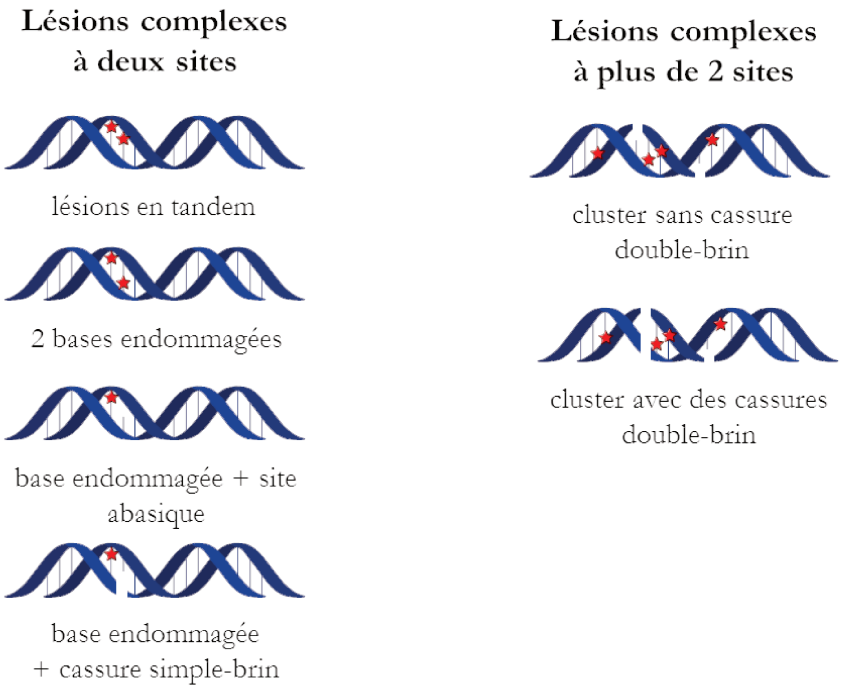


Figure 13 : Types de lésions complexes. Inspiré de [171].

2.3.2. Réparation

Les lésions complexes, même si elles sont composées de lésions simples d'oxydation, ne peuvent pas être considérées simplement comme un groupe des lésions individuelles. Elles ne sont réparées que très faiblement dans les cellules, et l'on peut détecter leur présence même plusieurs semaines après l'irradiation [172].

Afin d'analyser l'influence de la présence d'une lésion d'oxydation sur la réparation d'une autre, nous allons considérer la réparation des clusters composés uniquement de deux dommages. De nombreuses équipes de chercheurs ont analysé l'efficacité d'excision d'une base endommagée en fonction du dommage situé à proximité. Les lésions utilisées le plus souvent dans ces études

étaient : la 8-oxo-dG, la 5,6-dihydro-désoxythymine, le site abasique et la cassure simple-brin. Dans tous les cas, il a été démontré qu'une hiérarchie était présente dans la réparation de ces lésions complexes. Si deux bases endommagées sont localisées à proximité, l'excision de la première suivie de l'action d'une AP endonucléase va générer une cassure simple brin qui va, à son tour, inhiber la réparation du deuxième dommage [176]. Si le cluster est composé d'une base lésée et d'un site abasique ou d'une cassure simple-brin, cette deuxième va inhiber directement la réparation de la base endommagée [177]. De même, la prise en charge d'un site abasique va être empêchée par la présence d'un deuxième site abasique localisé à proximité [177]. La vitesse et l'efficacité de la réparation de la deuxième base dépend également de la distance entre les deux dommages : son inhibition est la plus forte si les deux lésions sont séparées d'un seul nucléotide, et elle diminue avec la distance croissante [176,177].

Cette inhibition ou ce retard de la réparation d'un des deux dommages faisant partie d'un complexe, ne sont pas accidentelles. En fait, la réparation simultanée d'une base oxydée et d'une cassure ou bien d'un site abasique pourrait mener à la formation des cassures double-brin et, par la suite, à la perte de l'information génétique. Ainsi, il a été postulé que les cellules pourraient se protéger contre ce dommage supplémentaire en retardant la réparation de la deuxième lésion [171].

Puisque la réparation des lésions complexes est problématique et mène souvent à une perte de l'information génétique, les cellules possèdent d'autres stratégies visant à éliminer les clusters de dommages. La première consiste en la réplication, qui peut séparer les deux lésions situées sur des brins opposés, en les transformant ainsi en deux lésions simples, qui peuvent ensuite être facilement prises en charge par les deux cellules filles. La deuxième stratégie est la mort cellulaire, qui sert simplement à éliminer les clusters irréparables [178]. Si, néanmoins, suite à une réparation parallèle de deux lésions, une cassure double-brin est formée, elle va être prise en charge par les systèmes de réparation propres à ce type de dommage. Ces différentes stratégies d'une cellule pour prendre en charge les lésions tandem sont présentées sur la Figure 14.

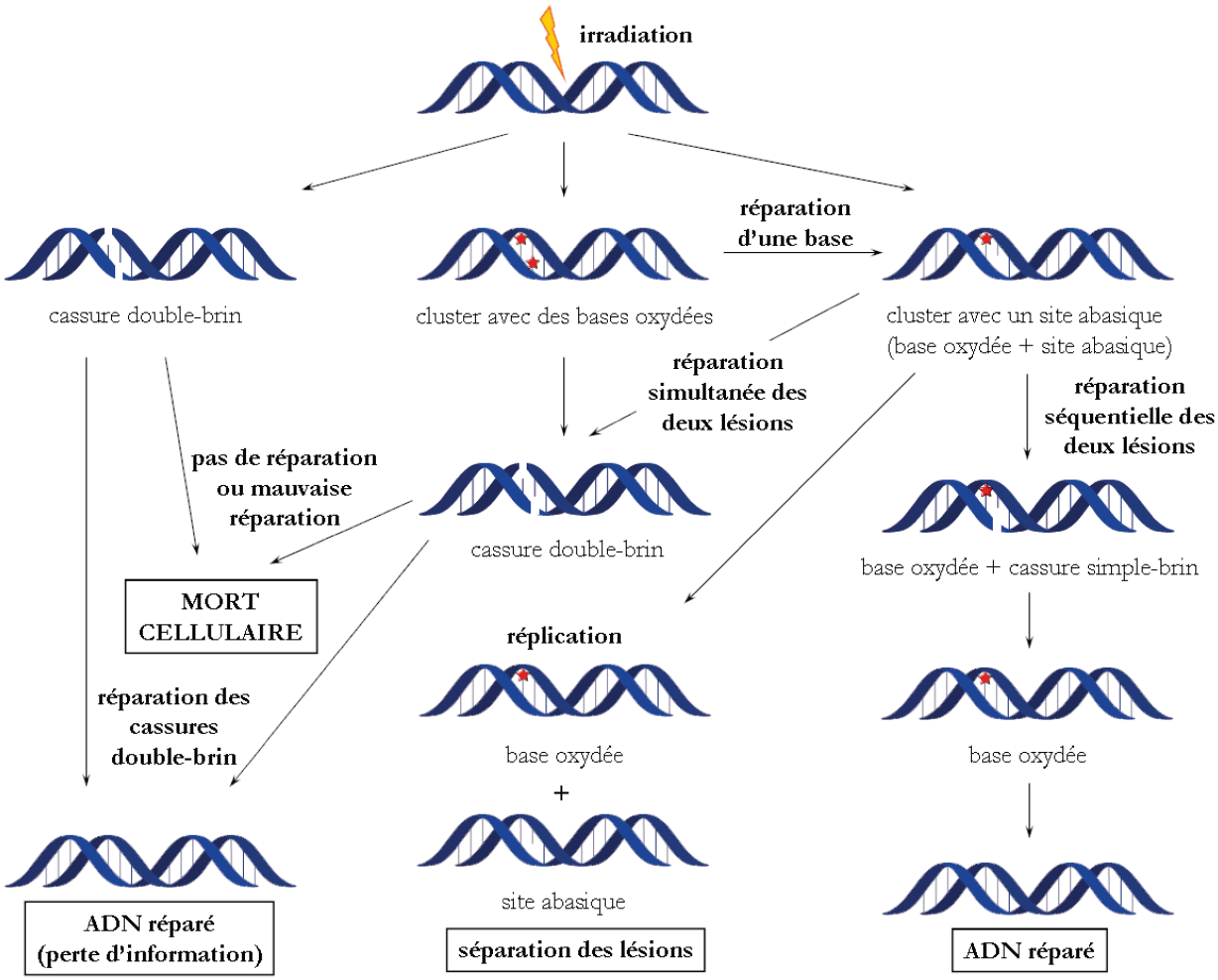


Figure 14 : Destin de différents types des clusters générés par le rayonnement ionisant, en fonction des types de dommages les composant. Inspiré de [171].

2.3.3. Conséquences

Comme cela a été présenté dans le paragraphe précédent, les lésions complexes sont difficilement prises en charge par les cellules et leur réparation est souvent incomplète. De ce fait, elles sont beaucoup plus néfastes pour les cellules que les lésions isolées. Ces lésions constituent un réel danger pour les cellules qui, dans le meilleur des cas, meurent suite au blocage de la réparation et de la transcription des gènes, ou qui subissent une perte d'information génétique qui peut déclencher la mutagénèse. Si une cellule ne meurt pas suite à l'irradiation qui engendre les lésions complexes de l'ADN, elle peut soit essayer de réparer ces dommages, soit, si elle en est incapable, attendre que ceux-ci « se diluent » au cours de la réplication. La tentative de réparation des clusters est risquée, puisque les intermédiaires de réparation qui sont formés, notamment les cassures simple- et double-brin, sont plus toxiques que les dommages initiaux [179].

L'efficacité des radiothérapies anti-tumorales est basée, entre autres, sur la formation des clusters de lésions dans les cellules cancéreuses. Le rayonnement à TEL élevé est utilisé de préférence, puisqu'il génère plus de clusters que le rayonnement à faible TEL, et le dépôt parallèle de l'énergie dans les cellules non-tumorales est relativement faible. Néanmoins, il existe un risque du développement d'un deuxième cancer dans les tissus avoisinant la tumeur, si un cluster est formé dans une cellule dont le suppresseur de tumeur a été muté. Le rapport bénéfice / risque reste toutefois plus avantageux pour les rayonnements à TEL élevé par rapport aux rayonnements à faible TEL. A faible TEL, moins de clusters de lésions sont formés dans le tissu cancéreux, tandis que les mutations dans les cellules avoisinantes sont plus fréquentes [179,180].

2.3.4. Cas particulier des cassures de l'ADN

Les cassures de l'ADN qui touchent en même temps les deux brins de l'hélice appartiennent aux lésions les plus cytotoxiques pour les cellules. Elles ont des origines très diverses – leur formation peut être due autant à des processus endogènes, comme la respiration mitochondriale ou la synthèse d'ADN à travers une cassure simple-brin, qu'à des facteurs exogènes, comme le rayonnement ionisant ou UV.

Les cassures double-brin font partie des lésions les plus toxiques pour les cellules, non seulement parce qu'elles sont difficilement réparables, mais également parce qu'une seule cassure non-réparée suffit pour déclencher l'apoptose [181,182]. En même temps, la réparation de ce type de dommages est souvent liée à une perte d'information génétique, à des mutations ou à des réarrangements chromosomiques [181]. C'est pour cela que la prise en charge de ces lésions et leur réparation correcte sont essentielles à la survie cellulaire.

La prise en charge des cassures double-brin ne consiste pas uniquement à recruter des protéines de réparation au site où est localisé le dommage. C'est un mécanisme beaucoup plus complexe, qui englobe la réparation et également la transmission du signal au sein de la cellule. A cause de son extrême toxicité, une cassure double-brin doit être réparée avant que la cellule ne se divise – pour donner du temps à la cellule d'effectuer la réparation, l'arrêt du cycle cellulaire a souvent lieu grâce à la présence de nombreux « postes de contrôle » (checkpoints), qui veillent à l'intégrité du génome [183].

Les cassures double-brin peuvent être prises en charge par l'un des deux systèmes de réparation suivants : Non-Homologous End-Joining (NHEJ) ou Recombinaison Homologue (HR) :

- NHEJ peut réparer tout type de cassures double-brin, puisqu'il ne nécessite pas la présence d'un brin non-endommagé servant de matrice – il colle les deux extrémités de la cassure se trouvant à proximité. Le désavantage de cette stratégie est que, si plusieurs cassures sont présentes au sein de la cellule, les extrémités des deux cassures différentes peuvent être collées ensemble, ce qui va engendrer des réarrangements chromosomiques. Aussi, même si les bonnes extrémités sont liguées ensemble, des délétions d'une partie de la séquence sont souvent observées [181,184–186].
- HR est une deuxième voie de réparation des cassures double-brin, favorisée surtout pendant la division mitotique. C'est un mécanisme qui utilise la chromatide-sœur en tant que matrice pour reconstituer correctement et fidèlement le brin endommagé [181,183,187,188]

Les cassures double-brin représentent un énorme danger pour les cellules, aussi bien lorsqu'elles ne sont pas réparées, que quand leur réparation n'est pas correcte. Si plusieurs cassures double-brin sont présentes au sein de la cellule, le risque consiste à suturer deux extrémités qui ne correspondent pas au même fragment du génome, ce qui va engendrer des réarrangements chromosomiques. Même si les bonnes extrémités sont collées ensemble, une partie de l'information génétique peut être perdue au cours de la réparation, ce qui est notamment le cas quand les cassures sont prises en charge par le système NHEJ, où aucun brin-matrice n'est présent. Les délétions ou les insertions peuvent ainsi survenir, ce qui peut mener à la tumorigénèse dans le cas où un oncogène ou un gène suppresseur de tumeur serait touché par ces altérations [181].

3. Cyclonucléosides et leurs conséquences pour les cellules

Comme il a été précisé dans le chapitre précédent, les cyclonucléosides font partie des lésions complexes de l'ADN. Nous allons nous concentrer spécifiquement sur eux, car ce sont les dommages qui ont été les plus étudiées au cours de cette thèse.

3.1. Caractéristiques chimiques des cyclonucléosides

3.1.1. Structure et mécanisme de formation

Les cyclonucléosides sont des lésions complexes générées surtout par le rayonnement γ, mais aussi par des agents radiomimétiques tels que la bléomycine. Ils sont formés dans la réaction de la cyclisation intramoléculaire entre le carbone C5' d'un désoxyribose et le carbone C8 d'une purine ou C6 d'une pyrimidine, leur particularité étant telle que le dommage touche à la fois au désoxyribose et à la base. [190] (Figure 15). Cette réaction a été découverte en 1968 par Keck, qui a décrit la cyclisation d'une molécule d'AMP et a suggéré que d'autres nucléotides pourraient suivre la même voie [191]. Ceci a été confirmé en 1976 par Mariaggi *et al.*, qui ont identifié la cyclodésoxyadénosine (cyclo-dA) en tant que principale lésion de l'adénosine générée au cours de la radiolyse [192]. La formation de la cyclodésoxyguanosine (cyclo-dG) n'a été, quant à elle, démontrée que dix ans plus tard par Dizdaroglu [190].

Les cyclonucléosides sont formés en deux étapes : dans un premier temps un atome d'hydrogène est enlevé du carbone C5' du désoxyribose par un radical $^{\bullet}OH$, ce qui donne naissance à un radical localisé sur le sucre, ce nouveau radical peut ensuite attaquer le carbone C8 de la purine pour former une liaison covalente supplémentaire entre le sucre et la base. Cette deuxième étape de la réaction de cyclisation est lente, et son rendement dépend du taux d'oxygénation de la solution. En effet, l'oxygène peut réagir avec le radical formé sur le sucre, pour donner un radical peroxyle, et ainsi empêcher la formation d'un cyclonucléoside [193]. Comme cela a été montré par Belmadoui *et al.*, le taux de cyclopurines formées diminue avec le pourcentage d'oxygène en solution. De plus, en absence d'oxygène, la formation des cyclo-dG est favorisée par rapport aux cyclo-dA avec un facteur de 1.7, mais cette différence diminue avec l'augmentation du taux d'oxygène en solution [194].

En fonction de la configuration du carbone C5' du sucre (orientation des groupes –H et –OH), deux diastéréoisomères de chaque cyclopurine peuvent être formés : R et S. La réaction de

cyclisation des purines est hautement stéréospécifique, et c'est toujours la formation du diastéréoisomère R qui est privilégiée : dans les solutions d'ADN saturées en N_2O, le ratio (5'R)/(5'S) est rapporté aller de 3 :1 à 8 :1 [190,194,195]. De plus, les expériences menées sur les cyclonucléosides isolés montrent que la forme (5'S) de cyclo-dA (5'S-cyclo-dA) pourrait se transformer en forme (5'R) (5'R-cyclo-dA) sous l'effet du rayonnement UVB [196].

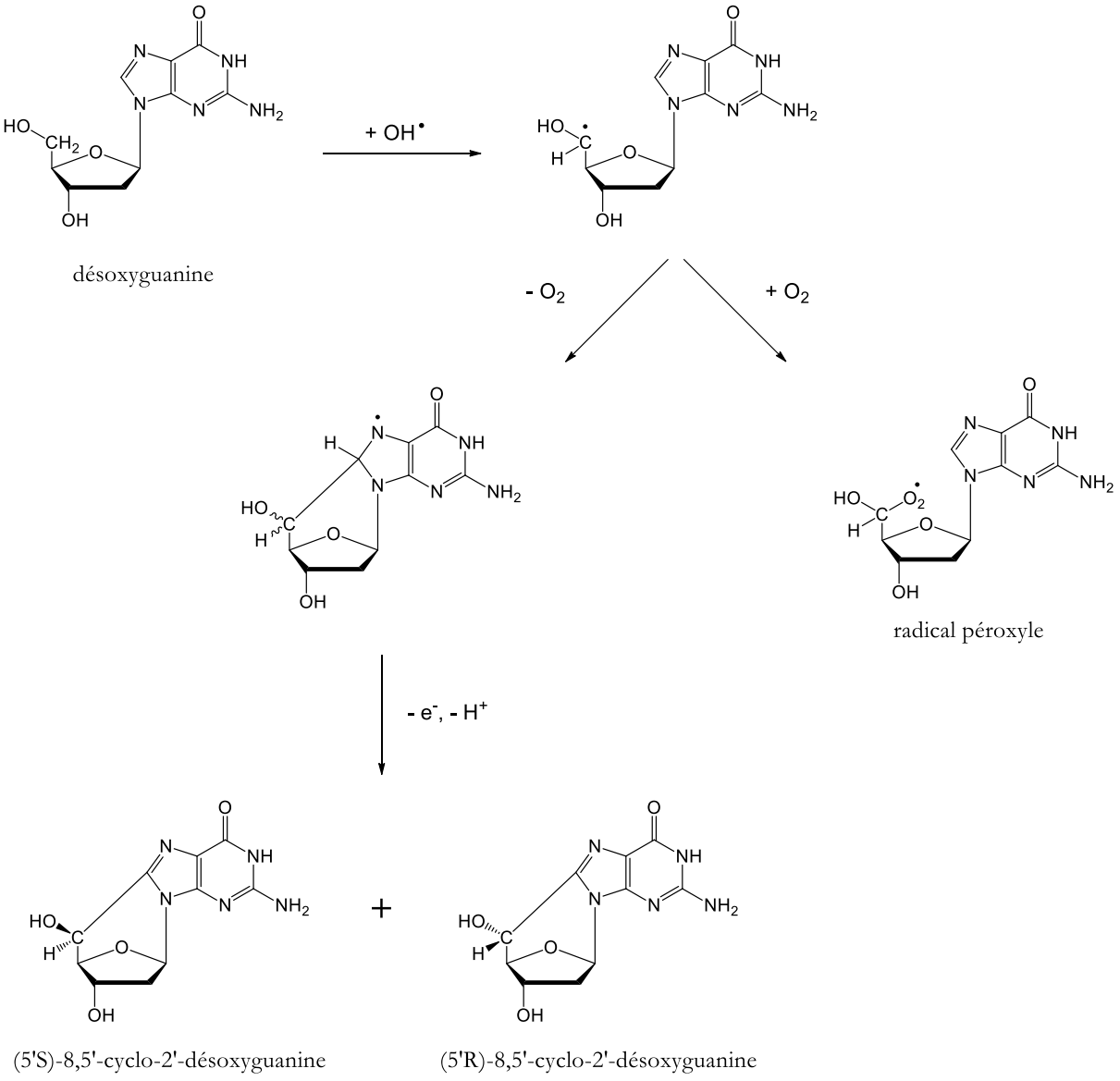


Figure 15 : Formation des cyclonucléosides, exemple de cyclo-dG. Sources : [193,197].

3.1.2. Formation des cyclonucléosides dans les cellules

La formation des cyclonucléosides dans l'ADN cellulaire, contrairement à leur formation dans l'ADN en solution, divise la communauté scientifique, et des valeurs très divergentes peuvent être retrouvées dans la littérature. Les premiers dosages des cyclonucléosides *in vivo* ont été effectués en utilisant la technique de post-marquage au ^{32}P. L'ADN du foie de rat a été digéré et marqué radioactivement, avant d'être séparé par chromatographie. Cette technique permet de détecter 1 à 5 nucléosides modifiés/10^{10} nucléosides normaux, sa sensibilité dépassant ainsi celles des méthodes spectroscopiques utilisées actuellement le plus couramment. Le niveau de cyclo-dA, mesuré avec cette technique, se situe autour de 180-320 lésions/cellule, soit 1 à 2 ordres de magnitudes plus bas que celui de la 8-oxo-dG, la lésion d'oxydation la plus étudiée [198]. Les autres études qui ont suivi sont basées sur la chromatographie liquide couplée à la spectrométrie de masse (LC/MS), combinée à la dilution isotopique. Celles-ci, même si elles font appel au même protocole, donnent des résultats très éloignés, dépendant de l'équipe effectuant les dosages. Le niveau des cyclodésoxyadénosines dans les cellules monocytiques THP1 non-irradiées a ainsi été identifié dans notre laboratoire au-dessous des limites de détection, estimées à 0.1 lésion/10^9 nucléosides normaux [194]. Cette valeur est en désaccord avec les résultats de l'équipe de Dizdaroglu, qui estime le niveau des 5'R-cyclo-dA et 5'S-cyclo-dA à environ 0.01–0.1 lésions/10^6 nucléosides normaux dans les cellules mammifères [199,200]. Quelques exemples de dosages de la 5'S-cyclo-dA dans différents tissus non-irradiés sont présentés dans le Tableau 4.

Tableau 4 : Exemples des niveaux de le 5'S-cyclo-dA mesurée par LC/MS dans différents tissus humains et animaux non-irradiés [200].

type de tissu	**nombre d'échantillons**	**5'S-cyclo-dA /10^7 nucléosides**	**référence**
foie de souris	3	1.92 ± 0.12	[201]
foie de souris	5	0.40 ± 0.14	[202]
foie de souris	10	0.48 ± 0.07	[200]
branchie de poisson	10	0.70 ± 0.19	[203]
sein humain	6	0.44 ± 0.19	[204]
sang humain	10	0.26 ± 0.07	[205]

3.1.3. Effet sur la structure de l'ADN

La présence d'un cyclonucléoside dans l'ADN va avoir des conséquences importantes pour sa structure. La longueur de la liaison covalente supplémentaire C8-C5' est similaire à celle d'une liaison covalente classique, ce sont par contre les autres liaisons du sucre qui deviennent plus longues, ce qui fait qu'il adopte une conformation inhabituelle. La structure de la 5'S-cyclo-dG appariée à une cytosine a récemment été résolue par la spectroscopie NMR. Il a été démontré que la liaison formée entre le carbone C8 de la guanine et le carbone C5' du sucre enfermait le nucléotide dans la conformation *anti* qui l'empêche de se retourner dans le duplexe. En parallèle, le sucre adopte une nouvelle conformation O4'-exo – l'atome d'oxygène O4' sort de sa configuration plate et se rapproche ainsi des atomes de la base – qui modifie les angles de torsions (Tableau 5, Figure 16) et déstabilise la structure de l'hélice. Elle est également défavorable d'un point de vue énergétique, puisqu'une baisse de température de fusion de 9°C a été constatée dans le duplexe modifié par rapport au duplexe non-modifié. Ceci pourrait avoir une influence sur la reconnaissance de la cyclo-dG par certaines enzymes de réparation. La cyclodésoxyguanosine participe toujours dans la formation des liaisons Watson-Crick, celles-ci semblent néanmoins affaiblies [197,206,207]

Tableau 5 : Modification des angles de torsion dans le sucre suite à la formation d'une 5'S-cyclo-dG [207].

angle de torsion	atomes impliqués	angle dans dG	angle dans 5'S-cyclo-dG
β	P – O5' – C5' – C4'	180°	- 87°
γ	O5' – C5' – C4' – C3'	50°	-67°
δ	C5' – C4' – C3' – O3'	120°	149°
ζ	C3' – O3' – P – O5'	-90°	-59°

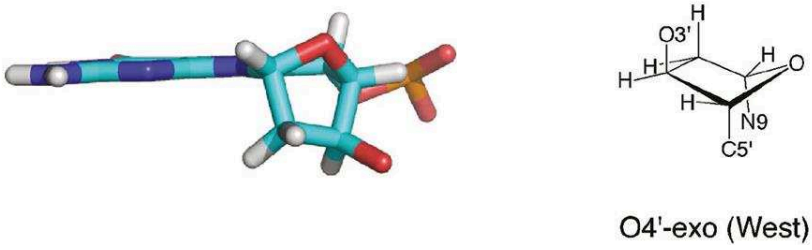


Figure 16 : Configuration adoptée par le sucre dans la 5'S-cyclo-dG [207].

3.2. Conséquences biologiques

Depuis l'année 2000, les articles se sont multipliés rapportant l'influence des cyclonucléosides sur de nombreuses fonctions cellulaires. Ces études effectuées à la fois *in vitro* et *in vivo*, mettent en avant le caractère hautement cytotoxique de ces lésions difficiles à réparer et inhibant l'action de nombreuses enzymes intervenant dans la réplication et la transcription de l'ADN.

3.2.1. Influence sur la synthèse de l'ADN

Une des premières études portant sur l'influence des cyclonucléosides sur la synthèse de l'ADN a été effectuée par l'équipe de Kuraoka. Ils ont travaillé avec deux cyclonucléosides : (5'S) et (5'R)-8,5'-cyclo-2'-désoxyadénosine (5'S/5'R-cyclo-dA) et ils ont dans un premier temps effectué un test d'extension d'amorces (primer extension assay), afin de déterminer si ces lésions constituent un obstacle pour les ADN polymérases Pol δ et T7. Ils ont alors constaté que les deux stéréoisomères de cyclo-dA inhibaient complètement l'activité de la polymérase Pol δ, qui arrête la synthèse de l'ADN juste avant la lésion. En ce qui concerne la T7 ADN polymérase, elle est capable d'insérer une base en face d'une cyclo-dA, et elle exprime également une faible activité dite « read-through », qui lui permet de continuer la synthèse de l'ADN – celle-ci est plus prononcée pour l'isomère R que le S. Ils ont démontré ainsi que les cyclo-dA étaient capables de bloquer la synthèse de l'ADN *in vitro* [208].

Dans une étude ultérieure, cette même équipe a démontré également l'inhibition par les cyclopurines de la synthèse de l'ADN par pol η : cette enzyme est capable d'insérer un nucléotide en face de la 5'S-cyclo-dA sans continuer la synthèse, tandis que cette synthèse peut continuer malgré la 5'R-cyclo-dA à forte concentration de pol η [209].

Des études plus approfondies portant sur la pol η humaine et également sur son homologue chez la levure, ont été menées par Swanson *et al.* L'objectif a été d'explorer la fiabilité et l'efficacité de ces enzymes, en ce qui concerne l'insertion d'une base en face de la lésion, et la poursuite de la synthèse. Pour cela, les auteurs ont effectué les tests d'extension d'amorces en utilisant les sondes contenant une 5'S-cyclo-dA ou une 5'R-cyclo-dG. Conformément aux résultats de Kuraoka, la polymérase humaine pol η est capable d'insérer seulement un ou deux nucléotides après la lésion, avant de se dissocier de l'ADN et de laisser la place à une autre polymérase qui peut continuer la synthèse. Ainsi, elle est moins processive quand la matrice

contient un cyclonucléoside par rapport au contrôle non-endommagé, ses efficacité et fidélité restent néanmoins du même niveau que la séquence soit lésée ou non. Il est intéressant de noter que l'homologue de la pol η chez la levure est moins fidèle quant à l'insertion des nucléotides en face des lésions, mais elle est pourtant capable de continuer la synthèse de l'ADN, sans que sa processivité soit diminuée par la présence de la lésion dans la matrice [210].

3.2.2. Influence sur la transcription de l'ADN

Pour étudier l'impact des cyclonucléosides sur la transcription de l'ADN, les chercheurs de l'équipe de Brooks ont conçu un système leur permettant d'analyser l'influence d'une lésion définie sur certaines fonctions de la cellule. Ils ont synthétisé des séquences d'ADN contenant une cyclo-dA dans un endroit spécifié du brin transcrit du promoteur du gène de la luciférase, et ils ont inséré cette séquence dans un plasmide non-réplicatif transfecté ensuite dans la cellule. La présence d'un site de coupure permet d'éliminer les plasmides dépourvus de la lésion, ce qui augmente la précision et la fiabilité des résultats obtenus. Ainsi, la présence d'une cyclo-dA dans le brin transcrit réduit fortement la transcription du gène dont le promoteur a été endommagé [211].

Dans une étude ultérieure, Marietta et Brooks ont également montré que l'ARN polymérase Pol II était capable d'effectuer la synthèse de l'ARN malgré la présence d'une cyclopurine, cette synthèse n'étant cependant pas sans erreur [213]. Ainsi, deux types de mutations au niveau des transcrits ont été détectés : l'insertion d'une rA en face de la dA directement après la cyclo-dA (en face de laquelle un uracile a été correctement inséré) – appelée mutation 5'A, ou bien une délétion de 7, 13 ou 21 nucléotides. Cette délétion serait due, selon l'hypothèse des auteurs, à un mécanisme de « froissement » de la matrice de l'ADN simple-brin, présenté pour la première fois dans une étude portant sur l'ARN polymérase d'*E. coli* [212]. Après l'insertion d'une base en face de la lésion, l'ARN polymérase pourrait ainsi « sauter » jusqu'à une prochaine dA dans la séquence, afin de recommencer la synthèse à cet endroit-là [213]. La mutation 5'A a également été constatée suite à une synthèse translésionnelle à travers une 5'S-cyclo-dG – cette modification pourrait ainsi être envisagée comme une signature des cyclopurines lors de la transcription de l'ADN lésé [214]. La présence des mutations 5'A a été confirmée pour la 5'S-cyclo-dA et démontré également pour la 5'S-cyclo-dG par You et *al.*. L'inhibition presque complète de l'activité de l'ARN polymérase II suite à la présence d'une 5'S-cyclo-dA ou 5'S-cyclo-dG a été constatée, aussi bien que l'inhibition partielle de l'ARN polymérase T7, qui était capable

de continuer la synthèse à travers la 5'S-cyclo-dA dans 28% et la 5'S-cyclo-dG dans 14% des cas. La synthèse effectuée par cette enzyme engendre néanmoins des erreurs, et la délétion d'un nucléotide après la lésion a notamment été détectée [214].

Tous ces résultats montrent clairement que la présence d'une cyclopurine dans le brin transcrit de l'ADN n'est pas anodine pour la cellule, puisqu'elle peut être à l'origine d'arrêts de transcription ou de mutations au niveau de l'ARN transcrit, dans le cas où sa synthèse n'est pas complètement abandonnée.

3.2.3. Influence sur l'activité d'autres enzymes

Les chercheurs de l'équipe de Kuraoka ont également démontré que les cyclonucléosides étaient capables de bloquer l'activité 3'→5' exonucléolytique de la T4 polymérase : une 5'S-cyclo-dA inhibe complètement la digestion à un nucléotide de la lésion, tandis que l'isoforme 5'R exerce un blocage plus faible, qui peut être surpassé par une concentration élevée de l'enzyme [209]. Ceci confirme les conclusions obtenues précédemment avec la T7 DNA polymérase, qui indiquent que le stéréoisomère 5'S est plus toxique pour les cellules que le 5'R. A l'inverse, la même expérience effectuée avec la DNase III montre l'inhibition totale de son activité à la fois par la 5'S que la 5'R-cyclo-dA. Il existe cependant des différences dans cette inhibition en fonction du stéréoisomère présent dans la séquence : la forme 5'S bloque l'activité exonucléolytique un nucléotide avant la lésion, tandis que cette digestion continue jusqu'à la lésion même dans le cas de la forme 5'R.

3.3. Réparation

Les cyclonucléosides sont classés parmi les lésions d'oxydation et c'est pourquoi le premier système de réparation qui a été envisagé pour leur prise en charge a été le BER. Néanmoins, jusqu'à maintenant, aucune glycosylase n'a été trouvée active sur ces lésions.

La structure des cyclonucléosides leur confère des caractéristiques particulières, ce qui est dû à la présence d'une liaison covalente supplémentaire entre la base et le désoxyribose. Ainsi, Theruvathu *et al.* ont analysé l'influence de cette liaison sur la résistance d'une 5'S-cyclo-dA à l'hydrolyse acide effectuée avec de l'acide formique. Ce traitement cause une dépurination en hydrolysant la liaison glycosydique. Les efficacités de la dépurination d'une désoxyadénosine et

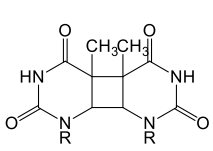
d'une 5'S-cyclo-dA ont été comparées : les résultats obtenus par ces chercheurs indiquent que la présence d'une liaison entre le carbone C8 de l'adénine et C5' du sucre rend la liaison glycosydique environ 40 fois plus stable à l'hydrolyse acide par rapport à la même liaison présente dans une adénine. L'explication possible de ce phénomène serait la formation d'un cycle à 6 atomes suite à l'apparition de la nouvelle liaison, ce cycle étant capable de renforcer la stabilité de la molécule entière [215].

L'équipe de Kuraoka a étudié la réparation des cyclopurines par le système BER. Leurs travaux portaient à vérifier si les cyclonucléosides pourraient être excisés d'une séquence de l'ADN à l'aide d'une glycosylase. Afin d'analyser cela, ils ont incubé des oligonucléotides contenant dans leurs séquences une 5'S ou 5'R-cyclo-dA avec des extraits protéiques de cellules HeLa. L'analyse sur gel des fragments obtenus n'a démontré aucune activité glycosylase, quelles que soient les conditions expérimentales, ce qui confirme que les cyclonucléosides ne peuvent pas être pris en charge par le BER [208]. Ceci pourrait être dû au renforcement de la stabilité de la liaison glycosydique par la présence d'une liaison covalente supplémentaire, comme décrit dans le paragraphe précédent.

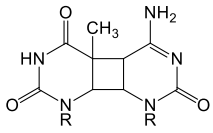
Au vu de l'incapacité du système BER à prendre en charge les cyclonucléosides, un intérêt grandissant se porte sur le système de réparation par excision des nucléotides (Nucleotide Excision Repair ; NER).

3.3.1. Prise en charge par le système de réparation NER

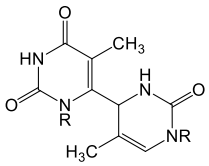
Le système NER est le plus souvent associé avec la réparation des lésions volumineuses. Ces dommages provenant de différentes sources ont des structures très variés, mais ils partagent une caractéristique commune, qui est de générer une distorsion importante de l'hélice. Quelques exemples de ces lésions sont montrés sur la Figure 17. On y retrouve les dimères et des pontages entre deux pyrimidines, générés par le rayonnement UV, des pontages formés par les dérivés du platine, ou encore des lésions dues à la fixation de métabolites de benzo[a]pyrène sur l'ADN. Des agents comme le gaz moutarde, le psoralène ou les aflatoxines peuvent également donner naissance à des lésions appartenant à ce vaste groupe [216–219].



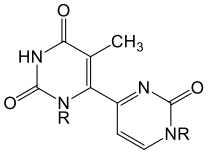
dimère de cyclobutane
(thymine - thymine)



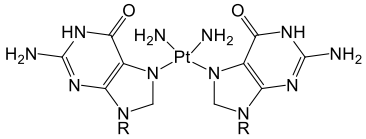
dimère de cyclobutane
(thymine - cytosine)



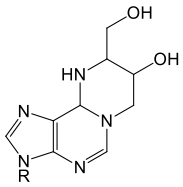
pontage 6-4
(thymine - thymine)



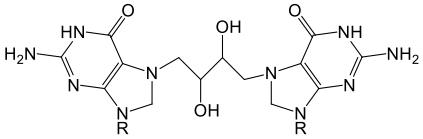
pontage 6-4
(thymine - cytosine)



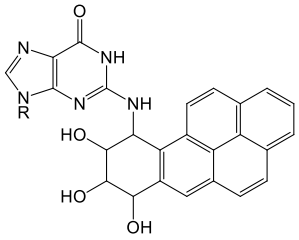
pontage 1,2-d(GpG)



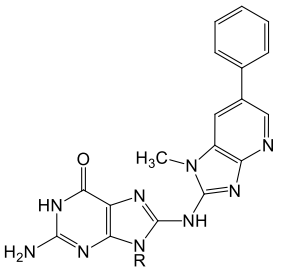
1,N6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxyadenosine (1, N6-aHMHP-dA)



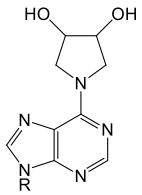
1,4-bis-(guan-7-yl)-2,3-butanediol
(bis-N7G-BD)



N2-BaP-dG



2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP)
dG-C8-PhIP



N6,N6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine
(N6, N6-DHB-dA)

Figure 17 : Exemples d'adduits volumineux étant les substrats du NER.

Les lésions réparées par le NER peuvent être localisées sur le brin transcrit de l'ADN, ou bien dans sa partie non transcrite. Suivant si elles bloquent la transcription ou pas, leur prise en charge ne va pas se dérouler de la même manière. Ainsi, il existe deux voies d'initiation de la réparation de ces dommages : global genome NER (GG-NER) et transcription-coupled NER (TC-NER) (Figure 18)

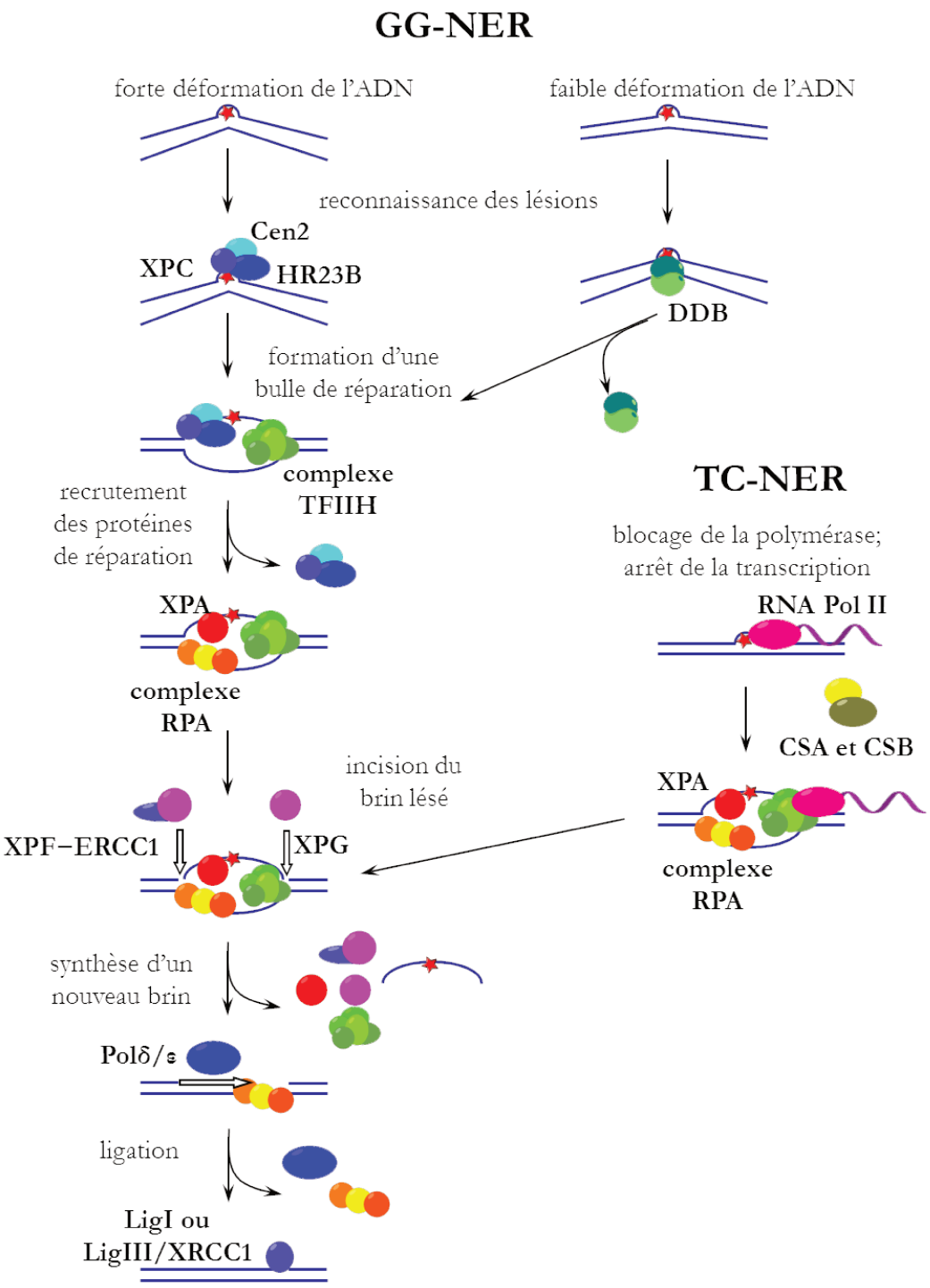


Figure 18 : Représentation schématique de la réparation de l'ADN par le système NER. Inspiré de [217].

Global genome NER (GG-NER)

GG-NER est une voie du NER qui sert à détecter et réparer les lésions volumineuses, quelle que soit leur localisation dans l'ADN. La reconnaissance initiale de la lésion est effectuée par le complexe XPC/hHR23 qui scanne la séquence d'ADN en cherchant ses distorsions. Ainsi, les lésions qui déforment fortement l'hélice, tels que les 6-4 photoproduits (6-4 PPs), vont être mieux reconnues et réparées plus efficacement que celles dont l'effet sur la structure de l'hélice est moins prononcé, comme par exemple les dimères de cyclobutane (cyclobutane pyrimidine dimers ; CPDs). Des deux partenaires, c'est la protéine XPC qui possède une affinité pour l'ADN, avec une préférence pour l'ADN endommagé, et cette affinité va être renforcée par son ubiquitinylation réversible [217,220,221]. Le rôle de XPC/hHR23 ne se limite pourtant pas à la seule reconnaissance du dommage, mais ce complexe va également déformer encore plus l'ADN autour de la lésion et recruter le complexe TFIIH, nécessaire pour la suite du processus de réparation [220,222]. Un troisième partenaire est souvent associé à XPC/hHR23 – la centrine-2, dont la présence n'est pas indispensable, mais qui stabilise le complexe [217,223].

Comme il a été indiqué, l'efficacité de la reconnaissance de la lésion par XPC dépend du degré de distorsion de l'hélice. Les lésions provoquant une faible déformation ne vont pas être détectées par le complexe XPC, et c'est pour cela qu'un autre complexe est nécessaire pour leur prise en charge – DDB. Ce complexe est composé de deux protéines : DDB1 et DDB2 (XPE), qui reconnaissent des lésions telles que les CPDs, se fixent à elles et déforment l'ADN, afin de faciliter le recrutement du complexe XPC [217,224,225].

Le complexe TFIIH qui est recruté par XPC, est composé de 10 protéines nécessaires à l'ouverture de la bulle formée auparavant autour de la lésion. Ainsi, les deux hélicases XPB et XPD déroulent l'ADN sur une distance d'approximativement 30 paires de bases autour de la lésion, tandis que les trois membres du sous-complexe CAK (cdk-activating kinase) : Mat1, Cdk7 et Cycline H phosphorylent le domaine carboxy-terminal de la grande sous-unité de la polymérase RNAPII, ce qui lui permet de continuer son activité d'élongation [217,226]. Les autres composants du complexe TFIIH : p34, p44, p62 et p52 forment avec XPB et XPD une structure circulaire, mais leur rôle précis dans le NER n'a pas encore été élucidé [227].

La protéine suivante qui participe à la prise en charge d'une lésion volumineuse est XPA. Elle a été initialement considérée comme celle qui détecte la lésion, mais il a été prouvé que son recrutement est précédé par le complexe TFIIH, et qu'elle déplace XPC, déjà présent sur le dommage [228,229]. D'autres rôles sont alors envisagés pour cette protéine indispensable au NER, tels que la vérification de la présence d'une lésion, l'aide au positionnement correct des

protéines de réparation, ou encore l'identification du brin sur lequel est situé le dommage [220,222]. XPA interagit avec de nombreuses protéines, comme ERCC1, XPF, RPA32, RPA70 et TFIIH, ce qui renforce son importance dans le processus du NER [222].

L'ouverture totale de la bulle formée autour du dommage nécessite une autre protéine – RPA. Ce complexe de trois sous-unités se fixe initialement sur un fragment d'environ 10 nucléotides, avant d'étendre son interaction sur une région de 30 nucléotides [230–233]. Ceci correspond à la taille de l'intermédiaire de réparation complètement ouvert, qui est stabilisé par la présence de RPA. Cette enzyme joue également un deuxième rôle dans la réparation, qui consiste à recruter des nucléases sur le site de la lésion : localisée sur le brin non-endommagé, RPA fixe ERCC1-XPF par son extrémité orientée vers 3', et XPG du côté 5'. Elle stimule aussi l'incision du brin endommagé par ERCC1-XPF, en améliorant ainsi sa spécificité [234].

Les deux endonucléases citées ci-dessus : XPG et ERCC1-XPF, sont structure-spécifiques et coupent le brin endommagé des deux côtés de la bulle, en reconnaissant l'endroit où l'ADN double-brin devient simple-brin. XPG fait l'incision du côté 3' de la lésion, tandis que ERCC1-XPF – du côté 5', et ainsi le fragment d'environ 25-30 nucléotides contenant la lésions peut être éliminé pour être ensuite remplacé par un brin nouvellement synthétisé [217,222,235,236].

La dernière étape de la réparation par le système NER consiste en une resynthèse du fragment qui a été excisé. Ceci est effectué par une des deux polymérases : delta (δ) ou epsilon (ε), qui sont associées au facteur de processivité PCNA. La synthèse est ensuite terminée par la suture des extrémités, faite essentiellement par la ligase III associée à son partenaire XRCC1 [217,237–239].

Transcription-coupled NER (TC-NER)

Dans le TC-NER la reconnaissance des lésions ne se fait pas à l'aide d'une enzyme « senseur », comme dans le cas du GG-NER, mais elle est déclenchée par le blocage de la polymérase RNA Pol II au niveau du dommage volumineux. Ainsi, cette voie de réparation est spécifique des lésions localisées sur le brin transcrit et qui bloquent la transcription. L'arrêt de la polymérase déclenche le recrutement des protéines CSA et CSB, et, par la suite, du complexe TFIIH [240]. A partir de ce moment, les deux voies du NER : GG-NER et TC-NER se déroulent de la même manière.

Les principales protéines du NER sont présentées dans le Tableau 6.

Tableau 6 : Principales protéines du système NER. Source : [222]

Facteur de NER		**Sous-unités**	
Nom	**fonction dans le NER**	**nom**	**remarques**
XPC-hHR23B	détection des lésions ; recrutement des facteurs de réparation ; (spécifique pour GG-NER, non-impliqué dans TC-NER)	XPC	affinité pour l'ADN endommagé
		hHR23A/B	*in vitro* : stimulation de l'activité de XPC
TFIIH	catalyse de la formation d'un complexe ouvert autour de la lésion ; facilitation de l'assemblage du complexe de réparation	XPB	hélicase 3'→5'
		XPD	hélicase 5'→3'
		p34	
		p44	
		p62	
		p52	
		Mat1	sous-complexe CAK
		Cdk7	
		Cycline H	
XPA	facilitation de l'assemblage du complexe de réparation		affinité pour l'ADN simple-brin et l'ADN endommagé
RPA	stabilisation du complexe ouvert ; positionnement des nucléases	RPA70	fixation de l'ADN simple-brin
		RPA32	
		RPA14	
XPG	l'incision en 3' ; stabilisation du complexe ouvert		membre de la famille des nucléases FEN-1
ERCC1-XPF	l'incision en 5'	ERCC1	endonucléase structure-spécifique
		XPF	

3.3.2. Conséquences de l'insuffisance du NER

Le système NER est extrêmement polyvalent et il prend en charge une multitude de lésions capables de déformer la structure de l'hélice d'ADN, indépendamment de leur source et de leur structure. C'est pour cela que la déficience du système NER peut avoir de graves conséquences : une forte prédisposition aux cancers, une photosensibilité accrue, ou encore des troubles neurodégénératifs. En fonction de la(des) protéine(s) déficiente(s), plusieurs maladies,

souvent autosomales et récessives, ont été identifiées, dont les plus étudiées sont *Xeroderma pigmentosum* et le Syndrome de Cockayne [240].

Xeroderma pigmentosum

Xeroderma pigmentosum (XP) est une maladie autosomale récessive, dont la principale caractéristique est une sensibilité extrême à la lumière solaire, présente dans 60% des cas. Ces malades développent des graves brûlures sur les parties du corps exposées au soleil, et ceci souvent dans les premières semaines de leur vie. Chez les 40% restant, les premiers symptômes se manifestent plus tard, vers l'âge de deux ans, avec l'apparition de nombreuses tâches de rougeur, surtout au niveau du visage. La gravité de ces altérations dépend beaucoup du type de la peau et de sa pigmentation, ainsi que du degré de l'exposition aux rayons UV et de la protection appliquée. La photosensibilité est souvent accompagnée de photophobie et de conjonctivites chroniques [241].

Les patients XP développent de nombreux cancers de la peau et des yeux, suite à la non-réparation des photoproduits dans l'ADN. Ces cancers apparaissent précocement, avant l'âge de dix ans, et sont 1000 fois plus fréquents chez ces patients que chez des individus sains. Il a aussi été estimé que le risque de développer un cancer cutané chez ces malades avant l'âge de 20 ans serait 10000 fois plus élevé pour un cancer non-mélanome, et 2000 fois plus élevé pour un mélanome [241]. Les symptômes cutanés sont dans 20 à 30 % des cas accompagnés des troubles neurologiques et développementaux – la neurodégénérescence progressive conduit à la surdité, l'incapacité à se nourrir et à se déplacer [241,242].

Syndrome de Cockayne

Le syndrome de Cockayne (CS) est une autre maladie autosomale récessive. Les enfants malades sont normaux à la naissance, mais leur développement post-natal est considérablement ralenti. Les symptômes du CS sont caractéristiques et distincts de ceux de *Xeroderma pigmentosum* : petite taille, retard physique et mental, microcéphalie, membres disproportionnellement longs, visage allongé avec des os prononcés et les yeux et les joues creux (visage ressemblant à un oiseau), ainsi que des troubles psychomoteurs et une surdité progressive. L'espérance de vie est raccourcie et peu de patients atteignent l'âge adulte. Les malades sont aussi photosensibles, mais,

contrairement aux patients XP, ils n'ont pas de brûlures ni de décoloration suite à l'exposition au soleil, et le risque plus élevé de cancer de la peau n'a pas été observé chez eux. L'absence du risque accru de cancer de la peau peut ainsi être expliquée par le fait que la voie de réparation GGR n'est pas touchée et qu'elle est suffisante pour éliminer les photoproduits de l'ADN des cellules cutanées [242–246]. Il arrive également que ces syndromes soient présents avec les syndromes de XP, dans le cas où plusieurs gènes seraient mutés en même temps [243,247].

3.3.3. NER et cyclonucléosides : tests de réparation *in vitro* et dans des modèles cellulaires

Comme ceci a été décrit dans le chapitre 3.3, les cyclonucléosides ne sont pas prises en charge par le système BER. Ainsi, les chercheurs de l'équipe de Kuraoka ont étudié la capacité du NER à réparer ce type de lésions. Des plasmides contenant une cyclopurine ont été incubés avec des extraits HeLa et des dNTP radioactivement marqués, pour pouvoir ensuite détecter le fragment de l'ADN provenant de la resynthèse du NER. Leurs résultats montrent que les cyclopurines sont effectivement des substrats pour le NER, mais aussi que leur réparation est beaucoup moins efficace que celle des lésions liées aux dérivés du platine – les substrats connus du NER utilisés en tant que contrôle. La conformation de la lésion influence également l'efficacité de sa réparation : le stéréoisomère 5'R de la cyclo-dA est réparé plus efficacement que le 5'S. Ces même chercheurs ont aussi démontré que la réparation des cyclopurines dépendait de la protéine RPA, ce qui renforce encore l'hypothèse de la prise en charge des cyclonucléosides par le NER [208].

Des résultats similaires ont été obtenus en parallèle par l'équipe de Brooks. Ils ont dans un premier temps incubé de courts oligonucléotides contenant une cyclo-dA avec des extraits de cerveau de rat afin de vérifier si cette lésion était un substrat pour le BER. Aucune coupure de l'ADN n'a été détectée, ce qui confirme que cette famille de lésions ne peut pas être prise en charge par le même système que la majorité des lésions simples d'oxydation. En parallèle, les oligonucléotides de taille plus importante ont été utilisés pour tester la réparation des cyclo-dA par le NER. Ainsi, l'excision d'un fragment de 24-30 nucléotides a été démontrée, ce qui correspond à l'excision effectuée lors de la réparation des lésions volumineuses par le NER. Ce fragment n'a pas pu être détecté après l'incubation des oligonucléotides avec des extraits déficients en l'une des protéines indispensables pour le NER : ERCC1 ou XPG, ce qui renforce l'hypothèse de l'implication de ce système dans la réparation des cyclopurines [211]. L'analyse

plus détaillée a montré que l'excision avait lieu à environ 5 nucléotides du côté 3' et à 21 nucléotides du côté 5' de la cyclo-dA – ce motif asymétrique étant déjà démontré pour d'autres substrats du NER [211,248].

Pour valider ces résultats, les mêmes chercheurs ont ensuite procédé à un test appelé « host-cell reactivation » dont l'objectif était d'étudier la réparation d'une lésion contenue dans un plasmide, qui a été transfecté dans les cellules. Ce plasmide contient un gène rapporteur de luciférase, dont l'expression est inhibée par la présence de la lésion. La réparation de celle-ci se manifeste par la restitution de l'activité de la luciférase, qui peut être facilement observée et quantifiée. Ainsi, avec cette technique et en utilisant les lignées cellulaires compétentes ou déficientes en NER, Brooks et *al.* ont confirmé que la cyclo-dA était prise en charge par le NER [211].

En 2007, Errico et *al.* ont testé la sensibilité des fibroblastes provenant des patients CSA. Une plus grande sensibilité envers le traitement par les agents oxydants a été constatée dans ces cellules par rapport aux fibroblastes provenant des donneurs sains. Néanmoins, ce même type d'expérience effectué sur les kératinocytes n'a montré qu'une très faible hypersensibilité aux agents oxydants, ce qui montre l'importance du modèle utilisé. Même si les kératinocytes CSA ne semblent pas avoir une sensibilité accrue au traitement oxydant, la réparation de 5'S-cyclo-dA dans ces cellules exposées à 5 Gy issus de rayons X est moins efficace que dans les kératinocytes provenant des donneurs sains [247].

3.3.4. NER et cyclonucléosides : tests de réparation sur les modèles animaux

Les résultats obtenus par de nombreux chercheurs utilisant les modèles cellulaires, ont également été confirmés sur des modèles animaux.

Un article de Kirkali et *al.* paru en 2009 démontre l'implication de la protéine CSB dans la réparation des cyclonucléosides, en utilisant pour la première fois un modèle animal. L'accumulation de 5'S-cyclo-dA dans trois organes : les reins, le foie et le cerveau de souris $csb^{-/-}$ a été mesurée et comparée aux résultats obtenus chez les souris wild-type, sans que ces animaux soient exposés au stress oxydant exogène. Il ressort de cette étude qu'en absence de la protéine CSB le niveau de la 5'S-cyclo-dA est beaucoup plus important dans les trois organes analysés, cette différence étant la plus remarquable dans le cerveau [202].

Une étude similaire a été menée également sur les souris $neil1^{-/-}$. Jaruga et *al.* ont démontré l'accumulation des deux stéréoisomères de cyclo-dA dans le foie des animaux déficients

en NEIL1, non-exposés à un stress oxydant supplémentaire. NEIL1 est surtout connu pour son rôle de glycosylase dans le système BER, mais puisque les cyclonucléosides ne peuvent pas être réparés par ce système, ces résultats suggèrent une autre activité de cette enzyme. Une preuve récente de coopération entre CSB et NEIL1 dans la réparation des formamidopirymidines [249] pourrait ainsi aider à expliquer le lien entre NEIL1 et accumulation des cyclo-dA. Aussi, on peut corréler les résultats provenant des deux études citées, en suggérant l'implication de NEIL1 dans le système NER, notamment par la reconnaissance de la lésion et la stimulation de sa réparation par CSB [250].

Néanmoins, il manque toujours des preuves directes de la prise en charge des cyclonucléosides par le NER *in vivo*. De plus, les études sur l'accumulation des lésions restent discutables à cause des désaccords parmi les chercheurs concernant la méthode de la quantification des lésions.

3.4. Implication dans les pathologies humaines

Depuis que les travaux sur les cyclonucléosides se sont accélérés dans les années 2000, de nombreuses hypothèses concernant le rôle des cyclonucléosides dans l'apparition et le développement de certaines maladies ont commencé à voir le jour. Vues leurs caractéristiques particulières qui les différencient des autres lésions d'oxydation, il a notamment été suggéré que ces dommages pourraient être impliqués dans la neurodégénérescence chez les malades XP et CS, et qu'ils seraient également les biomarqueurs de certains types de cancers, tels que le cancer de sein. Ces résultats doivent néanmoins être interprétés avec beaucoup de précautions, car ils sont basés sur le dosage des lésions au sein de l'ADN et qui reste discutable, comme ceci a été présenté dans les paragraphes précédents.

3.4.1. Maladies neurodégénératives (XP , CS)

Xeroderma Pigmentosum et le syndrome de Cockayne sont deux maladies génétiques rares liées au dysfonctionnement du NER, dont les caractéristiques ont été décrites dans le paragraphe 3.3.2. Dans les deux cas (XP et CS), l'hypersensibilité à la lumière et le manque de réparation des photoproduits ne peuvent pas expliquer le phénomène de la neurodégénérescence chez les malades, puisque les photoproduits ne peuvent pas être formés dans les neurones. Ainsi, il a été

suggéré déjà à la fin des années '70 que d'autres types de lésions devaient être responsables de l'apparition de ce phénotype [251,252].

Comme cela a été résumé par Brooks, pour pouvoir expliquer le phénomène de neurodégénérescence chez les patients XP, la lésion-candidat doit correspondre aux 4 critères :

- être réparée par le NER, qui est défectif chez les malades XP,
- bloquer la transcription par l'ARN polymérase II,
- provenir d'évènements endogènes,
- être stable dans le temps et s'accumuler dans les cellules [193,253].

Les cyclonucléosides semblent être les seules lésions qui remplissent toutes ces conditions, c'est pourquoi ils ont été proposés comme candidat idéal pour expliquer les mécanismes liés à ces deux maladies. Dans les articles publiés en parallèle par Brooks et *al.* [211] et Kuraoka et *al.* [208], cette hypothèse a été émise. Elle a commencé à gagner en popularité rapidement, quand de plus en plus de travaux ont mis en évidence le caractère cytotoxique des cyclonucléosides et le fait qu'ils seraient sujets à la réparation par le NER.

L'implication des cyclonucléosides dans le Syndrome de Cockayne est moins décrite que dans le cas de *Xeroderma Pigmentosum*, il existe néanmoins quelques études qui suggèrent leur rôle également dans cette maladie. Une publication parue en 2009 indique l'accumulation des cyclopurines dans les organes des souris $csb^{-/-}$ beaucoup plus importante que chez les souris saines, cette importance étant la plus remarquable au niveau du cerveau [202]. La preuve apportée par D'Errico et *al.* que la protéine CSA participerait également dans la réparation des cyclonucléosides [247], renforce l'hypothèse sur le rôle des cyclopurines dans la neurodégénérescence des patients CS.

Il faut néanmoins prendre en compte les différences qui existent entre ces deux maladies, notamment au niveau du processus menant à la neurodégénérescence. Chez les malades XP, ce sont les neurones qui sont touchés, dont une perte importante a souvent lieu. Au contraire, chez les malades CS, il s'agit des modifications au niveau des cellules gliales qui forment la myéline – substance qui protège les neurones. Ainsi, le phénotype CS provient non pas de la perte des neurones, mais plutôt de leur démyélination [253] – ce mécanisme étant moins direct, il a été moins exploré que celui observé dans le cas de *Xeroderma Pigmentosum*.

3.4.2. Cancers

Les cyclonucléosides, en plus de leur rôle potentiel dans les maladies neurodégénératives, semblent également être impliqués dans le développement de certains types de cancer. Une des premières études a été publiée en 2006, elle portait sur un groupe de 23 femmes non malades qui ont effectué une réduction mammaire, divisées en trois groupes d'âge,. Des cellules de stroma, épithélium et myoépithélium ont ensuite été isolées et le niveau de plusieurs lésions d'oxydation, dont 5'S-cyclo-dA, a été quantifié. Les auteurs ont ainsi trouvé que chez les femmes de la tranche d'âge intermédiaire (32-46 ans) le niveau basal de ces lésions était plus élevé que dans deux autres groupes : 17-30 ans et 50-62 ans. Ceci est en corrélation avec une plus grande incidence de cancer de sein chez les femmes âgées de 30 à 40 ans. Cette étude ne peut pas être considérée comme une étude clinique, vu le faible effectif dans chacun des groupes, mais elle suggère que les cyclopurines pourraient être un biomarqueur du cancer de sein [204].

Une autre étude a également été menée sur des lignées cellulaires provenant des cancers de sein, où les cellules ont été exposées au stress oxydant (H_2O_2), et le niveau des 5'S-cyclo-dA a été comparé à la fois avec les cellules non-exposées, et avec les cellules non-cancéreuses, exposées ou non au peroxyde d'hydrogène. La quantité des lésions dans les cellules cancéreuses stressées s'est avérée plus importante que dans les cellules contrôles, ce qui pourrait suggérer la déficience de certains systèmes de réparation de l'ADN dans les cellules cancéreuses de sein. Ces résultats ont été confirmés par un test de coupure effectué avec les extraits protéiques provenant des mêmes lignées, et qui n'a pas démontré de réparation de 5'S-cyclo-dA [254]. Si on considère le fait que certaines cellules cancéreuses sont capables de générer d'elles-mêmes des quantités élevées de peroxyde d'hydrogène [255], on pourrait supposer que le stress oxydant supplémentaire provenant d'une source exogène ne pourrait plus être géré par ces cellules, dont les systèmes de réparation sont déjà surchargés [254].

Rodriguez et *al.* ont mené une étude sur les lymphocytes prélevés chez les femmes diagnostiquées avec le cancer de sein et porteuses de mutation du gène BRCA1. La protéine BRCA1 est impliquée dans de nombreuses voies de réparation de l'ADN, notamment la recombinaison homologue par son association avec Rad51, mais aussi dans le NER [256–261]. Les cellules déficientes en cette enzyme présentent une sensibilité élevée vis-à-vis des agents oxydants tels que le peroxyde d'hydrogène et le rayonnement ionisant [262,263]. Le dosage des trois cyclopurines : 5'S-cyclo-dA, 5'S-cyclo-dG et 5'R-cyclo-dG dans les lymphocytes BRCA1-mutés, irradiés avec une faible dose de 5 Gy, montrent le niveau plus élevé de ces lésions après une heure de récupération, par rapport aux cellules non-irradiées et aussi par rapport aux cellules

non-porteuses de la mutation dans le gène BRCA1. Ceci confirme encore une fois la prise en charge des cyclopurines par le système NER, dans lequel BRCA1 participe, et montre également l'implication probable de ces lésions dans le cancer de sein [264].

Puisque chaque cancer est différent et que la réparation des lésions de l'ADN peut être dérégulée à la fois positivement ou négativement, dans certaines tumeurs on peut observer un taux de dommages inférieur à celui des tissus sains environnants. Ceci a été constaté dans le cas des 5'S-cyclo-dA mesurées dans les cellules provenant du cancer colorectal, comme rapporté par Kirkali et *al.* en 2011. Il a déjà été démontré auparavant que les mutations observées dans les cancers colorectaux et dans les cancers de sein n'étaient pas les mêmes, ce qui peut expliquer certaines différences notées entre ces deux types de tumeurs [265,266]. Ainsi, le fait que la quantité des cyclopurines dans les cellules du cancer colorectal soit inférieure à celle retrouvée dans les cellules saines, viendrait de la meilleure réparation de celles-ci. Les chercheurs suggèrent ainsi la surexpression dans les cellules du cancer colorectal de certaines protéines du NER, telles que NEIL1 ou CSA et CSB supposées participer à la réparation des cyclonucléosides, qui engendrerait une dérégulation positive de la réparation de l'ADN et procurerait ainsi à ces cellules une meilleure résistance aux traitements anticancéreux [267].

Au vu des nombreuses études mettant en avant l'implication des cyclopurines dans le développement des cancers, il a été suggéré que ces lésions pourraient être utilisées en tant que biomarqueurs pour la détection précoce des tumeurs et également pour le suivi de leur évolution et de l'efficacité des traitements. L'équipe de Dizdaroglu a réussi à doser les 5'S-cyclo-dA et 5'R-cyclo-dA dans les urines et a ainsi conclu que ces lésions étaient, après leur réparation par le NER, hydrolysées et éliminées dans les urines. Ceci en ferait des bons candidats pour biomarqueurs, vu la disponibilité du matériel biologique, et également parce qu'ils ont constaté le même niveau de ces lésions dans les tissus et dans les urines, indiquant leur élimination quantitative [199]. Cependant, comme il a été précisé dans le paragraphe 3.1.2, des discordances existent entre différentes équipes de chercheurs en ce qui concerne le dosage des cyclonucléosides et leur présence dans les cellules, que ce soit à l'état basal ou après l'induction d'un stress. C'est pourquoi les hypothèses concernant l'implication de ces lésions dans différentes pathologies, ainsi que leur utilisation comme biomarqueur, doivent être confirmées et ne peuvent pas, en l'état actuel des connaissances, être cautionnées par la totalité de la communauté des chercheurs.

II Stratégie et objectifs

1. Notion d'interactome et objectif de la thèse

L'interactome d'une lésion de l'ADN va être défini comme l'ensemble des protéines qui interagissent avec une lésion donnée. La nature de ces interactions et le rôle des protéines faisant partie de l'interactome peuvent être très variés. On peut y retrouver notamment : des protéines de réparation, des inhibiteurs de réparation, des protéines intervenant dans la modification de la structure chromatinienne ou encore des facteurs de transcription qui sont détournés de leur fonction habituelle.

L'objectif de cette thèse consiste en l'analyse de l'interactome de lésions simples ou complexes d'oxydation, dans le but de comprendre comment ces dommages sont pris en charge au niveau cellulaire et quelles sont les conséquences de leur présence au niveau de l'ADN. Cet objectif global va être réalisé par deux approches différentes :

- ⇒ **d'une part, le screening d'extraits protéiques va être réalisé afin de retrouver des protéines faisant partie de l'interactome des lésions complexes de l'ADN,**
- ⇒ **d'autre part des interactions des lésions spécifiques (simples ou complexes) avec des protéines isolées vont être étudiées.**

La 8-oxo-désoxyguanosine va être utilisée en tant que représentant principal des lésions simples d'oxydation (d'autres lésions simples vont également être étudiées), tandis que pour analyser les lésions complexes, nous allons prendre comme modèle les cyclonucléosides. Ces derniers ont été choisis, puisque parmi les lésions complexes de l'ADN, ils représentent un groupe homogène, contrairement aux lésions tandem ou en cluster, qui peuvent être composées de différentes lésions simples et qui, en fonction de leur composition, ne vont pas avoir le même comportement. Les cyclonucléosides sont des lésions originales et peu connues par rapport aux autres dommages complexes, peu de données existent quant à leurs interactions avec les protéines.

2. Stratégie globale

Le travail de cette thèse va être divisé en deux parties (Figure 19) :

- la première consiste à effectuer un screening des extraits protéiques provenant des cellules humaines HeLa, dans l'objectif d'identifier l'interactome des cyclonucléosides
- la deuxième concerne le travail sur les protéines isolées d'*Escherichia coli* ou de *Thermococcus gammatolerans*, afin d'étudier leurs interactions avec des lésions d'oxydation.

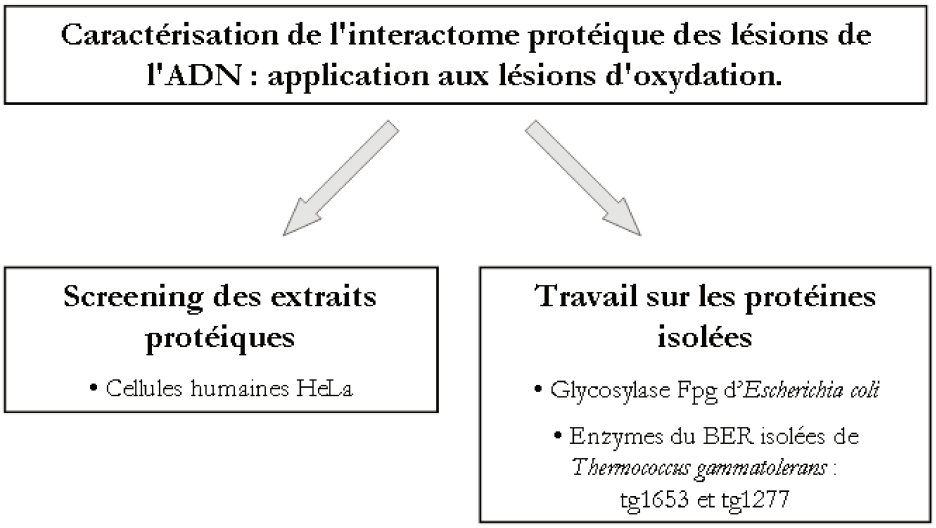


Figure 19 : Stratégie globale des travaux effectués.

2.1. Screening des protéines

Afin de déterminer l'interactome des cyclonucléosides, une stratégie consistant à piéger des protéines au sein d'extraits protéiques va être employée. Le piège utilisé est constitué des billes magnétiques fonctionnalisées avec de la streptavidine, sur lesquelles sont attachées les sondes nucléiques contenant dans leur séquence des lésions d'intérêt (son principe de fonctionnement va être décrit plus précisément dans le chapitre « Matériel et Méthodes »). Quatre cyclonucléosides sont utilisés au cours de cette étude : (5'S)-5'-8-cyclodésoxyadénosine (5'S-cyclo-dA), (5'R)-5'-8-cyclodésoxyadénosine (5'R-cyclo-dA), (5'S)-5'-8-cyclodésoxyguanosine (5'S-cyclo-dG) et (5'S)-5'-6-cyclodésoxydihydrothymidine (5'S-cyclo-dHT). En plus des sondes lésées, deux contrôles sont toujours utilisés en parallèle : les sondes de la même séquence, mais ne contenant pas de lésion, et les billes sans sonde. Le but de ces deux contrôles est de pouvoir comparer les protéines piégées dans différentes conditions et de ne pas prendre en compte celles retenues au niveau de l'ADN non-lésé ou des billes.

Après la première étape, qui consiste à piéger les protéines sur les sondes par la technique du ligand fishing, l'ensemble des protéines attrapées sont identifiées par analyse protéomique. Grâce à cette méthode puissante, toutes les protéines présentes dans un échantillon peuvent être déterminées.

Parmi les diverses protéines identifiées, quelques-unes seront ensuite sélectionnées comme intéressantes et pertinentes d'un point de vue biologique, et leurs interactions avec l'ADN et ses lésions vont être confirmées et caractérisées avec des techniques complémentaires. Pour cela, des techniques telles que le Western blot ou le gel shift (Electrophoretic Mobility Shift Assay ; EMSA) vont être utilisées.

Le schéma global de cette stratégie est présenté sur la Figure 20.

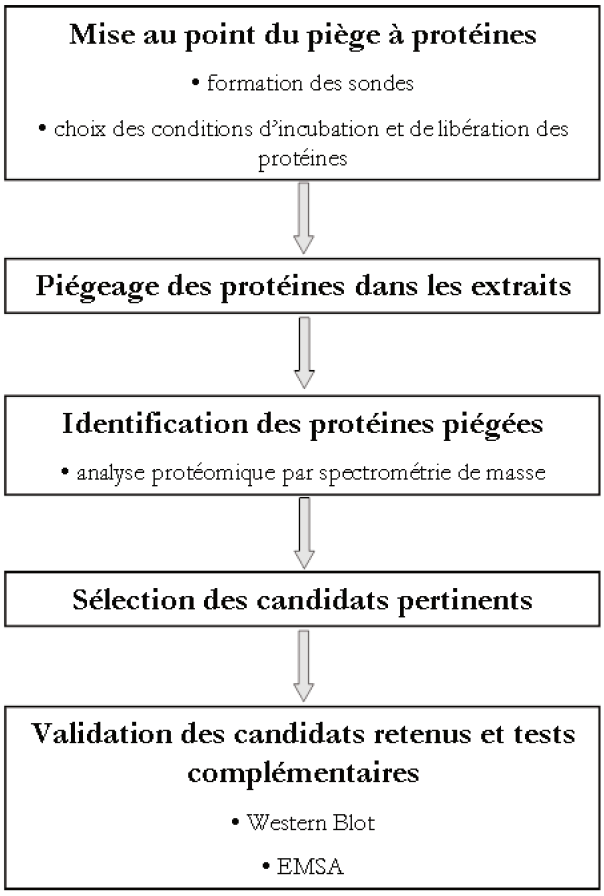


Figure 20 : Schéma général de la détermination de l'interactome des lésions complexes de l'ADN.

La technique du ligand fishing utilisée au sein de notre laboratoire permet d'identifier au moins une partie de l'interactome des lésions étudiées. Elle a déjà été appliquée avec succès par

notre équipe pour définir l'interactome des dommages provoqués par les dérivés du platine ainsi que pour identifier les protéines reconnaissant les modifications épigénétiques de l'ADN.

L'identification de l'interactome des lésions générées par le cisplatine, l'oxaliplatine ou le satraplatine a été réalisée en utilisant des plasmides linéarisés, exposés au préalable à l'agent génotoxique. Les deux dommages principaux générés ainsi sont les pontages entre deux guanines ou entre une adénine et une guanine adjacentes. Cette stratégie a permis d'établir une liste de 38 protéines interagissant avec les lésions platinylées, dont certaines étaient déjà connues pour leurs interactions avec ce type de dommages (HMGB1, hUBF). Trois nouvelles protéines ont été également identifiées comme étant membres de l'interactome des lésions générées par les dérivés du platine : TOX4 (TOX HMG box family member 4), PNUTS (phosphatase 1 nuclear targeting subunit) et WDR82 (WD repeat-containing protein 82) [268,269].

Afin de déterminer les protéines reconnaissant des modifications épigénétiques de l'ADN, un autre type de sondes a été utilisé. Celles-ci ont été synthétisées par la technique de PCR, et des méthylcytosines ou des hydroxyméthylcytosines ont été introduites à la place des cytosines pour obtenir les sondes modifiées. Le piégeage effectué avec les protéines extraites des trois modèles cellulaires : HeLa, SH-SY5Y et UT7-MPL, a permis l'identification d'une protéine qui reconnaît la méthylcytosine, mais pas l'hydroxyméthylcytosine. Cette protéine – ZBTB2 (zinc finger and BTB domain-containing protein 2) – pourrait jouer un rôle dans les processus épigénétiques impliqués dans la cancérogénèse [publication en cours].

2.2. Travail sur les protéines isolées

La deuxième partie des travaux va consister en l'analyse d'interactions entre des protéines isolées et des lésions d'oxydation, à la fois simples, à l'exemple de la 8-oxo-désoxyguanosine, et complexes, comme les cyclonucléosides déjà cités. Trois enzymes particulières vont être utilisées dans ces études : la formamidopyrimidine ADN glycosylase (Fpg) d'*Escherichia coli*, et deux enzymes provenant de *Thermococcus gammatolerans*, qui sont des homologues de séquence de la Fpg et de l'endonucléase III. Ces modèles vont être détaillés dans le paragraphe 3. Afin d'étudier l'activité de ces trois protéines, des tests de coupure seront effectués, ainsi que des tests d'interactions ayant pour but de déterminer leur affinité vis-à-vis des lésions ciblées.

3. Modèles utilisés

Trois types de modèles biologiques ont été utilisés pendant les travaux de cette thèse :

- cellules HeLa – ce modèle eucaryote a été utilisé pour le screening des protéines interagissant avec des lésions complexes d'oxydation, à la recherche de nouvelles interactions ;
- *Escherichia coli* – un modèle procaryote, dont la protéine Fpg a été étudiée pour ses interactions avec des lésions d'oxydation simples ou complexes ;
- *Thermococcus gammatolerans* – ce représentant des archées radiorésistantes a servi de modèle original afin d'étudier les protéines homologues des protéines de réparation connues chez d'autres organismes ;

Les trois modèles cités ci-dessus représentent ainsi les trois domaines : eucaryotes, procaryotes et archées, ce qui permet d'étudier les interactions observées au sein des groupes phylogénétiques très distants.

3.1. Lignée humaine HeLa

Les cellules HeLa ont été les premières lignées cancéreuses humaines à être cultivées dans un laboratoire. Isolées au début des années '50 d'un cancer cervical, elles se sont très vite répandues dans tous les laboratoires de recherche à travers le monde. 60 ans après leur mise en culture, elles restent le modèle de référence pour de nombreux chercheurs – plus de 75000 articles faisant référence à ce modèle cellulaire ont été publiés depuis leur découverte, et de nombreux processus cellulaires ont pu être élucidés grâce aux cellules HeLa. Malgré l'utilisation massive de cette lignée cellulaire dans la recherche biologique, le génome de HeLa n'a été séquencé que très récemment [270]. Aussi, il est possible que les cellules provenant de différents laboratoires ne possèdent plus les mêmes caractéristiques, suite à l'évolution indépendante des lignées au cours des années. En dépit de ces inconvénients évidents, les cellules HeLa restent le modèle le plus couramment utilisé dans le monde.

La disponibilité commerciale des extraits nucléaires de cellules HeLa a été un argument important pour le choix de ce modèle. La mise au point du « piège à protéines » a nécessité une grande quantité des protéines à concentration et qualité reproductibles, et c'est pourquoi l'utilisation de ces extraits a été jugée la plus pertinente. De plus, comme c'est un modèle très couramment utilisé et dont beaucoup de données sont disponibles dans la littérature, ce choix

nous a permis de confronter nos résultats aux informations déjà existantes sur les interactions entre protéines et lésions complexes d'oxydation.

3.2. Enzyme d'*Escherichia coli* : la Fpg

Le premier modèle utilisé au cours de mes expériences à partir d'enzymes purifiées a été une glycosylase d'origine bactérienne, provenant d'*Escherichia coli* – une bactérie gram-négative utilisée couramment dans de nombreuses études biologiques. Cette enzyme est la Fpg, qui est une glycosylase impliquée dans le BER. Elle a été identifiée et décrite pour la première fois en 1979 par l'équipe de Tomas Lindahl [271], qui a découvert son activité de coupure sur la 2-6-diamino-4-hydroxy-5N-methylformamidopyrimidine (Fapy-dG), un dérivé de la 7-methylguanine. La Fapy-dG est un inhibiteur puissant de la synthèse d'ADN, cette inhibition étant directement proportionnelle au nombre de ces dommages dans l'ADN. Elle est alors considérée comme une lésion létale, au même degré que les sites abasiques [272].

Puisque la Fapy-dG n'est pas trouvée en grande quantité dans l'ADN sous conditions physiologiques, et que les mutants défectueux en Fpg ne présentent pas de sensibilité accrue au traitement méthylant générant ce type de dommage [273,274], il a été postulé qu'elle n'était pas le substrat principal de la Fpg [271]. Cette hypothèse a été confirmée dix ans après la découverte de la première activité de la Fpg, quand l'équipe de Nishimura a trouvé que le substrat principal de la Fpg était la 8-oxo-dG [275]. Cette lésion, présente dans l'ADN en plus grande quantité que la Fapy-dG, est capable d'engendrer des mésappariements et, par la suite, des mutations génétiques, surtout des transversions de G−C vers T−A [276].

Depuis la découverte de la Fpg, de nombreuses études ont été effectuées, visant à analyser ses interactions avec différentes lésions simples de l'ADN. Ainsi, son affinité pour les sites abasiques, la Fapy-dG ou la 8-oxo-dG a été déterminée, ces résultats restant néanmoins très variables et dépendant des conditions expérimentales utilisées (voir : Tableau 18 dans le chapitre « Discussion »).

La Fpg est une protéine globulaire et monomérique, d'une masse moléculaire de 30.2 kDa, composée de 269 acides aminés et codée par un gène composé de 809 paires de bases [277,278]. C'est une métalloprotéine contenant une structure « doigt de zinc » C-terminale, cette structure rentrant en interaction avec l'ADN endommagé [279]. Elle est bifonctionnelle, car elle possède deux activités : ADN glycosylase et AP lyase [280]. Après l'excision de la base endommagée par son activité ADN glycosylase, la Fpg procède à la coupure du côté 3' (β-

élimination) et à la coupure du côté 5' (δ-élimination) du site abasique ainsi formé, effectuées par son activité lyase. La structure intermédiaire formée lors de ces étapes, appelée base de Schiff, peut être stabilisée à l'aide du borohydrure de sodium, la stratégie utilisée afin de cristalliser et d'analyser la structure du complexe Fpg−ADN (Figure 21). Celle-ci montre que les interactions entre la protéine et l'ADN se font par un pont hydrogène au niveau du petit sillon et que la base endommagée est éjectée à l'extérieur de l'hélice.

L'enzyme Fpg a été choisie pour ces études, car c'est un modèle couramment étudié et très bien caractérisé pour ses interactions avec les lésions simples d'oxydation, mais dont la reconnaissance des lésions complexes d'oxydation n'a pas été élucidée. De plus, des résultats intéressants concernant ses interactions avec les cyclonucléosides ont été obtenus il y a quelques années par les membres de notre équipe. En utilisant la technique SPRi (Surface Plasmon Resonance imaging), la reconnaissance de la 5'S-cyclo-dA par la protéine Fpg, sans activité de coupure, a été démontrée pour la première fois [281]. L'approfondissement des connaissances concernant les interactions spécifiques observées entre la Fpg et la 5'S-cyclo-dA, aussi bien que ses interactions avec d'autres cyclonucléosides dont nous disposions dans le laboratoire, constitue ainsi la suite de ce projet.

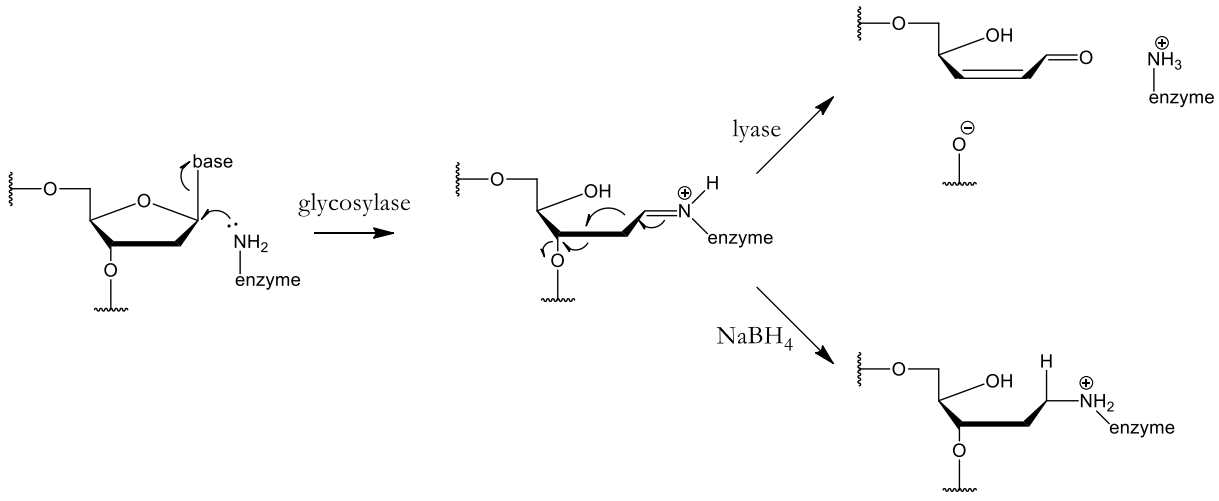


Figure 21 : Stabilisation de la base de Schiff formée entre la glycosylase et sa cible, à l'aide du borohydrure de sodium ($NaBH_4$). Inspiré de [282].

3.3. Enzymes de *Thermococcus gammatolerans*

Deux enzymes provenant d'un organisme original – *Thermococcus gammatolerans* – ont également été utilisées au cours de ces travaux, dans le but de caractériser leurs interactions avec différentes lésions d'oxydation. Ces enzymes, appelées dans ce manuscrit tg1653 et tg1277, sont susceptibles de participer à des voies de réparation des dommages de l'ADN, leur fonction a été prédite d'après les homologies de séquences avec d'autres protéines connues. Tg1653 est une protéine de 30.7 kDa composée de 263 acides aminés, et elle serait l'équivalent de la Fpg bactérienne. Tg1277 a une masse de 26.9 kDa et elle est composée par 230 acides aminés. Elle serait probablement un homologue de l'endonucléase III bactérienne [283].

3.3.1. Genre *Thermococcus*

Thermococcus gammatolerans appartient au genre *Thermococcus*, qui fait partie, avec *Pyrococcus* et *Palaeococcus*, de la famille des *Thermococcaceae*, membre de l'ordre des *Thermococcales*. Le genre *Thermococcus*, le plus grand groupe d'archées caractérisé, regroupe plus de 200 espèces répertoriées (http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2263). Ces archées ont été majoritairement isolées dans les fonds océaniques ou dans des cheminées hydrothermales. Elles sont toutes anaérobies, et elles peuvent vivre à partir de différents composés organiques, commes les sucres et/ou peptides. Elles croissent soit grâce à la fermentation, soit par la respiration en utilisant le soufre – cet élément est ainsi indispensable pour la majorité des *Thermococcus* [284].

3.3.2. Description de l'organisme et de ses conditions de vie

Thermococcus gammatolerans est une archée isolée dans les profondeurs du golfe de Californie (à plus de 2600 m sous l'eau) et a été décrite pour la première fois en 2003 par l'équipe de Daniel Prieur [284]. Ces anaérobies obligatoires vivent à des températures élevées : entre 55 et 95°C, avec un optimum de croissance à 88°C. Ils ont une forme de cocci de diamètre 0.6-1.4 µm.

Thermococcus gammatolerans peut être cultivé dans un milieu organique riche supplémenté de soufre, mais leur croissance a également été constatée dans de nombreux autres milieux, comme l'extrait de levure, le mélange de 20 acides aminés, la tryptone, la peptone, le pyruvate, l'amidon, le maltose ou le lactose. Les milieux de culture contenant des peptides et des acides aminés

nécessitent l'ajout de soufre élémentaire, par contre les cellules peuvent croitre à partir de sucres complexes (maltodextrines, amidon) ou des intermédiaires de la glycolyse comme le pyruvate sans S°. *Thermococcus gammatolerans* est auxotrophe pour neuf acides aminés : cystéine, leucine, lysine, méthionine, phénylalanine, tryptophane, valine, arginine et thyrosine, mais pour vivre sur ce milieu minimal, supplémenté en soufre, un de ces acides aminés doit être présent en plus grande quantité [283].

T. gammatolerans est caractérisée par sa radiorésistance exceptionnelle, à laquelle elle doit son nom : 100% de survie est constaté après l'irradiation à 5 kGy de rayons [283]. Plus d'informations sur les capacités de ces organismes à survivre après l'irradiation, seront données dans le paragraphe 3.3.5.

3.3.3. Génome et protéome de *T. gammatolerans*

En 2009, l'équipe de F. Confalonieri a publié une étude sur l'analyse du génome et du protéome de *T. gammatolerans*, et a comparé l'équipement génétique de cette espèce avec les autres génomes de *Thermococcales* qui avaient été préalablement séquencés. Cet organisme possède un génome circulaire de 2.045 Mpb dans lequel 2157 séquences codantes ont été identifiées, de longueur moyenne de 891 nucléotides. L'analyse protéomique à grande échelle a permis d'identifier 1101 protéines, parmi lesquelles la majorité sont communes pour toutes les *Thermococcales* [283].

Aucun nouveau gène lié à la réparation de l'ADN n'a été détecté, et il semblerait que *T. gammatolerans* disposerait des mêmes enzymes de réparation que les autres *Thermococcales*. Ses systèmes de réparation semblent aussi être plus proches de ceux des eucaryotes que des bactéries. Des protéines telles que les nucléases, les endonucléases, les glycosylases ou les AP-lyases, ainsi que certaines protéines impliquées dans la réparation des cassures double-brin ont été retrouvées dans le protéome analysé. Les fonctions de toutes ces protéines restent néanmoins à vérifier, puisque leur identification et annotation n'ont été faites que sur la base des homologies de séquences [283].

3.3.4. Systèmes de réparation de l'ADN chez les Archées thermophiles

Les conditions de vie des archées thermophiles, notamment leur température de vie optimale avoisinant souvent 80-90°C, sont particulièrement éprouvantes pour leur génome. La température élevée non seulement déstabilise la structure primaire et secondaire de l'ADN, mais accélère également les réactions spontanées menant à la formation de nombreuses lésions : la désamination, la dépurination, l'oxydation, la méthylation ainsi que la formation des cassures. Ainsi, à 100°C les taux de la dépurination de l'ADN et de la désamination de la cytosine seraient environ 3000 fois plus élevés que ceux observés à 37°C [285–287]. Malgré ces conditions de vie, les archées thermophiles se caractérisent par une bonne stabilité génomique, ce qui indique qu'elles doivent posséder des systèmes de réparation de l'ADN très efficaces [288,289].

De nombreuses études montrent que les protéines de réparation présentes chez les archées seraient plus proches des protéines eucaryotes que de leurs homologues bactériens [288]. La présence de plusieurs systèmes de réparation de l'ADN a ainsi été prouvée chez les archées thermophiles, notamment : le BER, la photo-réversion et la recombinaison homologue [285,286]. Par exemple, une O^6-alkylguanine-ADN transférase thermostable a été retrouvée chez *Sulfolobus acidocaldarius*, *Pyrobaculum islandicum*, *Pyrococcus furiosus* et *Thermococcus litoralis* [290–292]. Egalement, une uracile glycosylase thermostable a été identifiée chez plusieurs archées, ce qui aurait une importance particulière pour la réparation des cytosines désaminées [290,293,294].

Bien que la présence de nombreux systèmes de réparation de l'ADN chez les archées soit prouvée, l'existence du NER chez ces organismes reste non-élucidée. La majorité des archées possède les homologues de certaines protéines du NER eucaryote, comme les équivalents des nucléases XPF-ERCC1 et Fen1, ainsi que des hélicases XPB et XPD. Cependant, la nature et le fonctionnement exacts de ce système de réparation ne sont pas connus [288,295]. Par exemple, aucun homologue des protéines de reconnaissance des lésions volumineuses, telles que XPA/Rad4 ou XPC/Rad1 eucaryotes n'a à ce jour été découvert [285]. La présence d'un système alternatif de réparation des dommages volumineux chez les archées ne peut également pas être exclue. Par exemple, *Sulfolobus acidocaldarius* présente une courbe de survie suite aux irradiations UV qui suggère la prise en charge des lésions formées suite à ce stress [285,296].

Les gènes codant pour quatre protéines homologues de *Pyrococcus*, participant dans le système NER, ont été identifiés chez *T. gammatolerans* : XPB-rad25, XPD-rad3, XPF et XPG/Fen1-like, cependant seule cette dernière a également été retrouvés au niveau protéomique chez les organismes cultivés en absence de stress [283].

Un autre système de réparation qui semble manquer chez les archées thermophiles est la réparation des mésappariements (MMR), car les homologues de ses deux principales enzymes : MutS et MutL bactériennes, n'ont pas pu être identifiés. Il serait alors suggéré qu'un autre système puisse exister chez ces organismes, qui remplacerait en même temps le NER et le MMR [285].

3.3.5. Radiorésistance et événements dans l'organisme après l'irradiation

Comme son nom l'indique, *T. gammatolerans* se caractérise par une extrême résistance envers le rayonnement gamma, car aucune perte de viabilité après une irradiation à 5 kGy n'a été constatée [283]. D'après les expériences menées par l'équipe de F. Confalonieri, la phase de croissance n'a pas d'influence sur leur radiorésistance jusqu'à 7.5 kGy, si les cellules sont cultivées dans leurs conditions optimales. Cependant, la prolongation de la phase stationnaire sensibilise les cellules irradiées à 7.5 kGy, les rendant 100 fois plus sensible aux rayonnements γ [297].

Les variations de la radiorésistance en fonction du milieu de culture ont également été étudiées. Même en absence de stress, les cellules cultivées dans un milieu minéral minimal supplémenté en soufre, en vitamines et avec des 20 acides aminés, n'ont pas la même cinétique de croissance que les cellules cultivées dans un milieu riche. De plus, le milieu riche semble favoriser la radiorésistance de *T. gammatolerans*, quel que soit le milieu dans lequel les cellules ont été cultivées avant de subir les irradiations. Ces résultats confirment l'importance du type de milieu et de la disponibilité des nutriments pour la radiorésistance des *T. gammatolerans* [297].

Indépendamment de la survie suite aux irradiations à fortes doses, la capacité de *T. gammatolerans* à reconstruire son génome a également été étudiée. La reconstitution de l'intégralité du génome serait ainsi favorisée par le milieu riche. Les profils protéiques des organismes cultivés dans un milieu riche ou dans un milieu minimal ne sont pas les mêmes, ce qui peut jouer un rôle dans l'efficacité et la rapidité de la réparation de l'ADN. Il faut néanmoins remarquer que la fragmentation de l'ADN dans le milieu riche et dans le milieu minimal n'est pas la même, ce qui suggère aussi un rôle du milieu dans la protection de l'ADN contre les dommages générés par le rayonnement ionisant. L'ADN n'est cependant pas la seule molécule qui est endommagée dans les cellules irradiées : d'autres composants peuvent aussi subir les dommages, ce qui se traduit par un arrêt de croissance, qui ne reprend pas immédiatement après la reconstitution du génome. Ce décalage est dépendant de la dose d'irradiation, et cela indique

que la réparation de l'ADN n'est pas le seul phénomène qui contrôle la reprise de la croissance après un stress [297].

4. Structure du manuscrit

Dans la partie qui va suivre, les travaux effectués pendant les trois ans de cette thèse vont être détaillés. Dans un premier temps, les techniques utilisées vont être décrites : leur principe général va précéder à chaque fois le protocole détaillé permettant de reproduire les expériences présentées ici. Dans un deuxième temps, les résultats obtenus vont être décrits puis discutés : conformément aux objectifs énoncés dans le paragraphe 1, ceux-ci vont être divisés en deux parties :

- la mise au point du piège à protéines et le piégeage de celles-ci au sein d'extraits de cellules HeLa sur des sondes portant ou non des lésions complexes d'oxydation, ainsi que la confirmation des interactions identifiées ;
- l'analyse des interactions entre les protéines isolées d'*E. coli* ou de *T. gammatolerans*, et des lésions d'oxydation simples ou complexes, et l'étude de leur mode d'action.

La présentation des résultats va être suivie de discussion et terminée par une partie « conclusions et perspectives ».

III Matériel et méthodes

1. Screening des protéines : construction du piège à protéines

1.1. Principe de la technique du ligand fishing

Le ligand fishing est une technique conçue pour piéger, dans des extraits cellulaires, les protéines qui reconnaissent les sondes contenant dans leur structure des lésions cibles, afin de définir l'interactome de celles-ci. Les billes magnétiques fonctionnalisées sont exposées aux extraits protéiques, et après un temps d'incubation, le surnageant contenant des protéines non-associées aux pièges peut être facilement séparé des billes, à l'aide d'un aimant. Les protéines qui reconnaissent le piège peuvent le faire par interactions aspécifiques avec la surface des billes ou avec les parties non-endommagées des sondes, ou elles peuvent reconnaître les dommages localisés dans les sondes. En fonction du dommage que l'on souhaite analyser, la structure des sondes n'est pas toujours la même, puisque l'on peut disposer des sondes très courtes type oligonucléotides synthétiques ou des sondes longues, comme des plasmides linéarisés. Le choix de la sonde va être déterminant pour la construction du piège.

Chaque type de sonde va avoir ses avantages et ses inconvénients, qu'il faut prendre en compte lors de la mise en place du protocole. Les sondes plasmidiques sont faciles à obtenir en grande quantité et elles permettent de piéger une grande quantité de protéines grâce à leur taille (≈3 kpb) et au nombre de lésions qui peuvent y être formées [269]. Ces sondes ne se prêtent pourtant pas à tout type de lésions, puisque des dommages synthétisés chimiquement, ou ceux qui ont besoin d'être localisés dans un endroit bien précis de la séquence, ne peuvent pas y être insérés. L'usage des plasmides en tant que sondes est limité aux lésions qui peuvent être générées par exposition à l'agent toxique, tel que les irradiations UV ou bien certains produits chimiques (par exemple le cisplatine). Aussi, il faut prendre en compte qu'en exposant le plasmide à l'agent génotoxique, une panoplie de lésions va être créée. Ceci n'est pas gênant si un type de dommage est formé majoritairement, ou bien si l'on s'intéresse à l'ensemble des modifications engendrées par ce traitement précis.

Cependant, lorsque l'on souhaite analyser une lésion en particulier, il est nécessaire de travailler sur un tout autre type de sondes, notamment des oligonucléotides synthétisés chimiquement. Leur désavantage majeur est leur taille limitée par les possibilités de synthèse et de

purification, mais ils permettent l'introduction d'une lésion particulière dans un endroit défini de la séquence.

Une autre possibilité est de fabriquer les sondes à l'aide de la PCR – elles peuvent être alors plus longues que les oligonucléotides, mais cette application est limitée par la disponibilité des bases modifiées pouvant être insérées lors de la synthèse. Cette technique est toutefois utile pour l'étude de modifications épigénétiques telles que la méthylation, où l'on peut facilement insérer dans la séquence les bases modifiées à la place des bases non-modifiées.

Quel que soit le type des sondes, le principe général du fonctionnement des pièges utilisés dans notre laboratoire est le même. Les sondes sont fixées sur des billes magnétiques couvertes de streptavidine à l'aide des interactions entre celle-ci et la biotine située sur une extrémité de la sonde. Le système ainsi conçu est ensuite exposé aux extraits dans le but de piéger des protéines reconnaissant les lésions d'intérêt (Figure 22).

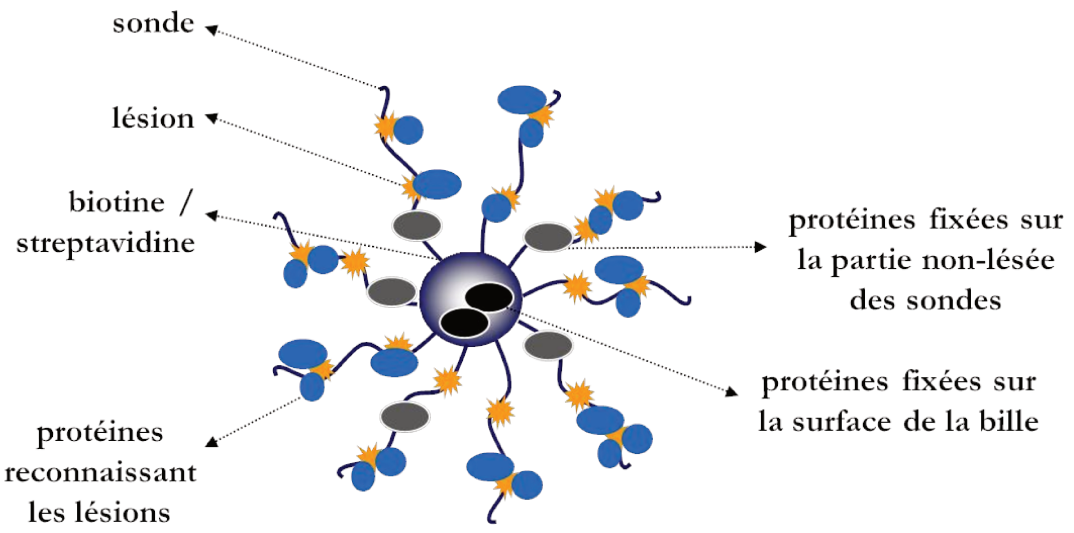


Figure 22 : Représentation schématique du piège à protéines.

Un inconvénient de ce système est que de nombreuses protéines présentes dans les extraits vont avoir une affinité pour les séquences non-lésées ou encore pour la surface des billes. Les rinçages effectués avec des tampons plus stringents que celui utilisé pour l'incubation sont censés éliminer ces interactions non-spécifiques sans pour autant influencer les interactions entre les protéines et les lésions d'intérêt, mais leur élimination totale n'est pas possible. C'est pour cela que la liste des protéines identifiées sur les pièges contenant les sondes lésées doit être toujours comparée à celle obtenu sur les pièges témoins (ne portant aucune lésion) ou sur les billes, pour

qu'un différentiel entre ces conditions puisse être établi et pour que les protéines non-spécifiques générant un bruit de fond ne soient pas prises en compte lors de l'analyse ultérieure. Malgré ces inconvénients, cette technique permet néanmoins d'identifier au moins une partie de l'interactome des lésions étudiées, comme cela a été décrit dans le chapitre « Stratégie et objectifs » pour les travaux effectués antérieurement au sein de notre laboratoire.

1.2. Etapes d'optimisation du piège

Comme il a été indiqué précédemment, la technique de « ligand fishing » a déjà été utilisée dans notre équipe pour l'analyse de l'interactome des lésions générées par le cisplatine [268, 269]. Dans cette étude, des plasmides linéarisés exposés à un agent génotoxique ont été utilisés. L'application de ce type de sondes aux cyclonucléosides n'est pas possible, car l'exposition d'un plasmide aux rayonnements ionisants générerait une panoplie de lésions, ce qui rendrait impossible l'analyse d'un seul type de dommage. C'est pour cette raison que nous avons utilisé, dans un premier temps, de courtes sondes oligonucléotidiques, dans lesquelles les cyclonucléosides ont été intégrés dans un site précis. Cependant, les premières expériences utilisant ces sondes ont démontré qu'elles étaient trop courtes et qu'elles ne possédaient pas suffisament de lésions pour permettre le piégeage efficace des protéines pertinentes. Ainsi, nous avons conçu un système de concatémères, c'est-à-dire des longues sondes composées de la même séquence oligonucléotidique répétée une dizaine de fois. La mise au point de ce type de sondes va être précisée dans les paragraphes qui suivent.

1.3. Protocoles détaillés

1.3.1. Synthèse des oligonucléotides

Tous les oligonucléotides, sauf les quatre séquences contenant les cyclonucléosides, ont été commandés chez Eurogentec. Les oligonucléotides: 5'S-cyclo-dA, 5'R-cyclo-dA, 5'S-cyclo-dG et 5'S-cyclo-dHT ont été synthétisés à l'aide d'un synthétiseur ABI 392 (Applied Biosystems) à l'échelle de 1 μmole, selon la voie phosphoramidite, comme décrit précédemment [298] (toutes les séquences sont présentées dans le Tableau 7). La synthèse a été effectuée sur le support, dans le sens 3' → 5'. Le principe de cette voie de synthèse est présenté sur la Figure 23. Tous les

synthons sont dilués dans l'acétonitrile anhydre extemporanément à la synthèse, pour obtenir la concentration finale de 0.1 M. La première base est attachée à des billes de verre par son extrémité 3', et son extrémité 5' est protégée par un groupement diméthoxytrityle (DMTr). Pour effectuer la première étape de la synthèse, cette base est déprotégée par un traitement à l'acide trichloroacétique, tandis que le monomère phosphoramidite suivant (Glen Research) est activé avec du tétrazole. La réaction de couplage a ensuite lieu, et elle est suivie de l'oxydation du phosphore III en phosphore V avec de l'iode. La dernière étape sert à inactiver les groupements qui n'ont pas réagi lors des étapes précédentes, avec l'anhydride phénoxyacétique et le méthylimidazole. Ce cycle de synthèse est ensuite répété le nombre de fois nécessaire pour atteindre la séquence désirée.

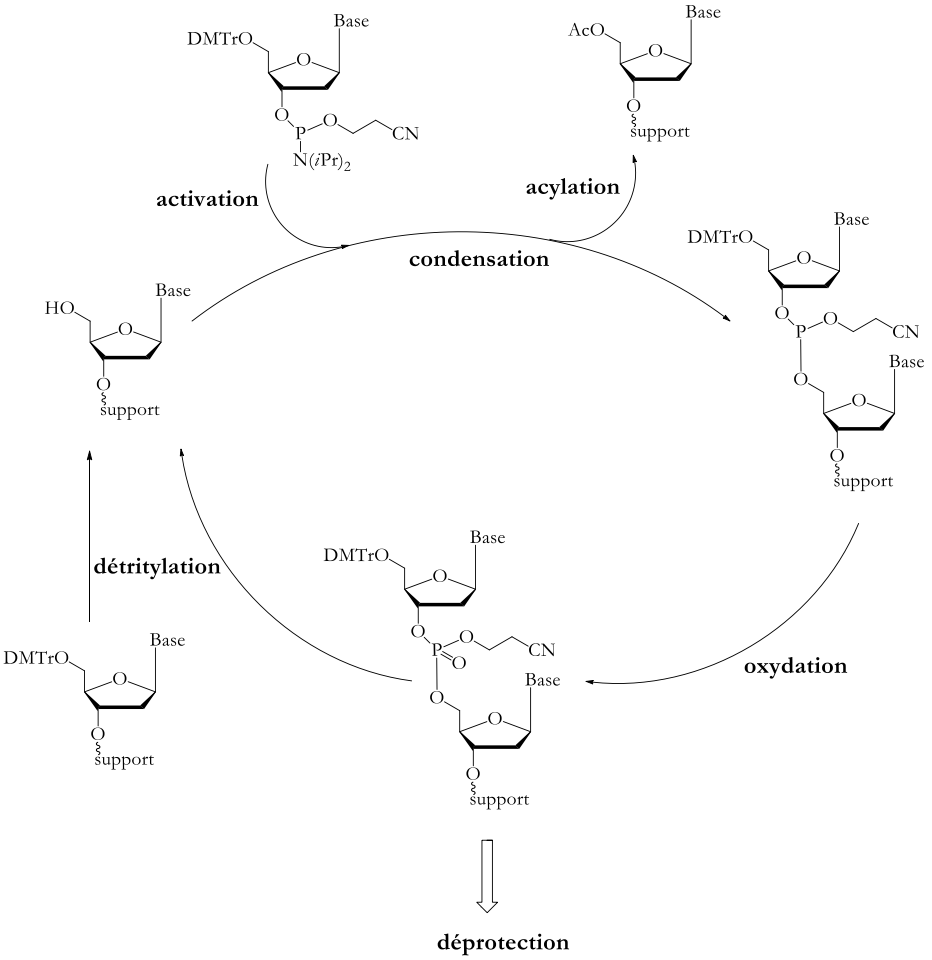


Figure 23 : Représentation schématique de la voie de synthèse d'ADN par la méthode phosphoramidite.

Une fois la synthèse terminée, tous les phosphates étant protégés par des groupements cyanoéthyles et les bases nucléiques par des groupements acyles, ceux-ci sont éliminés par un

traitement avec l'ammoniaque à 30% (1 mM d'ammoniaque pour 1 µmol d'oligonucléotide), pendant une nuit à 50°C. Après évaporation de l'ammoniaque, les oligonucléotides sont repris dans de l'eau ou du TEAA à 10 mM et filtrés, avant d'être purifiés par électrophorèse sur gel ou par chromatographie liquide haute performance.

Tableau 7 : Séquences et modifications des oligonucléotides utilisés au cours de ce travail. Les bases indiquées **X** représentent les sites où les modifications internes (qui correspondent aux noms des oligonucléotides) ont été insérées. biotine TEG et biotine PCB (avec un bras photoclivable) : voir Figure 25 ; comp = oligonucléotide complémentaire ; cis-Pt = oligonucléotide contenant un pontage entre deux guanines adjacentes, généré par le cisplatine ; NL = oligonucléotide non-lésée de la même séquence que cis-Pt ; comp-PCB = oligonucléotide complémentaire aux cis-Pt et NL, portant une biotine photoclivable.

nom de l'oligonucléotide	longueur [bases]	modification terminale	séquence (5' → 3')
concat-biot	16	5'-biotin TEG	AGA TCA GTC ACG ATC C
concat-PCB	16	5'-biotin PCB	AGA TCA GTC ACG ATC C
concat-1/ non-lésé	22	5'-phosphate	CAC TTC GGA TCG TGA CTG ATC T
concat-2	22	5'-phosphate	GAA GTG AGA TCA GTC ACG ATC C
5'S-cyclo-dA	22		CAC TTC GG**X** TCG TGA CTG ATC T
5'R-cyclo-dA	22		CAC TTC GG**X** TCG TGA CTG ATC T
5'S-cyclo-dG	22		CAC TTC G**X**A TCG TGA CTG ATC T
5'S-cyclo-dHT	22		CAC TTC GGA **X**CG TGA CTG ATC T
8-oxo-dA	22		CAC TTC G**X**A TCG TGA CTG ATC T
8-oxo-dG	22		CAC TTC G**X**A TCG TGA CTG ATC T
8-Me-dG	22		CAC TTC G**X**A TCG TGA CTG ATC T
8-oMe-dG	22		CAC TTC G**X**A TCG TGA CTG ATC T
8-Br-dG	22		CAC TTC G**X**A TCG TGA CTG ATC T
5-OH-dHyd	22		CAC TTC GGA **X**CG TGA CTG ATC T
5-OH-5-Me-dHyd	22		CAC TTC GGA **X**CG TGA CTG ATC T
comp-dA	22		AGA TCA GTC ACG AT**A** CGA AGT G
comp-dG	22		AGA TCA GTC ACG AT**G** CGA AGT G
comp-dT	22		AGA TCA GTC ACG AT**T** CGA AGT G
comp (comp-dC)	22		AGA TCA GTC ACG AT**C** CGA AGT G
THF	22	(5'-phosphate)	CAC TTC G**X**A TCG TGA CTG ATC T
fluorescent	15	3'-fluorescéine	CAC TTC GGA TCG TGA
14bases-1	14		TCG GAT CGT GAC TG
14bases-2	14		CAG TCA CGA TCC GA
44bases-1	44		CAC TTC GGA TCG TGA CTG ATC TCA CTT CGG ATC GTG ACT GAT CT
44bases-2	44		AGA TCA GTC ACG ATC CGA AGT GAG ATC AGT CAC GAT CCG AAG TG
comp-biot	22	5'-biotin TEG	AGA TCA GTC ACG ATC CGA AGT G
c1-scramble	22		AGA CGT TCG GTT ACT AGC TCT C
c2-scramble	22		GAG AGC TAG TAA CCG AAC GTC T
cis-Pt	16		CCT CTC T**GG** ACC TTC C
NL	16		CCT CTC TGG ACC TTC C
comp-PCB	21	5'-PCB	TTT TTG GAA GGT CCA GAG AGG

1.3.2. Purification des oligonucléotides

Les oligonucléotides sont repris dans de la formamide dans la proportion 50:50 (v/v) avec de l'eau, et séparés par électrophorèse sur un gel dénaturant d'acrylamide à 15%. Ils sont ensuite visualisés par UV-shadowing, et la bande correspondant au produit de synthèse recherché est découpée et reprise dans l'eau. L'extraction de l'ADN depuis l'acrylamide vers la phase aqueuse est effectuée pendant une nuit sous agitation, et le volume obtenu est ensuite réduit par évaporation au SpeedVac. Enfin, les oligonucléotides sont dessalés sur des colonnes NAP (GE Healthcare) et leur concentration est déterminée par une mesure UV au NanoDrop (Labtech France).

1.3.3. Phosphorylation des oligonucléotides

100 pmol d'oligonucléotides sont phosphorylés dans un volume final de 50 µL. Puisque les oligonucléotides n'ont pas tous la même concentration, leur volume doit être ajusté en fonction de celle-ci. Les concentrations des oligonucléotides synthétiques utilisés sont présentées dans le Tableau 8.

Tableau 8 : Concentrations des oligonucléotides portant des cyclonucléosides, utilisés pour préparer les pièges « ligand fishing ».

oligonucléotide	concentration [µM]	volume pour 100 pmol [µL]
5'S-cyclo-dA	3.8	26
5'R-cyclo-dA	1.8	56
5'S-cyclo-dG	32.9	3
5'S-cyclo-dHT	3.8	26

Puisque le volume correspondant à 100 pmol d'oligonucléotide 5'R-cyclodA dépasse le volume total de la réaction, il doit être séché au préalable pour que les 100 pmol soient comprises dans 33 µL au maximum. 2 U de kinase T4 (Life Technologies), 10 µL de son tampon 5x et 5 µL d'ATP (Sigma) à 10mM sont ajoutés dans le mélange réactionnel, qui est ensuite complété avec de l'eau jusqu'à 50 µL. Le tout est incubé 20 min à 37°C et l'enzyme est ensuite inactivée 10 min à 65°C. Afin d'éliminer les restes d'ATP et les sels présents dans le tampon, les oligonucléotides phosphorylés sont purifiés sur une colonne G-25 illustra MicroSpin (GE Healthcare).

1.3.4. Formation des concatémères sur les billes

Afin de former les concatémères sur les billes, trois types d'oligonucléotides sont nécessaires (Figure 24). Dans un premier temps, une séquence de 16 bases portant à son extrémité 5' une biotine TEG (structure de la biotine présentée sur la Figure 25), appelé « concat-biot » est attachée sur les billes magnétiques couvertes de streptavidine. Les deux autres oligonucléotides complémentaires, dont les séquences se chevauchent : « concat-1 » et « concat-2 » sont rajoutés et hybridés, pour être enfin ligués et former une longue sonde composée de la même séquence d'ADN répétée plusieurs fois. Dans des premières étapes de la mise au point nous avons également essayé d'introduire dans les sondes les séquences « stop » (« stop-1 » et « stop-2 »), qui s'hybrideraient avec les extrémités simple-brin des concatémères. Leur utilisation a néanmoins été abandonnée pour des raisons expliquées plus tard dans le chapitre « Résultats ».

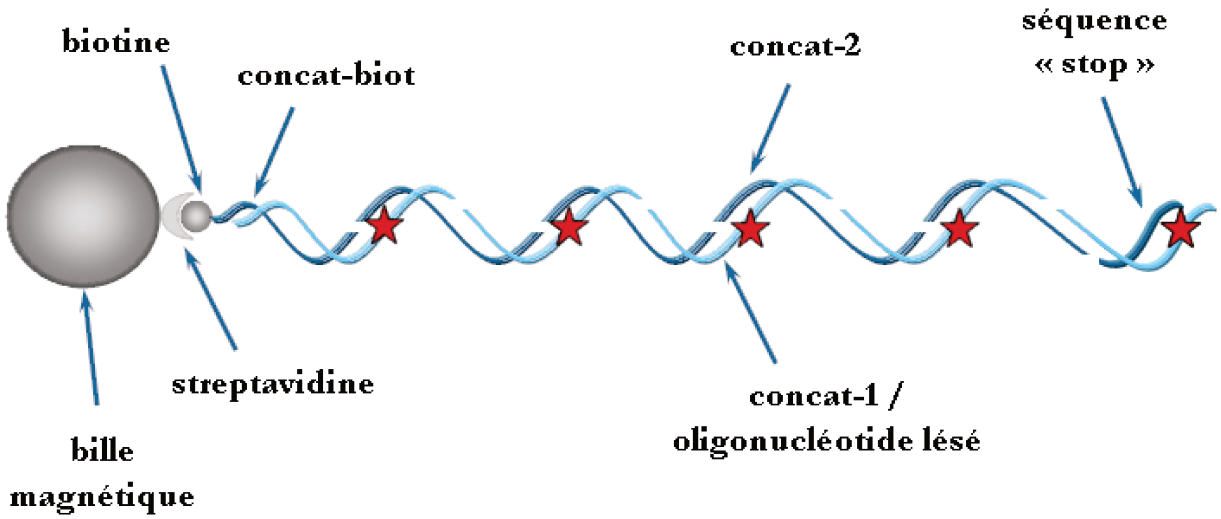


Figure 24 : Représentation schématique d'un piège à protéines (« ligand fishing »), basé sur une approche concatémères, avec ses composants majeurs.

La réaction de formation des concatémères a lieu dans des tubes ADN LoBind (Eppendorf), afin de maximiser la quantité des sondes fixées sur les billes en limitant les pertes des oligonucléotides par leur fixation non-spécifique sur les parois du tube. 150 µg (15 µL) de billes magnétiques couvertes de streptavidine et stockées dans un tampon contenant 0.1% de BSA (Dynabeads M-280 Streptavidin ; Life Technologies) sont rincés 2 fois dans 100 µL de tampon 1x de T4 DNA ligase (Life Technologies), afin d'éliminer les traces de leur tampon de stockage et de les équilibrer dans le tampon qui va être utilisé par la suite dans cette expérience. Elles sont ensuite reprises dans un volume final de 60 µL, contenant le tampon de T4 ligase à la

concentration de 1x et 30 pmol de « concat-biot », et incubées 10 min à température ambiante sous agitation régulière. L'agitation des billes est nécessaire pour éviter leur sédimentation au fond du tube et assurer l'homogénéité de la réaction. Ensuite, 300 pmol de chacun des deux oligonucléotides complémentaires sont ajoutés (« concat-2 » et « concat-1 »/oligonucléotide lésé), et le volume est ajusté à 255 µL. La concentration du tampon de la T4 ligase à ce moment de la réaction est légèrement supérieure à 1x (1.06x), mais celle-ci va être ajustée à 1x par l'ajout de la T4 ligase dans une étape ultérieure de cette réaction. Le mélange est ensuite chauffé 5 min à 70°C et laissé refroidir dans un bloc chauffant éteint, afin que l'hybridation des brins complémentaires s'effectue lentement. Après le refroidissement, 15 U de T4 DNA Ligase à 1 U/µL (Life Technologies) sont ajoutés – la concentration du tampon de T4 ligase est ajustée ainsi à 1x – et la réaction se poursuit 1 heure à 37°C sous agitation régulière. Une fois cette réaction terminée, le surnageant est éliminé et les billes sont rincées deux fois dans 100 µL de tampon 1x de T4 DNA ligase afin d'éliminer toute trace de l'enzyme ou des oligonucléotides non-hybridés, et transférées dans les tubes Protein LoBind (Eppendorf).

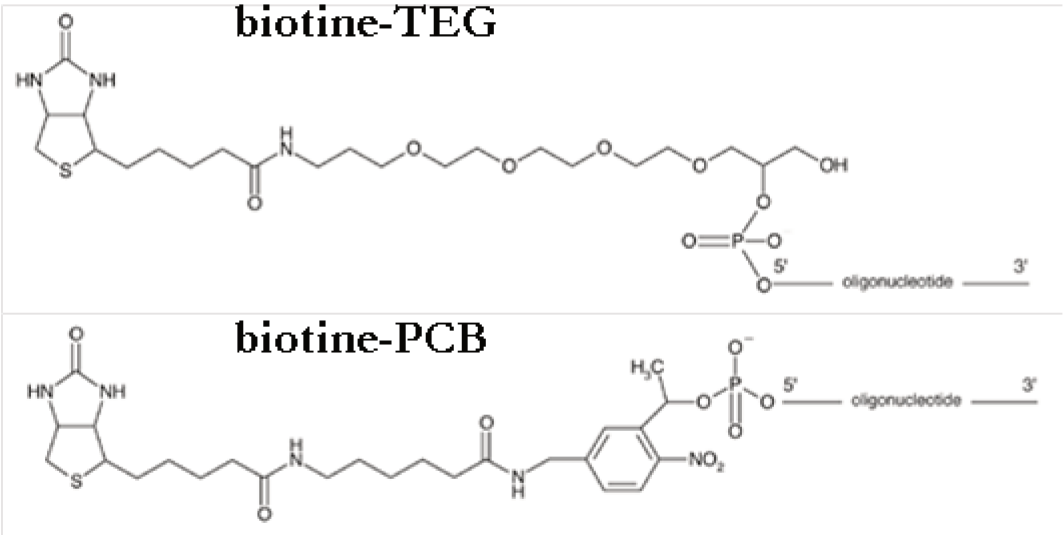


Figure 25 : Structure d'une biotine TEG et d'une biotine couplée à un bras photoclivable qui peuvent être incorporées à l'extrémité 5' d'un oligonucléotide. Source : Eurogentec.

Les pièges vont être exposés à deux types d'extraits : des extraits nucléaires commerciaux de cellules HeLa, ou des extraits bactériens d'*Escherichia coli*. Le protocole de préparation de ces derniers est présenté ci-dessous.

1.3.5. Culture d'*Escherichia coli*, extraction des protéines bactériennes et leur dosage

Les bactéries *Escherichia coli* (souche K12 fournie par Jean-François Collet de l'Université catholique de Louvain) sont cultivées dans le milieu LB (10 g de tryptone, 5 g d'extrait de levure et 10 g de NaCl (Sigma) dans 1 L d'eau) à 37°C sous agitation régulière. Une pré-culture est réalisée pendant une nuit dans 5 mL de milieu, qui sont transférés le lendemain dans 500 mL de milieu frais. Les bactéries sont ensuite cultivées jusqu'à ce que leur densité optique atteigne la valeur de 1.5. Elles sont ensuite centrifugées 10 min à 4000 g à 4°C, et le culot est repris dans 10 mL de tampon de lyse (20 mM Tris-HCl, 150 mM NaCl, antiprotéase – 1 pastille de complete mini (Roche) pour 20 mL de tampon) et vortexé. Les membranes cellulaires sont ensuite cassées par 3 cycles de choc de température : 3 min dans l'azote liquide suivies de 3 min au bain-marie à 37°C. Le lysozyme à 1% de concentration finale est ensuite ajouté et les tubes sont incubés sur la glace pendant 15 min. 10 cycles de sonication sont alors effectués : 20 sec de pulsations suivies de 20 sec de repos. Enfin, la solution est centrifugée pendant 50 min à 20000 g à 4°C, et le surnageant contenant les protéines est récupéré.

Les protéines solubles sont dosées d'après la méthode de Pierce, appelée également la méthode BCA. C'est un dosage colorimétrique, basé sur la réduction de Cu^{2+} en Cu^{+} par des protéines en milieu alcalin. L'absorbance du complexe de Cu^{+} formé lors de la réaction, qui est proportionnelle à la quantité des protéines dans l'échantillon, est ensuite dosée à 562 nm. Cette méthode permet le dosage des protéines à concentrations comprises entre 1 et 100 µg/mL.

Dans un premier temps, une gamme de concentration de la protéine témoin (BSA à 2 mg/mL ; Interchim) est préparée avec cinq points : 1, 5, 10, 15 et 20 µg/mL. La même protéine, ainsi que la solution à doser, sont diluées au 160^{e}, et trois dilutions sont ensuite préparées : 1/533^{e}, 1/400^{e} et 1/213^{e}. 100 µL de chacune de ces dilutions (la gamme étalon y comprise) sont placés dans un puits d'une plaque 96 puits, et 100 µL du réactif de Pierce (Interchim) y sont ajoutés. La plaque est incubée pendant 1 heure au bain-marie à 60°C, et après ce temps, l'absorbance de chaque puits est lue à 562 nm, avec un appareil SpectraMax M2 (Molecular Devices).

1.3.6. Exposition du piège aux extraits et récupération des protéines

L'exposition des pièges aux protéines a lieu dans les tubes Protein LoBind (Eppendorf), afin de minimiser la fixation non-spécifique des protéines sur les parois des tubes et d'optimiser ainsi leur piégeage. Avant d'incuber les pièges avec les extraits protéiques, ils sont équilibrés par deux rinçages dans 100 µL de tampon d'incubation. Au cours de l'optimisation du protocole de piégeage, différents tampons ont été testés, et leur composition est donnée dans le Tableau 9. Le tampon qui a été retenu est le tampon RBP – son choix sera expliqué dans la partie « Résultats ». Les pièges sont ensuite incubés avec 64 µL d'extraits nucléaires HeLa (CIL Biotech) à 10 mg/mL ou d'extraits *E.coli* à 11.36 mg/mL et 36 µL d'ADN non-spécifique de sperme de poisson à 10 mg/mL (Roche) dans un volume final de 300 µL à 1x de tampon d'incubation, pendant 20 min à 15°C, sous agitation régulière. Après incubation, les pièges sont rincés deux fois dans 100 µL de tampon RBP, afin d'éliminer les protéines non-piégées et l'ADN non-spécifique restant dans le mélange réactionnel. Afin de récupérer les protéines piégées, les billes sont reprises dans 30 µL de tampon Laemmli 1x et chauffées 10 min à 90°C – le surnageant, contenant les protéines dénaturées, est ensuite récolté.

Tableau 9 : Composition des tampons utilisés au cours de l'optimisation du piégeage des protéines.

nom du tampon	composition	tampon de rinçage
tampon « A » [268,269]	40 mM Hepes-KOH pH 7.8, 50 mM KCl ; 5 mM $MgCl_2$, 0.5 mM DTT	150 mM NaCl, 20 mM Hepes-NaOH pH 7.8, 5 mM $MgCl_2$, 0.025% NP-40
tampon « RBP » [299]	100 mM NaCl, 100 mM Tris-HCl pH 8.0, 0.25% NP-40, 1 mM DTT, antiprotéase	

1.3.7. Récupération des protéines en utilisant le bras photoclivable

L'oligonucléotide contenant un bras photoclivable est intégré dans le piège de la même manière que son homologue contenant une biotine TEG. La différence de son usage réside dans le mode d'élution qui va être appliqué, car le chauffage des pièges dans un tampon dénaturant dans le but de décrocher les protéines n'est plus nécessaire.

La séparation des sondes et donc des protéines de la surface des billes est effectuée sous effet du rayonnement UV. Après incubation avec les extraits protéiques et rinçages, les billes sont

reprises dans 20 µL de tampon de rinçage, déposées dans un puits d'une plaque 96-puits à fond blanc, « non-binding » (Greiner bio-one) et irradiées avec une dose de 54 J/cm² d'UVA (UVA 700L ; Waldmann) [300]. L'irradiation est réalisée sous agitation à 0°C. Les billes sont ensuite transférées dans les tubes, et le surnageant contenant les sondes et les protéines est séparé des billes. Lorsque les protéines récupérées sont déposées par la suite sur un gel SDS-PAGE, elles doivent subir en plus une dénaturation par chauffage 10 min à 90°C dans un tampon de type Laemmli.

1.3.8. Coloration au nitrate d'argent

Après la migration des protéines, les gels SDS-PAGE peuvent être colorés afin de visualiser la quantité des protéines dans les échantillons. Cette technique de coloration très sensible permet de visualiser toutes les protéines présentes sur le gel et ainsi comparer visuellement leurs quantités, même si aucune quantification précise ne peut être effectuée. Elle ne donne pas non plus accès aux informations qualitatives concernant les protéines visualisées, car toutes les protéines, indépendamment de leurs propriétés, sont détectées de la même manière.

La coloration au nitrate d'argent est effectuée avec le kit PlusOne Silver Staining Kit (GE Healthcare), d'après le protocole fourni.

1.3.9. Analyse protéomique

L'analyse protéomique des échantillons récupérés après le piégeage des protéines sur les sondes, a été effectuée dans l'un des deux laboratoires partenaires : les analyses faisant partie de l'optimisation des conditions de piégeage ont été faites dans le laboratoire de l'Etude de la Dynamique des Protéomes (iRTSV/BGE/EDyP ; CEA Grenoble), tandis que les analyses ultérieures ont été faites dans le Laboratoire de Biochimie des Systèmes Perturbés (DSV/iBEB/SBTN/LBSP ; CEA-ValRhô Marcoule) dans le cadre du projet ANR « Gammatolerans ». Le laboratoire qui a effectué les analyses protéomiques sera précisé dans le chapitre « Résultats ».

Analyse au sein du Laboratoire de l'Etude de la Dynamique des Protéomes

L'échantillon est déposé sur un gel NuPAGE (Life Technologies) et une courte migration est effectuée. Le gel est ensuite coloré au bleu de Coomassie et la partie contenant des protéines est découpée et hydrolysée à l'aide de la trypsine. La digestion est réalisée par le robot EVO150 (TECAN). Le mélange peptidique est dessalé sur une colonne C18 PepMap de 300 µm × 5 mm et séparé sur une colonne C18 de 75 µm × 150 mm (Phenomenex) avec un gradient de deux phases : A (0.1% acide formique, 97.9% eau, 2% ACN) et B (0.08% acide formique, 20% eau, 79.92% ACN). L'analyse LC-MS/MS est effectuée avec un LTQ-Orbitrap Discovery (Thermo Fisher) couplé à un système ultimate 3000 (Dionex). Les peptides sont identifiés avec le logiciel Mascot (Matrix Science) en utilisant la base de données Uniprot, et validés à l'aide du logiciel IRMa [269,301].

Analyse au sein du Laboratoire de Biochimie des Systèmes Perturbés

15 µL d'échantillon repris dans un tampon LDS (Life Technologies) sont déposés sur un gel 12% NuPAGE (Life Technologies) et une courte migration est effectuée. Le gel est ensuite coloré au bleu de Coomassie et la partie du gel contenant des protéines est découpée et hydrolysée à l'aide de la trypsine. Le mélange peptidique est dessalé sur une colonne C18 Pepmap en phase reverse (LC Packings) et résolu sur une colonne capillaire C18 Pepmap TM (LC Packings) avec un gradient de 90 minutes passant de 5% à 95% de mélange : 0.1% HCOOH/80% CH3CN dans 0.1% HCOOH. L'analyse LC-MS/MS est effectuée avec un spectromètre de masse LTQ-Orbitrap XL hybride (ThermoFisher) couplé à un système UltiMate 3000 LC (Dionex-LC Packings). Les peptides sont identifiés avec le logiciel Mascot (Matrix Science) en utilisant la base de données NCBInr, et validés à l'aide du logiciel IRMa [283].

2. Techniques de validation des interactions

Les interactions identifiées lors de l'analyse protéomique comme spécifiques pour un type de sondes, doivent être confirmées avec des techniques complémentaires. La première technique utilisée pour ceci est le Western blot. Les échantillons sont préparés de la même manière que ceux destinés à l'analyse protéomique, c'est-à-dire en utilisant les pièges à protéines. La deuxième

technique que nous allons utiliser pour vérifier les interactions entre les sondes et les protéines identifiées est le retard sur gel, qui permet de travailler sur une seule protéine purifiée, et non plus à partir d'extraits, à condition que celle-ci soit disponible.

2.1. Western blot

Les gels de polyacrylamide (acrylamide/polyacrylamide 29/1 ; Euromedex) sont composés de deux parties : la partie basse – le gel de séparation – est concentrée à 10% d'acrylamide (375 mM Tris-HCl pH 8.8, 0.1% SDS) et la partie haute – le gel de concentration – à 4% d'acrylamide (125 mM Tris-HCl pH 6.8, 0.2% SDS). Les échantillons récupérés après le piégeage des protéines y sont déposés et la migration est effectuée dans le tampon : 25 mM Tris, 192 mM glycine et 0.1% SDS, pendant environ 1 heure à 150 V.

Le transfert des protéines sur une membrane de nitrocellulose est effectué ensuite avec l'appareil Trans-Blot Turbo Transfer System (Bio-Rad). Les membranes sont conditionnées avec le tampon de transfert (Bio-Rad) et elles sont prêtes à utiliser en transfert semi-sec. Après le transfert, une visualisation des protéines à l'aide d'une solution de rouge neutre peut être réalisée, mais ceci n'est pas indispensable. La membrane est ensuite bloquée pendant au minimum 1 heure à température ambiante dans le tampon TBS (Euromedex) / 0.5% Tween-20 (Sigma) / 5% lait, avant d'être incubée pendant une nuit à 4°C avec l'anticorps primaire dilué, d'après le facteur souhaité, dans le même tampon. Le lendemain, la membrane est rincée 3 fois 5 minutes dans le TBS / 0.5% Tween-20 et incubée ensuite avec l'anticorps secondaire dilué dans le même tampon. Après cette incubation, la membrane est rincée 3 fois 5 minutes dans le TBS / 0.5% Tween-20. Les dilutions des anticorps primaires et secondaires utilisés sont données dans le Tableau 10.

Afin de révéler la membrane, celle-ci est tout d'abord essuyée pour enlever les restes du tampon. Le réactif de révélation ECL Plus (GE Healthcare) est ensuite appliqué pendant 2 minutes. La membrane est alors essuyée et emballée dans un film plastique, avant d'être exposée contre un film photographique Hyperfilm ECL (Amersham ; GE Healthcare) – la durée de l'exposition dépend de la qualité et de la dilution de l'anticorps, ainsi que de l'intensité de la bande souhaitée, et elle peut varier de quelques secondes à une dizaine de minutes. Le film est ensuite incubé 2 minutes dans un bain révélateur (Kodak) et 2 minutes dans un bain fixateur (Kodak), afin de visualiser et stabiliser l'image obtenue.

Après la révélation, la membrane est deshybridée pendant 30 min dans le tampon : 25 mM glycine, 1% SDS, pH 2.0 ajusté avec HCl, et rincée 3 fois 5 min dans le TBS / 0.5% Tween-20. Ensuite, le blocage de la membrane est effectué comme décrit précédemment, et l'anticorps anti-actine dilué dans TBS / 0.5% Tween-20 / 5% lait est appliqué, pendant 1 heure à température ambiante. L'application de l'anticorps secondaire et la révélation de la membrane sont faites d'après le protocole déjà décrit.

Lors des analyses antérieures effectuées au sein du laboratoire, nous avons remarqué la présence de la β-actine dans les extraits nucléaires des cellules HeLa [269]. Elle est sans doute un contaminant de nos extraits, dont nous avons toujours utilisé le même lot, car c'est une protéine cytoplasmique. La β-actine semble s'associer de manière aspécifique à la surface des billes, et elle va être alors utilisée comme contrôle de charge. Sa détection va servir à vérifier si, après l'étape du piégeage, la même quantité des protéines est présente dans tous les échantillons.

Tableau 10 : Anticorps utilisés pour les expériences de Western Blot.

anticorps	type	source	dilution	fabricant	référence
anti-PARP1	primaire	lapin	1/1000	Cell Signalling	9542
anti-DREF / (anti-ZBED1)	primaire monoclonal	souris	1/2000	Abcam	ab131594
anti-actine	primaire	souris	1/10000	Sigma	A2228
anti-souris	secondaire	chèvre	1/20000	Amersham Biosciences	NA931
anti-lapin	secondaire	singe	1/20000	Amersham Biosciences	NA934
anti-NONO	primaire monoclonal	lapin	1/20000	Abcam	ab109511
anti-SFPQ	primaire monoclonal	souris	1/2000	Abcam	ab76955
anti-HMGB1	primaire polyclonal	lapin	1/2500	Abcam	ab18256

2.2. EMSA

2.2.1. Principe

La technique de retard sur gel, appelée également gel-shift ou EMSA (Electrophoretic Mobility Shift Assay), est la méthode la plus couramment utilisée dans la caractérisation des

interactions entre l'ADN et les protéines. Son principe est basé sur les différences de vitesses de migration entre un fragment d'ADN libre, et ce même fragment lié à une protéine. La masse et la charge du complexe jouent sur la mobilité électrophorétique des molécules, et ainsi un complexe ADN-protéine va se déplacer sur gel beaucoup plus lentement que l'ADN non lié à la protéine. Cette technique permet alors la mise en évidence des interactions entre une séquence d'ADN donnée et une protéine, et de suivre également la formation des complexes, en incubant la même quantité de sonde avec des quantités croissantes de la protéine. Ainsi, l'EMSA permet de déterminer l'affinité d'une protéine pour différentes séquences d'ADN, ou bien de différentes protéines pour une même sonde nucléique.

2.2.2. Protocole détaillé

Marquage

10 pmol d'oligonucléotides non-phosphorylés à l'extrémité 5' sont marqués radioactivement avec le phosphore-32 (Perkin Elmer France), dans un volume total de 15 µL. Pour la facilité des expériences, leur volume est ajusté à 5 µL. On prépare ensuite le mix qui contient par échantillon : 1 µL d'ATP-32, 3 µl de tampon kinase 5x, 1 U de T4 kinase (Life Technologies) et H_2O jusqu'à 10 µL. Il est ensuite ajouté à l'oligonucléotide, puis incubé pendant 1 heure à 37°C. Après l'incubation, 25 µL de H_2O sont ajoutés à ce mélange, qui est ensuite purifié sur une G-25 illustra MicroSpin (GE Healthcare). On obtient ainsi l'oligonucléotide simple brin marqué radioactivement, à 250 nM, dans un volume de 40 µL.

L'oligonucléotide marqué est hybridé avec son complémentaire non-radioactif, qui est ajouté en excès afin d'assurer 100% d'hybridation. Ainsi, 60 pmol d'oligonucléotide complémentaire est ajouté à l'oligonucléotide marqué, et le volume est ajusté à 50 µL avec H_2O et le tampon d'incubation (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 150 mM KCl). Après un chauffage pendant 3 minutes à 90°C, le mélange est refroidi afin de permettre l'hybridation des séquences complémentaires. On obtient ainsi 50 µL d'oligonucléotide double brin à 200 nM, que l'on dilue ensuite 4 fois, pour obtenir une concentration de 50 nM.

Incubation

L'incubation de l'oligonucléotide double brin avec la protéine a lieu dans un volume total de 12 µL, dans le tampon d'incubation à composition finale : 20 mM Tris-HCl pH 7.5, 1 mM EDTA, 150 mM KCl, et avec 1 µL de glycérol (8% dans le mélange réactionnel). 100 fmol (8.33 nM dans le volume final) d'oligonucléotide sont ainsi incubées avec différentes concentrations de la protéine, pendant 20 min à température ambiante, avant d'être déposées sur un gel préalablement équilibré.

Une modification de ce protocole a été apportée pour l'étude des interactions avec la protéine DREF. 480 fmol d'oligonucléotide marqué sont incubées avec différentes concentrations de protéine dans un volume total de 5 µL et dans un tampon de composition finale : 20 mM HEPES, 150 mM KCl 0.1 mM EDTA, 0.5 mM DTT, 10% de glycérol, pH 7.6.

Le Tableau 11 présente les protéines recombinantes qui ont été utilisées pour effectuer les expériences de retard sur gel.

Tableau 11 : Protéines recombinantes utilisées lors des expériences de retard sur gel.

protéine	**source**	**référence**
PARP1	Abcam	ab123834
DREF	Abcam	ab132527
Fpg	Bertrand Castaing (CNRS Orléans)	[278]
tg1653	Fabrice Confalonieri (Université Paris Sud)	
tg1277		

Migration

Les gels sur lesquels va être effectuée la migration, sont constitués d'acrylamide/bisacrylamide 29 :1 (Sigma), dilué à la concentration voulue : 4% ou 8% dans le tampon TBE (Life Technologies) à 1x au final, en fonction du degré de séparation souhaité, et de la taille du gel. Deux tailles de gels ont été utilisées : les « petits » gels (8 cm de hauteur) à 4% d'acrylamide ont été surtout utilisés pour l'optimisation des conditions, tandis que les « grands » gels (40 cm de hauteur) à 8% d'acrylamide ont été utilisés pour les expériences définitives, en raison de la très bonne séparation des bandes obtenues après une longue migration (5 heures environ).

La présence du glycérol dans le mélange réactionnel "alourdit" l'échantillon et facilite son dépôt dans les puits, qui peut se faire sans ajout de formamide dénaturante – le bleu de charge est déposé indépendamment dans un puits séparé, afin de suivre la migration. Celle-ci est effectuée à 100 V pendant 1 heure sur les petits gels, et à 500 V pendant 4 heures sur les grands gels, dans le tampon TBE 1x.

Après la migration, le gel est démoulé et exposé pendant une nuit contre une plaque (Kodak), celle-ci enregistrant à sa surface le rayonnement provenant du phosphore-32. La lecture de cette dernière est faite le lendemain en utilisant un appareil Personal Molecular Imager (Bio-Rad).

2.2.3. Analyse des résultats et calculs de K_d

La quantification et l'analyse des images des gels sont effectuées à l'aide du logiciel Quantity One (Bio-Rad), qui permet de mesurer l'intensité relative des bandes visibles sur le gel et de calculer le pourcentage de cette intensité par rapport à la totalité du signal. Un exemple d'une telle quantification est présenté sur la Figure 26 :

- dans un premier temps, le long de chaque piste de migration, on trace une ligne qui traverse toutes les bandes, en passant par le milieu de chaque bande ou par sa partie la plus intense, voir : Figure 26 A ;
- on peut ensuite visualiser le profil de chaque piste, qui représente sous forme de pics la force du signal correspondant aux bandes que l'on voit sur le gel (voir : Figure 26 B) ; ce profil permet de définir le segment de la ligne qui va représenter chaque bande et sur lequel la quantification du signal va avoir lieu ; ces segments sont marqués sur le gel avec les crochets ;
- le signal total mesuré sur chaque piste est assigné comme 100%, et cette valeur est ensuite distribuée aux bandes déterminées préalablement en fonction de l'intensité du signal qui y a été détectée ; on obtient ainsi les pourcentages relatifs de la quantité de sonde libre et de sonde sous forme de complexe.

Les pourcentages relatifs sont ensuite tracés en fonction de la concentration de la protéine, ce qui peut donner deux types de graphiques : le pourcentage de l'oligonucléotide libre en fonction de la concentration de la protéine ou le pourcentage du complexe formé en fonction de la concentration de la protéine – les deux représentent la même information et sont complémentaires. Nous avons choisi la deuxième représentation, c'est-à-dire la quantité des

complexes en fonction de la quantité de la protéine. Les graphiques ont été tracés en utilisant le logiciel Origin 7.5, qui retrouve la formule mathématique la plus adaptée aux points expérimentaux et trace la courbe d'après celle-ci. Les détails des calculs effectués seront présentés dans le chapitre « Résultats ».

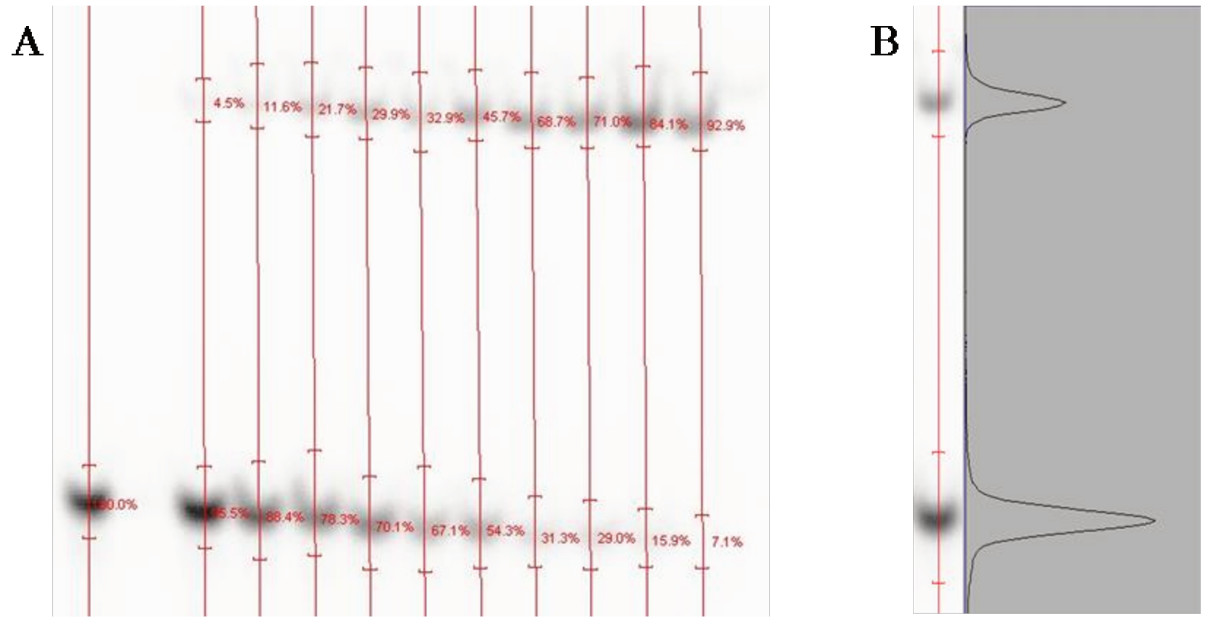


Figure 26 : A) Exemple d'un gel quantifié. Les quantités relatives des oligonucléotides libres et complexés sont présentées comme pourcentage de la quantité totale du signal détecté sur chaque piste de migration ; B) exemple de profil d'une bande ; ici : interactions entre la Fpg et la sonde 5'S-cyclo-dHT.

3. Tests d'activité enzymatique

3.1. Test de coupure sur gel

3.1.1. Principe

Le test d'activité enzymatique sur gel est une technique permettant de voir si une enzyme est capable d'effectuer une action d'excision/coupure sur un substrat donné. Des oligonucléotides contenant dans leur séquence des lésions d'intérêt, qui ont été au préalable marqués radioactivement, peuvent être incubés soit avec des protéines purifiées, soit avec des extraits protéiques, et ils subissent ensuite une migration électrophorétique sur un gel dénaturant. Si une coupure a eu lieu, un fragment court de la sonde va être généré, résultat d'une activité de l'enzyme testée.

3.1.2. Protocole détaillé

Pour effectuer les tests de coupure sur gel, les oligonucléotides sont marqués radioactivement et hybridés suivant le même protocole que celui utilisé pour l'EMSA (paragraphe 2.2.2 de ce chapitre). L'incubation avec les protéines purifiées est réalisée dans un volume de 10 µL, dans un tampon propre à chaque enzyme testée (les compositions des tampons utilisés pour toutes les enzymes sont données dans le Tableau 12), en présence de 600 fmol d'oligonucléotide marqué. L'incubation est effectuée à 37°C pendant 1 heure, à l'exception des deux enzymes provenant de *Thermococcus gammatolerans*, pour lesquelles des conditions particulières ont été adaptées : l'incubation est initiée à 37°C, puis après 30 minutes la température est augmentée jusqu'à 65°C et la réaction se poursuit encore pendant 10 minutes. Ce protocole a été conçu pour optimiser l'activité des enzymes thermophiles : il n'est pas possible de commencer l'incubation à 65°C, puisqu'à cette température-là, les oligonucléotides ne sont plus sous forme double-brin. Ainsi, la première étape effectuée à 37°C permet la reconnaissance initiale de la lésion et une certaine stabilisation du complexe ADN-protéine, avant d'augmenter progressivement la température dans le but d'améliorer éventuellement l'activité de l'enzyme.

Tableau 12 : Enzymes utilisées pour effectuer les tests de coupure, et conditions de leur incubation.

enzyme	conditions d'incubation		fournisseur / provenance
	quantité	tampon	
Fpg	3.2 U	NEBuffer 1 : 10 mM bis-tris-propane-HCl, 10 mM $MgCl_2$, 1 mM DTT pH 7	New England Biolabs
Fpg	100 ng (+50 fmol d'oligo)	« Fpg » : 20 mM Tris-HCl, 1 mM EDTA, 150 mM KCl, pH 7.8	B. Castaing
OGG1	1.6 U	NEBuffer 2 : 50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9	New England Biolabs
Endo III	3 U	Nth buffer: 20 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0	New England Biolabs
Endo IV	2.7 U	NEBuffer 3 : 100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9	New England Biolabs
Endo VIII	2.8 U	50 mM Tris, 2 mM EDTA, 50 mM KCl, pH 7.6	Sigma
APE1	3 U	50 mM KCl, 25 mM HEPES, 0.5 mM $MgCl_2$, pH 6.8	New England Biolabs
tg1653	2 pmol	« Fpg »	F. Confalonieri (Université Paris Sud)
tg1277			

Après l'incubation, 4 µL de tampon de charge (2.5 mg de bleu de bromophénol, 2.5 mg de Xylène cyanol, 200 µL de TBE 2x, qsp 1.1 mL de formamide) sont ajoutés dans chaque échantillon, qui est ensuite chauffé 3 minutes à 90°C afin de dénaturer les complexes ADN-protéine et deshybrider les oligonucléotides double-brin. Les échantillons préparés de cette manière sont alors déposés immédiatement sur gel, ou conservés à -20°C jusqu'à leur dépôt.

3.2. Analyse de masse par MALDI-TOF

3.2.1. Principe

La technique MALDI-TOF (Matrix Assisted Laser Desorption Ionisation – Time of Flight) a été décrite pour la première fois dans les années '80, et elle s'est popularisée dans les années '90 avec l'apparition des appareils commerciaux. Elle permet de déterminer la masse de nombreuses biomolécules, telles que les protéines, les polypeptides, les sucres ou encore les acides nucléiques. Son principe est basé sur l'ionisation de la molécule à analyser avec un faisceau laser, l'accélération des ions ainsi formés par un champ électrique, et ensuite leur détection.

La première étape consiste en la préparation de la cible – celle-ci doit être mélangée avec la matrice, qui sert à protéger la molécule et également à faciliter son ionisation. Le type de matrice est déterminé par la nature des molécules qui sont analysées, et son choix va dépendre du type de composé (protéines, oligosaccharides, oligonucléotides etc.) et également du genre d'information souhaitée. Pour l'analyse des oligonucléotides, trois types de matrices sont couramment utilisés : 2,4,6-trihydroxyacétophénone (THAP), acide 3-hydroxypicolinique (3-HPA) ou 6-aza-2-thiothymine (ATT).

L'échantillon est préparé en utilisant la technique de goutte séchée : après avoir été mélangée avec la matrice en phase liquide, une goutte de cette préparation est déposée sur la surface de la cible et séchée afin que tous ces ingrédients co-cristallisent. Le laser va être ensuite dirigé sur ces cristaux et il va les vaporiser et ioniser, le plus souvent par l'addition d'un proton – ainsi, des ions monochargés et, plus rarement, multichargés, vont être formés. Ces ions sont ensuite accélérés par un champ électrique et le temps qui leur est nécessaire pour atteindre la source est mesuré – ce temps dépend du rapport masse moléculaire/charge ($[M-H]^{-1}$), et ce sont les grosses molécules qui vont avoir une moindre vitesse. Ainsi, la masse précise de la molécule, ou des molécules si l'on procède à l'analyse d'un mélange, peut être déterminée.

Dans nos expériences, cette technique a été utilisée pour déterminer la taille des fragments d'un oligonucléotide obtenus après la coupure effectuée par une enzyme. Connaissant la masse initiale de l'oligonucléotide et celle des fragments hypothétiques, la détermination de la taille des fragments effectivement obtenus et sa comparaison avec les masses théoriques calculées, permet d'identifier le mécanisme de coupure qui est utilisée par la protéine en question.

3.2.2. Protocole détaillé

Le mécanisme de coupure de deux enzymes de *T.gammatolerans* : tg1653 et tg1277 a été étudié avec la technique MALDI-TOF/MS. Deux autres enzymes : Fpg et Endo III d'*E.coli* ont été utilisées en tant que contrôles.

Afin de vérifier la taille des fragments obtenus après la coupure des oligonucléotides lésés par différentes enzymes, les expériences ont été menée sur quatre sondes : oligonucléotides contenant la 8-oxo-dG, le THF, la 5-OH-dHyd (5-hydroxy-désoxyhydantoïne ; cf. chapitre 2.2.1.3 de l'« Analyse bibliographique ») et la 5-OH-5-Me-dHyd (5-méthyl-5-hydroxy-désoxyhydantoïne ; cf. chapitre 2.2.1.2 de l'« Analyse bibliographique »), avec l'utilisation de quatres enzymes : tg1653, tg1277, Fpg et Endo III. 10 pmol d'oligonucléotide sont incubées avec 20 pmol de protéine dans le tampon « Fpg » (20 mM Tris-HCl, 1 mM EDTA, 150 mM KCl, pH 7.8) et dans un volume final de 20 µL. La réaction est réalisée pour les quatre protéines dans les mêmes conditions que celles utilisées pour les tests de coupure des enzymes de *T.gammatolerans* sur gel : 30 min à 37°C, ensuite la température est augmentée progressivement à 65°C, puis maintenue 10 min à 65°C. Les échantillons sont stockés à -20°C en attendant leur analyse.

Pour pouvoir analyser un échantillon par MALDI-TOF/MS, il est nécessaire d'éliminer les sels qui sont présents dans le mélange. Les ions tels que Na^+ ou K^+ compliquent l'analyse de la masse des composés, qu'ils rendent également moins volatils. Afin de faciliter l'analyse, ces ions doivent être échangés contre les ions NH_4^+, ce qui peut être effectué en utilisant une résine échangeuse d'ions. La résine que nous utilisons est de type Dowex (50WX8 100-200 ; Sigma-Aldrich) : elle est présente sous forme H^+, et elle doit être tout d'abord échangée sous forme NH_4^+, avant de pouvoir être utilisée pour dessaler les échantillons. L'échange est effectué en équilibrant les billes de la résine dans une solution d'acétate d'ammonium à 0.1 M, en les rinçant ensuite dans de l'isopropanol et en les laissant sécher à l'air pendant 1 nuit. Une résine préparée de cette manière peut être stockée et utilisée ultérieurement.

La matrice utilisée pour l'analyse des oligonucléotides est l'acide 3-hydroxy-picolinique. 30 à 40 mg de cette matrice sont pesés et solubilisés dans 1 mL de citrate d'ammonium à 5 mM. La solution est ensuite traitée pendant 5 min aux ultrasons afin d'éliminer des éventuels cristaux ou agrégats.

10 µL d'échantillon à dessaler sont déposés sur les billes de résine préalablement équilibrées et laissés ainsi pendant 10 min environ. 1 µL d'échantillon est ensuite prélevé et déposé sur la surface de la cible. Le même volume de la matrice préparée comme décrit ci-dessous, est déposé par-dessus et ce mélange est laissé sécher, avant d'être analysé.

L'analyse est effectuée avec un appareil Microflex (Bruker) en mode linéaire négatif. Les échantillons sont désorbés et ionisés à l'aide d'un laser d'azote à la longueur d'onde de 337 nm, et les ions sont ensuite accélérés avec une tension de 20 kV.

4. Dosage des lésions de l'ADN

4.1. Principe général

Le dosage des lésions au sein de l'ADN est effectué à l'aide d'une séparation chromatographique couplée à un spectromètre de masse en tandem : HPLC-MS/MS (High Performance Liquid Chromatography/ tandem mass spectrometry). Cette technique permet à la fois d'identifier le dommage et de le quantifier avec une grande spécificité et une grande sensibilité.

Afin de pouvoir doser les lésions au sein de l'ADN, celui-ci doit être extrait des cellules. Les protocoles d'extraction varient en fonction des laboratoires et, surtout, des organismes utilisés. Dans le cas du dosage des lésions d'oxydation, un soin particulier doit être apporté à la préparation de l'échantillon, pour éviter une oxydation artefactuelle de l'ADN au cours de son extraction et de sa digestion. Afin de l'éviter, l'ajout de déféroxamine dans les tampons utilisés peut être pratique : la déféroxamine est un chélateur de fer qui permet de limiter les réactions d'oxydation catalysées par celui-ci. La digestion de l'ADN sert à obtenir un mélange de nucléosides endommagés ou non (« normaux »), dont l'identification va ensuite pouvoir être réalisée.

La première étape de cette identification consiste en une séparation des molécules présentes dans le mélange digéré au préalable, en utilisant une colonne HPLC. Les nucléosides normaux ainsi que les nucléosides modifiés sont ainsi séparés en fonction de leur temps de

rétention sur la colonne, qui dépend des caractéristiques physico-chimiques des molécules analysées.

Le dosage des nucléosides non-endommagés s'effectue avec un détecteur UV, tandis que les nucléosides modifiés nécessitent une méthode de détection plus sensible et plus spécifique, due à leur relativement faible quantité dans les mélanges. Ainsi, leur quantification est réalisée à l'aide d'un spectromètre de masse. Dans un premier temps, l'échantillon est ionisé à sa sortie de la colonne chromatographique. L'ion que l'on souhaite analyser est identifié grâce à son rapport masse/charge, et séparé des autres ions à l'aide des électrodes constituant un quadripôle. Cet ion est par la suite accéléré par l'application d'un champ électromagnétique, et il est dirigé vers le détecteur. Dans un deuxième temps, si l'on dispose d'un spectromètre de masse en tandem (MS/MS), une fragmentation de l'ion sélectionné (ion parent) en ions fils a lieu, et ces derniers sont détectés de la même manière que précédemment décrit. Cette deuxième fragmentation permet de confirmer l'identité de l'ion parent, en analysant les fragments générés, qui constituent sa « signature », et également d'améliorer la sensibilité de la détection, en passant d'1 base modifiée /10^5 nucléosides normaux à 1 base modifiée /10^7 nucléosides normaux.

Il faut néanmoins préciser que toutes les modifications de l'ADN ne peuvent pas être dosées. La disponibilité d'un standard est un prérequis, et celui-ci doit souvent être synthétisé, notamment quand il s'agit d'une lésion rare. Ainsi, l'analyse des lésions originales n'est accessible qu'à des laboratoires possédant des connaissances approfondies en chimie de synthèse des molécules standards, en plus de leur équipement en systèmes de détection des dommages.

4.2. Protocoles détaillés

4.2.1. Culture de *Thermococcus gammatolerans*

Thermococcus gammatolerans a été cultivé par l'équipe de Fabrice Confalonieri dans le laboratoire de Génomique des Archaea à l'Université Paris-Sud, comme décrit précédemment [297].

Les cellules sont cultivées en anaérobie à 85°C dans le milieu complexe VSM (20 g/L de NaCl, 0.25 g/L de KCl, 0.05 g/L de NaBr, 0.02 g/L d'acide borique, 0.01 g/L de $SrCl_2$•$6H_2O$, 0.5 g/L de citrate de trisodium, 3 g/L de Pipes, 1 g/L d'extrait de levure, 4 g/L de bactotryptone, 5 mL de $MgSO_4$ à 20%, 1 mL de $CaCl_2$ à 5% et 1 mL de KH_2PO_4 à 5%, pH 6.8, supplémenté en

2 g/L de soufre inorganique). Les cultures sont inoculées à 2×10^5 cellules/mL dans 25 mL de milieu et elles atteignent la phase exponentielle à 2×10^7 cellules/mL.

Les irradiations sont effectuées en phase exponentielle, dans le milieu complexe VSM, en anaérobie et sur la glace. Une dose de 2.5 ou de 5.0 kGy est délivrée par la source de ^{137}Cs (IBL637 Cis Bio International, Institut Curie, Orsay) avec le débit de 30 Gy/min. Les contrôles non-irradiés sont incubés sur la glace pendant les temps équivalents aux temps d'irradiations. Les cellules sont soit récoltées immédiatement après l'irradiation, soit remises en culture pendant 1 heure. Elles sont ensuite centrifugées 20 min à 2000g et le culot sec est stocké à -80°C.

4.2.2. Extraction et digestion de l'ADN de *Thermococcus gammatolerans*

Le protocole de l'extraction d'ADN utilisé au sein de notre laboratoire a été optimisé pour des cellules eucaryotes et il est présenté ci-dessous. Les premières tentatives d'extraction de l'ADN de *T.gammatolerans* en utilisant ce protocole ont cependant montré qu'il n'était pas adapté aux organismes type archées. Ainsi, il a dû être modifié, surtout en ce qui concerne la composition du tampon de lyse et la quantité des enzymes utilisées. Après de nombreux essais, qui vont être détaillés dans le chapitre « Résultats », un protocole définitif a été établi.

Extraction de l'ADN des cellules eucaryotes

Les cellules sont transférées dans des tubes et centrifugées 5 minutes à 1500 g, ensuite le surnageant est éliminé et le culot sec est préservé à -80°C jusqu'à l'extraction de l'ADN. L'ADN est extrait suivant le protocole mis au point dans le laboratoire. Le culot est repris dans 750 µL de tampon de lyse A (320 mM sucrose, 5 mM MgCl2, 10 mM Tris-HCl, 0.1 mM déféroxamine, 1% Triton X100, pH 7.5), centrifugé 5 minutes à 1500 g, et cette opération est répétée une fois. Ensuite le culot est repris dans 300 µL de tampon de lyse B (5 mM EDTA-Na2, 10 mM Tris-HCl, 0.15 mM déféroxamine, pH 8) et vortexé, après quoi 18 µL de SDS à 10% sont ajoutés et le mélange est vortexé encore une fois. 1.5 µL de RNAse A à 100 mg/mL (Sigma) et 3.5 µL de RNAse T1 à 1U/µL (Sigma) sont ensuite ajoutés et le tout est incubé pendant 15 minutes à 50°C. Après ce temps, 15 µL de protéase à 20 mg/mL (Qiagen) sont ajoutés et l'incubation se poursuit pendant 1 heure à 37°C. L'ADN est précipité par l'ajout de 600 µL de solution de précipitation (20 mM EDTA-Na2, 7.6 M NaI, 40 mM Tris-HCl, 0.30 mM déféroxamine) et de 1 mL

d'isopropanol à 40%. Les tubes sont doucement agités, jusqu'à l'apparition du culot blanc d'ADN. Après 5 minutes de centrifugation à 5000 g, celui-ci est ensuite rincé avec 500 µL d'isopropanol à 40%, centrifugé 5 minutes à 5000 g et rincé dans 500 µL d'éthanol à 70%. Après la dernière centrifugation de 5 minutes à 5000 g, le culot est repris dans 50 µL de déféroxamine à 0.1 mM – il peut être ensuite congelé à -20°C, ou bien être utilisé dans la continuité pour effectuer la digestion de l'ADN.

Extraction de l'ADN de *Thermococcus gammatolerans*

10 mg de culot sont suspendus dans 300 µL de tampon de lyse (100 mM Tris-HCl, 50 mM EDTA, 100 mM NaCl, 0.15 mM déféroxamine, pH 8) et homogénéisés. 30 µL de SDS à 10% et 30 µL de sarcosyl à 10% sont ensuite ajoutés et le tout est vortexé. 1.5 µL de RNAse A à 100 mg/mL (Sigma) et 3.5 µL de RNAse T1 à 1U/µL (Sigma) sont ensuite ajoutés et le tout est incubé pendant 15 minutes à 50°C. 30 µL de protéase à 20 mg/mL (Quiagen) sont ajoutés par la suite, le mélange est homogénéisé et incubé pendant 1.5 heures à 37°C. La précipitation et la purification de l'ADN sont effectuées comme décrit précédemment pour les cellules eucaryotes.

Digestion de l'ADN

La digestion de l'ADN préalablement extrait est effectuée d'après le protocole mis au point dans le laboratoire. Dans chaque échantillon d'ADN solubilisé dans la déféroxamine, sont ajoutés : 0.25 µL de phosphodiestérase II (Sigma) à 0.1 U/µL, 0.25 µL de DNAse II (Sigma) à 10 U/µL, 2.5 µL de nucléase P1 (Sigma) à 0.2 U/µL et 2.5 µL de tampon MNSPDE (200 mM acide succinique, 100 mM $CaCl_2$, pH 6). L'incubation se poursuit pendant 2 heures à 37°C, avec une homogénéisation du mélange après 30 minutes. Dans la deuxième étape de la digestion, sont ajoutés : 6 µL de tampon de phosphatase alcaline (10x : 500 mM Tris, 1 mM EDTA, pH 8), 0.5 µL de phosphodiestérase I (Worthington) à 0.03 U/µL et 2 U de phosphatase alcaline (Sigma). Ensuite, l'incubation de 2 heures à 37°C a lieu, et après ce temps la digestion est arrêtée par l'ajout de 3.5 µL de HCl à 0.1 N. L'échantillon est ensuite centrifugé pendant 5 minutes à 5000 g afin de séparer d'éventuels fragment non-solubles, et le surnageant contenant de l'ADN digéré est transféré dans des vials HPLC. Si le dosage des cyclonucléosides est envisagé par la

suite, la répétition de la première étape de la digestion (phosphodiestérase II, DNAse II, nucléase P1) avant d'ajouter le HCl, est nécessaire.

4.2.3. Dosage des lésions

Deux types de lésions de l'ADN ont été dosés au cours de ces travaux de thèse : la 8-oxo-dG et les cyclonucléosides. Ces dommages possèdent des caractéristiques bien distinctes, et leur détection et quantification ne sont pas effectuées sur le même appareil. C'est pourquoi les deux protocoles de leurs dosages vont être présentés séparément.

4.2.3.1. Dosage de la 8-oxo-désoxyguanosine

La 8-oxo-dG a été détectée avec un système HPLC couplé à un spectromètre de masse en mode tandem. La séparation des mélanges digérés a été effectuée avec le système HPLC Agilent series 1100 sur les colonnes Uptispher 2 × 150 mm de gel de silicium octadecylsilyle (Interchim) avec des particules de 5 μm. La détection des nucléosides normaux a été effectuée avec un détecteur UV à 260 nm. Les nucléosides modifiés ont été détectés avec un détecteur de masse API 3000 (PerkinElmer/SCIEX) équipé d'une source d'ionisation par électrospray. Les nucléosides ont été séparés dans un gradient linéaire de 2 à 10% de méthanol dans le formiate d'ammonium pendant 25 minutes et un deuxième gradient linéaire de 10 à 25% de méthanol dans le formiate d'ammonium pendant 10 minutes. Le temps de rétention de la 8-oxo-dG dans ces conditions est de 25 minutes environ. Les ions ont ensuite été détectés en utilisant la transition 284 → 168, qui correspond à la cassure de la liaison C1'−N7 entre la base et le sucre (Figure 27).

La quantification a été effectuée grâce aux solutions standards : 8-oxo-dG à 0.05 et 0.1 μM, et le mélange de dT et dG à 0.25, 0.5 et 1 mM.

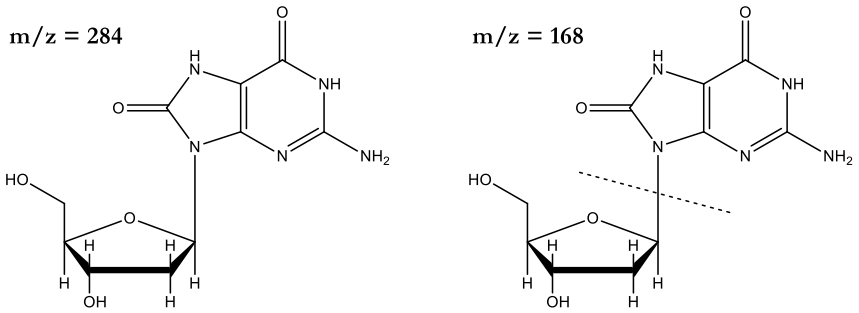


Figure 27 : Fragmentation de l'ion parent de la 8-oxo-dG.

4.2.3.2. Dosage des cyclonucléosides

Les 5',8-cyclo-2'-désoxyribonucléosides ont été détectés avec un système HPLC couplé à un spectromètre de masse en tandem. La séparation des mélanges digérés a été effectuée avec le système HPLC Accela (Thermo Corporation) sur les colonnes Uptispher 2 × 150 mm de gel de silicium octadecylsilyle (Interchim) avec des particules de 3 µm. La détection des nucléosides normaux a été effectuée avec un détecteur UV à 260 nm. Les nucléosides modifiés ont été détectés à l'aide d'un détecteur de masse TSQ Quantum Ultra équipé d'une source d'ionisation par électrospray. Les nucléosides ont été séparés par gradient linéaire de 0 à 15% d'acétonitrile dans le formiate d'ammonium pendant 30 minutes, les temps de rétention des cyclonucléosides dans ces conditions sont les suivants : 6.8 min pour 5'R-cyclo-dG, 10.6 min pour 5'S-cyclo-dG, 9.3 min pour 5'R-cyclo-dA et 14.3 min pour 5'S-cyclo-dA. Les ions ont ensuite été détectés en utilisant leurs transitions les plus intenses : 250 → 164 pour cyclo-dA et 266 → 180 pour cyclo-dG (quels que soient leurs isomères) en utilisant l'énergie de collision de 18 eV. Elle correspond à la cassure des liaisons C1'−N7 et C4'−C5' entre la base et le sucre. Une deuxième fragmentation, bien que mineure, est également utilisée : 250 → 134 pour cyclo-dA et 266 → 150 pour cyclo-dG. Elle est obtenue pour l'énergie de collision de 38 eV et elle correspond à la cassure des liaisons C1'−N7 et C5'−C8 entre la base et le sucre (Figure 28).

La quantification des quatre 5',8-cyclo-2'-désoxyribonucléosides a été effectuée grâce aux solutions standards. Les limites de quantification sont de 5 fmol pour les deux stéréoisomères de cyclo-dA et de 100 fmol pour la cyclo-dG [194].

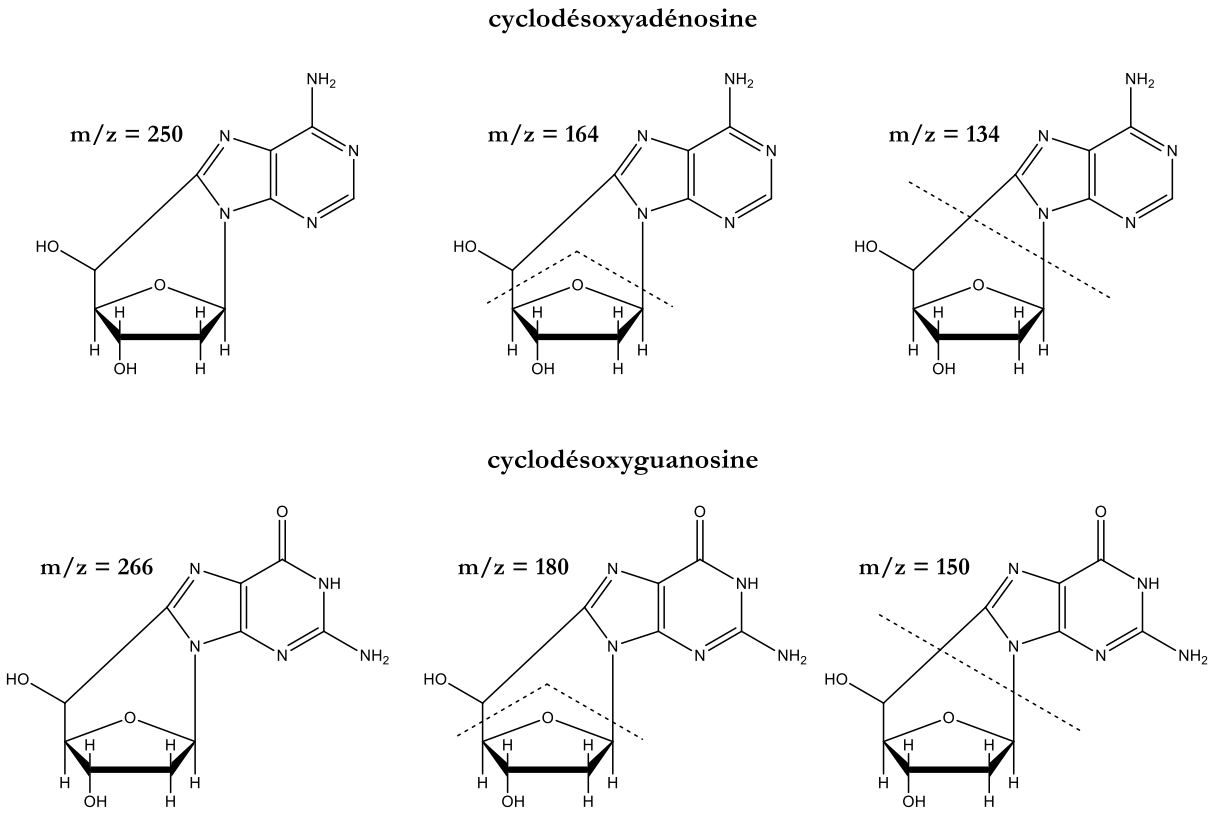


Figure 28 : Fragmentation des ions parents de la cyclodésoxyguanosine et de la cyclodésoxyadénosine.

4.2.4. Tests statistiques

Afin de comparer les niveaux des lésions dosés dans différentes conditions (cellules irradiées vs contrôles, ou temps 0 h vs 1 h), le test non-paramétrique de Mann-Whitney a été appliqué. Les calculs ont été effectués à l'aide du logiciel Statistica 8. La différence entre deux valeurs est considérée comme statistiquement significative pour $p < 0.05$.

IV Résultats

1. Screening des extraits protéiques dans le but d'identifier l'interactome des cyclonucléosides

La première partie des travaux décrits ici a pour objectif d'identifier l'interactome des cyclonucléosides, à savoir les protéines qui sont capables de s'associer avec ces dommages particuliers de l'ADN. Pour faire ceci, les extraits protéiques ont été incubés avec les pièges composés des billes magnétiques sur lesquelles différentes sondes sont greffées, et les protéines piégées de cette manière ont été identifiées via une analyse par spectrométrie de masse.

1.1. Premier format de piège testé : courtes sondes oligonucléotidiques

Le premier format des pièges qui a été testé, contient des courtes sondes synthétisées chimiquement, dans lesquelles des lésions d'intérêt sont incorporées.

Les expériences décrites ci-dessous ont été réalisées avec les extraits nucléaires HeLa.

1.1.1. Mise au point du bras photoclivable

Le piège à protéines que nous avons utilisé, dans sa version modifiée, a été mis au point au sein du laboratoire pour analyser les interactions entre les dommages de l'ADN générés par les dérivés du platine, et les protéines les reconnaissant [269]. Malgré son efficacité dans l'identification des nouvelles protéines prenant en charge des dommages de l'ADN, ce piège présente un inconvénient, qui consiste en la récupération d'une grande quantité des protéines non-spécifiques. Ces protéines se fixent sur la surface des billes, et sont ensuite éluées avec les protéines présentes sur les sondes, car l'élution s'effectue par le chauffage à 90°C dans un tampon dénaturant du type Laemmli, ce qui libère toutes les protéines quelle que soit leur localisation. Le bruit de fond qui est ainsi généré peut nuire à la qualité de l'identification de nouveaux candidats intéressants, car les protéines minoritaires ou à faibles interactions avec l'ADN peuvent être cachées par l'abondance des protéines non-spécifiques. Ce problème, déjà présent pour les longues sondes plasmidiques permettant le piégeage d'une grande quantité des protéines

spécifiques, va être encore plus important quand les sondes courtes vont être appliquées. Le rapport entre les protéines pertinentes et le bruit de fond va devenir plus défavorable, et l'identification des protéines intéressantes plus difficile. C'est pour cette raison que nous avons envisagé d'introduire dans nos pièges un « bras photoclivable ». Cette structure, présentée sur la Figure 29, est facilement dégradable par irradiation UV. Sa présence entre la biotine, permettant la fixation de la sonde sur la surface de la bille, et la sonde elle-même, rend possible la libération uniquement de l'oligonucléotide et des protéines fixées dessus, sans récupérer les protéines présentes sur la surface des billes.

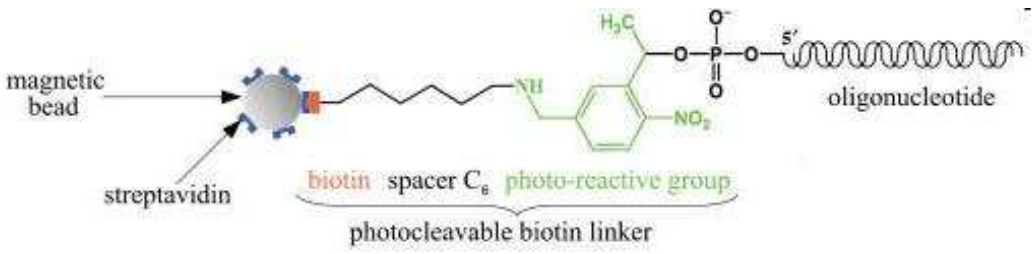


Figure 29 : Structure d'un bras photoclivable introduit entre la biotine et la séquence de l'oligonucléotide ; source : [300]

Afin de vérifier l'efficacité de la libération des sondes de la surface des billes, nous avons suivi le protocole décrit précédemment [300], en faisant varier la concentration des billes lors de l'irradiation. Pour chacune des deux concentrations testées (10 µg/µL et 3 µg/µL), nous avons effectué 5 irradiations consécutives : après chaque irradiation, l'éluat contenant les sondes décrochées a été récupéré, et les billes ont été reprises dans un nouveau volume du tampon et ont subi une nouvelle irradiation. Toutes les élutions ont ensuite été déposées sur un gel d'agarose. La visualisation des sondes a été faite avec du bromure d'éthidium (Figure 30).

Nous pouvons constater que l'efficacité de la rupture du bras photoclivable dépend de la concentration des billes. Si celles-ci sont plus concentrées, 2 voire 3 irradiations sont nécessaires afin de décrocher la totalité des sondes de la surface des billes. Quand les billes sont reprises dans un plus grand volume du tampon, leur disponibilité aux irradiations semble accrue, et toutes les sondes sont décrochées de leur support après la première irradiation. Ainsi, c'est cette deuxième condition qui a été retenue pour la suite des expériences.

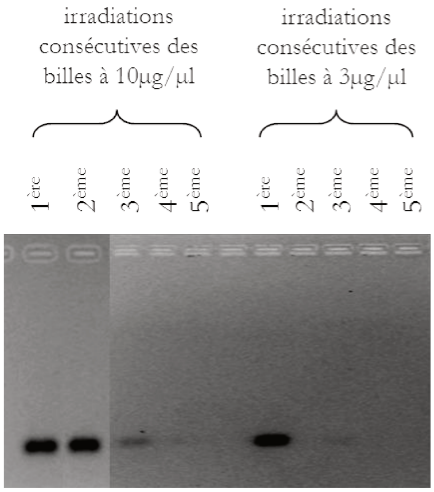


Figure 30 : Efficacité du décrochage des sondes contenant un bras photoclivable à l'aide des irradiations UV. Les billes ont été irradiées dans différents volumes de tampon, et 5 irradiations consécutives ont été effectuées. Migration sur un gel d'agarose, visualisation avec du bromure d'éthidium.

1.1.2. Essais de piégeage de HMGB1 sur les adduits du cisplatine

Afin de vérifier si le piège contenant de courtes sondes oligonucléotidiques était fonctionnel, nous l'avons appliqué à une lésion que nous avions déjà analysée au sein du laboratoire, à savoir un pontage entre deux désoxyguanosines adjacentes généré par le cisplatine [268,269]. Dans cette étude ultérieure, de longues sondes plasmidiques contenant une quarantaine de lésions générées par les dérivés du platine ont permis de déterminer la reconnaissance de ces dommages par la protéine HMGB1. Ainsi, cette protéine nous a servi de contrôle positif pour la validation de nos pièges.

Nous avons effectué un piégeage des protéines en utilisant les mêmes conditions (temps et température de l'incubation, composition du tampon) que celles utilisées pour les sondes plasmidiques [269]. Trois types de sondes oligonucléotidiques ont été utilisés :

- cis-Pt, contenant un pontage entre deux guanosines adjacentes
- non-lésées, et
- les billes sans sonde.

Les sondes ont été attachées à la surface des billes fonctionnalisées à l'aide de la biotine, et entre la biotine et la séquence de la sonde, un bras photoclivable a été introduit. Ainsi, les protéines ont été récupérées avec une irradiation UVA, dans les conditions optimisées d'après le paragraphe 1.1.1 de ce chapitre.

Les protéines récupérées ont ensuite été déposées sur un gel SDS-PAGE. Après la migration, une détection par la technique de western blot de la protéine HMGB1 a été effectuée (Figure 31). Nous pouvons constater que cette protéine a été piégée uniquement sur les sondes contenant une lésion, et nous ne la détectons pas sur les sondes non-lésées ou sur les billes. Ceci confirme la validité de notre piège, qui nous a permis d'attraper une protéine dont les interactions étaient connues. Ainsi, on va pouvoir l'utiliser pour identifier des nouvelles protéines étant capables de reconnaître d'autres dommages de l'ADN, que l'on peut insérer dans les oligonucléotides.

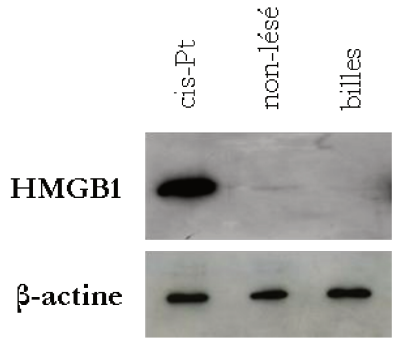


Figure 31 : Identification par western blot de la protéine HMGB1 piégée sur la sonde oligonucléotidique contenant une lésion générée par le cisplatine (cis-Pt) ; les sondes non-lésées et les billes servent de contrôles négatifs ; la β-actine est utilisée pour vérifier la charge égale de tous les échantillons en protéines.

1.1.3. Analyse des protéines piégées sur les cyclonucléosides

Le piégeage de la protéine HMGB1 sur les sondes oligonucléotidiques contenant une lésion générée par le cisplatine, nous a permis de confirmer la validité de notre piège. Nous avons alors procédé à l'exposition des sondes contenant un cyclonucléoside aux extraits protéiques, afin d'identifier les protéines capables de reconnaître ces dommages. Trois types de pièges ont été utilisés pour ceci : les sondes 5'S-cyclo-dG, les sondes non-lésées et les billes. Le bras photoclivable a été introduit entre la biotine et la séquence de la sonde, et les protéines ont été récupérées par irradiation UVA.

Dans un premier temps, nous avons exposé ces pièges aux extraits nucléaires des cellules HeLa, et l'identité des protéines récupérées sur chacune des sondes a été analysée par la spectrométrie de masse (analyses effectuées par le laboratoire EDyP). La comparaison des protéines identifiées dans cette analyse est présentée sur la Figure 32.

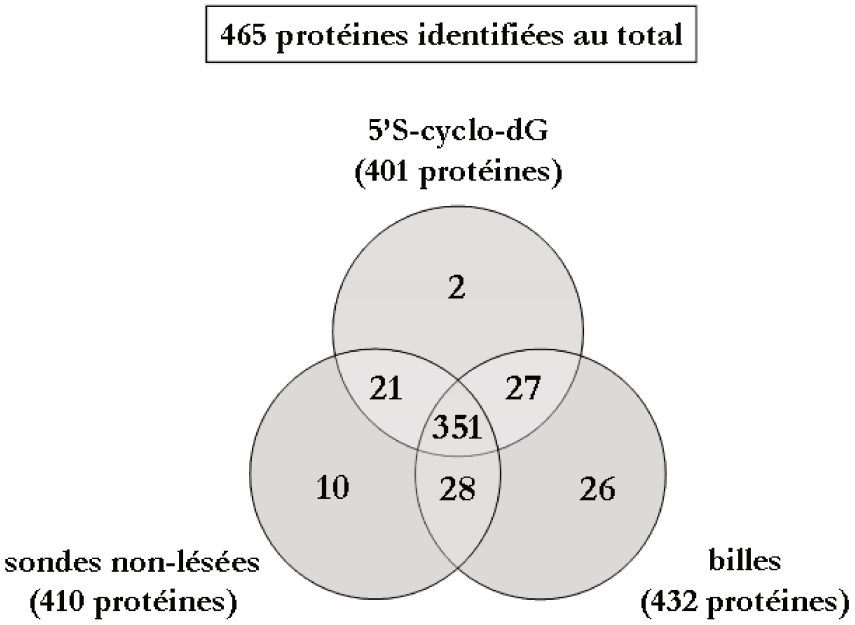


Figure 32 : Comparaison des protéines provenant des extraits nucléaires HeLa, identifiées sur les sondes 5'S-cyclo-dG ou non-lésées, récupérées par irradiation UVA.

Parmi les 465 protéines identifiées, 75% (351) ont été retrouvées sur tous les trois pièges, ce qui indique le caractère non-spécifique de leurs interactions. Parmi les 25% restant, 26 protéines sont présentes uniquement sur les billes, 21 sont retrouvées sur les sondes lésées ou non, et seulement 2 et 10 protéines sont identifiées comme spécifiques pour une sonde 5'S-cyclo-dG ou une sonde non-lésée, respectivement.

Pour chaque protéine identifiée, deux valeurs importantes doivent être considérées : la première est le nombre de peptides qui ont servi à son identification, et la deuxième est le score Mascot – un chiffre qui correspond à la fiabilité de son identification. Plus ces deux valeurs sont élevées et plus l'identification de la protéine est fiable. Une valeur seuil pour la fiabilité de l'identification des protéines a été fixée à 2 peptides et un score Mascot de 70 au minimum [302].

Les deux protéines spécifiques pour les sondes 5'S-cyclo-dG sont : cytochrome b5 identifié avec 2 peptides et un score Mascot de 130.72, et cleavage stimulation factor subunit 1 identifié avec 2 peptides et un score de 117.37. Ces protéines ne sont cependant pas pertinentes du point de vue biologique, car elles ne jouent aucun rôle dans la reconnaissance ou la prise en charge des dommages de l'ADN. On en conclut alors que ces deux candidats ont été retrouvés sur les sondes lésées de manière aléatoire et nous n'allons pas poursuivre l'analyse de leurs interactions avec les cyclonucléosides.

Parmi les 10 protéines dont les interactions avec l'ADN non-lésé ont été constatées et qui semblent ne plus interagir avec celui-ci en présence d'un cyclonucléoside, aucune ne présente un rôle biologique lié à la prise en charge des dommages de l'ADN. Comme nous l'avions déjà constaté pour les protéines identifiées uniquement sur les sondes lésées, ici aussi aucun candidat pertinent n'a été retrouvé. Le système de piège avec une courte sonde nucléotidique n'est ainsi pas adapté au screening des protéines interagissant avec ce type de dommages de l'ADN. Il s'était avéré fonctionnel pour les sondes portant un pontage généré par le cisplatine, et il nous a permis de retrouver une protéine connue pour ses interactions avec ce type de dommages (HMGB1), mais il est possible que les protéines minoritaires dans les extraits nucléaires, ou celles dont l'affinité pour le dommage donné est faible, ne peuvent pas être identifiées par ce biais. Une des explications possibles de ce phénomène serait un faible nombre de lésions (1 par sonde), qui limite considérablement la quantité globale des protéines qui peuvent être fixées sur les pièges. Aussi, les sondes que nous utilisons sont relativement courtes (22 paires de bases) par rapport aux plasmides utilisées dans les expériences précédentes [269] (3000 paires de bases environ), ce qui va limiter, voire rendre impossible, les interactions des protéines qui nécessitent une plus longue séquence de l'ADN afin d'exercer leur action.

Les résultats provenant de cette expérience de piégeage nous ont incités à repenser le système des pièges que nous utilisons. Les deux problèmes majeurs dont nous devons nous affranchir sont la petite taille des sondes et un faible nombre de lésions.

1.2. Deuxième format de piège testé : concatémères

Comme ceci a été présenté dans le chapitre « stratégie et objectifs », nous allons concevoir des sondes plus longues, appelées « concatémères », qui vont être constituées de plusieurs oligonucléotides collés les uns aux autres. Ainsi, la même séquence nucléotidique va être répétée plusieurs fois au sein de la même sonde, ce qui va non seulement augmenter sa longueur, mais également multiplier le nombre des cyclonucléosides introduits.

1.2.1. Mise au point des conditions de la phosphorylation des oligonucléotides

Afin de pouvoir générer les concatémères, les oligonucléotides qui vont les former doivent posséder à leurs extrémités 5' un groupement phosphate. Les sondes « concat-1 », « concat-2 », « stop-1 », « stop-2 » et « concat-biot » (voir : Figure 24 dans le chapitre « Stratégie et objectifs ») ont été synthétisées directement avec une modification 5'-phosphate, cependant les sondes contenant des cyclonucléosides ayant déjà été synthétisées avec un groupement –OH en 5', elles ont nécessité une phosphorylation enzymatique. Nous avons ainsi décidé de phosphoryler ces oligonucléotides par action de la T4 kinase, de manière analogue au marquage radioactif. Quand les oligonucléotides sont marqués avec le phosphore-32, il suffit qu'une partie de ceux-ci soit effectivement marquée, puisque cette fraction va être représentative du comportement de tous les oligonucléotides, qu'ils soient marqués ou non. C'est pourquoi l'efficacité du marquage n'est pas prioritaire pour ce type d'expérience. Au contraire, pour former les concatémères, il faut impérativement que toutes les sondes contiennent un groupement 5'-phosphate. Dans le cas opposé, la ligation des sondes ne va pas pouvoir se réaliser sur toute la longueur des concatémères, ce qui va générer des cassures simple- ou double-brin dans leur séquence.

Pour trouver des conditions optimales, permettant d'obtenir 100% de phosphorylation des oligonucléotides, nous avons testé différentes conditions, en faisant varier notamment :

- le temps d'incubation : 20 minutes, 1 heure ou 2 heures
- la température d'incubation (23°C ou 37°C)
- la concentration de l'ATP : 10 μM, 1 mM ou 2 mM
- la quantité de kinase : entre 0.001 U et 1 U pour 1 pmol d'oligonucléotide.

La phosphorylation modifie la masse et la charge de l'oligonucléotide. Ainsi, les formes phosphorylée et non-phosphorylée peuvent être séparées sur un gel dénaturant d'acrylamide. Pour pouvoir visualiser l'efficacité de la phosphorylation, nous avons utilisé un oligonucléotide contenant une fluorescéine en extrémité 3' (voir : Tableau 7 dans le chapitre « Matériel et méthodes »). L'usage de la radioactivité a été exclu, puisque le marquage au ^{32}P s'effectue sur la même extrémité de la sonde que la phosphorylation que nous avons voulu évaluer.

Sur la Figure 33 est montrée la dernière étape de cette mise au point, avec la température et la quantité de l'ATP optimisées comme précisé dans le protocole (37°C, 10 mM d'ATP ; voir paragraphe 1.3.3 dans le chapitre « Matériel et méthodes »), en faisant varier la quantité de kinase ainsi que le temps d'incubation.

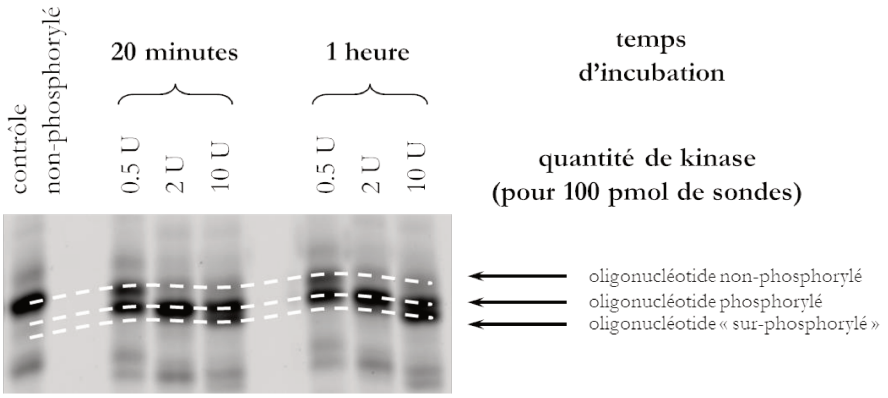


Figure 33 : Migration des oligonucléotides marqués à la fluorescéine à leur extrémité 3', et phosphorylés à l'extrémité 5' avec différentes quantités de la T4 kinase ; deux temps de marquage ont également été testés.

Nous voyons dans un premier temps que le profil de la phosphorylation est identique quel que soit le temps de l'incubation. Les modifications obtenues après 20 minutes d'action de la T4 kinase ne semblent pas être altérées par une incubation prolongée. C'est pour cette raison que nous avons décidé de sélectionner le temps plus court, suffisant pour l'obtention de la phosphorylation souhaitée. Nous pouvons ensuite constater que différents produits, seuls ou en mélanges, sont visibles sur le gel en fonction de la quantité de l'enzyme appliquée. Pour la plus petite concentration de kinase – 0.5 U pour 100 pmol d'oligonucléotides – le produit non-phosphorylé de départ est toujours présent, même si une partie des oligonucléotides a été phosphorylée (les oligonucléotides phosphorylés migrent plus vite malgré leur masse plus importante, car le groupement phosphate leur confère une charge négative). Quand 2 U de kinase sont appliquées, on constate la disparition totale de la sonde de départ et la présence d'un seul produit – l'oligonucléotide phosphorylé. Si l'on augmente encore la quantité de kinase – 10 U – un nouveau produit apparaît en plus de la sonde phosphorylée. Celui-ci migre encore plus vite sur le gel, ce qui signifierait la présence d'une double charge (oligonucléotide bi-phosphorylé) ou d'une autre modification dont nous ignorons la nature.

Afin de confirmer la phosphorylation des oligonucléotides visualisés sur le gel, nous avons analysé la masse de ceux-ci à l'aide d'un spectromètre de masse MALDI-TOF/MS. Dans un premier temps, nous avons vérifié l'intégrité de l'oligonucléotide fluorescent non-phosphorylé. Son spectre (Figure 34) montre la présence de deux pics, à $[M-H]^-$ 5149.2 et 5164.8, soit une différence de 16 Da environ (pour une masse théorique calculée de 5165.6 Da). Nous ignorons la nature de ce dédoublement de pic. Ceci ne nous empêche pas de poursuivre la mise au point des conditions de phosphorylation, même si ces deux pics ont été obtenus tout au long des analyses de masse par MALDI-TOF/MS.

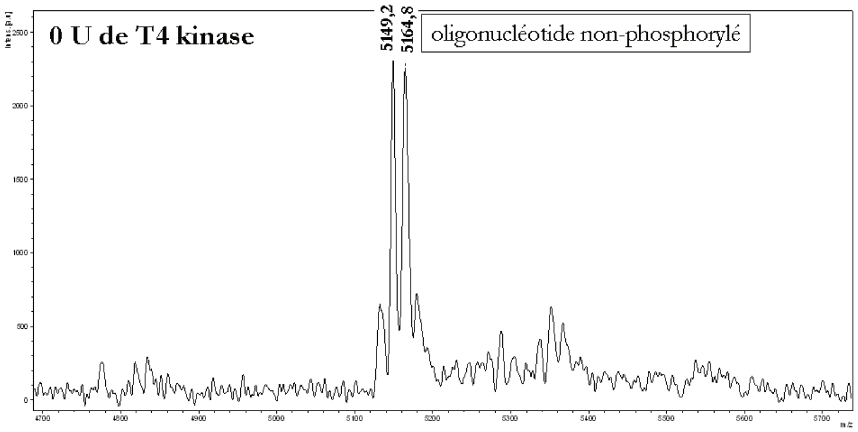


Figure 34 : Spectre de masse MALDI-TOF/MS de l'oligonucléotide simple-brin contenant à son extrémité 3' une fluorescéine.

Les oligonucléotides phosphorylés avec les trois quantités de la kinase utilisées auparavant sur gel d'acrylamide, ont été analysés par la suite par MALDI-TOF/MS. Ces analyses confirment les résultats obtenus précédemment sur gel. Ainsi, nous pouvons voir sur les spectres, les pics correspondant aux produits phosphorylés ou non :

- pour 0.5 U de kinase (Figure 35) : 4 pics regroupés par 2, à $[M-H]^-$ 5150.2 et 5166.2 pour le premier massif, représentant l'oligonucléotide non-modifié, et 5231.4 et 5247.1 pour le deuxième massif ; l'écart entre les deux pics de chaque massif est toujours de 16 Da, et la différence entre les pics correspondants dans les deux massifs est de 81 Da, ce qui représente un groupement phosphate (masse théorique de 80 Da) ; le deuxième massif représente alors bien l'oligonucléotide phosphorylé ;
- pour 2 U de kinase (Figure 36) : l'oligonucléotide non-phosphorylé n'est pas présent ; un massif à $[M-H]^-$ 5230.7 et 5246.0 est visible, qui représente l'oligonucléotide phosphorylé, et l'on aperçoit également un deuxième massif, beaucoup moins intense, avec les pics à $[M-H]^-$ 5311.8 et 5327.0 – la différence de masse entre les pics correspondant du premier et du deuxième massif est 81 Da, ce qui indique la présence d'un deuxième groupement phosphate (nous ignorons comment celui-ci peut se fixer sur un oligonucléotide déjà phosphorylé) ;
- pour 10 U de kinase (Figure 37) : encore une fois deux massifs de pics sont présents, à $[M-H]^-$ 5231.5 et 5246.2, et 5309.5 et 5326.7 – leur masses sont analogues à celles identifiées pour le spectre précédent, mais le deuxième massif est beaucoup plus prononcé que pour la condition de 2 U.

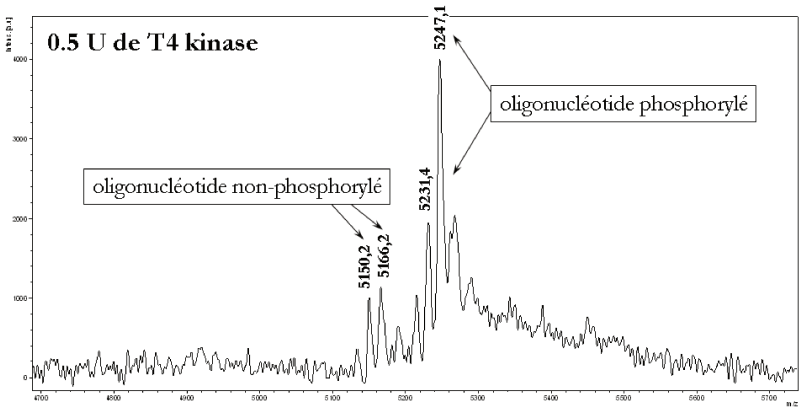


Figure 35 : Spectre de masse MALDI-TOF/MS de l'oligonucléotide simple-brin contenant sur son extrémité 3' une fluorescéine, phosphorylé avec 0.5 U de T4 kinase pour 100 pmol d'oligonucléotide.

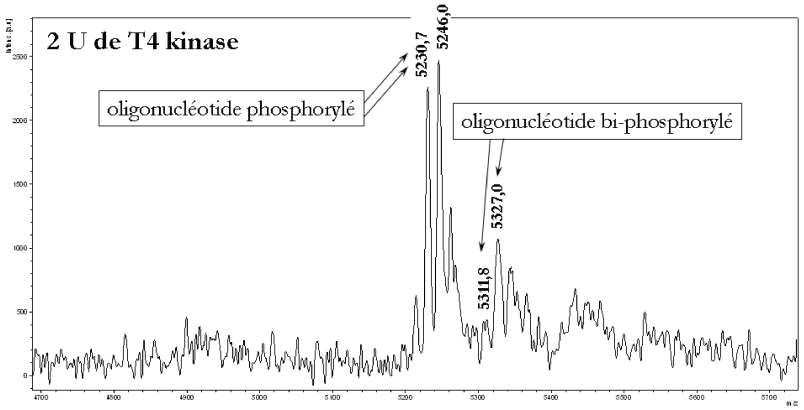


Figure 36 : Spectre de masse MALDI-TOF/MS de l'oligonucléotide simple-brin contenant sur son extrémité 3' une fluorescéine, phosphorylé avec 2 U de T4 kinase pour 100 pmol d'oligonucléotide.

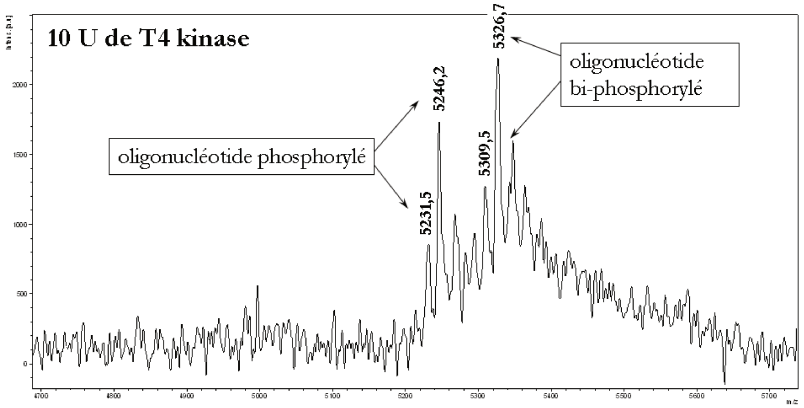


Figure 37 : Spectre de masse MALDI-TOF/MS de l'oligonucléotide simple-brin contenant sur son extrémité 3' une fluorescéine, phosphorylé avec 10 U de T4 kinase pour 100 pmol d'oligonucléotide.

L'analyse par spectromètre de masse MALDI-TOF/MS confirme ainsi l'identité des oligonucléotides identifiés précédemment sur gel. Ces deux expériences nous ont ainsi permis de choisir les conditions les mieux adaptées pour obtenir 100% d'oligonucléotides phosphorylés – la quantité de kinase doit être suffisante pour éliminer complètement le produit de départ, tout en prenant en compte que son excès va engendrer la formation des espèces doublement phosphorylées. Nous avons sélectionné la condition intermédiaire parmi les trois testées, soit 2 U de T4 kinase pour 100 pmol d'oligonucléotide – bien que l'analyse par spectromètre de masse MALDI-TOF/MS a démontré que des espèces doublement phosphorylées soient formées à cette concentration, leur quantité est relativement petite et nous la considérons comme négligeable. La spectrométrie de masse MALDI-TOF/MS n'est pas quantitative, et sur le gel d'acrylamide la bande correspondante à ces oligonucléotides est très faible, constituant un infime pourcentage de l'oligonucléotide recherché portant un seul groupement phosphate.

Après avoir optimisé les conditions de la phosphorylation en utilisant un oligonucléotide fluorescent, nous avons testé les mêmes paramètres pour un oligonucléotide lésé que nous allons utiliser pour la formation de nos pièges. Nous n'avons pas pu vérifier sa phosphorylation sur le gel d'acrylamide, à cause de l'absence du marquage fluorescent à l'extrémité 3', et nous avons alors eu recours à la spectrométrie de masse MALDI-TOF/MS pour détecter la modification de la masse liée à la présence d'un groupement phosphate. Nous n'avons testé qu'une condition de phosphorylation, celle qui s'était avérée optimale pour obtenir 100% de sondes phosphorylées dans les expériences antérieures utilisant les sondes fluorescentes.

La Figure 38 présente le spectre de masse de l'oligonucléotide 5'S-cyclo-dG avant la phosphorylation. Les deux pics détectés correspondent à la forme mono- et di-chargée de celui-ci, avec $[M\text{-}H]^-$ 6699.0 et $[M\text{-}2H]^-/2$ 3349.9 pour sa masse théorique de 6699.4 Da. Aucun dédoublement n'est présent, ce qui indique que la présence des deux pics dans les expériences précédentes était propre à l'oligonucléotide fluorescent et que ceux-ci provenaient probablement d'une dégradation partielle de la fluorescéine.

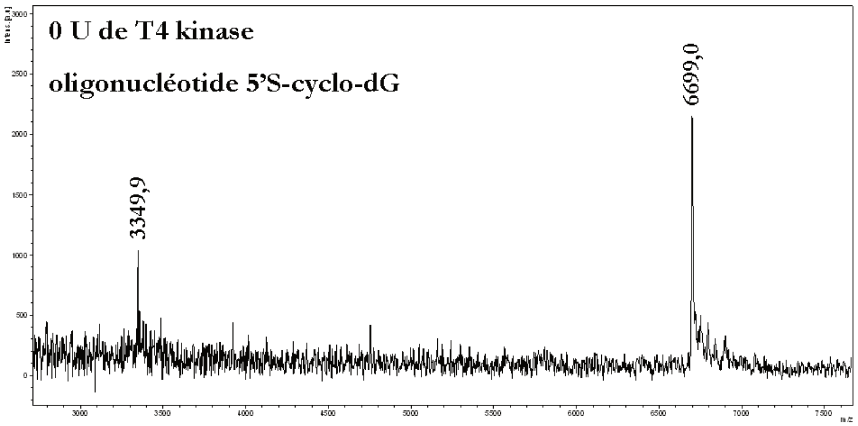


Figure 38 : Spectre de masse MALDI-TOF/MS de l'oligonucléotide 5'S-cyclo-dG simple-brin.

Après la phosphorylation de 100 pmol d'oligonucléotide 5'S-cyclo-dG avec 2 U de kinase, sa forme non-phosphorylée n'est plus détectable (Figure 39). Deux nouveaux pics sont visibles, à $[M-H]^-$ 6778.3 et $[M-2H]^-/2$ 3389.5. Cette analyse confirme la phosphorylation quantitative de l'oligonucléotide (masse augmentée de 78.9). Aucune espèce doublement phosphorylée n'est détectée, ce qui confirme que dans ces conditions nous obtenons une phosphorylation complète de l'oligonucléotide, nous pourrons donc utiliser ce protocole pour la formation des concatémères.

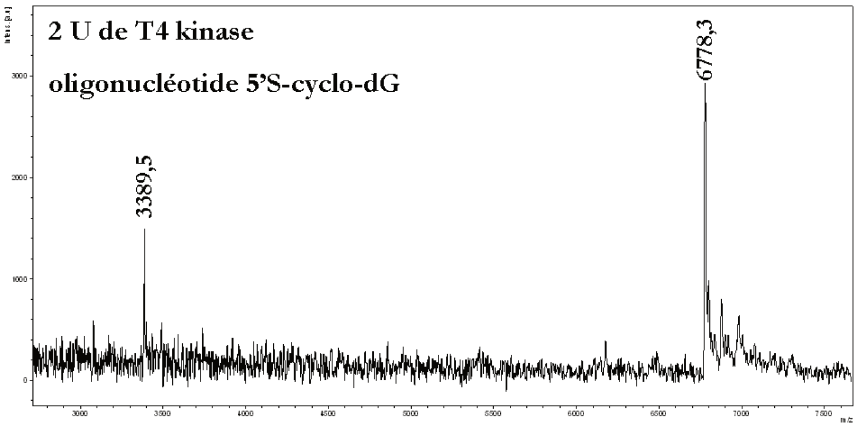


Figure 39 : Spectre de masse MALDI-TOF/MS de l'oligonucléotide 5'S-cyclo-dG simple-brin, phosphorylé avec 2 U de T4 kinase pour 100 pmol d'oligonucléotide.

Ainsi, pour effectuer la phosphorylation de nos sondes, nous allons les incuber pendant 20 minutes à 37°C, en appliquant 2 unités de T4 kinase pour 100 pmol d'oligonucléotides.

1.2.2. Mise au point des conditions de la ligation des oligonucléotides

Une fois que les conditions de phosphorylation des oligonucléotides ont été mises au point, nous avons procédé à l'optimisation de l'étape de la ligation, pour pouvoir former les concatémères et déterminer leur taille. Voici la liste des conditions que nous avons testées :

- la composition du tampon : tampon commercial de la T4 ligase (Life Technologies) ou tampon « Fpg » utilisé pour les expériences EMSA, supplémenté de l'ATP ;
- les conditions de l'hybridation des oligonucléotides : après le chauffage à 90°C pendant 5 min, les tubes peuvent être soit sortis du bloc chauffant et subir le refroidissement rapide sur la paillasse (« hybridation rapide »), soit laissés dans le bloc chauffant éteint qui descend lentement (3 heures environ) à température ambiante (« hybridation lente ») ;
- le temps de la ligation : 20 minutes, 1 heure ou 4 heures
- la température de la ligation : 23°C ou 37°C
- la quantité de la ligase : une gamme allant de 0.001 U à 0.1 U pour 1 pmol d'oligonucléotide a été testée ;
- la quantité des séquences « stop » – puisque les concatémères sont composés des séquences qui se chevauchent, ces courtes séquences oligonucléotidiques ont été conçues, comme cela a été précisé dans le paragraphe 1.3.4 du chapitre « Matériel et méthodes », pour terminer les sondes et éviter la présence des extrémités simple-brin ; en faisant varier leur quantité relative par rapport à la quantité des sondes « concat-1 » et « concat-2 », on pourrait aussi contrôler la longueur des concatémères, car l'hybridation d'une séquence « stop » rend impossible la prolongation de la séquence : nous avons vérifié la formation des concatémères en absence des séquences « stop », et en leur présence à 1/10 ou 1/20 de la concentration des oligonucléotides « concat-1 » et « concat-2 ».

Dans un premier temps, la mise au point de ces conditions a été effectuée en solution. Ceci a été plus aisé pour tester de nombreux paramètres, en s'affranchissant du problème de la libération des concatémères de la surface de la bille. Une fois que les conditions optimales en solution seront définies, celles-ci devront cependant être confirmées par les expériences effectuées sur les billes.

Une des étapes dans cette mise au point est la comparaison de l'efficacité de la formation des concatémères en fonction du type de l'hybridation (« rapide » ou « lente ») et de la température de la ligation (Figure 40). Nous pouvons constater dans un premier temps que l'influence des conditions de l'hybridation sur l'efficacité de la formation des concatémères est

très faible. Le mode « lent » de l'hybridation semble permettre une meilleure formation des concatémères, bien que cette différence soit relativement faible. Des écarts peuvent pourtant être constatés pour différentes températures de la ligation, même si l'estimation exacte de la taille des sondes n'est pas possible à cause du faible pouvoir résolutif du gel. Il apparait cependant que quand la ligation se déroule à 4°C, la majorité des oligonucléotides ne sont pas ligués et le signal est accumulé en bas du gel, représentant les oligonucléotides courts. Nous pouvons aussi voir que plus la température de la ligation augmente, plus le signal des oligonucléotides courts s'estompe pour se déplacer vers le haut du gel, la réaction étant la plus efficace à 37°C.

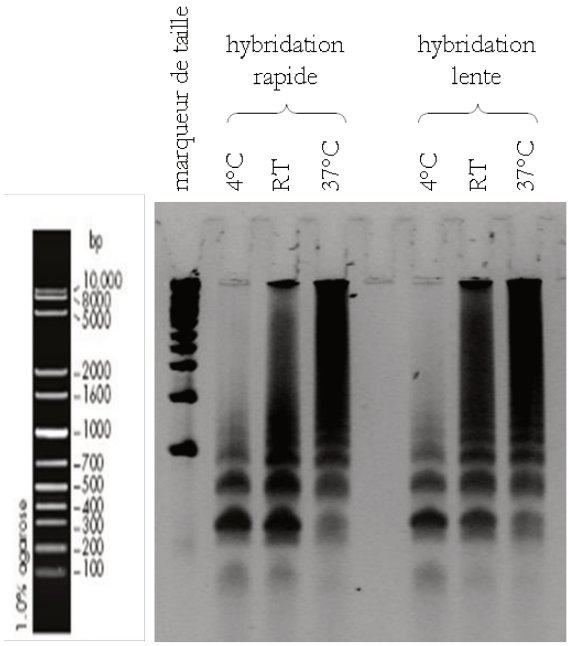


Figure 40 : Formation des concatémères en solution : 2 conditions d'hybridation et 3 températures de ligation ont été comparées. Migration effectuée sur un gel dénaturant d'acrylamide, visualisation de l'ADN avec SYBR Green.

Pour la suite de nos expériences nous choisirons de former les concatémères par hybridation « lente » suivie d'une ligation à 37°C.

Afin de déterminer plus précisément la taille des concatémères formés, nous devons les faire migrer sur un gel d'agarose. L'avantage de ces gels par rapport aux gels d'acrylamide est leur plus grand pouvoir résolutif. Les oligonucléotides de taille élevée ne migrent pas sur les gels d'acrylamide, destinés plutôt à la séparation des petits fragments d'ADN, et nous avons pu le constater dans l'expérience précédente, où la majorité du signal était accumulée dans le haut du gel. Les concatémères de quelques centaines de paires de bases, que l'on souhaite obtenir, ne peuvent pas être bien séparés par cette technique. Les gels d'agarose ne sont pas appropriés à des

petits oligonucléotides, mais ils permettent la migration et la séparation des longs fragments de l'ADN et même des génomes entiers. Cette méthode a alors été plus adaptée à nos besoins. Nous avons cependant remarqué que la visualisation de l'ADN était moins efficace, probablement à cause de la plus grande épaisseur du gel, ce qui nous oblige à y déposer plus de matériel.

Le gel d'agarose présenté sur la Figure 41 montre l'une des dernières étapes de l'optimisation des conditions de la ligation. Nous avons testé 2 paramètres : la quantité de la ligase (1 U ou 5 U pour 100 pmol d'oligonucléotides) et la quantité relative des séquences stop (0, 1/10 ou 1/20). La quantité de ligase joue un rôle important car pour 5 U de celle-ci, les concatémères sont non seulement mieux visibles sur gel, mais ils sont également plus longs. La présence ou l'absence des séquences « stop », quant à elle, ne semble pas avoir d'importance sur la longueur des concatémères, ou du moins notre technique de visualisation n'est pas suffisamment sensible pour détecter ces différences.

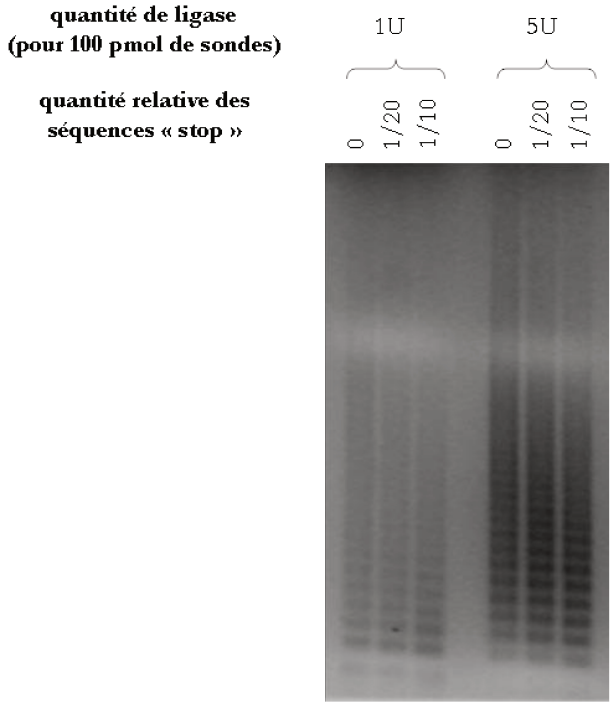


Figure 41 : Formation des concatémères en solution, en fonction de la quantité de la T4 ligase (1 U ou 5 U) et de la quantité relative des séquences « stop » (0, 1/10 ou 1/20) ; les échantillons sont analysés sur un gel natif d'agarose et ont été visualisés avec le bromure d'éthidium.

Ne pouvant pas déterminer précisément l'influence de la présence et de la quantité de ces petites séquences sur la formation et la longueur des concatémères, nous avons alors décidé d'abandonner leur utilisation. Ainsi, toutes nos sondes vont contenir des extrémités simple-brin plus ou moins longues, en fonction de l'oligonucléotide localisé en dernière position. Nous estimons néanmoins que ceci ne devrait pas être gênant pour nos

analyses ultérieures, puisque des sondes non-lésées sont toujours utilisées en tant que témoin, et ainsi toutes les protéines qui pourraient reconnaître les extrémités simple-brin vont être retrouvées à la fois dans les conditions lésées et non-lésées, et ainsi écartées de l'analyse ultérieure.

1.2.3. Formation des concatémères sur les billes

Après avoir mis au point les conditions de la formation des concatémères en solution, nous avons dû vérifier si ces mêmes conditions s'appliquent également quand les concatémères sont formés directement sur les billes. Nous ne pouvons pas greffer les concatémères formés en solution sur les billes, puisque la fixation des sondes à la surface des celles-ci doit s'effectuer via un oligonucléotide biotinylé. C'est pour cela que la synthèse des concatémères directement sur la surface des billes s'impose.

Contrairement à l'hybridation en solution, où la longueur des sondes formées n'est pas limitée, les billes présentent une contrainte : une extrémité du concatémère est déjà immobilisée et sa formation ne peut se poursuivre que dans un sens, ce qui peut réduire sa longueur. Un autre facteur à considérer est la densité de greffage des oligonucléotides initiaux (« concat-biot ») sur la surface des billes – leur quantité va être déterminante pour la longueur des sondes, puisque pour la même quantité des oligonucléotides « concat-1 » et « concat-2 », on va pouvoir obtenir plus ou moins de sondes courtes et de sondes longues.

La mise au point et la vérification des paramètres optimaux pour la formation des concatémères sur les billes vont être réalisées globalement de la même manière que ce qui a été décrit dans le paragraphe précédent. La difficulté supplémentaire liée à la présence des billes consiste en la nécessité de libérer les sondes de leur support avant de les faire migrer sur gel d'agarose. Pour s'affranchir de ce problème, nous avons utilisé le bras photoclivable, que nous avions testé lors des premiers essais de piégeage des protéines. Ainsi, l'oligonucléotide greffé sur les billes n'est plus « concat-biot », mais « concat-PCB », où la biotine classique est remplacée par la biotine photosensible. A la fin de la réaction, les billes sont irradiées, comme décrit dans le paragraphe 1.1.1 de ce chapitre et dans le chapitre « Matériel et méthodes », et les oligonucléotides récupérés en solution sont ensuite déposés sur gel.

L'utilisation des irradiations pour casser la structure photoclivable attachant la sonde à son support présente un inconvénient supplémentaire : elle peut provoquer les cassures de l'ADN, ce qui va être accompagné de la fragmentation des concatémères. Ceci n'est pas gênant si les irradiations sont effectuées afin de récupérer les protéines déjà fixées sur les sondes, car dans

ce cas, la formation des lésions supplémentaires ne va pas influencer le piégeage, toutes les protéines non fixées étant déjà éliminées par les rinçages. Néanmoins, cette dégradation peut rendre impossible l'estimation de la taille des concatémères formés sur les billes. Pour nous affranchir de cette problématique, nous avons comparé la taille des concatémères formés en solution qui ont été par la suite irradiés ou non, et que nous avons fait migrer sur un gel d'agarose. Nous avons déposé en parallèle un échantillon de concatémères formés directement sur les billes et libérés par la même irradiation que celle appliquée aux sondes générées en solution (Figure 42).

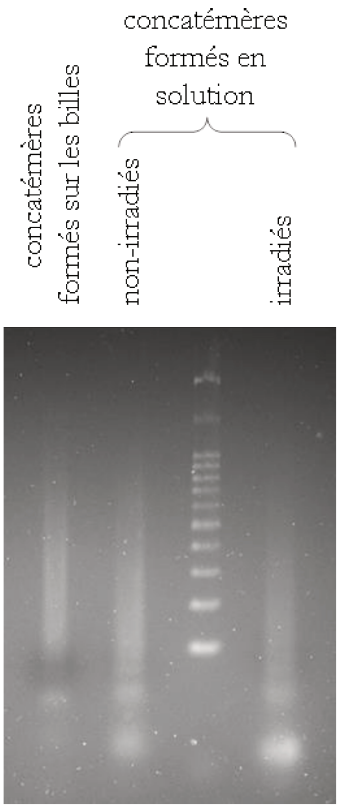


Figure 42 : Vérification de la dégradation des concatémères formés en solution ou sur les billes, par les irradiations utilisées pour la rupture du bras photoclivable. Migration effectuée sur un gel natif d'agarose, visualisation avec SYBR Green.

Comme nous pouvons le voir sur la photo du gel, il n'existe pas de grandes différences de taille entre les concatémères irradiés ou non en solution. Il semblerait que les sondes qui ont subi une irradiation soient plus courtes, car un signal plus intense vers le bas du gel, représentant les fragments très courts de l'ADN, est visible. Les sondes formées directement sur les billes ne semblent pas être endommagées par les irradiations, car leur longueur est comparable à celle des concatémères formés en solution. Nous ne pouvons pas définir la taille précise des concatémères

à cause de la faible résolution du gel, mais nous observons tout de même que les irradiations utilisées pour casser le bras photoclivable ne dégradent pas, ou peu, l'ADN fixé sur les billes.

La dernière expérience que nous avons effectuée avant d'exposer nos pièges aux extraits protéiques, a consisté à déterminer la taille des concatémères formés sur les billes. Pour cela, nous les avons libérés de leur support par irradiation et les avons fait migrer sur un gel d'acrylamide, que nous avons ensuite coloré au SYBR Green. Bien que la résolution obtenue sur un gel d'acrylamide soit plus faible que celle qui peut être observée sur un gel d'agarose, surtout pour les grands fragments de l'ADN, la visualisation est de meilleure qualité grâce à une moindre épaisseur du gel. Comme nous pouvons le voir sur la Figure 43, les concatémères sont bien séparés après une longue migration, car nous pouvons voir des bandes qui correspondent à différentes sondes, dont la taille varie d'un oligonucléotide. Le signal le plus intense est observé dans la partie basse du gel, qui correspond à la longueur entre 100 et 350 paires de base.

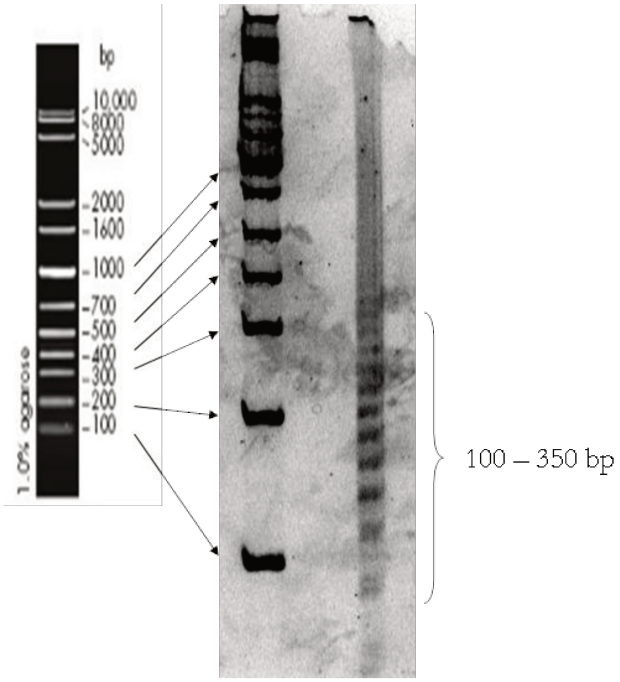


Figure 43 : Détermination de la taille des concatémères formés directement sur les billes. Après l'irradiation du bras photoclivable, les concatémères sont analysés sur un gel natif d'acrylamide qui a été ensuite coloré au SYBR Green.

Ainsi, nous constatons que les concatémères formés sur les billes sont composés de 5 à 15 oligonucléotides environ, et que le nombre des cyclonucléosides sur les sondes lésées sera ainsi compris, lui aussi, entre ces mêmes valeurs.

1.2.4. Comparaison de différentes conditions de récupération des protéines

Après la mise au point du piège, nous avons vérifié les différences d'efficacité de la récupération des protéines en fonction de leur mode d'élution. Le bras photoclivable, dont les conditions d'utilisation avaient été optimisées dans le paragraphe 1.1.1 de ce chapitre, a été utilisé, ou alternativement les protéines ont été éluées par le chauffage des billes pendant quelques minutes à 90°C dans un tampon dénaturant du type Laemmli. Cette expérience a été effectuée en utilisant les sondes non-lésées ou les billes, pour comparer l'importance du bruit de fond et l'impact du mode d'élution sur la quantité des protéines récupérées. Les pièges ont été incubés avec des extraits nucléaires HeLa et les protéines analysées sur un gel SDS-PAGE qui a ensuite été coloré au nitrate d'argent, pour visualiser la totalité des protéines (Figure 44).

Nous pouvons remarquer que l'élution des protéines par le chauffage est plus efficace que leur élution par irradiation, car plus de protéines sont présentes sur le gel. Ceci est valable également pour les billes, ce qui confirme que l'élution par irradiation est plus « propre ». Néanmoins, l'élution par irradiation, bien qu'elle soit plus propre, ne permet pas de récupérer toutes les protéines qui ont interagi avec les sondes. On risque alors de ne pas détecter un candidat intéressant interagissant avec les cyclonucléosides.

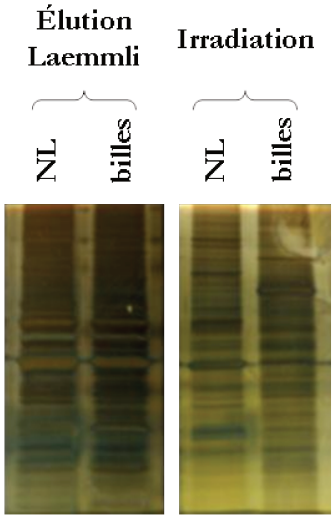


Figure 44 : Comparaison de la quantité des protéines récupérée après le piégeage, sur les sondes non-lésées ou sur les billes, en effectuant une élution dénaturante (Laemmli) ou une irradiation d'un bras photoclivable.

Pour la suite de nos expériences en ligand fishing nous opterons pour l'élution dénaturante dans le tampon Laemmli, qui permet l'identification de plus de protéines sur les sondes, au risque d'augmenter le bruit de fond.

1.3. Exposition des pièges aux extraits protéiques

Les pièges préparés d'après le protocole décrit dans les paragraphes précédents ont été ensuite exposés aux extraits protéiques humains (HeLa). Puisque la quantité et la qualité des protéines piégées dépend largement des conditions d'incubation, nous avons testé différents temps et températures d'incubation, ainsi que différentes compositions du tampon réactionnel. Le temps et la température de la réaction sont d'importance particulière, car nos sondes ne sont pas protégées contre les exonucléases présentes dans tous les extraits protéiques – ainsi, il est important de diminuer leur activité, par exemple en baissant la température de l'incubation et en limitant sa durée, tout en gardant des conditions favorables à la reconnaissance de l'ADN par des protéines spécifiques.

Le piégeage des protéines sur les sondes non-lésées ou sur les billes a été effectué dans deux tampons : A ou RBP (compositions données dans le chapitre « Matériel et méthodes »), et deux temps d'incubation : 20 minutes ou 2 heures ont été testés (Figure 45). Les protéines ont été récupérées par chauffage du piège dans un tampon dénaturant (Laemmli). Les échantillons récupérés dans chacune de ces conditions ont ensuite subi une migration sur un gel SDS-PAGE, qui a été coloré, après la migration, au nitrate d'argent.

Nous pouvons constater qu'il existe des différences dans la quantité des protéines piégées dans les différentes conditions testées. On peut voir que le temps d'incubation ne semble pas jouer de rôle important dans l'efficacité du piégeage, du moins si son caractère quantitatif est considéré. Aucune différence n'est constatée non plus entre la quantité des protéines piégées sur les sondes non-lésées ou sur les billes. Nous pouvons voir, cependant, qu'en fonction du tampon qui a été utilisé pour effectuer l'exposition des pièges aux extraits protéiques, une quantité de protéines très variable est visible sur le gel. De manière générale, quel que soit le type de sondes utilisées, beaucoup plus de protéines sont piégées dans le tampon A que dans le tampon RBP. Pour le tampon RBP, au contraire, beaucoup de protéines de haute masse moléculaire sont visibles, tandis que le signal correspondant aux protéines de petit poids moléculaire est plus faible. Le tampon RBP étant beaucoup plus salin que le tampon A, il ne permet pas la reconnaissance des sondes par les protéines de faible affinité, ce qui se traduit par un nombre réduit de protéines piégées.

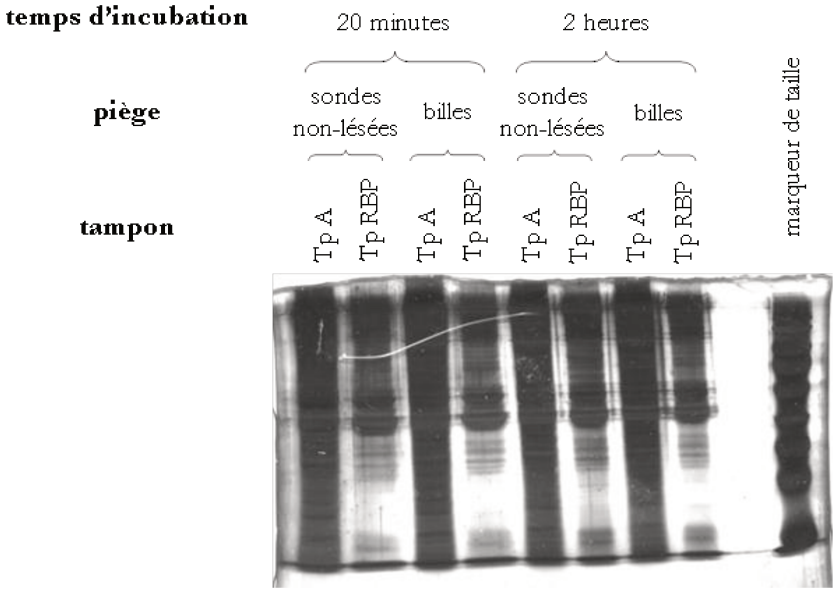


Figure 45 : Protéines humaines (extraits HeLa) piégées sur les sondes non-lésées ou sur les billes, dans différentes conditions d'incubation (20 minutes/2 heures d'exposition, tampon A/tampon RBP). Les protéines sont analysées sur un gel SDS-PAGE, coloré ensuite au nitrate d'argent.

Cette analyse semi-quantitative ne nous permet pas néanmoins de choisir la meilleure condition d'incubation. Le piégeage d'un grand nombre de protéines augmente la probabilité de retrouver parmi celles-ci un candidat intéressant, qui pourrait être éliminé par l'utilisation des conditions plus restrictives, surtout si ses interactions avec la sonde sont faibles. Mais d'autre part, la réduction du nombre des protéines totales réduit le bruit de fond et permet d'identifier uniquement les protéines dont l'affinité pour l'ADN et ses lésions est forte, sans prendre en compte les protéines moins pertinentes.

Pour définir les meilleures conditions d'incubation, une expérience de piégeage des protéines va être réalisée à nouveau et suivie d'une analyse des résultats par la spectrométrie de masse. Un seul temps d'incubation va tout de même être choisi, celui de 20 minutes, afin de limiter la dégradation probable des sondes par les nucléases présentes dans les extraits.

1.4. Identification par analyse protéomique des protéines piégées

Trois analyses protéomiques ont été effectuées afin d'identifier les protéines interagissant avec les cyclonucléosides et de les confirmer par la suite :

- 1ère analyse – elle a servi au choix définitif des conditions d'incubation
- 2ème analyse – elle a servi à confirmer les interactions déterminées lors de la 1ère analyse
- 3ème analyse – est une analyse complète utilisant les sondes contenant les quatre cyclonucléosides dont nous disposons.

Ces trois analyses ont été effectuées par le laboratoire IBEB.

Les aspects quantitatifs de ces trois analyses vont être exposés dans un premier temps, avant de poursuivre avec l'analyse qualitative des protéines identifiées.

1.4.1. Aspects quantitatifs des trois analyses protéomiques

Dans un premier temps, la quantité des protéines piégées dans différentes conditions va être analysée. Les conditions réactionnelles permettant le piégeage d'un grand nombre de protéines, tout en limitant le bruit de fond, seront sélectionnées. Ensuite, les deux autres analyses vont permettre d'identifier les interactions protéines-cyclonucléosides sur un grand panel de lésions.

1.4.1.1. Première analyse et choix des conditions d'incubation

Dans le paragraphe 1.3 de ce chapitre nous avons montré les différences dans la quantité des protéines piégées sur les sondes en fonction du tampon utilisé pour effectuer l'incubation avec les extraits. Afin de vérifier ces résultats et également pour mettre en évidence les différences qualitatives qu'il peut y avoir entre les protéines piégées dans différentes conditions, nous avons réalisé une analyse par spectrométrie de masse des échantillons récupérés sur les sondes 5'S-cyclo-dG, non-lésées ou sur les billes. Les pièges ont été exposés aux extraits dans deux tampons testés précédemment, pendant 20 minutes. Une comparaison à la fois quantitative et qualitative a été effectuée afin de choisir les conditions optimales, qui permettent de piéger un grand nombre de protéines sur les sondes et récupérer un minimum de bruit de fond généré par les protéines non-spécifiques (notamment celles qui se fixent sur la surface de billes) (Figure 46).

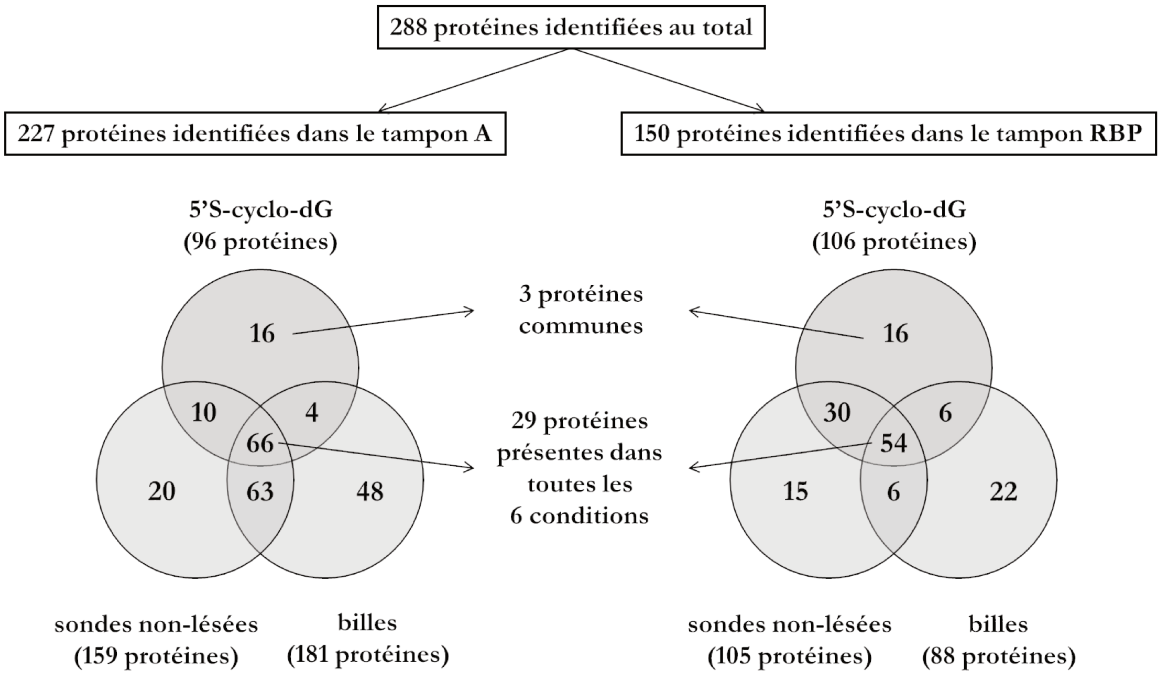


Figure 46 : Comparaison des protéines piégées sur les sondes 5'S-cyclo-dG, non-lésées ou sur les billes, après une incubation de 20 min avec des extraits nucléaires HeLa dans un des deux tampons : A ou RBP (composition donnée dans le chapitre « Matériel et méthodes »).

288 protéines différentes ont été identifiées dans l'ensemble des échantillons, dont 227 ont été retrouvées dans les échantillons préparés dans le tampon A, et seulement 150 dans ceux qui ont été incubés dans le tampon RBP. Ceci rejoint les observations du paragraphe 1.3 de ce chapitre, où nous avons vu sur un gel d'acrylamide coloré au nitrate d'argent, que le tampon RBP semblait être beaucoup plus stringent et retenir moins de protéines sur les pièges que le tampon A. Quel que soit le tampon, nous pouvons remarquer qu'un tiers environ des protéines est identifié sur les trois pièges.

Nous pouvons analyser le nombre de protéines piégées sur différentes sondes dans chacune des deux conditions testées. Si l'on regarde uniquement les échantillons provenant de l'incubation dans le tampon A, on peut voir que parmi les 227 protéines identifiées, 181 sont retrouvées sur les billes, soit 80%. Cette proportion baisse à 59% dans les échantillons préparés dans le tampon RBP. Comme les protéines présentes sur les billes génèrent le bruit de fond et qu'elles ne vont pas être prises en compte dans la recherche des candidats pertinents, la diminution de leur quantité est importante pour la réussite de notre étude.

En ce qui concerne les protéines présentes uniquement et spécifiquement sur un type de sondes : lésées ou non-lésées, leur quantité est comparable quel que soit le tampon utilisé, et elle se situe entre 15 et 20. Parmi ces protéines, très peu sont communes pour les deux conditions

d'incubation (3 sur les sondes lésées et une seule sur la sonde non-lésée). On remarque bien alors l'importance du choix des conditions d'incubation pour la qualité des protéines piégées, car en fonction de la composition du tampon utilisé, deux ensembles distincts des protéines ont été identifiés.

Après avoir comparé le nombre des protéines identifiées sur différentes sondes dans les deux conditions testées, nous avons décidé de retenir pour les expériences suivantes le tampon RBP. Il est plus stringent que le tampon A et il permet de réduire de manière significative le nombre de protéines qui reconnaissent les pièges, et surtout la surface des billes magnétiques, de manière non-spécifique, et ainsi de limiter le bruit de fond. Bien que le nombre de protéines-candidats soit réduit par la stringence du tampon, nous pouvons espérer que ce sont les protéines plus pertinentes qui vont être retenues sur les sondes, mais nous ne pouvons pas écarter le risque de ne pas détecter une protéine intéressante, dont l'affinité pour nos sondes serait faible et nécessiterait les conditions d'incubation plus délicates.

1.4.1.2. Deuxième analyse : sondes 5'S-cyclo-dG et 5'R-cyclo-dA

La deuxième analyse protéomique a été effectuée uniquement dans les conditions retenues d'après le paragraphe 1.4.1.1. Deux sondes lésées : 5'S-cyclo-dG et 5'R-cyclo-dA ont été utilisées. De la même manière que nous l'avons fait pour la première analyse protéomique, pour cette deuxième analyse nous avons effectué une comparaison du nombre des protéines identifiées sur nos pièges (Figure 47). 180 protéines au total ont été identifiées, dont 25% environ (46 sur 180) étaient présentes sur tous les pièges, et 72% (129 sur 180) sur les billes. 7 et 8 protéines ont été identifiées comme spécifiques pour les sondes 5'R-cyclo-dA et 5'S-cyclo-dG, respectivement, et 7 autres protéines ont été retrouvées comme communes et spécifiques pour ces deux sondes lésées. 14 protéines ont été identifiées comme spécifiques pour les sondes non-lésées.

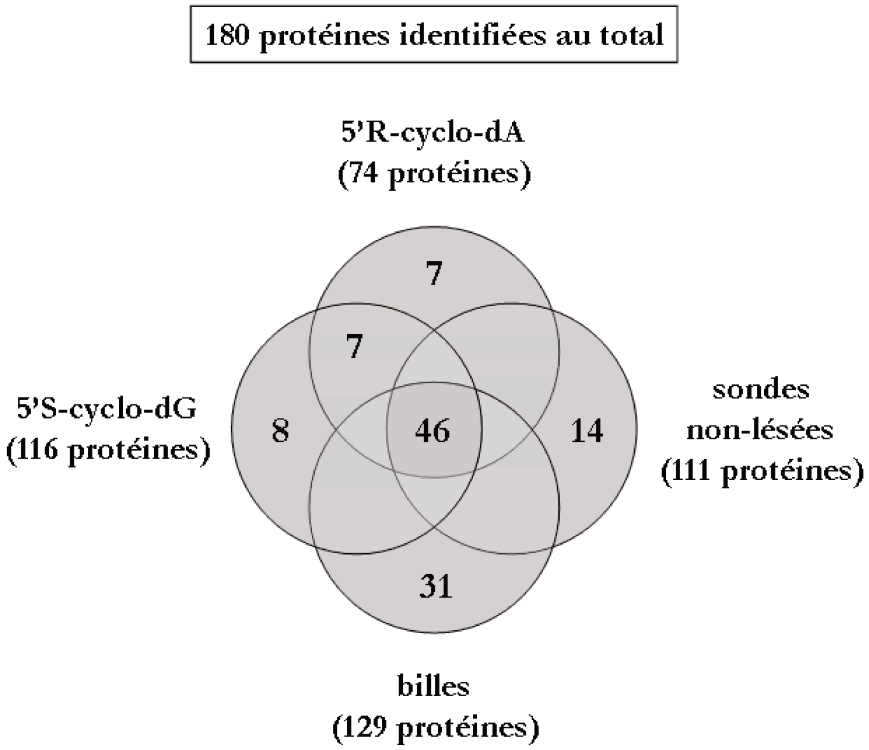


Figure 47 : Comparaison des protéines piégées sur les sondes 5'R-cyclo-dA, 5'S-cyclo-dG, non-lésées ou sur les billes.

1.4.1.3. Troisième analyse : utilisation du panel complet des cyclonucléosides

La troisième analyse que nous avons effectuée a été la plus complexe, car les sondes portant les quatre cyclonucléosides ont été utilisées (Figure 48). Nous avons identifié 151 protéines au total, dont 13 étaient présentes sur tous les pièges et 50 ont été retrouvées sur les billes. Entre 5 et 15 protéines sont spécifiques pour une sonde lésée, 26 sont retrouvées sur 2 ou 3 sondes lésées, et 4 sont communes pour les quatre cyclonucléosides. 9 protéines ont été identifiées uniquement sur les sondes ne portant pas de lésions.

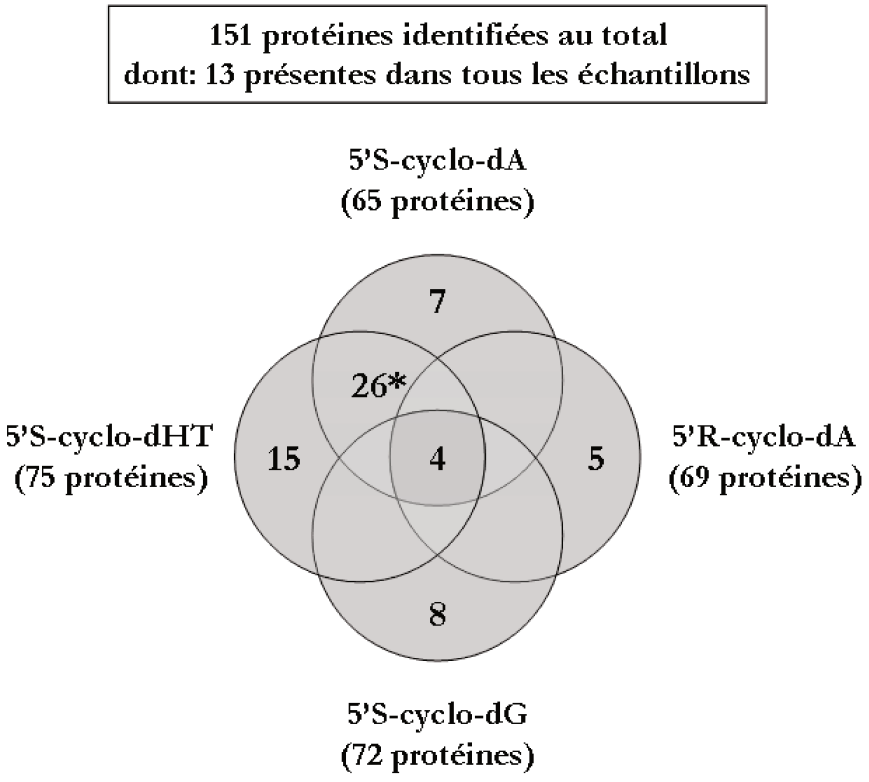


Figure 48 : Comparaison des protéines piégées sur les sondes 5'S-cyclo-dA, 5'R-cyclo-dA, 5'S-cyclo-dG, 5'S-cyclo-dHT, non-lésées ou sur les billes. (*) = ce chiffre correspond à la somme des protéines identifiées sur 2 ou 3 sondes lésées, quelle que soit la lésion.

1.4.2. Analyse qualitative des protéines identifiées

Après l'analyse comparative du nombre des protéines identifiées par le biais de la protéomique, nous allons maintenant nous pencher sur leur aspect qualitatif. Nous avons vu sur les figures 46 à 48, qu'une dizaine de protéines a été retrouvée uniquement sur les sondes lésées ou uniquement sur les sondes non-lésées. Elles sont présentées dans le Tableau A 1 des annexes, qui compare les résultats de l'identification dans les trois analyses effectuées. Pour les analyses qui ont employé plusieurs sondes lésées, seules les protéines présentes sur toutes les lésions ont été prises en compte. Pour chaque protéine identifiée, deux valeurs sont données : la première est le nombre de peptides qui ont servi à son identification, et la deuxième est le score Mascot. Une valeur seuil pour la fiabilité de l'identification des protéines a été fixée à 2 peptides et un score mascot de 70 au minimum [302]. Nous allons donc présenter et prendre en compte dans la suite de notre analyse uniquement les protéines qui répondent à ces deux conditions. Ces valeurs ne

nous permettent pas de faire une analyse quantitative, néanmoins si un grand écart est constaté entre le nombre des peptides et le score Mascot pour deux conditions différentes, on peut supposer que ceci soit lié aux différences d'abondance de la protéine.

Nous pouvons constater que même parmi les protéines identifiées de manière spécifique sur un type de sonde, certaines n'ont aucune pertinence biologique, malgré leur score Mascot élevé. Tel est le cas par exemple de la kératine 18 identifiée uniquement sur les sondes 5'S-cyclo-dG lors de la première analyse. Egalement, nombreuses sont les protéines identifiées comme spécifiques pour une condition dans une des analyses, mais dont la présence unique dans cette condition n'a pas été confirmée par les deux autres analyses. C'est notamment le cas d'une protéine que nous avons sélectionnée dans un premier temps comme candidat intéressant au vu de son rôle biologique et qui est NONO. Ses fonctions biologiques et les raisons de son choix comme candidat seront exposées dans le paragraphe 1.4.3.1. Lors de la première analyse NONO a été identifié uniquement sur les sondes portant un 5'S-cyclo-dG, son identification étant effectué avec 16 peptides et un score Mascot de 400.75. Ces valeurs élevées, bien qu'encourageantes, n'ont pas été retrouvées dans les analyses suivantes :

- la deuxième analyse a identifié NONO sur les sondes lésées et non-lésées, et également sur les billes ;
- la troisième analyse l'a identifié sur trois sondes lésées des quatre, mais avec des très faibles scores Mascot : 38.66 (1 peptide), 41.28 (1 peptide) et 82.88 (2 peptides) pour les sondes 5'S-cyclo-dA, 5'R-cyclo-dA et 5'S-cyclo-dHT (Tableau 13)

Une deuxième protéine dont la détection n'a pas été confirmée est SFPQ. Ce partenaire de NONO (voir chapitre 1.4.3.2) n'a pas été identifié uniquement sur les sondes lésées. On le retrouve :

- dans la première analyse : dans toutes les conditions avec 21 peptides et un score de 521.54 sur les sondes 5'S-cyclo-dG, 14 peptides et un score de 363.06 sur les sondes non-lésées et 11 peptides et un score de 296.57 sur les billes ;
- dans la deuxième analyse : dans toutes les conditions, avec un score le plus élevé sur les sondes non-lésées ;
- dans la troisième analyse : sur les quatre sondes lésées et sur les billes, mais avec de faibes scores, à l'exception de la sonde 5'R-cyclo-dA : 2 peptides et un score 97.61 (Tableau 13).

Nous l'avons néanmoins sélectionné pour des analyses complémentaires surtout à cause de ses interactions connues avec NONO.

Aucune autre protéine biologiquement pertinente n'a été identifié de manière fiable et répétée uniquement sur les sondes contenant des cyclonucléosides.

Deux protéines : PARP1 et DREF, ont été identifiées de façon reproductible uniquement sur les sondes non-lésées avec les scores relativement élevés :

- PARP1 : 15 peptides et un score de 477.19 dans la première analyse, 10 peptides et un score de 380.47 dans la deuxième (identifié également sur les sondes 5'S-cyclo-dG, mais avec 2 peptides et un score de 73.07 seulement), et 8 peptides et un score de 257.61 dans la troisième ;
- DREF : 6 peptides et un score de 244.82 dans la première analyse, 11 peptides et un score de 386.37 dans la deuxième, et 8 peptides et un score de 287.48 dans la troisième (Tableau 13).

La présence majoritaire de ces deux protéines sur les sondes non-lésées a ainsi été confirmée par les trois analyses protéomiques, dont la dernière effectuée avec quatre cyclonucléosides.

Au cours des trois analyses protéomiques nous n'avons identifié aucune protéine reconnaissant les cyclonucléosides. Néanmoins, les expériences complémentaires effectuées sur deux candidats initialement sélectionnées : NONO et SFPQ, vont être présentées dans les paragraphes suivants.

D'autre part, les interactions des deux protéines PARP1 et DREF avec l'ADN semblent être inhibées par la présence des cyclonucléosides. Ainsi, les expériences complémentaires vont être effectuées avec ces deux candidats.

Tableau 13 : Analyse protéomique des 4 protéines d'intérêt : NONO, SFPQ, PARP1 et DREF. Pour chaque protéine, le nombre de peptides ayant servi à son identification est donné, suivi de son score Mascot.

entrée (Uniprot)	**nom de la protéine**	**nom du gène**	**1ère analyse**	**2ème analyse**	**3ème analyse**
NONO_HUMAN	Non-POU domain-containing octamer-binding protein	NONO	**5'S-cyclo-dG : 16/400.75** non-lésé : --- billes : ---	5'R-cyclo-dA : 9/254.55 5'S-cyclo-dG : 11/361.55 non-lésé : 8/210.43 billes : 20/500.74	5'S-cyclo-dA : 1/38.66 5'R-cyclo-dA : 1/41.28 5'S-cyclo-dG: --- 5'S-cyclo-dHT : 2/82.88 non-lésé : --- billes : ---
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich	SFPQ	5'S-cyclo-dG : 21/521.54 non-lésé : 14/363.06 billes : 11/296.57	5'R-cyclo-dA : --- 5'S-cyclo-dG : --- non-lésé : 10/327.41 billes : ---	5'S-cyclo-dA : 2/62.78 5'R-cyclo-dA : 2/97.61 5'S-cyclo-dG: 1/51.65 5'S-cyclo-dHT : 2/39.49 non-lésé : --- billes : 2/46.48
PARP1_HUMAN	Poly [ADP-ribose] polymerase 1	PARP1	5'S-cyclo-dG : --- **non-lésé : 15/477.19** billes : ---	5'R-cyclo-dA : --- 5'S-cyclo-dG : 2/73.07 **non-lésé : 10/380.47** billes : ---	5'S-cyclo-dA : --- 5'R-cyclo-dA : --- 5'S-cyclo-dG: --- 5'S-cyclo-dHT : --- **non-lésé : 8/257.61** billes : ---
ZBED1_HUMAN	Zinc finger BED domain-containing protein 1	ZBED1 (DREF)	5'S-cyclo-dG : --- **non-lésé : 6/244.82** billes : ---	5'R-cyclo-dA : --- 5'S-cyclo-dG : --- **non-lésé : 11/386.37** billes : ---	5'S-cyclo-dA : --- 5'R-cyclo-dA : --- 5'S-cyclo-dG: --- 5'S-cyclo-dHT : --- **non-lésé : 8/287.48** billes : ---

1.4.3. Choix des candidats pertinents

Après la première analyse protéomique, nous avons sélectionné quelques protéines qui nous paraissaient intéressantes au regard de leur rôle biologique. Les protéines sur lesquelles nous avons porté notre choix sont présentées ci-dessous.

1.4.3.1. NONO

Cette protéine a été retrouvée de manière spécifique sur les sondes 5'S-cyclo-dG lors de la première analyse. Les deux analyses suivantes n'ont pourtant pas confirmé ces résultats. Néanmoins, nous allons présenter les caractéristiques de cette protéine, car, après les résultats de la première analyse protéomique, des expériences complémentaires ont été réalisées.

NONO, appelé également p54(nrb) est une protéine nucléaire de 54 kDa, appartenant à la famille RRM, car elle contient des domaines RRM : RNA recognition motif [303]. Il serait impliqué dans des mécanismes tels que l'épissage [304]. Néanmoins, de plus en plus d'études évoquent d'autres rôles, par exemple dans la réponse aux dommages à l'ADN et surtout dans la réparation des cassures double-brin [303–305]. Cette réponse serait effectuée par un hétérodimère que NONO forme avec une autre protéine : SFPQ [303,305]. Le complexe NONO/SFPQ est rapidement relocalisé vers les lésions de l'ADN générées *in cellulo* par un laser [306]. De plus, la réduction de l'expression de NONO implique une déficience dans la réparation des cassures double-brin, augmente le nombre des aberrations chromosomiques induites par le rayonnement et confère également une radiosensibilité accrue [303,307]. Son implication dans la réparation des cyclonucléosides, qui sont générés par le rayonnement ionisant, semble ainsi probable et nous avons décidé de poursuivre cette piste en effectuant des tests complémentaires sur la protéine NONO.

1.4.3.2. SFPQ

SFPQ n'a pas été retrouvé de manière spécifique sur les sondes lésées ou non-lésées. Nous avons cependant décidé de vérifier ses interactions avec l'ADN endommagé, par des techniques complémentaires, car cette protéine est un partenaire de NONO, comme ceci a été

décrit dans le paragraphe précédent, et après la première analyse protéomique elle paraissait tout de même plus abondante sur les sondes portant une lésion, que sur les sondes contrôles.

SFPQ appartient à la même famille des protéines RRM que son partenaire NONO, avec qui elle partage aussi 71% d'identité [305]. Il interagit également avec RAD51 – une protéine centrale dans la recombinaison homologue [305], et son implication dans la réparation des cassures est renforcée par sa capacité à générer une boucle-D par invasion de l'ADN double-brin [304]. De plus, la réduction de son expression ralentit la réparation des cassures double-brin et augmente la sensibilité cellulaire aux agents pontant ou alkylant [307,308]. SFPQ est également impliqué dans le mécanisme de radiorésistance et il a été démontré qu'il pouvait reconnaître l'ADN de manière non-spécifique, avant d'être relocalisé vers la lésion [306].

1.4.3.3. PARP1

Contrairement aux protéines NONO et SFPQ, qui semblaient reconnaître d'une manière spécifique les cyclonucléosides, PARP1 a été retrouvé uniquement sur les sondes non-lésées.

PARP1 est le membre le mieux connu de la famille PARP contenant 18 protéines. Il joue un rôle dans de nombreux processus cellulaires tels que la réparation de l'ADN, la signalisation cellulaire, la régulation de l'expression des gènes ou la modification de la chromatine. PARP1 possède un domaine de fixation à l'ADN contenant deux doigts de zinc, et un domaine BRCT qui est caractéristique des protéines de réparation [309–311].

PARP1 est associé majoritairement avec le système de réparation BER, mais son implication dans la réparation des cassures de l'ADN a également été démontrée. Son rôle consiste à retrouver et reconnaître les lésions, et y recruter les protéines qui vont accomplir la réparation de celles-ci. Ce recrutement est effectué par la poly(ADP-ribos)ylation (PARylation) des protéines-effecteurs – PARP1 produit des longs polymères d'ADP-ribose en utilisant de l'ATP et NAD^+ comme substrats. En plus de modifier d'autres protéines, PARP1 peut subir une auto-modification, qui lui permet de se dissocier de l'ADN une fois que son rôle a été joué [309,310].

Le fait que nous ayons identifié PARP1 sur les sondes non-lésées uniquement, et non pas sur celles qui contiennent un cyclonucléoside, présente une piste intéressante pour un nouveau mode d'action de cette protéine impliquée de différentes manières dans de nombreux processus se déroulant dans les cellules.

1.4.3.4. DREF

DREF, de la même manière que PARP1, a été identifié dans nos analyses uniquement sur les sondes non-lésées.

Cette protéine a été découverte en 2003 par l'équipe de Fumiko Hirose, qui l'a identifiée dans un premier temps en recherchant des régions conservées à l'aide de BLAST. Les régions ciblées étaient trois domaines conservés chez deux espèces de *Drosophila* : CR1, CR2 et CR3. La région CR1 est nécessaire pour les interactions de DREF avec l'ADN et pour la formation des homodimères et c'est elle qui a permis l'identification de hDREF. La conservation des deux autres domaines a également été confirmée et l'existence de hDREF a ensuite été prouvée [312–315].

HDREF est homologue de DmDREF de *Drosophila melanogaster*, avec qui elle partage 21.3% d'identité et 41.0% de similarité, et dont elle partage également les fonctions. DREF est localisé principalement dans le noyau, où il est associé de manière stable avec les structures granulaires [312].

La séquence consensus de fixation de DREF a été déterminée comme 5'-TGTCG(C/T)GA(C/T)A-3', dont la partie centrale de 6 paires de bases 5'-TCG(C/T)GA-3' est de plus grande importance. Elle est retrouvée dans les régions promotrices de nombreux gènes liés à la prolifération cellulaire, qui sont impliqués notamment dans la synthèse ou la réparation de l'ADN, la transcription, la synthèse des protéines, la régulation de la structure de la chromatine ou encore la régulation du cycle cellulaire. Elle a notamment été retrouvée dans la région promotrice du gène de l'histone H1 dont il a été démontré que hDREF régulait positivement l'expression [312]. La séquence de fixation de hDREF est aussi présente dans 22 des 79 gènes de protéines ribosomales. La régulation de ceux-ci par hDREF est dépendante du cycle cellulaire, avec l'induction maximale pendant la transition G1/S, et elle a un impact sur la prolifération cellulaire, même si le mécanisme exact de ces interactions n'a pas été élucidé [316].

1.5. Validation des candidats identifiés

Les quatre protéines retenues comme pertinentes et intéressantes d'un point de vue biologique : PARP1, DREF, NONO et SFPQ, ont fait objet d'analyses complémentaires, dont le but était de confirmer ou d'infirmer leur association aux sondes lésées ou non-lésées.

1.5.1. Western Blot

La première technique utilisée pour vérifier la pertinence des protéines identifiées par la spectrométrie de masse, a été le Western blot. Les échantillons ont également été préparés par ligand fishing. L'utilisation de cette méthode nous permet d'avoir recours à une autre technique de détection, mais s'il existe des biais au niveau du piégeage même des protéines, ceci ne va pas pouvoir être constaté. Pour chaque protéine, trois conditions ont été testées ; sonde 5'S-cyclo-dG, sonde non-lésée et billes, un contrôle d'extraits nucléaires HeLa non soumis au piégeage a également été déposé sur le même gel.

NONO

Le Western Blot que nous avons effectué pour détecter NONO montre la présence d'une bande de taille attendue à 54 kDa (Figure 49), que ce soit dans les extraits non soumis au piégeage, ou dans une des trois conditions testées. Hormis un léger dédoublement de cette bande dans la condition 5'S-cyclo-dG, nous ne constatons pas de différences d'intensité des bandes entre les conditions, ce qui indique que NONO a été piégé sur les sondes lésées et non-lésées, et également sur les billes. Nous ne pouvons alors pas confirmer les résultats de la première analyse protéomique, qui suggérait l'affinité spécifique de NONO pour les sondes contenant un 5'S-cyclo-dG.

Au vu des résultats des analyses protéomiques qui n'ont pas confirmé la présence spécifique de NONO sur les sondes lésées, et des résultats de la détection par le Western blot, nous devons alors écarter l'hypothèse de l'implication de cette protéine dans la reconnaissance et l'éventuelle prise en charge des cyclonucléosides.

SFPQ

La deuxième protéine dont nous avons voulu confirmer la détection par Western blot est SFPQ (Figure 49). Nous ne constatons aucune différence entre les profils des échantillons : 5'S-cyclo-dG, non-lésés et billes, car les mêmes bandes sont visibles pour les trois conditions de piégeage. Nous en concluons alors que la protéine SFPQ, tout comme son partenaire NONO, n'est pas spécifique des cyclonucléosides et qu'elle reconnaît les pièges de manière aspécifique.

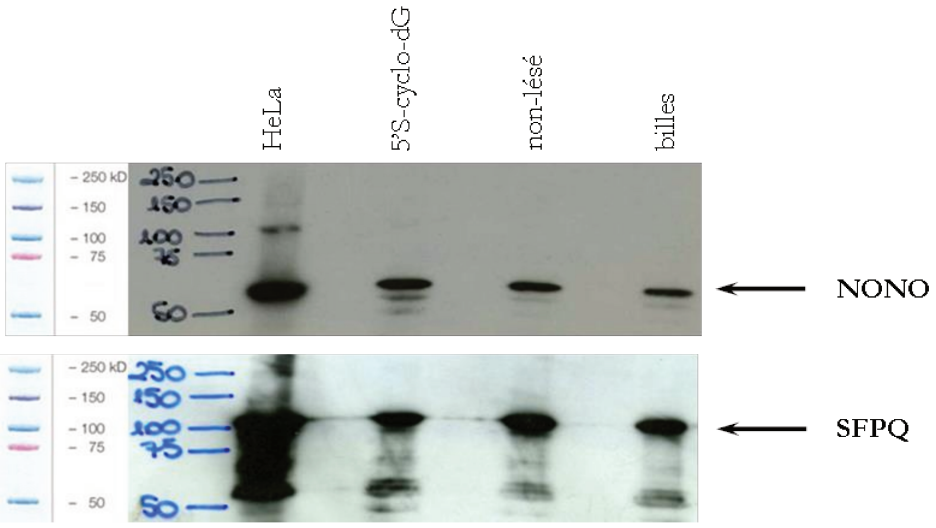


Figure 49 : Identification par Western blot des protéines NONO et SFPQ piégées ou non sur les sondes 5'S-cyclo-dG, non-lésées ou sur les billes ; les extraits nucléaires HeLa ont été utilisées en tant que contrôle.

Les deux autres protéines dont nous avons testé en Western blot l'affinité pour nos sondes sont DREF et PARP1. Elles ont été identifiées par la spectométrie de masse principalement sur les sondes non-lésées. Ainsi, elles ne seraient pas spécifiques des cyclonucléosides, mais, au contraire, la présence de ces lésions dans la séquence de l'ADN les empêcherait de se fixer sur celui-ci.

DREF

La protéine DREF devrait être détectée par son anticorps autour de 78 kDa. Comme nous pouvons le constater sur la Figure 50, cette bande est absente des contrôles non-exposés aux pièges. Néanmoins, une bande à 75 kDa environ apparaît dans l'échantillon provenant du piégeage sur les sondes non-lésées, correspondant ainsi à DREF. Son apparition dans un échantillon provenant du piégeage pourrait s'expliquer par la faible quantité de la protéine dans les extraits : l'anticorps n'est pas suffisamment sensible pour pouvoir la détecter dans 20 µg d'extraits nucléaires, mais elle devient plus abondante dans l'échantillon provenant de l'incubation des sondes avec quelques centaines de µg de ces extraits. C'est alors l'enrichissement des extraits en protéines minoritaires qui rend la détection de la protéine DREF possible.

Le Western blot que nous avons effectué permet de constater que DREF reconnaît nos sondes non-lésées, et que cette reconnaissance est rendue impossible par la présence d'un

cyclonucléoside. Ce résultat a été reproductible dans trois analyses protéomiques consécutives, et la détection de DREF par son anticorps renforce ces résultats. Pour analyser plus en détails les interactions qui existent entre cette protéine et l'ADN lésé ou non, nous allons procéder par la suite à des tests complémentaires, notamment l'EMSA.

PARP1

La dernière protéine qui a fait l'objet des analyses en Western blot est PARP1. Son anticorps doit reconnaître la forme entière de la protéine à 116 kDa, ainsi qu'un fragment à 89 kDa. Dans les extraits HeLa une bande de forte intensité est détectée au dessus de 100 kDa, et deux bandes moins intenses sont également visibles à 100 kDa et à 60 kDa environ (Figure 50). Dans les échantillons provenant du piégeage, les bandes minoritaires deviennent plus faibles, et quelques petites bandes de faible taille apparaissent également. La bande qui était la plus intense dans les extraits, est présente dans la condition « non-lésée ». Elle est absente des échantillons récupérés sur les billes, et très faible sur les sondes lésées.

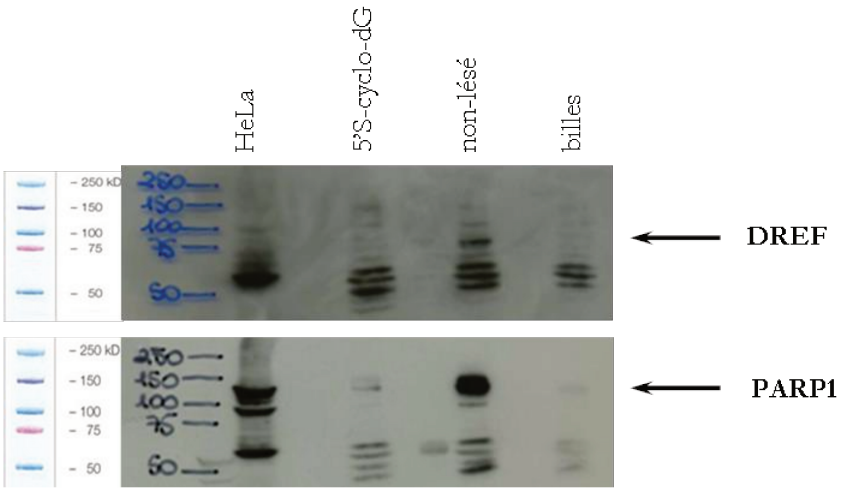


Figure 50 : Identification par Western blot des protéines DREF et PARP1 piégées ou non sur les sondes 5'S-cyclo-dG, non-lésées ou sur les billes ; les extraits nucléaires HeLa ont été utilisées en tant que contrôle.

Ce test nous permet de confirmer la présence de PARP1 principalement sur les sondes non-lésées par rapport aux sondes contenant des cyclonucléosides. Nous avons voulu vérifier si la présence des cyclonucléosides dans la séquence d'ADN empêchait la reconnaissance de celle-ci par PARP1, ou s'il s'agissait d'une modification des interactions dans le temps. Nous avons pour

cela effectué des incubations à temps variables, allant de 5 à 20 minutes, et nous avons ensuite procédé à la détection de PARP1 dans les échantillons récupérés, comme décrit précédemment. Nous avons utilisé les trois conditions : sondes non-lésées, sondes 5'R-cyclo-dA et billes. Les résultats de la détection de PARP1 sont présentés sur la Figure 51. Tout d'abord, nous pouvons voir que PARP1 n'a pas été piégé sur les billes, quel que soit le temps de l'incubation. Le profil de la détection change avec le temps, et nous pouvons constater la présence de PARP1 à la fois sur les sondes non-lésées et lésées, quand l'incubation avec des extraits ne dure que 5 minutes. Si cette incubation est prolongée jusqu'à 10 minutes, le signal visible pour les sondes contenant des cyclonucléosides est toujours présent, mais il est plus faible qu'à 5 minutes et que celui des sondes non-lésées. Enfin, à 20 minutes d'incubation, on retrouve le même profil que celui que nous avons observé auparavant : PARP1 est présent uniquement sur les sondes non-lésées, et on ne le détecte plus sur les sondes contenant un cyclonucléoside. Ceci semble montrer que la présence d'un cyclonucléoside dans l'ADN n'empêche pas la reconnaissance de celui-ci par PARP1, mais qu'elle favorise plutôt sa dissociation après une reconnaissance initiale.

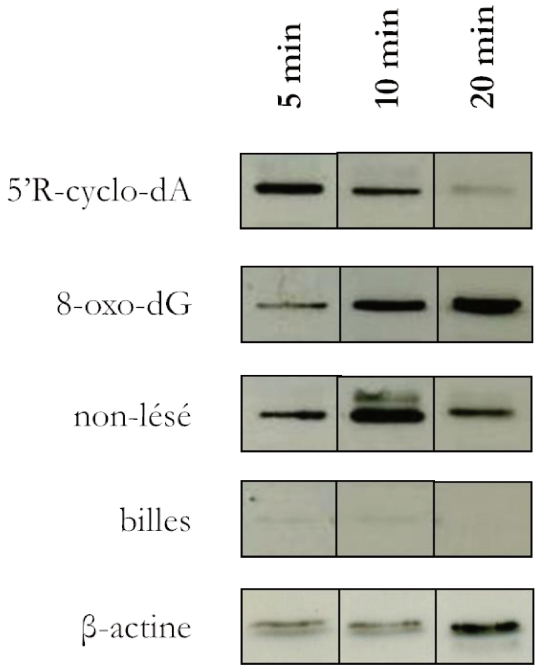


Figure 51 : Cinétique du piégeage de PARP1 sur les sondes 5'R-cyclo-dA, 8-oxo-dG, non-lésées et sur les billes, à 5, 10 ou 20 minutes d'incubation. Un contrôle de β-actine a également été effectué.

Afin de vérifier si les lésions simples d'oxydation avaient la même influence que les cyclonucléosides sur la liaison de PARP1 à nos sondes, nous avons incubé les pièges contenant

les 8-oxo-dG avec les extraits HeLa, en faisant varier le temps de l'incubation de 5 à 20 minutes. Le profil observé dans ce cas est inverse à celui que nous avons vu pour les sondes 5'R-cyclo-dA, car sa quantité ne diminue pas avec le temps sur les sondes contenant une 8-oxo-dG (Figure 51).

Les expériences effectuées en Western blot n'ont pas confirmé la présence des interactions entre les cyclonucléosides et les protéines NONO et SFPQ.

L'abolition des interactions de PARP1 et de DREF avec l'ADN non-lésé par la présence des cyclonucléosides a été confirmée par ces expériences complémentaires, les interactions PARP1−ADN se trouvant, de plus, dépendantes du temps de la réaction.

1.5.2. EMSA

La deuxième technique que nous avons utilisée pour confirmer les interactions observées entre l'ADN et PARP1 ou DREF, est l'EMSA. Contrairement aux expériences que nous avons effectuées sur les pièges, où des sondes contenant plusieurs lésions ont été incubées avec des extraits complexes, dans l'EMSA une courte séquence d'ADN contenant une seule lésion va être utilisée. L'inconvénient de cette approche est que si des facteurs supplémentaires sont nécessaires pour l'association d'une protéine donnée à l'ADN, ou bien si elle interagit au sein d'un complexe avec d'autres protéines, ces relations ne pourront pas être analysées.

DREF

Les sondes non-lésées et les sondes contenant une 5'S-cyclo-dG ont été incubées avec la protéine purifiée DREF, dont une gamme de concentration a été appliquée. Comme nous pouvons le voir sur la Figure 52, une faible formation des complexes entre l'ADN non-lésé et DREF est visible – ces complexes n'engagent pas la totalité de l'oligonucléotide disponible, puisque la majorité des sondes reste à l'état libre, même si la concentration de la protéine augmente. Ainsi on voit que les interactions de DREF avec nos sondes ne sont pas limitées par la disponibilité de celles-ci, mais plutôt par l'affinité de la protéine.

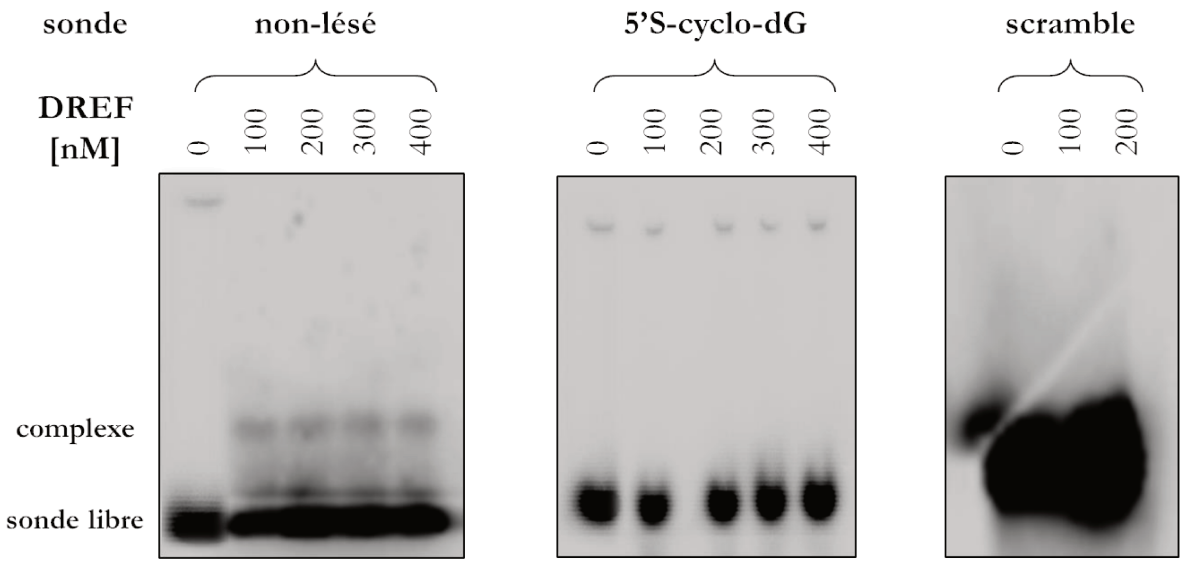


Figure 52 : Formation des complexes entre la protéine recombinante DREF et les sondes : non-lésée, 5'S-cyclo-dG ou scramble.

La raison probable du piégeage de DREF sur nos sondes non-lésées était la présence de sa séquence cible dans celles-ci. Afin de vérifier si la fixation de DREF sur les sondes non-lésées est effectivement séquence-dépendante, nous avons conçu des sondes appelée « scramble ». Ces sondes ont la même composition en nucléotides que les sondes non-lésées, mais leur séquence est différente. Le retard sur gel effectué avec ces sondes montre que, contrairement aux sondes non-lésées, aucun complexe n'est formé avec les sondes « scramble », ce qui confirme la fixation séquence-spécifique de DREF (Figure 52).

PARP1

Nous avons ensuite effectué le même type d'expérience avec la protéine PARP1, dont une gamme de concentration a été incubée avec les sondes non-lésées ou 5'S-cyclo-dG (Figure 53). Nous pouvons constater une formation des complexes sur les deux types d'oligonucléotides, à partir de 300 nM d'enzyme. Une légère différence est visible à cette dose entre les deux sondes, car plus de complexes sont visibles sur les sondes lésées que sur les sondes non-lésées. Ainsi, cette expérience ne confirme pas les résultats obtenus par l'analyse protéomique et par le Western blot. Ceci peut être expliqué par le fait que pour ces expériences une protéine purifiée a été utilisée, et qu'elle est alors dépourvue de ses partenaires qui étaient présents dans les extraits nucléaires. Cette explication va être développée dans le chapitre « Discussion ».

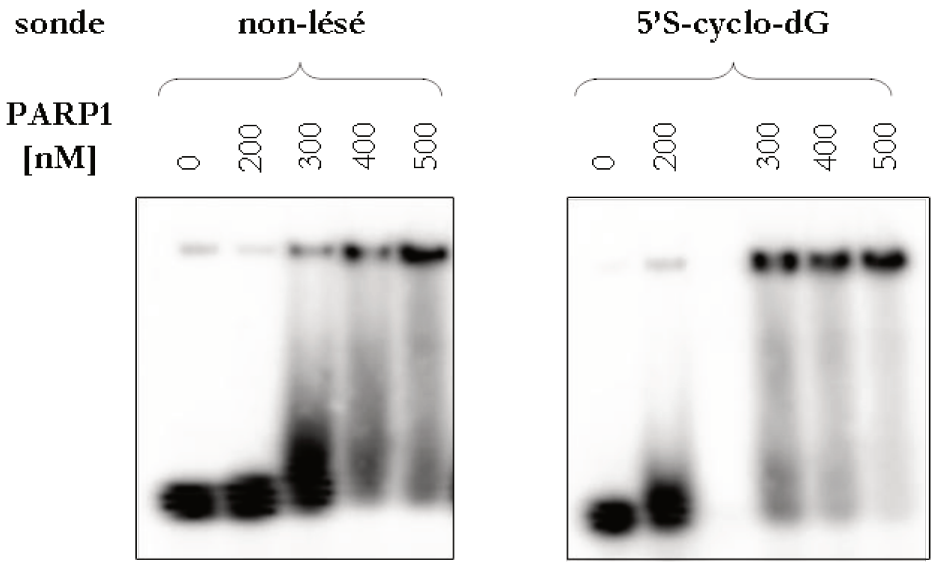


Figure 53 : Formation des complexes entre la protéine recombinante PARP1 et les sondes : non-lésée ou 5'S-cyclo-dG.

Les expériences du retard sur gel ont confirmé la reconnaissance des sondes non-lésées par DREF, cette reconnaissance étant inhibée par la présence des cyclonucléosides.

Le même résultat n'est cependant pas obtenu pour PARP1, dont les interactions avec la sonde ne semblent pas perturbées par la présence d'une lésion.

1.6. Identification par analyse protéomique des protéines bactériennes piégées sur les cyclonucléosides

Les essais de piégeage des protéines sur les cyclonucléosides ont également été effectués sur les extraits protéiques d'*Escherichia coli*. Les expériences antérieures de notre équipe ayant démontré une interaction entre la Fpg et les cyclonucléosides, le but premier était de valider notre piège en essayant de piéger cette glycosylase. Puis, si ce but est atteint, l'objectif était d'identifier d'éventuelles autres protéines capables de reconnaître ces dommages de l'ADN. Les mêmes stratégies que celles décrites pour les extraits humains ont été utilisées, à savoir le piégeage sur les sondes courtes et sur les concatémères.

1.6.1. Premier format de piège : sondes courtes

La première analyse protéomique a été effectuée sur les échantillons provenant du piégeage sur les sondes courtes (premier format des pièges décrit dans le paragraphe 1.1, analyse effectuée par le laboratoire EDyP). Dans cette analyse, uniquement les sondes 5'S-cyclo-dG et les sondes non-lésées ont été utilisées, et nous n'avons pas effectué un contrôle sur les billes. La Figure 54 résume de manière quantitative les résultats de ce piégeage.

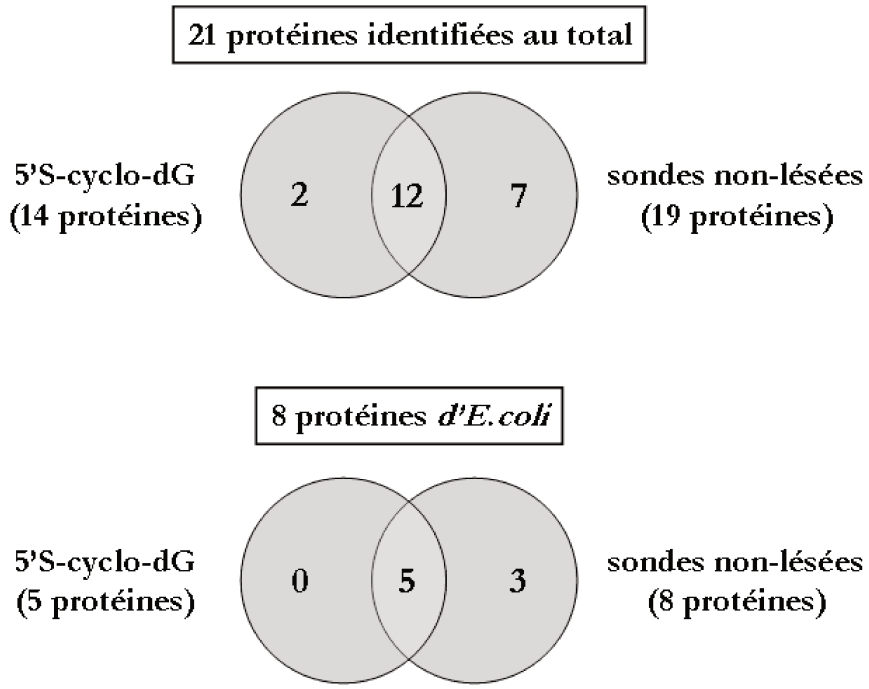


Figure 54 : Analyse quantitative des protéines provenant des extraits *E.coli*, identifiées sur les sondes 5'S-cyclo-dG ou non-lésées, récupérées par irradiation UVA.

De manière générale, très peu de protéines ont été piégées : seulement 21, ce qui représente un infime pourcentage de la quantité des protéines piégées dans les mêmes conditions dans les extraits humains (environ 400). De plus, parmi ces protéines, seulement 8 proviennent d'*E.coli*, les autres étant dues à des contaminations, notamment par des kératines humaines. Parmi ces 8 protéines, aucune n'est spécifique des sondes lésées, et aucune n'a de pertinence biologique dans la prise en charge de dommages de l'ADN. Ces résultats montrent une extrêmement faible capacité de ces pièges à capturer les protéines bactériennes.

Contrairement aux expériences avec des extraits nucléaires humains, où de nombreuses protéines ont été fixées sur les pièges, très peu de protéines associées aux sondes ont pu être identifiées dans les extraits bactériens. Nous avons alors changé de stratégie de piégeage et utilisé les sondes longues afin de favoriser les interactions ADN−protéines.

1.6.2. Deuxième format de piège : concatémères

Le deuxième format de pièges – les pièges contenant les concatémères, dont l'optimisation a été décrite dans le paragraphe 1.2 – a été appliqué aux extraits bactériens.

De manière analogue aux expériences effectuées sur les extraits nucléaires des cellules HeLa, nous avons voulu déterminer les meilleures conditions d'incubation. Deux temps d'incubation : 20 minutes et 2 heures, ainsi que deux tampons : A et RBP, ont été comparés.

Les protéines piégées sur les sondes non-lésées ou sur les billes ont été visualisées sur un gel SDS-PAGE coloré au nitrate d'argent (Figure 55).

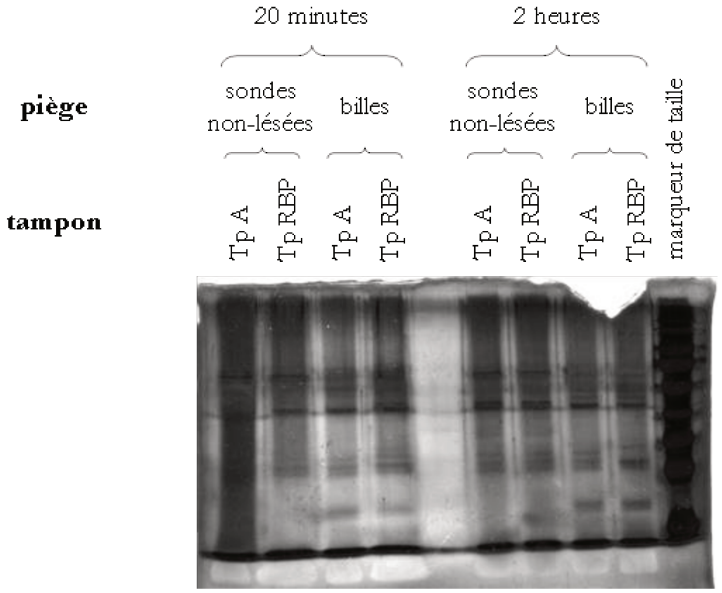


Figure 55 : Protéines bactériennes (extraits *E.coli*) piégées sur les sondes non-lésées ou sur les billes, dans différentes conditions d'incubation (20 minutes/2 heures d'exposition, tampon A/tampon RBP). Les protéines sont analysées sur un gel SDS-PAGE, coloré ensuite au nitrate d'argent.

La comparaison des conditions d'incubation des pièges avec les extraits protéiques d'origine bactérienne (*E.coli*) ne donne pas le même type de résultats que ce que nous avons observé pour les extraits humains. De manière générale nous pouvons voir que la quantité des protéines piégées est faible, quelle que soit la sonde, et ceci indépendamment du tampon utilisé ou du temps de l'incubation.

Afin de vérifier les résultats obtenus sur gel, une analyse protéomique avec trois types de pièges a été réalisée : 5'S-cyclo-dG, sondes non-lésées et billes (Tableau 14, analyse effectuée par le laboratoire IBEB). Elle confirme le faible nombre des protéines piégées, quelle que soit la condition de l'incubation. Entre 2 et 12 protéines ont été identifiées par échantillon. De plus, beaucoup de ces protéines ont été identifiées avec un seul peptide, ou avec un score Mascot inférieur à 70, ce qui ne permet pas de considérer leur identification comme fiable. Ainsi, si on ne prend en compte que les protéines dont le score Mascot dépasse le seuil de 70, on peut constater qu'entre 1 et 5 protéines seulement sont présentes par échantillon. De plus, dans 11 échantillons sur 12 la protéine la plus abondante, et souvent la seule identifiée de manière fiable, est le facteur d'élongation EF-Tu 2 (dans le 12ème échantillon il a été identifié avec un faible score Mascot). Les autres protéines ne possèdent pas non plus de rôle dans le traitement ou la réparation de l'ADN. Ainsi, nous constatons que le piège que nous avons conçu ne nous permet pas d'identifier des protéines bactériennes impliquées dans la prise en charge des cyclonucléosides.

Tableau 14 : Nombre des protéines bactériennes identifiées sur les pièges dans différentes conditions de l'incubation.

temps	**20 minutes**						**2 heures**					
tampon	**tampon A**			**tampon RBP**			**tampon A**			**tampon RBP**		
sonde	5'Cyclo-dG	non-lésées	billes	5'Cyclo-dG	non-lésées	billes	5'Cyclo-dG	non-lésées	billes	5'Cyclo-dG	non-lésées	billes
nombre de protéines identifiées	3	2	8	4	6	12	4	5	9	7	10	7
dont : celles avec un score Mascot > 70	1	1	2	1	2	5	1	1	2	1	2	4

Vu la très faible quantité de protéines bactériennes identifiées sur nos pièges, nous avons abandonné l'utilisation des extraits d'*E.coli* et limité nos analyses aux extraits des cellules HeLa, qui semblent être plus adaptés à ces pièges.

2. Travail sur les protéines isolées

La deuxième partie des travaux présentés ici concerne le travail sur les protéines isolées et non plus les extraits protéiques. La protéine bactérienne Fpg et deux protéines de *Thermococcus gammatolerans* : tg1653 et tg1277 ont été étudiées. Pour chacune de ces protéines nous avons effectué des tests de coupure sur différentes lésions simples ou complexes, et leurs interactions avec les cyclonucléosides ont également été étudiées.

La première partie de ce chapitre est consacrée au travail sur la Fpg bactérienne, la seconde à l'étude des enzymes de *T.gammatolerans*.

2.1. Etude biochimique de la protéine Fpg

2.1.1. Activité de coupure

Dans la première partie du chapitre consacré à la protéine Fpg, son activité de coupure sur différentes lésions simples ou complexes va être présentée.

2.1.1.1. Test d'activité de la Fpg sur les lésions simples

Afin d'étudier l'activité enzymatique de la Fpg sur les sondes lésées, celles-ci ont été marquées radioactivement et incubées avec l'enzyme. Les fragments entiers et coupés de l'oligonucléotide résultats de l'activité de l'enzyme ont été ensuite séparés sur un gel d'acrylamide.

A titre de comparaison, une enzyme possédant un mécanisme d'action différent de celui de la Fpg, la N-glycosylase/AP-lyase Endo III d'*Escherichia coli*, a également été étudiée. Cette enzyme, qui exerce une coupure via un mécanisme de β-élimination, est active sur les lésions suivantes : urée, 5,6-dihydroxythymine, 6-hydroxy-5,6-dihydrothymine, glycol de thymine, glycol d'uracile, 5-hydroxyhydantoïne et 5-hydroxy-5-méthylhydantoïne.

Chacune de ces enzymes a été incubée avec un oligonucléotide portant dans sa séquence une 8-oxo-dG. D'autres lésions simples ou analogues : 8-méthyl-désoxyguanosine (8-Me-dG), 8-méthoxy-désoxyguanosine (8-oMe-dG) et 8-bromo-désoxyguanosine (8-Br-dG) ont également été étudiées, pour vérifier si une activité de coupure est exercée sur ces dommages originaux. De

plus, l'Endo III a également été incubée avec une sonde contenant la 5-hydroxy-désoxyhydantoïne (5-OH-dHyd), qui a servi de témoin positif afin de vérifier son activité, l'activité de l'Endo III sur la 5-OH-dHyd et sur la 5-méthyl-5-hydroxy-désoxyhydantoïne (5-Me-5-OH-dHyd) ayant été démontrée il y a quelques années [317,318]. Les structures des lésions étudiées sont présentées sur la Figure 56. Une sonde non-lésée a également été testée dans cette étude afin d'exclure l'activité aspécifique des enzymes. Tous les oligonucléotides ont été incubés en absence d'enzymes : (1) pour confirmer qu'il n'y a pas de dégradation spontanée des sondes et (2) pour servir de témoin de taille.

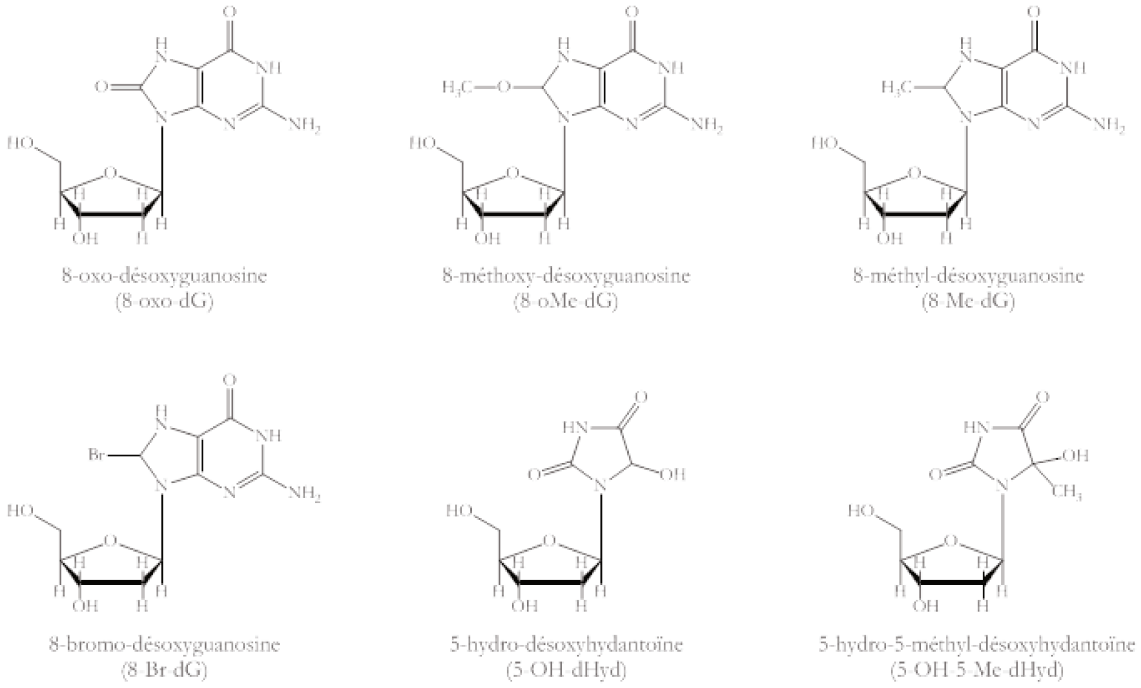


Figure 56 : Structures des lésions simples ou analogues utilisées lors des tests d'activité enzymatique.

L'incubation des six sondes en absence d'enzymes prouve qu'il n'y a pas de dégradation spontanée des oligonucléotides due à la simple présence de la base modifiée dans leur séquence. La légère différence entre les vitesses de migration des différentes sondes provient de la présence des lésions – chaque modification apportée à la séquence influence sa masse et sa charge, et par conséquence sa mobilité éléctrophorétique (Figure 57).

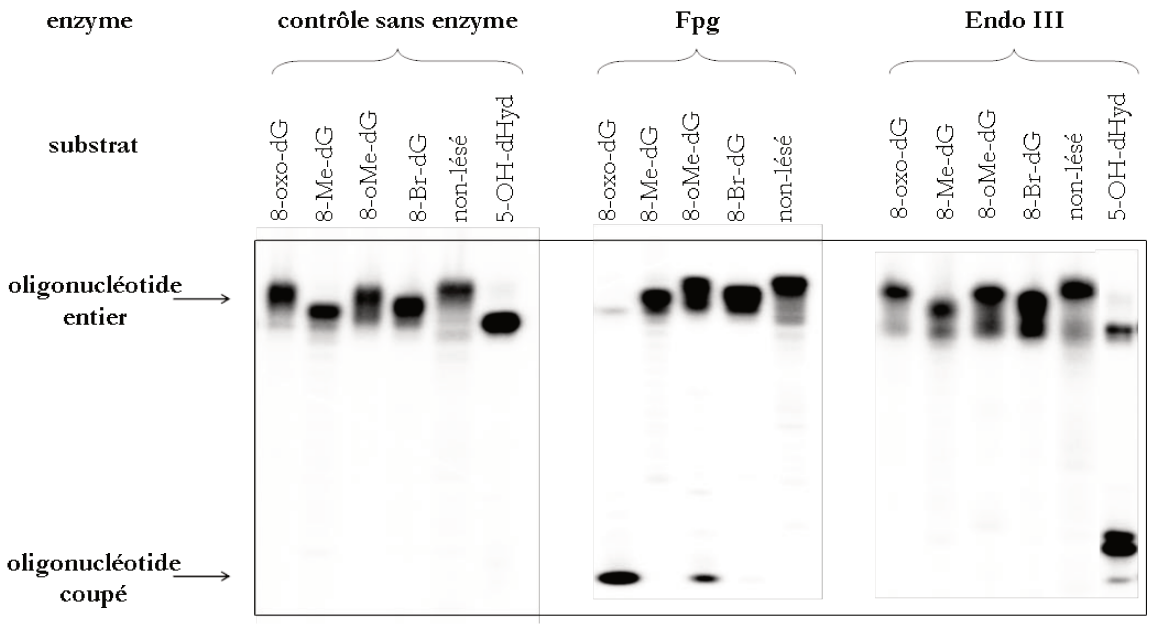


Figure 57 : Activité de coupure des enzymes bactériennes : glycosylase Fpg et endonucléase Endo III sur les oligonucléotides contenant une lésion simple d'oxydation ou un analogue.

Quand ces mêmes sondes sont incubées avec la Fpg, nous remarquons dans un premier temps la coupure au niveau de la 8-oxo-dG, cette lésion servant de témoin positif pour vérifier l'activité enzymatique et le bon choix de conditions d'incubation. Une faible coupure de la sonde contenant une 8-oMe-dG est également constatée. L'excision de la 8-oMe-dG peut signifier que la Fpg aurait une faible activité au niveau de cette lésion, mais il est également possible que la sonde 8-oMe-dG contienne des traces de 8-oxo-dG et que ce soit cette deuxième lésion « contaminante » qui est excisée.

Lorsque les sondes contenant une lésion simple sont incubées avec l'Endo III, la coupure au niveau de la 8-oxo-dG n'est plus observée, puisque cette lésion ne fait pas partie de ses substrats. La lésion utilisée en tant que contrôle positif de l'activité enzymatique est alors la 5-OH-dHyd. Nous pouvons constater également, que le fragment obtenu lors de la coupure par l'Endo III ne migre pas à la même vitesse que celui généré par l'action de la Fpg. Ceci est dû aux différences dans leurs mécanismes d'action, qui génèrent des fragments contenant ou pas le résidu du sucre.

Les autres lésions testées ne semblent pas être des substrats pour la Fpg et pour l'Endo III, puisqu'aucune coupure, même partielle, n'a été constatée dans nos conditions d'incubation.

2.1.1.2. Test d'activité de la Fpg sur les lésions complexes d'oxydation

Après avoir testé l'activité de la Fpg et de l'Endo III sur les lésions simples, nous avons procédé au même type d'expériences avec des sondes contenant cette fois-ci des lésions complexes. Les lésions complexes choisies pour cette analyse sont les quatre cyclonucléosides suivants : 5'S-cyclo-dA, 5'R-cyclo-dA, 5'S-cyclo-dG et 5'S-cyclo-dHT.

Quelle que soit l'enzyme appliquée, on ne constate aucune activité enzymatique de glycosylase ni d'endonucléase au niveau des cyclonucléosides (Figure 58). Ceci confirme que ces lésions, bien que faisant partie des lésions d'oxydation, ne sont pas prises en charge par le système BER.

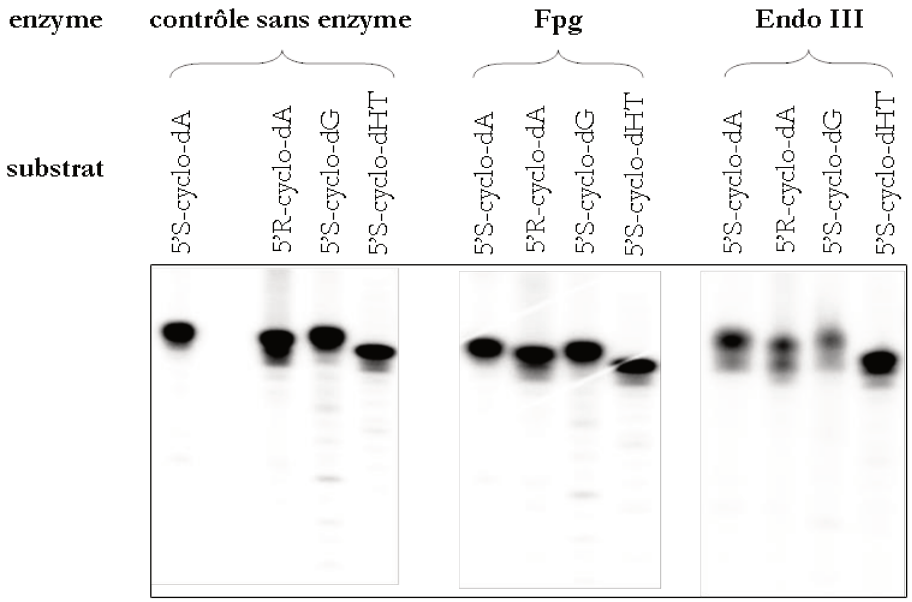


Figure 58 : Activité de coupure des enzymes bactériennes : glycosylase Fpg et endonucléase Endo III sur les oligonucléotides contenant un cyclonucléoside.

2.1.2. Analyse des interactions entre les cyclonucléosides et la Fpg

Comme cela a été présenté ci-dessus, les enzymes du BER ne sont pas capables d'exciser les cyclonucléosides de l'ADN. Néanmoins, une expérience menée en 2008 au sein de notre laboratoire a prouvé qu'une reconnaissance des cyclonucléosides par la Fpg avait lieu, sans qu'une incision soit effectuée [281]. Nous avons alors décidé d'analyser plus précisément ces interactions par l'EMSA.

Nous avons étudié les interactions entre la Fpg purifiée et les oligonucléotides contenant l'un des quatre cyclonucléosides utilisés précédemment dans les tests de coupure : 5'S-cyclo-dA,

5'R-cyclo-dA, 5'S-cyclo-dG et 5'S-cyclo-dHT. En tant que témoin négatif, nous avons utilisé une sonde non-lésée, qui représente les interactions non-spécifiques qu'il peut y avoir entre une séquence quelconque d'ADN et la Fpg. Le témoin positif que nous avons choisi est l'oligonucléotide contenant un site abasique modifié : le tétrahydrofurane (THF). Il est connu que la Fpg possède une forte affinité pour les sites abasiques [319,320]. Ainsi, le THF est reconnu comme un site abasique « classique », sans toutefois être excisé par la suite. Quand un site abasique est formé dans la séquence d'ADN, un groupement -OH est introduit à la place de la base pour donner un hémiacétal. Cette forme fermée du sucre peut facilement être ouverte et un équilibre entre trois formes est observé : hémiacétal, aldéhyde (sucre ouvert) et aldéhyde hydraté (Figure 59).

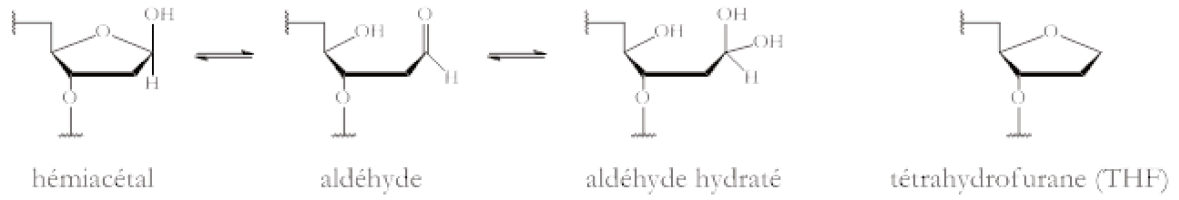


Figure 59 : Différences structurelles entre le site abasique « naturel » et son homologue THF. Inspiré de [139].

La présence de la forme ouverte du désoxyribose est nécessaire pour que la coupure puisse ensuite être exercée par les enzymes du BER. Le tétrahydrofurane est un analogue de site abasique inactif, puisqu'il ne possède pas de groupement hydroxyle sur le carbone C1', et il ne peut pas être transformé en forme ouverte. Il est alors reconnu par les ADN glycosylases en tant qu'un site abasique, mais il n'est pas coupé par celles-ci. De cette manière, seule l'étape de reconnaissance de la lésion par la glycosylase est analysée, servant de contrôle positif afin d'évaluer la force d'interaction de la même protéine avec d'autres lésions de l'ADN [139].

Dans un premier temps, l'optimisation des conditions d'incubation a été nécessaire. Celle-ci consistait à adapter notamment : le temps d'incubation, la température des interactions, la composition du tampon et la quantité relative de la sonde et de la protéine.

2.1.2.1. Optimisation des conditions de l'EMSA

Pour commencer la mise au point des conditions expérimentales, nous avons procédé à un screening des paramètres, en testant notamment différentes températures d'incubation et différentes concentrations en protéines.

La première difficulté que nous avons rencontrée lors de l'optimisation des conditions d'interactions était la formation de bandes multiples, indiquant la présence de plusieurs complexes entre les sondes et la protéine (Figure 60). Ces complexes sont obtenus à la fois avec l'ADN non-lésé et avec les sondes lésées. La présence de complexes hétérogènes a été exclue, puisque l'enzyme que nous avons utilisée est purifiée et ne contient pas de traces d'autres protéines [278]. Ainsi, les deux complexes visibles doivent être composés de la même protéine, avec une ou deux de ses molécules fixées sur l'oligonucléotide. Il a déjà été démontré que la Fpg formait des complexes avec l'ADN lésé en ratio 1 :1 [321] – le deuxième complexe que l'on observe doit ainsi correspondre à des interactions non-spécifiques. Ces interactions peuvent être dues à la concentration trop élevée en protéine, ou à des conditions d'incubation qui ne sont pas suffisamment stringentes pour favoriser la présence des interactions spécifiques uniquement. La première hypothèse a été écartée, puisqu'on peut voir que le deuxième complexe se forme pratiquement en même temps que le premier, et avant que toutes les sondes soient impliquées dans les interactions avec la protéine. S'il s'agissait d'un complexe formé en excès de l'enzyme, le profil de sa formation serait différent : on observerait dans un premier temps la formation progressive d'un premier complexe et la disparition en parallèle de la sonde libre, et l'apparition du deuxième complexe plus tardivement, lors de la disparition totale des sondes libres.

La formation des doubles complexes sur nos sondes pourrait donc être due aux conditions d'incubation qui ne seraient pas assez sélectives et qui permettraient la fixation plus ou moins spécifique de plusieurs molécules de la protéine sur le même oligonucléotide. Le même phénomène est observé pour les sondes lésées et non-lésées, ce qui renforce cette hypothèse. La température ne semble pas jouer un rôle dans la formation des complexes dans les conditions testées, et nous avons donc choisi de continuer l'optimisation à température ambiante. Les modifications qui ont été appliquées consistaient à faire varier le temps de l'incubation et la composition du tampon. Le temps prolongé de l'incubation pourrait favoriser la formation des complexes aspécifiques. La salinité du tampon est un deuxième facteur qui joue un rôle extrêmement important dans les interactions ADN-protéines. Toutes les protéines nécessitent la présence de sels pour pouvoir exercer leur activité et également pour pouvoir interagir avec

l'ADN, mais l'excès de ceux-ci empêche la formation des complexes. Ainsi, chaque protéine possède des conditions optimales de liaison, qui doivent être adaptées à chaque fois qu'une nouvelle enzyme est testée. Afin d'éliminer les complexes formés de manière aspécifique, nous avons alors augmenté la salinité du tampon utilisé pour l'incubation, tout en diminuant le temps de celle-ci. Ainsi, nos conditions d'incubation sont passées d'1 heure dans un tampon à 20 mM de KCl, à 20 min dans un tampon à 150 mM de KCl. Nous avons également éliminé le bleu de dépôt, qui a été remplacé par l'ajout de 8% de glycérol dans le mélange – le glycérol joue ainsi un double rôle en stabilisant les complexes et en alourdissant les échantillons, facilitant leur dépôt dans les puits.

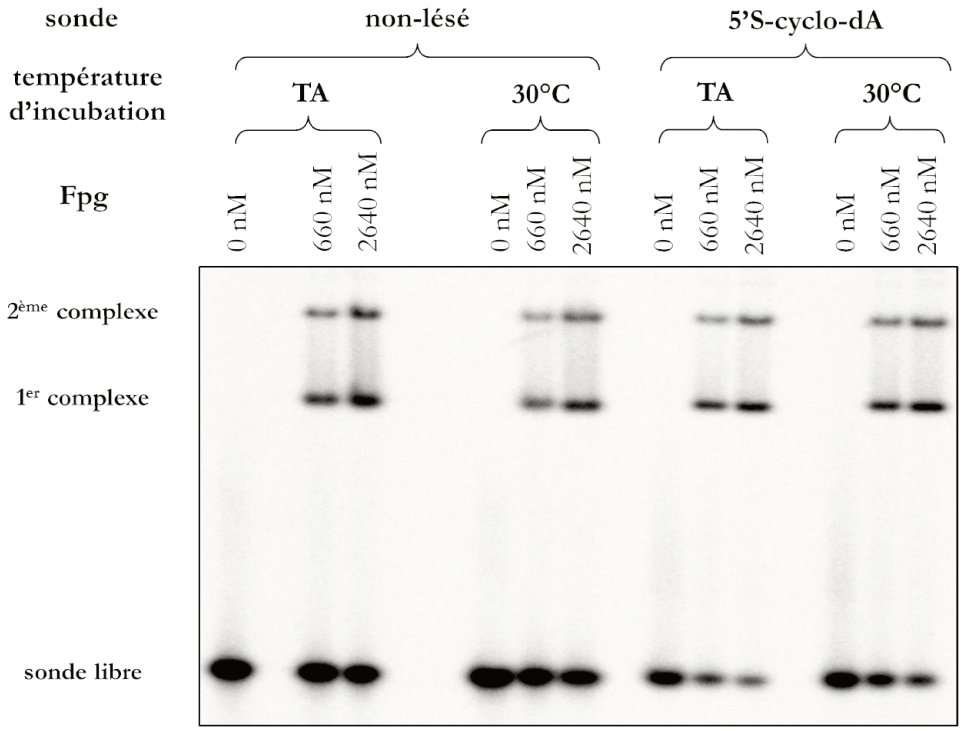


Figure 60 : Formation simultanée des complexes spécifiques et non-spécifiques entre la Fpg et les sondes lésées (5'S-cyclo-dA) ou non. TA = température ambiante.

Comme ceci est visible sur la Figure 61, la modification des conditions citées ci-dessus a mené à une amélioration des résultats. Plusieurs changements par rapport aux observations précédentes peuvent être constatés :

1) la bande correspondant à l'oligonucléotide libre disparaît complètement même à la plus petite concentration de l'enzyme appliquée ; une faible trace est visible uniquement pour le premier point de la gamme de concentration avec la sonde non-lésée – ceci indique que

nous avons atteint 100% des complexes et que les concentrations de la Fpg pour lesquelles on pourrait établir une courbe de formation de ces complexes sont inférieures à 550 nM ;

2) on voit toujours l'apparition d'un deuxième complexe, cette fois-ci formé plus tardivement ; dans un premier temps un seul complexe est créé avec la sonde non-lésée, et le deuxième apparaît progressivement ; pour la sonde lésée, les deux complexes sont formés dès la première concentration de la protéine, mais il y en a un qui est majoritaire, tandis que le deuxième devient plus marqué avec l'augmentation de quantité de Fpg, en remplaçant le premier. Ce deuxième complexe est plus lourd et il est probablement constitué de deux protéines.

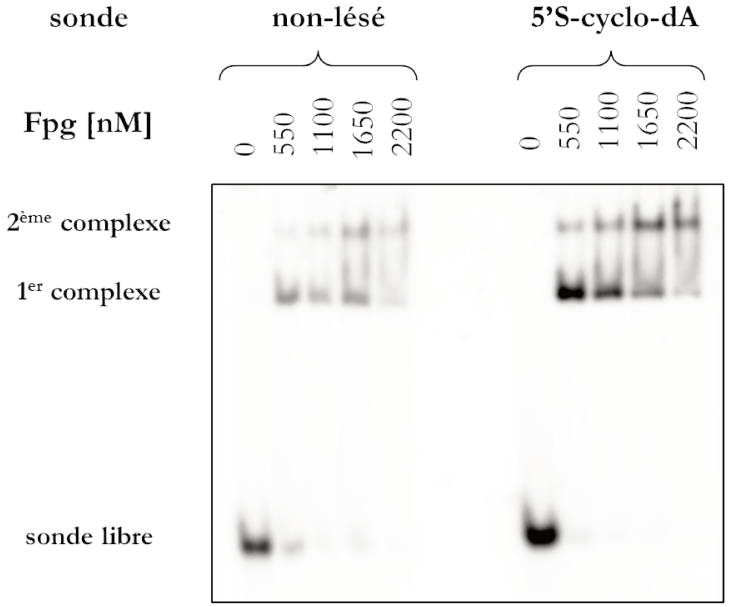


Figure 61 : Formation des complexes spécifiques et non-spécifiques entre la Fpg et les sondes lésées (5'S-cyclo-dA) ou non, après modification des conditions d'incubation : diminution du temps et augmentation de la salinité du tampon.

Nous avons réussi à optimiser les conditions d'incubation pour les interactions Fpg-cyclonucléosides. La dernière étape a consisté à sélectionner une gamme de concentrations de la protéine pour laquelle on obtient la formation progressive des complexes jusqu'à la disparition complète de la sonde seule. **Cette gamme a été déterminée entre 55 et 550 nM, avec les points distribués régulièrement, tous les 55 nM. Elle permet d'obtenir 100% de complexes entre la Fpg et le témoin positif, qui est la sonde portant un THF, et presque 100% de complexes avec les autres sondes.** Pour la sonde non-lésée, la disparition totale des sondes libres n'est pas obtenue, mais le choix de ces conditions permet d'observer la formation des complexes presque jusqu'à l'épuisement des sondes, sans qu'il y ait l'apparition des complexes

secondaires, dont la présence rendrait difficile la quantification et l'interprétation des résultats. Toutes les expériences ont été répétées trois fois, pour confirmer les résultats et vérifier leur reproductibilité.

2.1.2.2. Formation des complexes entre la Fpg et les sondes

Les figures 62 à 67 présentent les résultats obtenus pour la gamme sélectionnée de Fpg et les six sondes analysées. Sur chaque image, correspondant à un type de sonde, un gel est montré à titre d'exemple ainsi qu'un graphique représentant le pourcentage des complexes formés en fonction de la concentration en protéine. Les moyennes des trois expériences indépendantes ont été présentées avec leurs écart-types, et la courbe suit au plus proche leur tendance. Les calculs qui ont permis de tracer ces courbes et de déterminer la constante de dissociation vont être précisés par la suite.

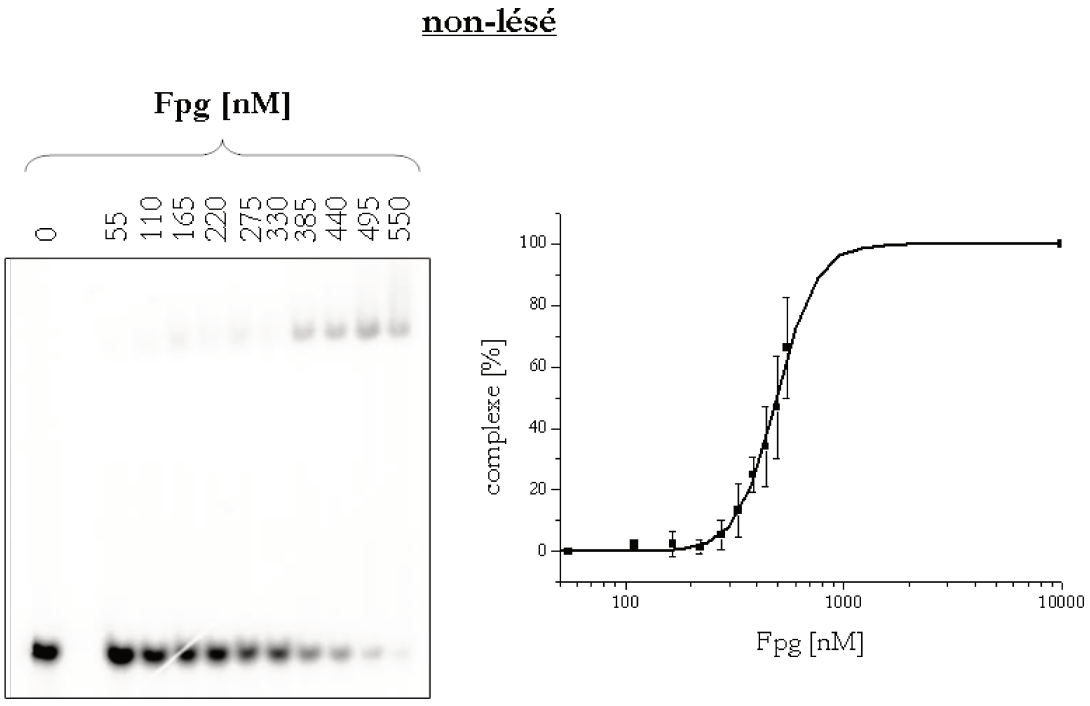


Figure 62 : Interactions entre un oligonucléotide non-lésé et la protéine Fpg ; gauche : exemple d'un gel shift, droite : courbe de formation des complexes obtenue avec trois expériences indépendantes.

5'S-cyclo-dA

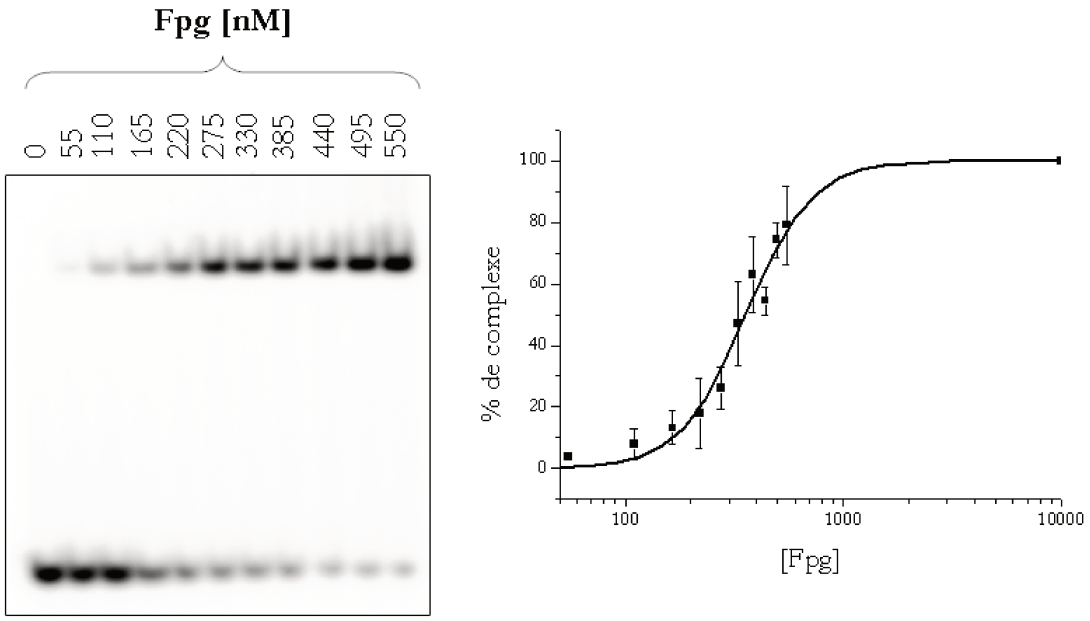


Figure 63 : Interactions entre un oligonucléotide contenant une 5'S-cyclo-dA et la protéine Fpg ; gauche : exemple d'un gel shift, droite : courbe de formation des complexes obtenue avec trois expériences indépendantes.

5'R-cyclo-dA

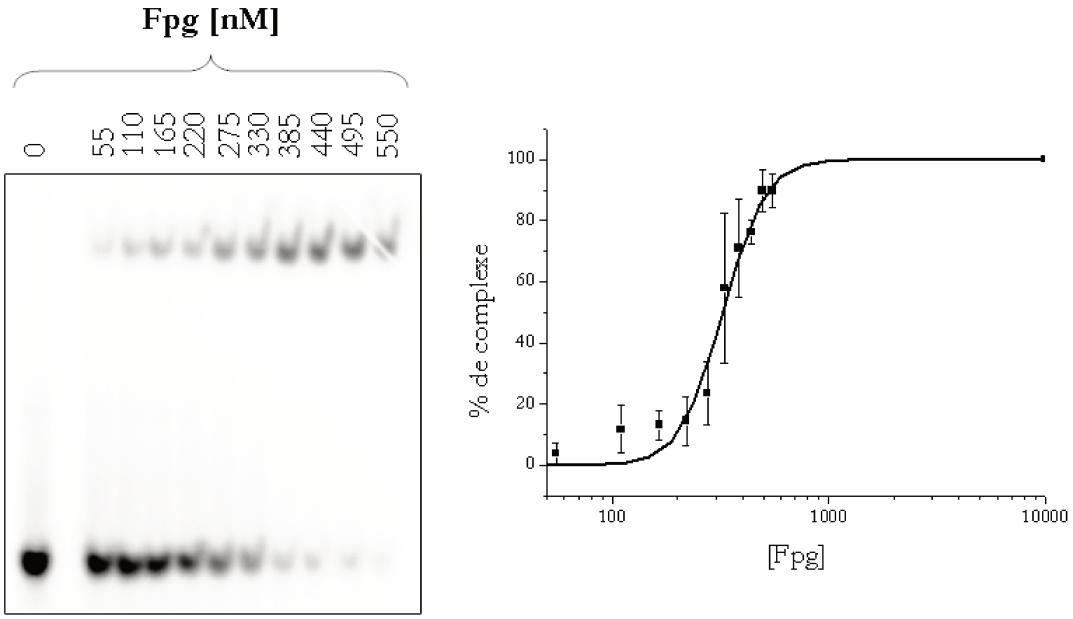


Figure 64 : Interactions entre un oligonucléotide contenant une 5'R-cyclo-dA et la protéine Fpg ; gauche : exemple d'un gel shift, droite : courbe de formation des complexes obtenue avec trois expériences indépendantes.

5'S-cyclo-dG

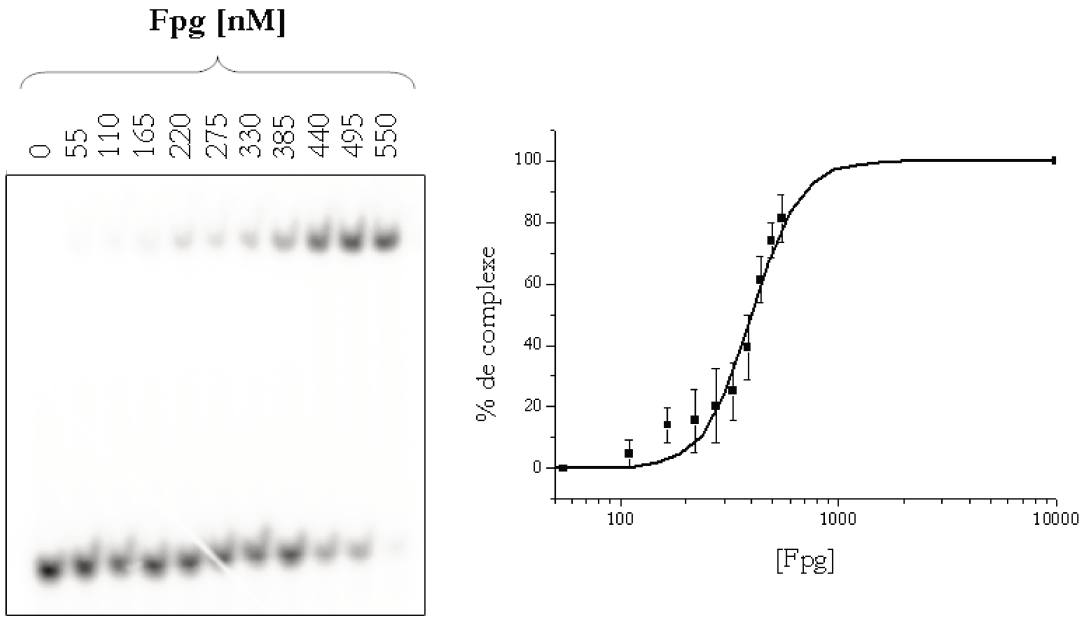


Figure 65 : Interactions entre un oligonucléotide contenant une 5'S-cyclo-dG et la protéine Fpg ; gauche : exemple d'un gel shift, droite : courbe de formation des complexes obtenue avec trois expériences indépendantes.

5'R-cyclo-dHT

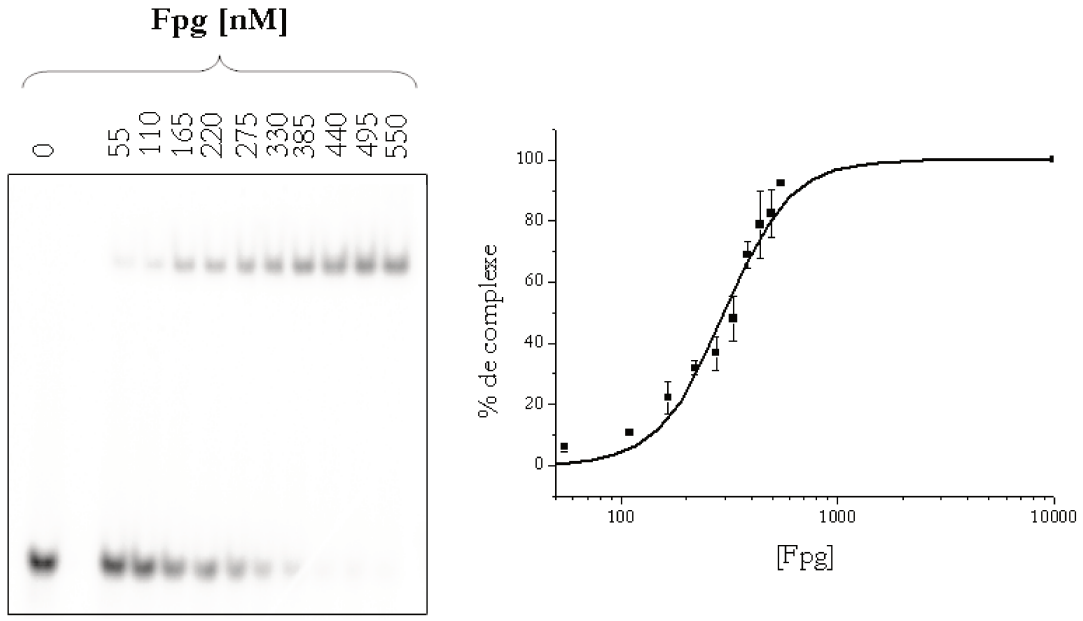


Figure 66 : Interactions entre un oligonucléotide contenant une 5'S-cyclo-dHT et la protéine Fpg ; gauche : exemple d'un gel shift, droite : courbe de formation des complexes obtenue avec trois expériences indépendantes.

THF

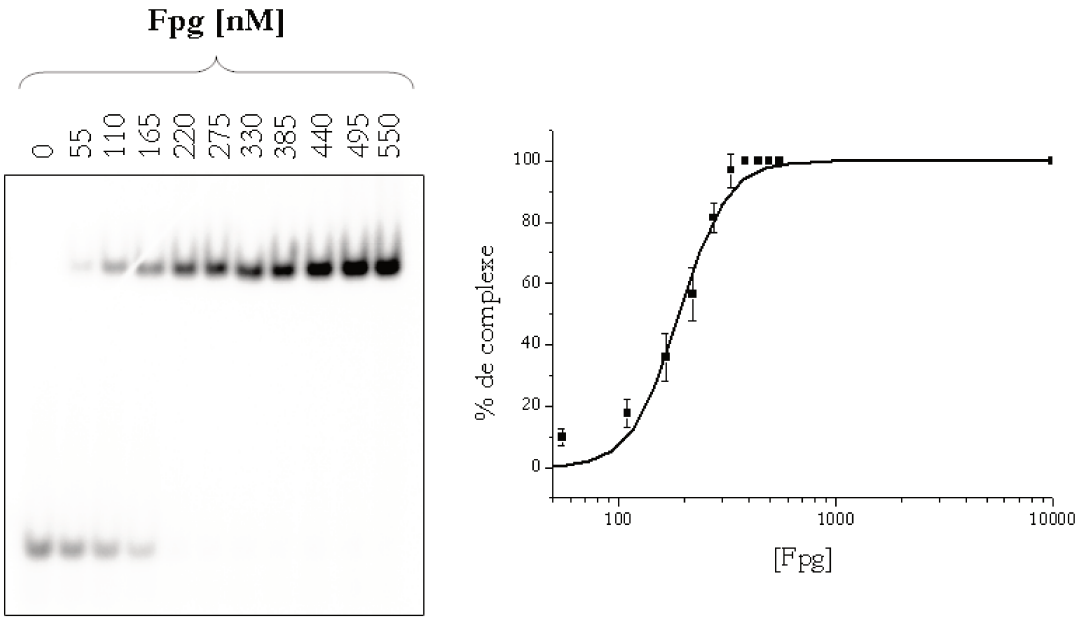


Figure 67 : Interactions entre un oligonucléotide contenant un THF et la protéine Fpg ; gauche : exemple d'un gel shift, droite : courbe de formation des complexes obtenue avec trois expériences indépendantes.

Pour chaque sonde, trois expériences indépendantes ont été effectuées et chaque gel a ensuite été quantifié. Pour chaque point de la gamme nous avons alors calculé le pourcentage moyen des complexes et les écart-types. Puisque pour aucune sonde, à l'exception de la sonde THF, nous n'avons pu obtenir 100% de complexes, cette valeur a été ajoutée sur le graphique de manière artificielle. Nous avons ainsi estimé qu'à 10 000 nM de Fpg – 20 fois plus que le dernier point de la gamme – toutes les sondes sont sous forme des complexes.

Après avoir tracé les points expérimentaux représentant le pourcentage moyen des complexes formés entre chacune des sondes et la Fpg, en fonction de la concentration de cette dernière, nous avons procédé aux calculs de la constante de dissociation. Cette valeur représente la quantité de la protéine pour laquelle 50% de sondes sont libres et 50% sont sous forme de complexes. Elle peut être déterminée simplement sur la base du graphique, mais elle peut être également calculée, si un modèle mathématique, permettant de tracer une courbe qui décrit la tendance des points expérimentaux, est disponible.

La première équation que nous avons voulu mettre en œuvre est dérivée de l'équilibre de formation des complexes entre la sonde et la protéine Fpg, qui sont créés avec une molécule de Fpg pour une molécule de sonde.

(1) $[\mathrm{Fpg}]+[\mathrm{sonde}]\leftrightarrow[\mathrm{complexe}]$

Ainsi, la constante de dissociation de la formation de ce complexe va être définie de manière suivante :

(2) $\mathrm{K_d}=\dfrac{[\mathrm{Fpg}]\times[\mathrm{sonde}]}{[\mathrm{complexe}]}$

De plus, d'après la loi de conservation de la matière on obtient deux relations :

(3) $[\mathrm{sonde}]_0=[\mathrm{sonde}]+[\mathrm{complexe}]$ et (4) $[\mathrm{Fpg}]_0=[\mathrm{Fpg}]+[\mathrm{complexe}]$

L'équation (2) peut être transformée en :

(5) $[\mathrm{complexe}]=\dfrac{[\mathrm{Fpg}]\times[\mathrm{sonde}]}{\mathrm{K_d}}$

qui est ensuite reporté dans l'équation (4) :

(6) $[\mathrm{Fpg}]_0=[\mathrm{Fpg}]\times\left\{1+\dfrac{[\mathrm{sonde}]}{\mathrm{K_d}}\right\}$ soit (7) $[\mathrm{Fpg}]=\dfrac{\mathrm{K_d}}{\mathrm{K_d}+[\mathrm{sonde}]}\times[\mathrm{Fpg}]_0$

Nous pouvons ensuite combiner les équations (5) et (7), ce qui va donner :

(8) $[\mathrm{complexe}]=\dfrac{[\mathrm{sonde}]}{\mathrm{K_d}+[\mathrm{sonde}]}\times[\mathrm{Fpg}]_0$

ou autrement : (9) $[\mathrm{sonde}]=\mathrm{K_d}\times\dfrac{[\mathrm{complexe}]}{[\mathrm{Fpg}]_0-[\mathrm{complexe}]}$

Nous pouvons ensuite introduire l'équation (9) dans l'équation (3) et obtenir :

$$(10)\ [\text{sonde}]_0 = [\text{complexe}] \times \left\{ 1 + \frac{K_d}{[\text{Fpg}]_0 - [\text{complexe}]} \right\}$$

Après transformation, on obtient une équation de deuxième degré :

$$(11)\ [\text{complexe}]^2 - \{[\text{sonde}]_0 + [\text{Fpg}]_0 + K_d\} \times [\text{complexe}] + [\text{sonde}]_0 \times [\text{Fpg}]_0 = 0$$

dont la solution est suivante : (12)

$$[\text{complexe}] = \frac{1}{2} \times \left\{ ([\text{sonde}]_0 + [\text{Fpg}]_0 + K_d) \pm \sqrt{([\text{sonde}]_0 + [\text{Fpg}]_0 + K_d)^2 - 4[\text{sonde}]_0[\text{Fpg}]_0} \right\}$$

Pour tracer une dépendance du pourcentage de complexes en fonction de la concentration de la Fpg, on va introduire des nouvelle inconnues :

$x = [\text{Fpg}]_0$,

$y = \frac{[\text{complexe}]}{[\text{sonde}]_0} \times 100$, qui est la valeur quantifiée sur le gel,

et $[\text{sonde}]_0 = 8.33\text{nM}$,

ce qui donne :

$$y = 6 \times \left[(8.33 + x + K_d) \pm \sqrt{(8.33 + x + K_d)^2 - 33.32x} \right]$$

Comme chaque équation de deuxième degré, celle-ci possède deux solutions (indiquées par « ± »). Pour en éliminer une, nous allons effectuer les calculs pour le cas le plus simple, soit en absence de protéine. Dans ce cas-là il n'y a pas de complexes, alors : x = 0 et y = 0. Quand on intègre ces valeurs dans l'équation finale, on obtient alors :

- pour (13) $y = 6 \times \left[(8.33 + x + K_d) + \sqrt{(8.33 + x + K_d)^2 - 33.32x} \right] \rightarrow \qquad K_d = -99.96\text{nM}$

- pour (14) $y = 6 \times \left[(8.33 + x + K_d) - \sqrt{(8.33 + x + K_d)^2 - 33.32x} \right] \rightarrow \quad 0 = 0$

Puisque la valeur de la constante de dissociation ne peut pas être négative, l'équation que nous allons utiliser pour tracer les courbes représentant la formation des complexes en fonction de la concentration de la protéine est :

(14) $y = 6 \times \left[(8.33 + x + K_d) - \sqrt{(8.33 + x + K_d)^2 - 33.32x} \right]$

En utilisant le logiciel Origin 7.5 nous avons alors essayé de tracer les courbes en utilisant les valeurs connues, c'est-à-dire : les pourcentages quantifiés des complexes (y) et la concentration initiale de la protéine (x). L'équation précise assignée à chacune de ces courbes permet d'obtenir comme solution la valeur de la constante de dissociation K_d pour chaque sonde utilisée. La Figure 68 A montre l'exemple d'une courbe tracée pour les interactions de la Fpg avec les sondes THF.

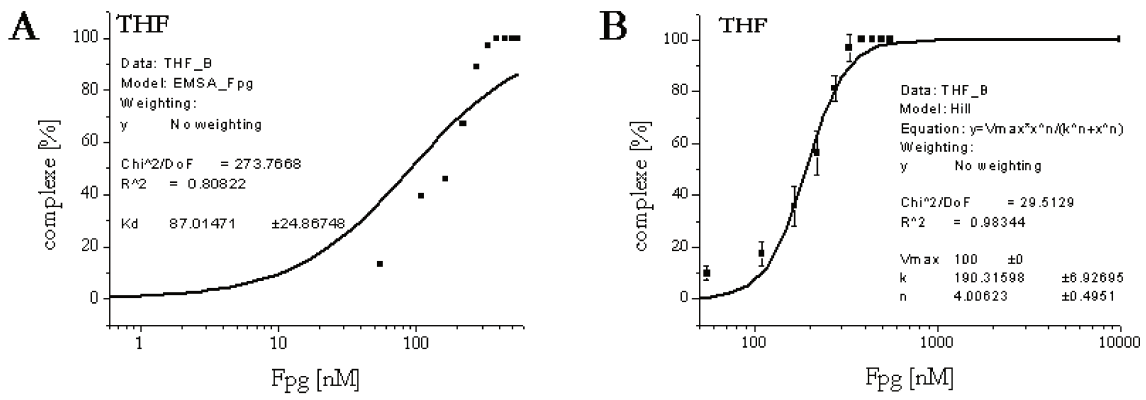


Figure 68 : Courbe tracée à l'aide du logiciel Origin représentant la tendance suivie par les points expérimentaux. A) courbe tracée d'après l'équation (14) ; B) courbe tracée d'après l'équation de Hill. Exemple concernant les interactions de la Fpg avec le sonde THF.

Nous pouvons constater, que la courbe qui a été tracée d'après l'équation (14), ne correspond pas aux points expérimentaux : elle ne les traverse pas et le coefficient de corrélation est de 0.8 seulement. Pour les autres sondes nous avons obtenu le même décalage. Il est alors évident que cette équation ne peut pas être appliquée à notre modèle expérimental, puisqu'elle ne reproduit pas les données empiriques. Nous n'avons pas pu expliquer la raison de ces divergences, mais il a été évident que nous étions obligés de chercher un autre modèle mathématique qui serait représentatif de nos résultats.

La recherche d'une représentation graphique des données expérimentales que nous avons obtenues, nous a conduits à utiliser le modèle de Hill :

(15) $y = 100 \times \frac{x^n}{x^n + K_d^n}$

où n = coefficient de Hill

y = pourcentage des complexes

x = concentration de la Fpg

Cette formule n'a pas, contrairement à la formule testée auparavant, de signification biologique, mais c'est un modèle qui s'est avéré le mieux adapté à nos conditions expérimentales et qui nous a permis de tracer des courbes en corrélation avec les données empiriques. Tous les coefficients de corrélation dépassent le seuil de 0.95 et ils se rapprochent même de 0.98, ce qui est une valeur entièrement satisfaisante, permettant de constater que ce modèle reflète la réalité expérimentale (voir Figure 68 B et Tableau 15).

L'équation de Hill que nous avons appliquée à nos résultats expérimentaux nous a permis également de déterminer la valeur de la constante de dissociation. Cette valeur peut être lue directement sur le graphique, mais sa détermination avec un modèle mathématique est plus précise. La constante de dissociation permet ainsi de comparer l'affinité de la même protéine pour différentes sondes, des différentes protéines pour la même sonde, ou encore des variations de l'affinité en fonction des conditions d'incubation. Dans notre cas, elle est utilisée pour évaluer la différence entre la reconnaissance par la Fpg d'une séquence d'ADN comportant des cyclonucléosides et de la même séquence qui ne comporte pas de lésion. Comme nous l'avons précisé avant, la sonde THF sert de référence, pour pouvoir ensuite évaluer les interactions avec les cyclonucléosides, par rapport à une lésion connue.

Nous pouvons constater (Figure 69 et Tableau 15) que l'affinité de la Fpg pour les quatre sondes portant l'un des cyclonucléosides est plus grande que son affinité pour la sonde non-lésée. Ceci signifie qu'il existe une reconnaissance de ces lésions par la Fpg, même si le rôle de ces interactions n'a pas été élucidé (comme nous l'avons démontré dans le paragraphe 2.1.1.2 de ce chapitre, la Fpg n'exerce pas d'activité de coupure sur les cyclonucléosides). Son affinité pour les cyclonucléosides reste néanmoins moins importante que celle pour le THF.

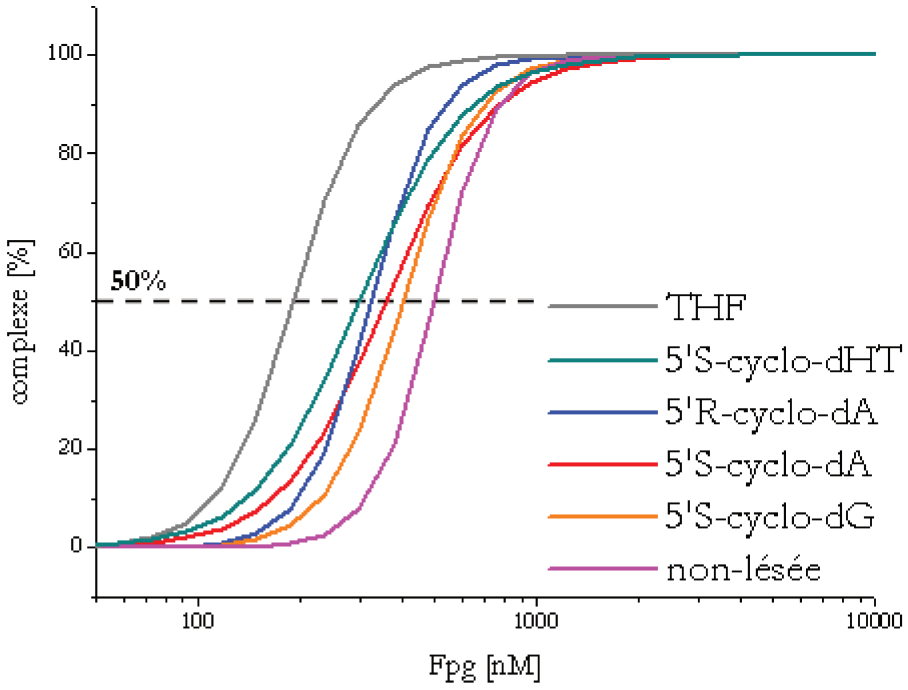


Figure 69 : Formation des complexes entre les sondes lésées ou non, et la protéine Fpg. Les courbes représentant chacune une sonde ont été tracées à l'aide du logiciel Origin, et montrent la tendance suivie par les points expérimentaux.

Tableau 15 : Valeurs du coefficient de corrélation, de la constante de dissociation et du coefficient de coopérativité obtenues avec le modèle de Hill, pour les interactions entre la Fpg et les sondes contenant un cyclonucléoside, la sonde THF et la sonde non-lésée.

sonde	$\mathbf{R^2}$ (coefficient de corrélation)	$\mathbf{K_d}$ (constante de dissociation) [nM]	**n** (coefficient de Hill)
non-lésée	0.99638	494.01 ± 4.77	4.89 ± 0.28
5'S-cyclo-dA	0.97897	357.47 ± 11.17	2.86 ± 0.30
5'R-cyclo-dA	0.97854	323.68 ± 9.21	4.50 ± 0.54
5'S-cyclo-dG	0.97594	400.27 ± 10.83	4.01 ± 0.49
5'S-cyclo-dHT	0.97749	299.18 ± 10.80	2.84 ± 0.31
THF	0.98344	190.32 ± 6.93	4.01 ± 0.50

Il existe également des différences entre les quatre sondes contenant des cyclonucléosides. Parmi elles, la Fpg semble avoir la plus grande affinité pour la 5'S-cyclo-dHT, qui est la seule cyclopyrimidine utilisée au cours de nos études. Cependant, pour pouvoir conclure avec certitude que la Fpg possède la plus grande affinité pour les cyclopyrimidines par rapport aux cyclopurines il serait nécessaire de refaire le même type d'expérience avec des sondes contenant d'autres cyclopyrimidines. Parmi les cyclopurines, la Fpg possède une plus faible affinité pour la 5'S-cyclo-dG, et des deux cyclo-dA, c'est la forme 5'R qui est le mieux reconnue.

Nous avons démontré que la protéine bactérienne Fpg avait une affinité particulière pour les cyclonucléosides, et que cette affinité était plus grande que pour l'ADN non-lésé, mais plus faible que pour le THF. Il semble également que la Fpg distingue les quatre lésions testées, dans l'ordre d'affinité suivant : 5'S-cyclo-dHT > 5'R-cyclo-dA > 5'S-cyclo-dA > 5'S-cyclo-dG.

2.1.3. Influence de la longueur de la sonde sur les interactions non-spécifiques avec la Fpg

Comme nous l'avons indiqué dans le Tableau 15 du paragraphe précédent, dans nos expériences d'interactions entre la Fpg et les cyclonucléosides nous avons obtenu la valeur de coéfficient de Hill « n » supérieure à 1 pour les 6 oligonucléotides testés, qu'ils soient lésés ou non. Ceci ne correspond pas aux résultats observés sur les gels, où un seul complexe est obtenu à chaque fois.

L'objectif des expériences présentées dans ce paragraphe consiste ainsi à élucider le mode de fixation de la Fpg sur l'ADN non-lésé. Nous avons vérifié si plusieurs molécules de la protéine Fpg étaient capables de se fixer de manière non-spécifique sur les sondes et si le nombre des complexes était dépendant de la longueur de l'oligonucléotide. Afin d'éviter les biais liés à la reconnaissance spécifique des lésions, nous avons utilisé uniquement les sondes non-lésées, en gardant dans la mesure du possible la même séquence, pour laquelle trois tailles ont été testées :

- 22 bases – la sonde non-lésée utilisée dans les expériences de gel shift,
- 44 bases – la sonde deux fois plus longue avec une répétition de la séquence 22 bases,
- 14 bases – 3 bases de chaque côté de la sonde 22 bases ont été retirées en préservant le centre de la séquence (Figure 70 A).

Puisque dans nos conditions expérimentales nous n'avions pas observé de complexes multiples, nous avons élargi la gamme de concentration de la protéine, pour favorises la formation de ceux-ci : le dernier point de la gamme est alors à 1375 nM contrairement à 550 nM pour les expériences précédentes, pour la même concentration de la sonde. L'élargissement de la gamme se justifie également par la plus faible affinité de la Fpg pour les oligonucléotides non-lésés, ce qui nécessite l'application d'une quantité plus importante de la protéine afin d'observer la complexation totale de la sonde libre (Figure 70 B). Ces expériences ont été répétées deux fois, mais la quantification d'une seule série de gels est présentée ici.

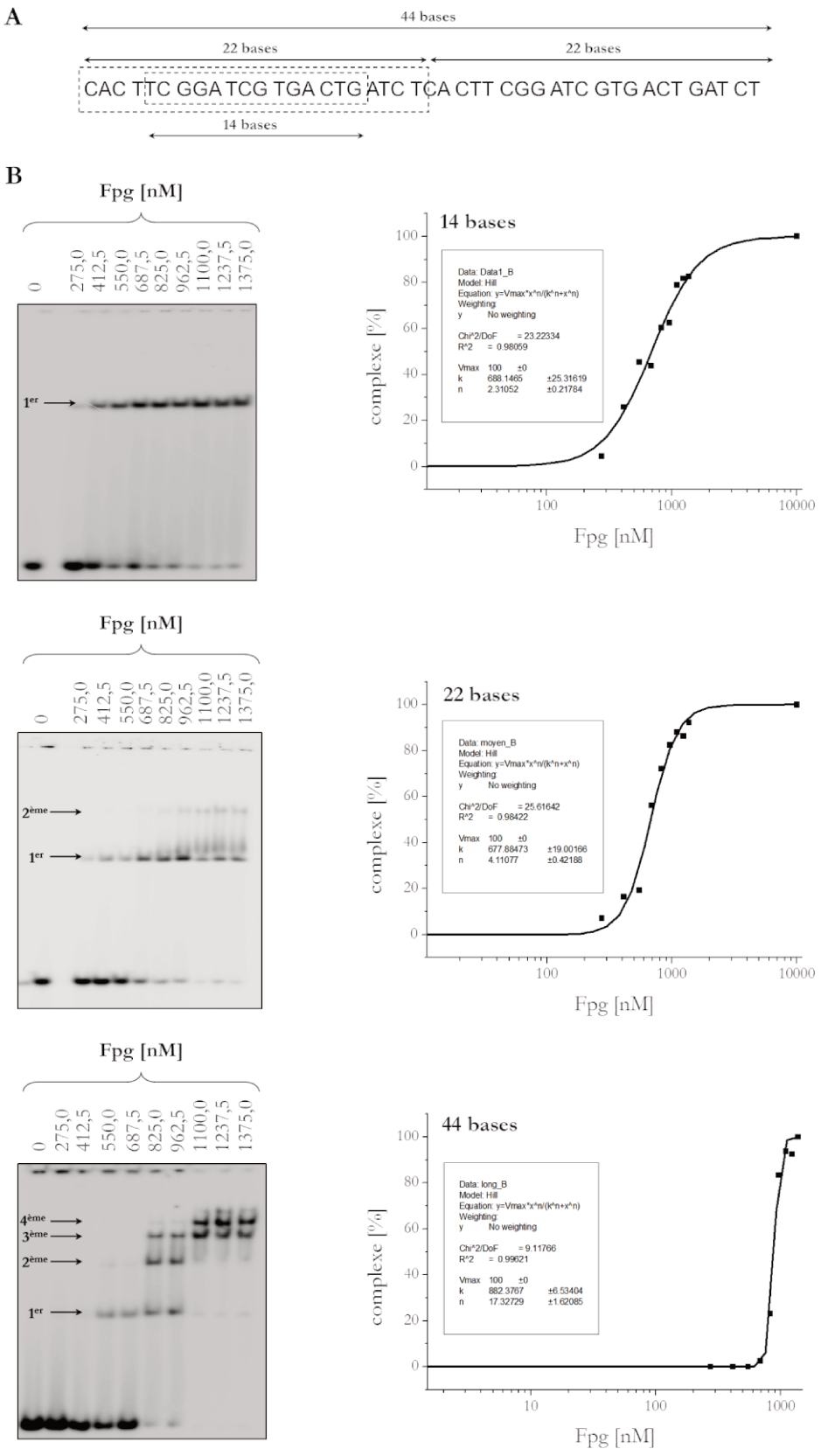


Figure 70 : Formation des complexes spécifiques et non-spécifiques entre la protéine Fpg et les sondes non-lésées de différentes longueurs ; A : séquences et longueurs des trois sondes ; B : interactions entre les trois sondes et la protéine Fpg ; gauche : image d'un gel shift, droite : courbe de formation des complexes.

Nous pouvons observer sur les gels que les complexes sont formés sur les trois sondes utilisées, et que le nombre de ceux-ci dépend effectivement de la longueur de l'oligonucléotide. Sur l'oligonucléotide le plus court (14 bases), un seul complexe est visible même à très haute concentration de la protéine. Sur les sondes de taille intermédiaire (22 bases), on voit dans un premier temps la formation d'un premier complexe, et ensuite, quand il n'y a plus de sondes libres disponibles, la génération progressive d'un deuxième complexe. Enfin, pour la sonde la plus longue (44 bases), jusqu'à quatre complexes sont observés : le deuxième complexe apparaît pour des concentrations plus basses de Fpg par rapport aux complexes obtenus avec la sonde 22 bases, et les complexes suivants sont formés ensuite très rapidement, en remplaçant les plus petits complexes qui disparaissent progressivement.

Nous avons quantifié les gels en procédant de la même manière que pour les expériences avec les sondes lésées. Dans le cas où plusieurs complexes sont formés, nous les avons quantifiés ensemble, pour disposer uniquement de deux valeurs : pourcentage de la sonde libre et pourcentage de la forme complexée, quels que soient la taille et le nombre des complexes dans lesquels elle est impliquée. Nous avons ensuite tracé les courbes d'interaction en utilisant toujours l'équation de Hill (Figure 70 B). Le coefficient de corrélation est élevé – autour de 0.98-0.99 (voir : Tableau 16) – et les courbes suivent parfaitement la tendance des points empiriques.

Nous pouvons dans un premier temps constater que la constante de dissociation ne varie pas entre les deux sondes courtes : 14 bases et 22 bases. Elle est cependant plus élevée pour la sonde de 44 bases, ce qui indique que l'affinité de la Fpg pour l'oligonucléotide long serait inférieure à celle déterminée pour les oligonucléotides plus courts.

Un deuxième paramètre qui peut être évalué est l'inclinaison de la courbe. Elle fournit des informations sur la formation des complexes entre la protéine et sa cible. En effet, elle exprime la vitesse de la formation des complexes et plus elle est pentue, plus la gamme de concentrations pour laquelle les complexes sont générés, est serrée. L'existence d'interactions allostériques entre plusieurs complexes formés avec la même sonde peut également avoir une influence sur l'inclinaison de la courbe des interactions.

Comme nous pouvons le constater sur la Figure 70 B, la courbe devient plus pentue avec l'augmentation de la longueur de la sonde. Ainsi, pour la sonde de 44 bases on observe la formation du premier complexe plus tardivement que pour les sondes 14 bases et 22 bases, mais en même temps on constate la disparition totale de la sonde libre plus tôt que dans les deux autres cas. La gamme de concentrations de la Fpg dans laquelle les interactions entre la sonde et la protéine sont observées, est alors plus réduite que pour les sondes plus courtes. La même

comparaison peut également être effectuée entre les sondes 14 bases et 22 bases, et c'est cette deuxième qui commence à former les complexes plus tardivement, mais en atteignant leur saturation plus tôt.

Finalement, nous pouvons regarder aussi la variation du coefficient de Hill « n » entre les différentes sondes testées. Ce coefficient indique le nombre de molécules de ligand capables de se fixer sur la sonde. Les valeurs de « n » calculées pour nos trois courbes sont présentées dans le Tableau 16. On peut constater que pour toutes les sondes, le coefficient de Hill est supérieur à 1, ce qui indiquerait que plusieurs molécules de la protéine Fpg interagissent avec l'oligonucléotide. Ainsi, pour la sonde 14 bases il est égal à ≈ 2, pour 22 bases ≈4 et pour 44 bases ≈17. Cependant, ces valeurs ne correspondent pas au nombre des complexes qui sont visibles sur les gels : 1, 2 et 4 pour les sondes 14, 22 et 44 bases. Nous ne pouvons pas expliquer ces disparités entre les résultats empiriques et les calculs : il existe une relation entre le coefficient de Hill (n) et le nombre de complexes (x), qui peut être présenté d'une manière approximative par l'équation $n = 2^x$.

Tableau 16 : Valeurs du coefficient de corrélation, de la constante de dissociation et du coefficient de coopérativité obtenues avec le modèle de Hill, pour les interactions entre la Fpg et les sondes non-lésées de différentes longueurs.

sonde	$\mathbf{R^2}$ (coefficient de corrélation)	$\mathbf{K_d}$ (constante de dissociation) [nM]	**n** (coefficient de Hill)	nombre de complexes sur le gel
14 bases	0.98059	688.15 ± 25.32	2.31 ± 0.22	1
22 bases	0.98422	677.88 ± 19.00	4.11 ± 0.42	2
44 bases	0.99621	882.38 ± 6.53	17.33 ± 1.62	4

Nous pouvons constater que le nombre des complexes formés entre la Fpg et une séquence non-spécifique de l'ADN dépend de la longueur de cette dernière, et on peut estimer que la fixation d'une molécule de la protéine est possible toutes les 10 bases environ. Il semble également possible que des mécanismes de coopérativité soient impliqués dans la formation des complexes, et que celle-ci soit facilitée par la fixation de la première molécule de Fpg sur la sonde.

2.2. Etude biochimique des protéines isolées chez *Thermococcus gammatolerans*

Après analyse de la protéine bactérienne Fpg, les mêmes expériences ont été effectuées sur deux enzymes isolées de *Thermococcus gammatolerans*. Néanmoins, avant d'étudier la réparation des lésions simples et complexes de l'ADN, nous avons voulu savoir si elles existaient dans ces organismes, et à quel niveau. C'est pour cela que le dosage des dommages d'oxydation dans leur ADN a été effectué.

2.2.1. Etude d'un modèle original : *Thermococcus gammatolerans*

Le troisième modèle étudié au cours de cette thèse, après le modèle eucaryote représenté par les cellules HeLa et celui procaryote représenté par *E.coli*, est un organisme original : *Thermococcus gammatolerans*. Sachant que les cyclonucléosides sont des lésions formées à fortes doses de radiations ionisantes et en absence d'oxygène, *Thermococcus gammatolerans* (anaérobie radiorésistante) est un organisme chez lequel on pouvait espérer retrouver une quantité plus importante de ces dommages que chez les eucaryotes, ce qui pourrait fournir des nouvelles informations sur la prise en charge de ces lésions. La 8-oxo-dG et les cyclonucléosides ont été mesurés dans l'ADN de *T.gammatolerans* au niveau basal et après une irradiation.

2.2.1.1. Mise au point du protocole d'extraction de l'ADN de *Thermococcus gammatolerans*

Afin de pouvoir procéder au dosage des lésions d'oxydation dans l'ADN, nous avons dû mettre au point un protocole d'extraction de l'ADN des cellules de *Thermococcus gammatolerans*. Le protocole couramment employé au sein de notre laboratoire n'a pas pu être appliqué ici, car il est adapté aux cellules eucaryotes possédant un noyau, ce qui n'est pas le cas des archées. Nous avons ainsi conçu un nouveau protocole, en nous basant sur celui-ci et également sur les protocoles utilisés pour d'autres organismes extremophiles [322,323]. Parmi les paramètres qui ont dû être mis au point, on trouve notamment :

- la quantité de culot cellulaire
- la composition du tampon de lyse
- le temps et la température de la lyse

- le type et la quantité des agents de lyse : SDS et sarcosyl
- la quantité et le type de protéinase (protéinase K ou protéase Qiagen), ainsi que les conditions de son application (temps et température)
- la quantité de RNAse et les conditions de son application : dans le tampon de lyse / changement de tampon nécessitant une précipitation de l'ADN
- volume dans lequel l'ADN est précipité.

L'optimisation de ces paramètres a été effectuée sur les organismes qui n'ont subi aucun stress. Chaque modification d'un ou de plusieurs paramètres dans le protocole a été suivie d'un dosage des nucléosides normaux, comme décrit dans le paragraphe 4.2.3 du chapitre « Matériel et méthodes », avec un détecteur UV. Ainsi, la qualité de l'ADN extrait, surtout sa pureté, et sa quantité ont pu être estimées, avant que le dosage des lésions soit, éventuellement, effectué.

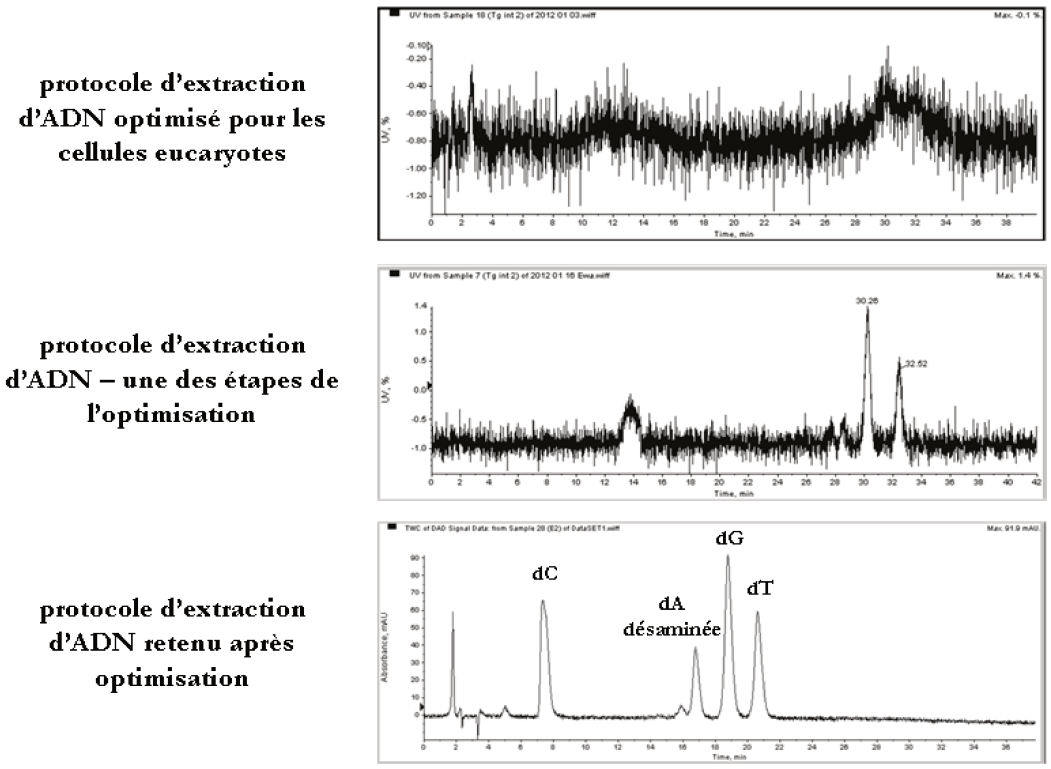


Figure 71 : Optimisation du protocole d'extraction de l'ADN de *Thermococcus gammatolerans* ; les trois chromatogrammes UV montrent l'amélioration de la qualité de l'ADN extrait.

Sur la Figure 71 nous pouvons voir les étapes consécutives d'amélioration du protocole d'extraction de l'ADN. Au début (chromatogramme du haut), aucun pic des nucléosides normaux ne sortait du bruit de fond, ce qui indiquait l'échec de l'extraction. Sur le chromatogramme du milieu, deux pics représentant la thymine et la guanine sont visibles, le niveau du bruit de fond restant toutefois important, témoignant d'une mauvaise qualité de

l'ADN. Enfin, comme on peut le constater en observant le chromatogramme du bas, nous avons réussi à obtenir de l'ADN propre et de bonne qualité, qui permet d'effectuer le dosage des nucléosides modifiés.

2.2.1.2. Dosage de la 8-oxo-désoxyguanosine

La 8-oxo-dG est la lésion d'oxydation de référence, c'est pourquoi nous avons procédé à son dosage chez *Thermococcus gammatolerans*. Son taux a été mesuré dans les cellules qui n'ont subi aucun stress, ainsi que dans les cellules irradiées à forte dose de rayonnement ionisant :

- 2500 Gy, cellules récoltées immédiatement après le traitement (2500 Gy, 0 hr) ou remises 1 heure en culture (2500 Gy, 1 hr)
- 5000 Gy, cellules récoltées immédiatement après le traitement (5000 Gy, 0 hr) ou remises 1 heure en culture (5000 Gy, 1 hr)
- cellules non-irradiées, mais gardées sur la glace le temps de l'irradiation, récoltées immédiatement (NI 2500 Gy, 0 hr ou NI 5000 Gy, 0 hr) ou remises 1 heure en culture (NI 2500 Gy, 1 hr ou NI 5000 Gy, 1 hr)

Un exemple de chromatogrammes obtenus lors du dosage de la 8-oxo-dG est présenté sur la Figure 72.

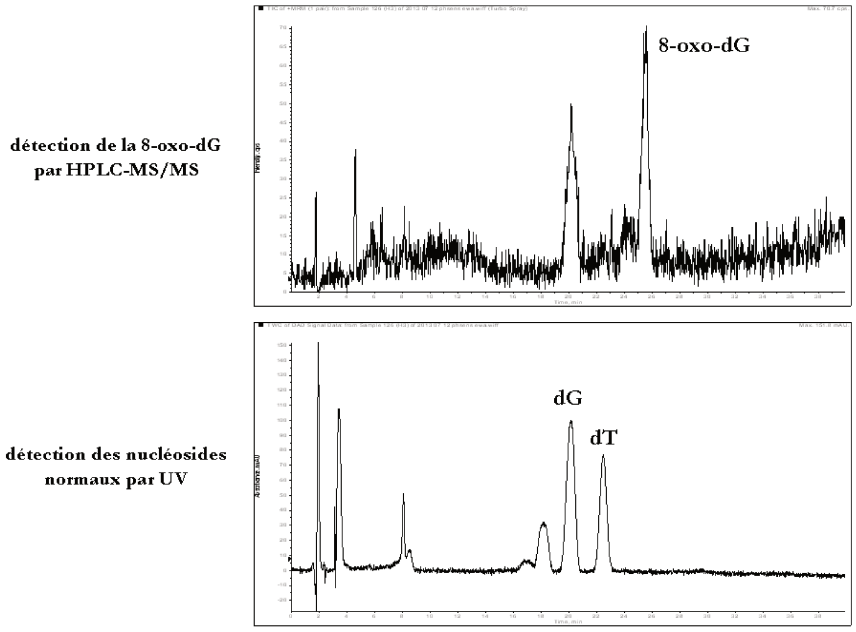


Figure 72 : Dosage de la 8-oxo-dG. Exemple des chromatogrammes obtenus pour l'échantillon irradié à 2.5 kGy et récolté immédiatement après le traitement.

Cette expérience a été réalisée deux fois. Le premier dosage a été effectué avec des duplicatas et le deuxième – avec des triplicatas techniques : deux ou trois extractions et digestions indépendantes ont été réalisées à partir de chaque culot cellulaire. Une moyenne des cinq mesures a été calculée pour chaque condition (Figure 73). Le taux de la 8-oxo-dG dans les cellules non-irradiées et non-incubées sur la glace (« témoin ») a également été calculé et il est estimé à 10 lésions environ par million de nucléosides normaux.

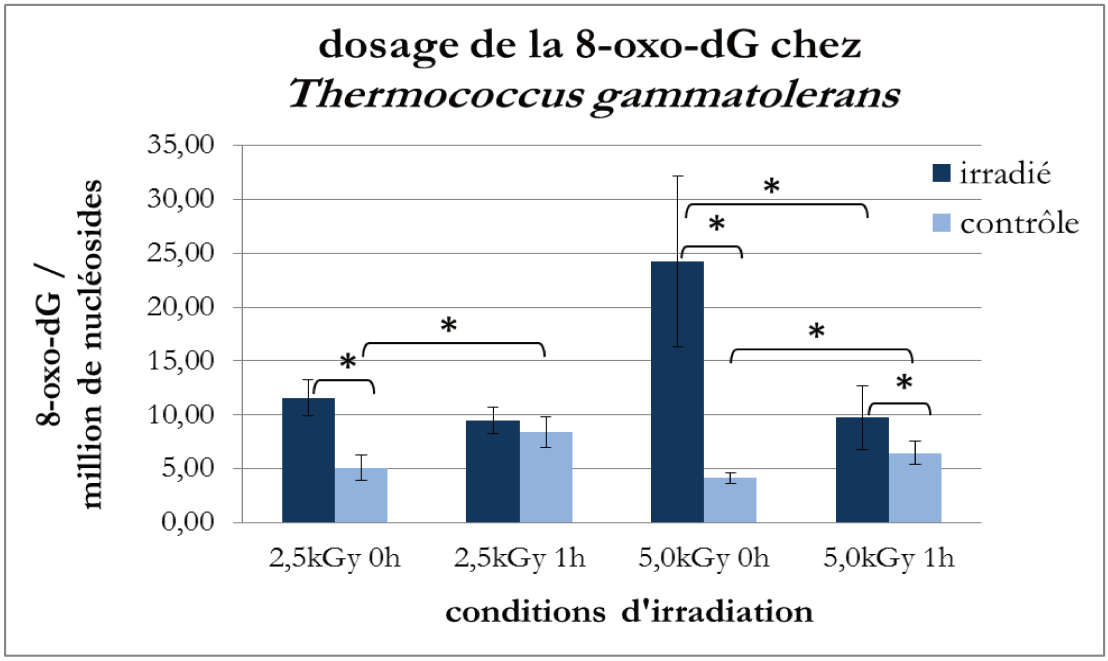


Figure 73 : Dosage de la 8-oxo-dG dans l'ADN de *Thermococcus gammatolerans* irradiés à 2,5 kGy ou 5,0 kGy, récoltés immédiatement après le traitement (0h) ou remis en culture pendant 1 heure (1h). Les valeurs sont présentées en nombre de lésions pour 1 million de bases normales. Les tests statistiques Mann-Whitney ont été réalisés afin de comparer le niveau de lésions entre différentes conditions. (*) = une différence entre deux valeurs a été constatée avec $p < 0.05$.

Dans un premier temps, nous pouvons constater une augmentation du taux de 8-oxo-dG dans les cellules irradiées par rapport aux contrôles incubés sur la glace. Cette différence est particulièrement importante pour la dose de 5 kGy de rayonnement γ, car le niveau des dommages est environ 6 fois plus élevé dans les cellules irradiées que dans les cellules contrôles. Le taux de 8-oxo-dG dosé dans les cellules contrôles est inférieur à celui trouvé dans les témoins non-incubés sur la glace.

La remise des cellules en incubation pendant 1 heure suivant le traitement n'a pas le même effet sur les cellules irradiées et non-irradiées. Dans l'ADN des cellules irradiées, une diminution du niveau de la 8-oxo-dG est constatée après 1 heure d'incubation. Quelle que soit la dose du rayonnement, le taux de lésion descend au niveau retrouvé dans les cellules témoins. Au

contraire, dans les cellules incubées sur la glace mais non-irradiées, la remise en incubation provoque l'augmentation du niveau de 8-oxo-dG, qui rejoint également le niveau des cellules témoins. Les différences observées entre les organismes irradiés et non-irradiés immédiatement après l'irradiation diminuent si les cellules sont remises en culture pendant 1 heure.

2.2.1.3. Dosage des cyclonucléosides

Le dosage des divers cyclo-dA a été effectué dans les cellules qui n'ont pas subi d'irradiation, ainsi que dans les cellules irradiées à 5 kGy en anaérobie et récoltées immédiatement après ce traitement, afin de voir l'induction de ces dommages dans les conditions favorables à leur formation.

Dans les deux conditions, aucune détection des cyclo-dA n'a pu être réalisée, les niveaux de ces lésions sont alors estimés comme étant inférieurs à 1.3 lésions pour 10^6 nucléosides normaux, ce qui est la limite de leur détection (Figure 74).

Le dosage des lésions d'oxydation effectué dans l'ADN de *Thermococcus gammatolerans* au niveau basal ainsi qu'après irradiation à forte dose, nous a permis de mieux connaître cet organisme original. Le taux basal de la 8-oxo-dG est environ dix fois plus élevé chez ces Archées que chez la majorité des eucaryotes. Sa quantité augmente suite à l'exposition aux rayonnements γ, la 8-oxo-dG est néanmoins rapidement réparée après la remise des cellules en culture. Au contraire, les cyclonucléosides ne sont pas détectés chez *Thermococcus gammatolerans*, ni au niveau basal, ni après exposition à 5 kGy de rayons γ.

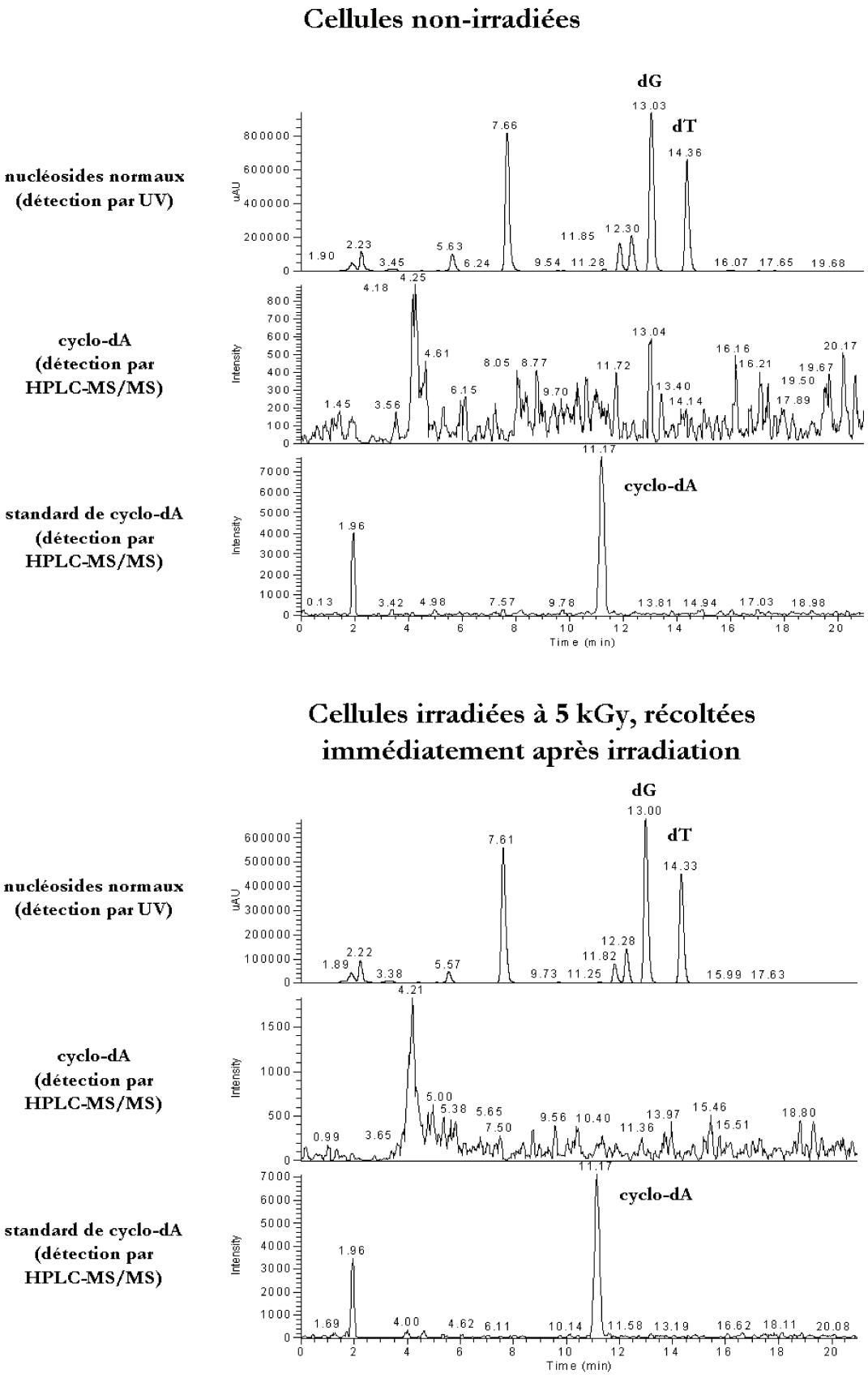


Figure 74 : Dosage des cyclo-dA dans l'ADN de *Thermococcus gammatolerans* non-irradié ou irradié à 5 kGy et récolté immédiatement après le traitement.

2.2.2. Activité de coupure

Comme décrit précédemment pour la Fpg, les tests d'activité de coupure vont être réalisés avec les deux protéines provenant de *Thermococcus gammatolerans*, leur objectif étant de trouver leurs substrats, mais également d'identifier leur mécanisme d'action.

2.2.2.1. Test d'activité sur gel

L'activité des deux enzymes isolées de l'Archée *Thermococcus gammatolerans* : tg1653 et tg1277, qui sont les homologues probables de la 8-oxo-dG ADN glycosylase et de l'endonucléase III, a été testée vis-à-vis des lésions simples. Elles ont été incubées avec des oligonucléotides double-brin contenant sur un des brins l'une des lésions suivantes : 8-oxo-dG, 8-Me-dG, 8-oMe-dG, 8-Br-dG, 5-OH-dHyd ou 5-OH-5-Me-dHyd. La glycosylase bactérienne Fpg a été utilisée en parallèle à titre comparatif (Figure 75).

L'activité de coupure de la tg1653 a été mise en évidence au niveau de la 8-oxo-dG, mais aucune coupure n'a été détectée sur les oligonucléotides contenant d'autres lésions d'oxydation ou analogues. Une faible activité d'incision de la 8-oMe-dG est également observée, mais la pertinence de ce résultat est discutable, comme cela a déjà été présenté dans le paragraphe 2.1.1.1. La comparaison des séquences des protéines avait permis de prévoir que la tg1653 pourrait être une enzyme homologue de la Fpg bactérienne. Néanmoins, on peut constater que le spectre d'activité de ces deux enzymes n'est pas identique. Tg1653 a un spectre d'activité plus restreint que la Fpg, puisqu'elle n'est pas capable d'exciser les hydantoïnes modifiées, telles que la 5-OH-dHyd ou la 5-OH-5-Me-dHyd. A l'heure actuelle, la 8-oxo-dG est le seul substrat identifié de tg1653.

L'enzyme tg1277, quant à elle, est capable d'exercer une activité à la fois sur la 8-oxo-dG, la 5-OH-dHyd et la 5-OH-5-Me-dHyd. Son spectre d'activité est alors plus large que celui de la tg1653 et, si l'on prend on compte les substrats testés, il se superpose à celui de la Fpg. Initialement, la tg1277 a été isolée et identifiée en tant que l'homologue probable de l'Endo III. Cependant, l'Endo III, dont l'activité a été examinée auparavant (voir le paragraphe 2.1.1.1 de ce chapitre, Figure 57), n'exerce pas d'activité d'incision au niveau de la 8-oxo-dG. Ainsi, le spectre d'activité de la tg1277 semble plus proche de celui de la Fpg, car parmi les lésions testées ici, ces deux enzymes sont actives vis-à-vis des mêmes modifications de l'ADN.

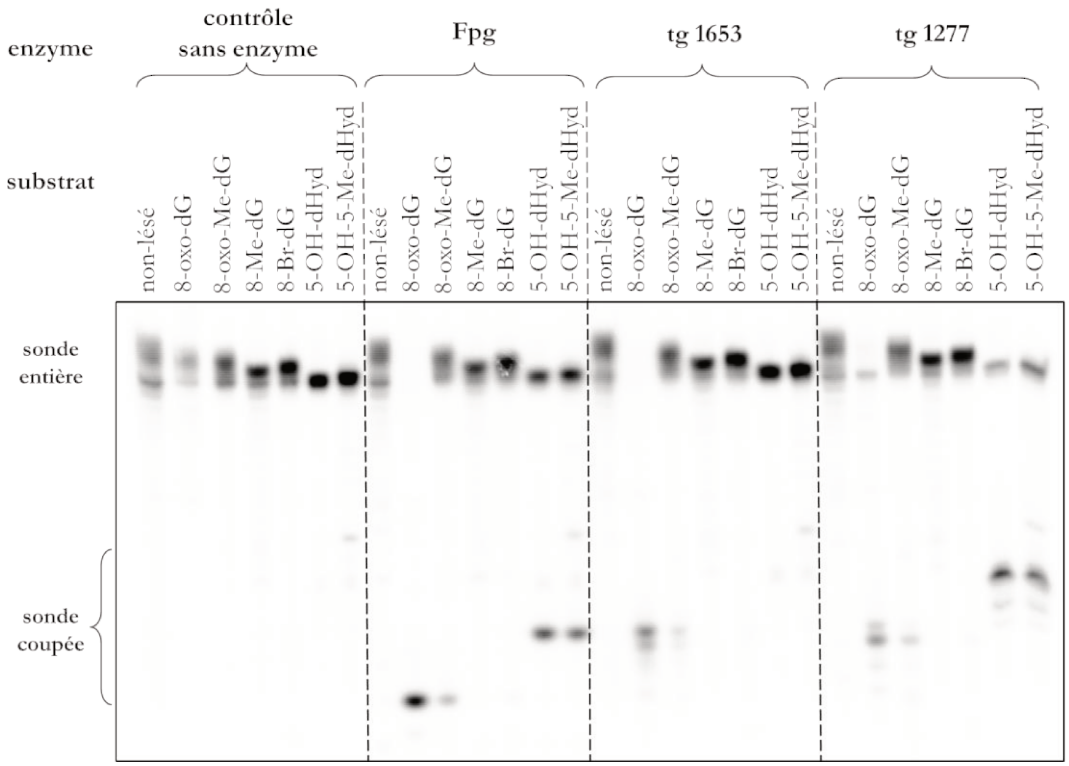


Figure 75 : Activité de coupure des enzymes isolées de *Thermococcus gammatolerans* sur les oligonucléotides contenant une lésion simple d'oxydation ou analogue.

Une autre particularité a été remarquée concernant l'activité des deux enzymes de *T.gammatolerans*. La taille des fragments de la sonde obtenus après la coupure n'est pas la même pour les trois protéines : Fpg, tg1653 et tg1277. Pour la même enzyme, la différence observée entre les fragments générés suite à la coupure au niveau de la 8-oxo-dG ou d'une hydantoïne modifiée, s'explique par la localisation différente de ces lésions dans la séquence – la coupure n'est alors pas effectuée au même endroit de l'oligonucléotide et les fragments générés sont plus ou moins longs. Cependant, la différence constatée pour la même sonde coupée par deux enzymes différentes doit avoir une toute autre origine. L'hypothèse émise à ce stade est que cette coupure n'est pas effectuée via le même mécanisme. Les deux mécanismes principaux par lesquels la coupure peut être effectuée sont la β-élimination et la β,δ-élimination, dont l'exploration va être détaillée dans le paragraphe suivant.

Une seule lésion substrat pour tg1653 a pu être identifiée, à savoir la 8-oxo-dG. Le spectre d'activité de tg1277 est plus large, car il comprend également les dérivés de l'hydantoïne : la 5-OH-dHyd et la 5-OH-5-Me-dHyd. De plus, ces deux enzymes n'exercent pas leur activité de coupure par le même mécanisme que la Fpg.

2.2.2.2. Vérification du mécanisme de coupure par MALDI-TOF/MS

Après excision de la base endommagée, qui est la première étape de la réparation d'une lésion simple de l'ADN, un site abasique est formé, qui contient dans sa structure un sucre ouvert. La coupure effectuée ensuite au niveau de ce site abasique va générer une cassure au niveau du brin endommagé. Cette coupure peut être exercée par un mécanisme de β-élimination ou de β,δ-élimination (Figure 76).

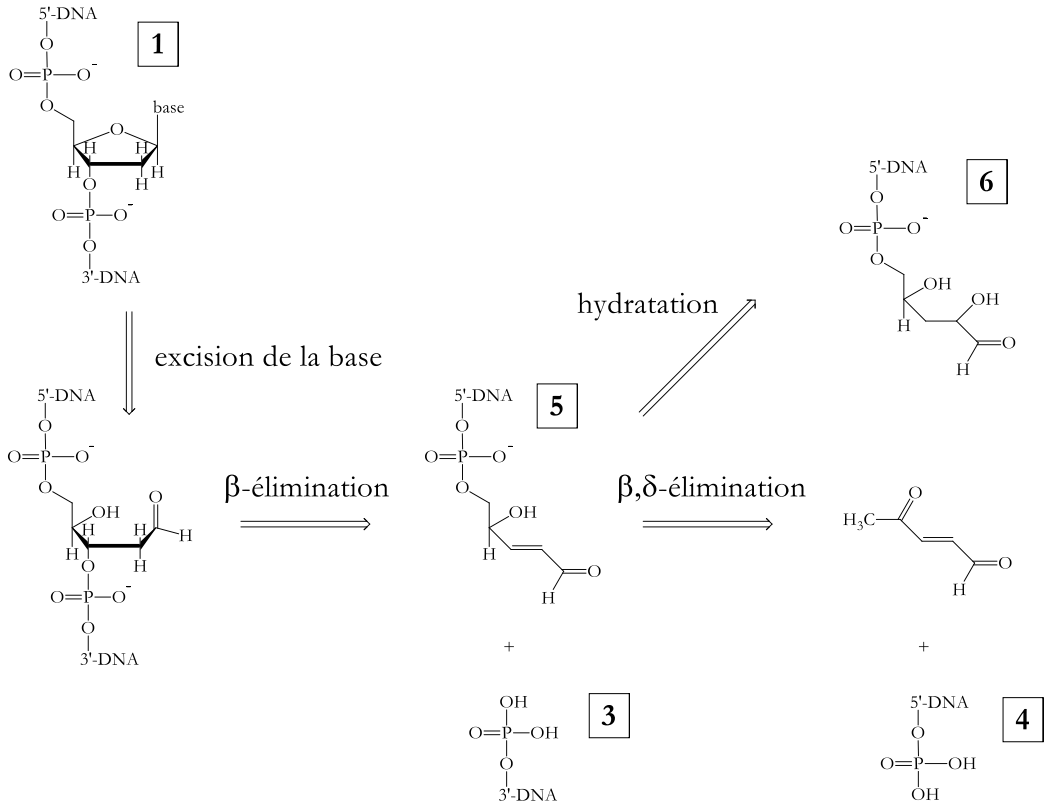


Figure 76 : Schéma du mécanisme de coupure d'un site abasique par β- et δ-élimination. Les différents composés formés sont identifiés avec des numéros qui correspondent aux fragments déterminés ultérieurement lors de l'analyse par MALDI-TOF/MS (voir : Figures 77-81). Le composé numéro 2 est absent car il représente le brin complémentaire.

La première étape de la coupure consiste à inciser la liaison du côté 3' du site abasique, au niveau du carbone C3' du désoxyribose et du groupement phosphate. L'action enzymatique à ce niveau va générer deux fragments de l'ADN : un du côté 3' de la lésion, contenant un groupement phosphate (**3**), et l'autre du côté 5' de la lésion, qui arbore le résidu de sucre ouvert (**5**). Cette étape, appelée la β-élimination, peut terminer la réaction de coupure, comme c'est le cas pour l'Endo III, ou bien elle peut être suivie par d'autres évènements. L'hydratation au niveau de

la double liaison du sucre peut advenir (**6**), cependant le mécanisme le plus courant est celui de la δ-élimination, comme c'est par exemple le cas pour la Fpg. La coupure est effectuée alors du côté 5' du sucre généré par la β-élimination, et le fragment arborant un groupement phosphate est ainsi séparé du résidu de sucre, qui est éliminé (**4**) (Figure 76) [138].

2.2.2.2.1. Coupure de la 8-oxo-dG

Afin de vérifier si pour les trois enzymes étudiées (Fpg, tg1653 et tg1277), il existe effectivement différents mécanismes de coupure, un autre type de test a été effectué : la taille des fragments obtenus après la réaction enzymatique a été déterminée avec une grande précision, en utilisant la technique MALDI-TOF/MS. Contrairement à l'analyse sur gel, qui permet de visualiser uniquement le fragment de l'oligonucléotide du côté de la lésion qui était au préalable marqué radioactivement, l'analyse par spectrométrie MALDI-TOF/MS permet d'identifier tous les fragments présents dans le mélange et de déterminer précisément leur masse.

Dans un premier temps, l'oligonucléotide contenant une 8-oxo-dG a été incubé comme décrit dans le « Matériel et méthodes », avec les enzymes Fpg, tg1653 et tg1277. Un contrôle non-digéré a également été effectué pour vérifier l'intégrité de la sonde. Tous les pics majeurs ont été attribués aux fragments de séquence correspondants, comme visualisé sur les spectres (Figure 77). Chaque pic a été identifié grâce à sa masse qui a été comparée à la masse théorique calculée pour les divers produits attendus – à chaque pic obtenu sur les spectres nous avons attribué un numéro qui correspond à la structure attendue après la coupure (Figure 76 à 78).

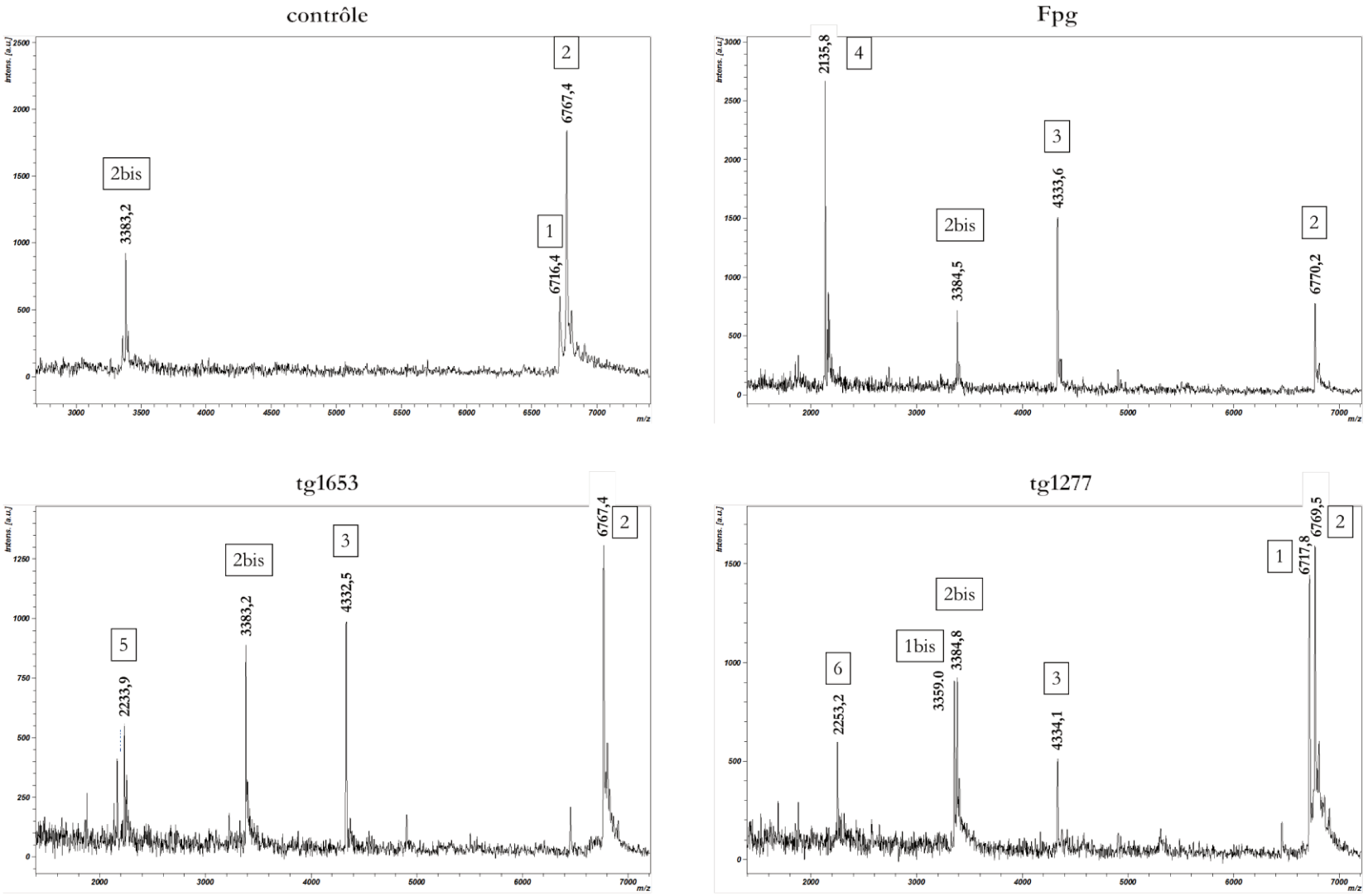


Figure 77 : Spectres MALDI-TOF/MS de l'oligonucléotide double-brin, contenant sur un brin une 8-oxo-dG, incubé avec la Fpg ou avec une des deux enzymes isolée de *T.gammatolerans* : tg1653 ou tg1277. Contrôle = oligonucléotide duplexe non-digéré.

Ainsi, nous pouvons constater tout d'abord que l'oligonucléotide duplexe contenant la 8-oxo-dG n'a pas subi de dégradation spontanée, puisque seuls les pics correspondant aux deux brins complémentaires ont été détectés sur son spectre (Figure 77). Le brin comportant une 8-oxo-dG a été détecté à $[M-H]^-$ 6716.4 pour une masse théorique de 6717.33 Da (pic 1), et son brin complémentaire à $[M-H]^-$ 6767.4 pour une masse théorique de 6768.40 Da (pic 2). Un troisième pic, appelé pic 2bis, a également été détecté, qui correspond au brin complémentaire doublement chargé – son rapport masse/charge est alors divisé par deux par rapport à son équivalent monochargé et on le retrouve à $[M-H]^-/2$ 3383.2.

Les spectres obtenus après l'analyse des produits de coupure enzymatique permettent de comparer la taille des fragments générés par différentes protéines et ainsi vérifier l'hypothèse que nous avons proposée suite aux résultats des tests d'activité sur gel. Bien que les spectres MALDI-TOF/MS ne donnent pas la structure du fragment de l'ADN analysé, mais uniquement sa masse, la comparaison de celle-ci avec des prévisions théoriques et également avec des produits connus, permet de les identifier de manière fiable.

Le premier test a été effectué sur la protéine Fpg – son mécanisme d'action est connu et elle nous sert alors de témoin auquel vont être comparées les deux autres enzymes. Le spectre montre quatre pics bien distincts (Figure 77). Nous constatons l'absence du produit de départ, ce qui indique la coupure complète de l'oligonucléotide lésé – la même observation a été faite lors de l'analyse de la coupure sur gel, où la bande de l'oligonucléotide entier n'était plus détectable après l'incubation avec la Fpg. Nous pouvons toujours distinguer les deux pics : 2 et 2 bis, correspondant à l'oligonucléotide complémentaire non-lésé, sous forme d'ion mono- et dichargé. Le brin complémentaire n'a pas été coupé ni digéré de manière aléatoire.

Deux autres pics ont été identifiés sur ce spectre, à $[M-H]^-$ 4333.6 et 2135.8. Nous avons comparé ces valeurs aux masses molaires calculées pour les fragments prévus et qui correspondent aux côté 3' et côté 5' de la lésion incisée avec le mécanisme de β,δ-élimination. Nous leur avons assigné les numéros 3 et 4 qui correspondent à ces produits de coupure, comme présenté sur le schéma général de la Figure 76 et, plus en détail, sur la Figure 78. Ainsi, le pic 3 représente le fragment monochargé du côté 3' de la lésion, portant un phosphate à son extrémité 5', et dont la masse moléculaire théorique est de 4333.76 Da. Le pic 4 représente le fragment monochargé situé de l'autre côté de la lésion, dont la masse moléculaire est 2136.36 Da. Il porte un phosphate à son extrémité 3' et il résulte de la deuxième étape de coupure, c'est-à-dire la δ-élimination.

Ces résultats nous montrent que la technique MALDI-TOF/MS est suffisamment précise pour permettre de déterminer de manière univoque la taille des fragments d'oligonucléotides et, par déduction, le mécanisme enzymatique par lequel ils ont été générés, comme cela a été fait ici pour la Fpg.

Sur le spectre représentant les produits de coupure de la 8-oxo-guanine par tg1653 nous avons identifié 4 pics. De même manière que pour Fpg, nous ne voyons plus le pic de l'oligonucléotide de départ, mais nous retrouvons les deux pics de l'oligonucléotide complémentaire : mono- et dichargé. Le produit 3 formé suite à la β-élimination, comme cela a été observé pour Fpg, est également retrouvé à $[M-H]^-$ 4332.5. Ce fragment est généré indépendamment de la suite de la réaction de coupure et il n'est donc pas surprenant de l'identifier sur tous les spectres, que le mécanisme de coupure soit β- ou β-δ-élimination. La différence entre ce spectre et celui de la Fpg consiste en la taille du quatrième pic identifié. Dans le cas de la Fpg, il était à $[M-H]^-$ 2135.8, tandis que pour tg1653 on le retrouve à 2233.9, ce qui signifie qu'il ne s'agit pas du même fragment – nous lui avons alors attribué le numéro 5. En effet, la comparaison de son $[M-H]^-$ avec les calculs théoriques indique qu'il correspondrait au segment libéré du côté 5' de la coupure dans l'étape de la β-élimination, et dont la masse est attendue à 2234.46 Da. Ainsi, le mécanisme de coupure exercée par tg1653 serait probablement la β-élimination. Cette protéine, que l'on supposait être un homologue de la Fpg, n'exerce alors pas son activité via le même mécanisme, en plus de ne pas avoir le même spectre d'activité.

Le spectre obtenu pour tg1277 est plus complexe que les spectres analysés précédemment, puisqu'on peut y constater la présence de plus de pics. Nous détectons d'abord la présence du pic 1 à $[M-H]^-$ 6717.8 – cela montre que, contrairement à la Fpg et tg1653, cette enzyme n'induit pas 100% de coupure dans les mêmes conditions d'incubation. Ceci est en accord avec les résultats du test de coupure réalisé sur gel (Figure 75), où une bande correspondant à l'oligonucléotide non-coupé était toujours présente pour la tg1277. De plus, un produit dichargé provenant du même oligonucléotide est présent (1 bis), à $[M-H]^-/2$ 3359.0 Les deux pics correspondants à l'oligonucléotide complémentaire, ainsi que le pic 3 représentant le produit généré dans la première étape de la coupure, sont également présents.

L'analyse sur gel des produits de coupure avait montré que les fragments générés par tg1653 et tg1277 avaient la même taille, puisqu'ils migraient à la même vitesse. Néanmoins, la détermination de la masse montre que le pic désigné sur le spectre par le numéro 6 n'a pas la même masse que le pic 5 identifié pour tg1653 : $[M-H]^-$ 2253.2, soit une masse supérieure de

18 Da par rapport à la masse du pic 5. On en déduit que ce produit serait généré par une réaction de β-élimination suivie de l'ajout d'une molécule d'eau au niveau de la double liaison dans le sucre ouvert (voir : Figure 76 et Figure 78). Cette étape d'hydrolyse peut être liée à l'activité enzymatique de tg1277, mais elle peut également avoir lieu pendant le traitement de l'échantillon. Le mécanisme d'action de tg1277 semble donc être le même que celui de tg1653, à savoir la β-élimination.

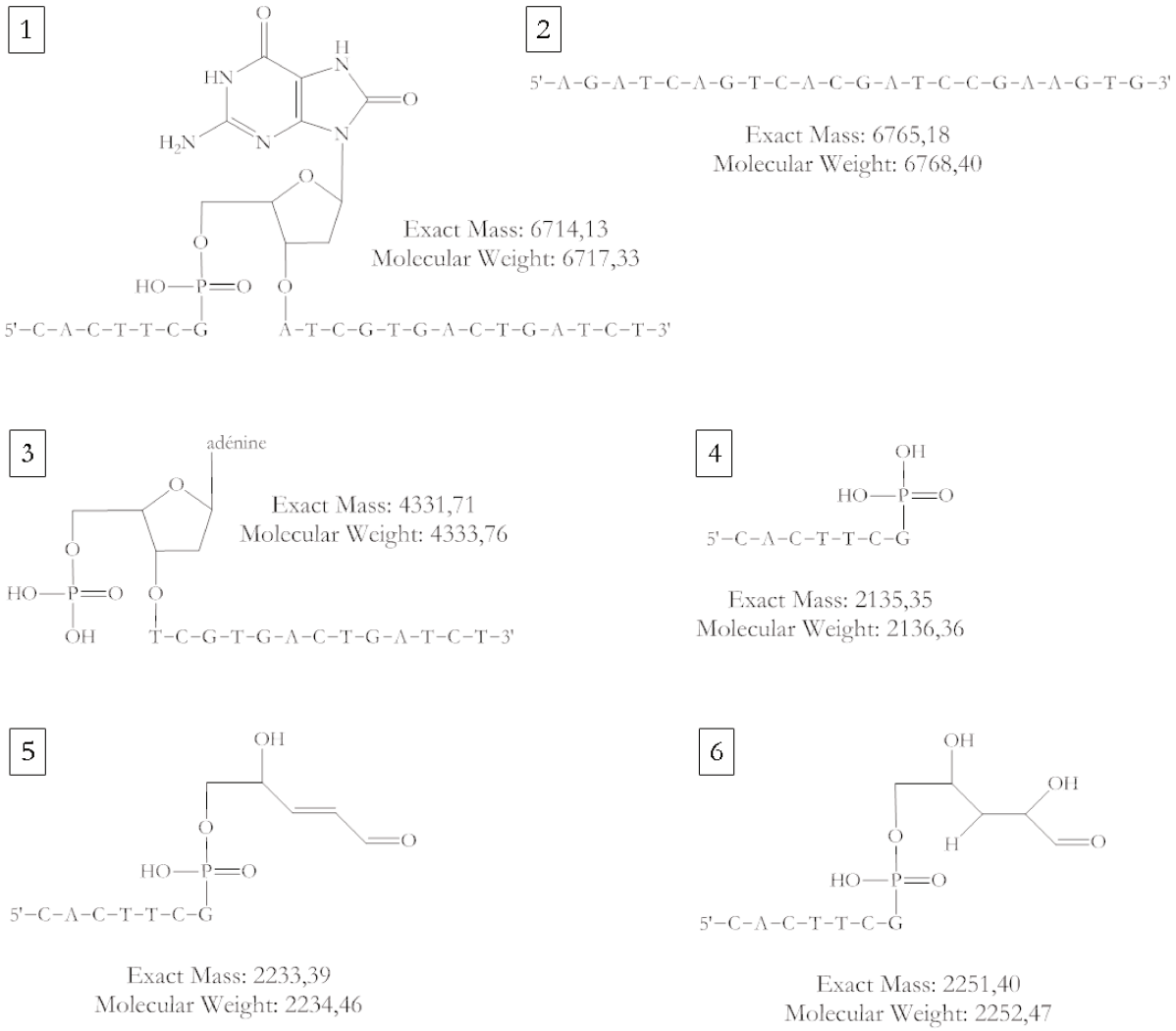


Figure 78 : Structure de l'oligonucléotide portant une 8-oxo-dG et des fragments obtenus après coupure par les différentes enzymes de réparation.

2.2.2.2.2. Coupure de la 5-OH-5-Me-dHyd

Afin de confirmer le mécanisme de coupure de l'enzyme tg1277 et de vérifier qu'il s'agit d'une β-élimination – le schéma caractéristique pour l'endonucléase Endo III – les

oligonucléotides contenant une 5-OH-5-Me-dHyd ont été incubés avec les enzymes suivantes : Fpg, Endo III, tg1653 et tg1277, puis analysés par MALDI-TOF/MS. Nous ne pouvions pas confirmer ce mécanisme en utilisant les sondes 8-oxo-dG, puisqu'elles ne sont pas substrats pour l'Endo III. Un contrôle non-digéré a également été préparé pour vérifier l'intégrité de la sonde avant l'étape de la digestion.

Dans un premier temps, nous avons vérifié l'intégrité de la sonde. La sonde 5-OH-5-Me-dHyd a été préférée à la sonde 5-OH-dHyd, également substrat pour l'Endo III, à cause de sa plus grande stabilité : nous avons constaté des difficultés dans la détection du pic correspondant à l'oligonucléotide entier dans le cas de la 5-OH-dHyd. Ceci est probablement dû à la dégradation de cette base modifiée sous l'effet du rayon laser utilisé pour ioniser les analytes. Sur le spectre correspondant au duplexe 5-OH-5-Me-dHyd (Figure 79), trois pics sont visibles :

1 – $[M-H]^-$ = 6704.0

2 – $[M-H]^-$ = 6767.9

2bis – $[M-H]^-/2$ = 3382.9.

Le pic 1 correspond à l'oligonucléotide entier qui porte la modification, et dont la masse moléculaire calculée est de 6705.32 Da, et le pic 2 représente l'oligonucléotide complémentaire, dont la masse était calculée à 6768.40 Da (Figure 81). Ces deux pics sont monochargés, contrairement au pic 2bis, qui correspond aussi à l'oligonucléotide complémentaire, mais avec une double charge – son rapport masse/charge est alors deux fois plus petit que celui déterminé pour le pic 2.

Les spectres MALDI-TOF présentant les fragments des oligonucléotides obtenus après la digestion par les enzymes : Fpg, Endo III et tg1277 sont montrés sur la Figure 79, et le spectre correspondant à tg1653 est montré sur la Figure 80. Dans tous ces cas, sauf pour la coupure par tg1277, on détecte un faible pic de l'oligonucléotide lésé qui n'a pas été digéré (pic 1 qui était également visible sur le spectre de l'oligonucléotide non-digéré). Sur tous les quatre spectres, les pics de l'oligonucléotide complémentaire (avec une ou deux charges) sont également visibles, ce qui indique qu'aucune coupure au niveau de la séquence non-lésée n'a été effectuée.

L'analyse MALDI-TOF/MS confirme les résultats observés précédemment, c'est-à-dire l'absence de coupure de la 5-OH-5-Me-dHyd par tg1653 – aucun nouveau pic correspondant à un fragment de l'oligonucléotide de départ n'est détecté dans ce mélange (Figure 80), et son spectre se superpose à celui de l'oligonucléotide non-digéré. On peut ainsi confirmer que le seul substrat de cette enzyme trouvé à ce jour est la 8-oxo-dG.

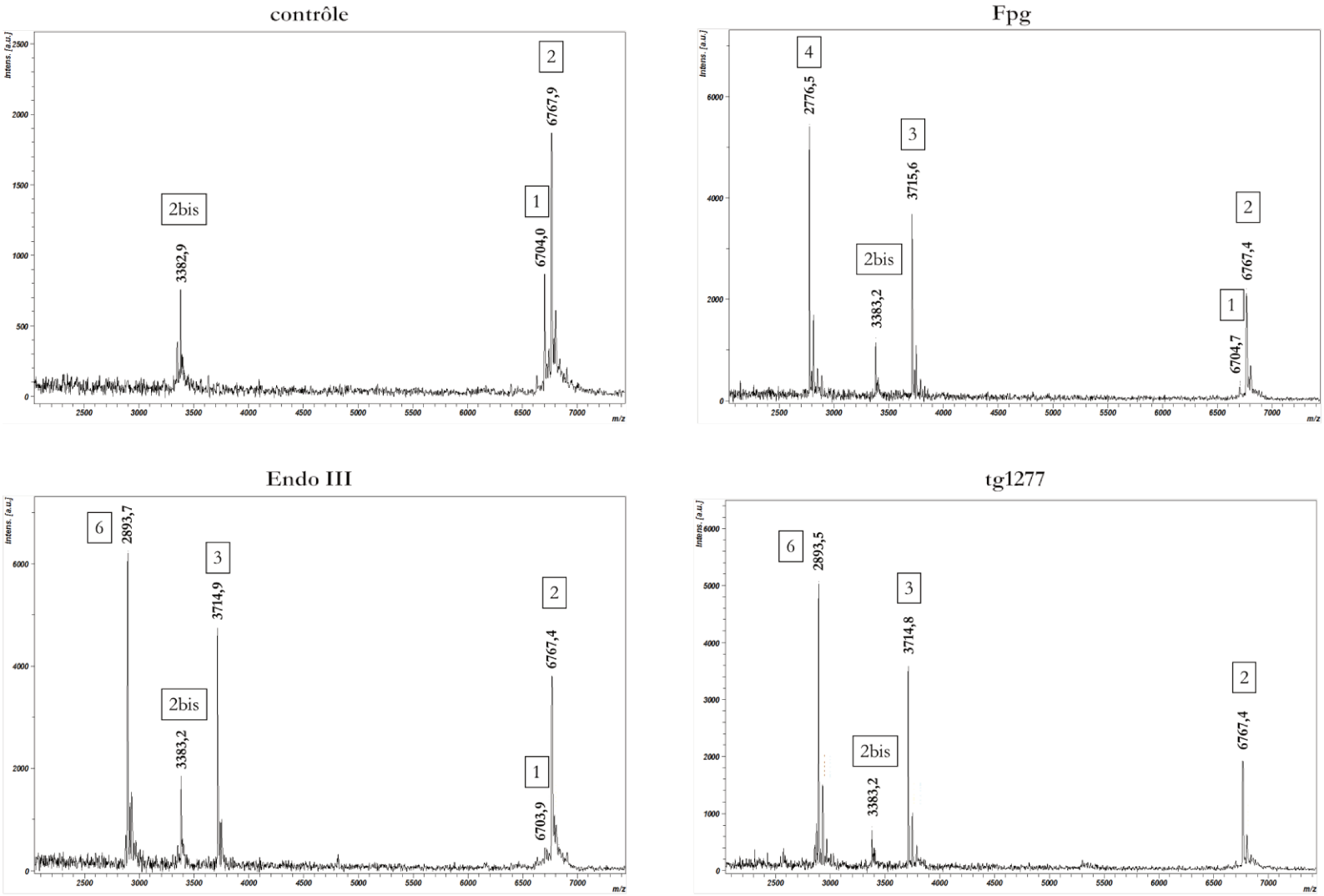


Figure 79 : Spectres MALDI-TOF/MS de l'oligonucléotide double-brin, contenant sur un brin une 5-OH-5-Me-dHyd, incubés avec la Fpg, par l'Endo III ou par l'enzyme tg1277 isolée de *T.gammatolerans*. Contrôle = oligonucléotide duplexe non-digéré.

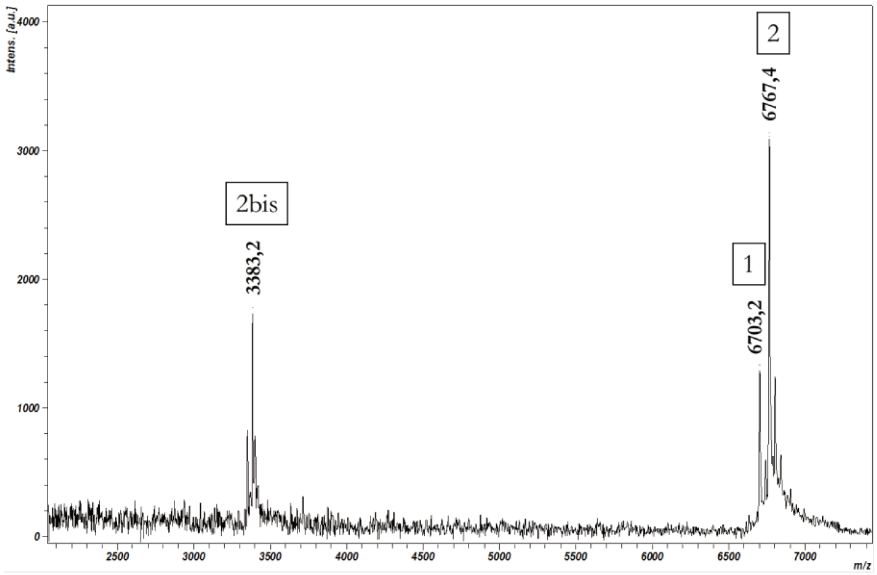


Figure 80 : Spectre MALDI-TOF/MS de l'oligonucléotide double-brin contenant sur un brin une 5-OH-5-Me-dHyd, incubé avec l'enzyme tg1653 isolée de *T.gammatolerans*.

Sur les trois spectres : coupure par la Fpg, l'Endo III et la tg1277, nous voyons l'apparition de nouveaux pics, qui correspondent aux fragments de la sonde obtenus après la coupure au niveau de la modification. Dans les trois cas on observe un pic appelé pic 3, visible à $[M-H]^-$ autour de 3715 : 3715.6 (Fpg), 3714.9 (Endo III) et 3714.8 (tg1277), pour une masse théorique de 3716.36 Da. Il correspond au fragment obtenu suite à la première étape de la coupure, c'est-à-dire la β-élimination, et sa structure est similaire au pic 3 identifié auparavant sur les spectres de la 8-oxo-dG. Ce pic n'a pas la même masse que celui identifié pour la coupure de la sonde 8-oxo-dG, puisque la lésion n'est pas localisée au même site dans la séquence, mais il est issu du même mécanisme d'action enzymatique.

Le dernier pic identifié sur le spectre de digestion de la sonde 5-OH-5-Me-dHyd par la Fpg est analogue au pic observé pour la sonde 8-oxo-dG, c'est pourquoi nous lui avons attitré le même numéro – 4. Sa masse théorique est de 2778.77 Da et nous l'identifions sur le spectre à $[M-H]^-$ 2776.5, ce qui confirme qu'il correspond au fragment de la sonde obtenu du côté 5' de la lésion suite à une réaction de δ-élimination.

Le spectre de l'Endo III nous permet de voir l'autre mécanisme de coupure, celui de la β-élimination. C'est le pic 6 qui constitue la différence par rapport au spectre de la Fpg, puisque son $[M-H]^-$ est de 2893.7, soit supérieur de 117 à la masse du pic 4 ($[M-H]^{-1}$ de 2776.5). Cette différence correspond à la masse d'un résidu de sucre hydraté au niveau de sa double liaison, c'est

alors un produit de la β-élimination auquel a été ajouté une molécule d'eau. La présence du résidu de sucre sur le site de coupure confirme le mécanisme d'action de l'Endo III.

Nous disposons à présent des contrôles des deux mécanismes de coupure qui peuvent être observés pour les enzymes participant dans l'initiation du système BER : β,δ-élimination pour la Fpg et β-élimination pour l'Endo III. La comparaison du spectre obtenu pour tg1277 avec les spectres de la Fpg et de l'Endo III va nous permettre maintenant de confirmer son mécanisme d'action.

En effet, le spectre de tg1277 se superpose parfaitement à celui de l'Endo III, avec comme seule différence le manque de pic 1, ce qui prouve une coupure plus efficace effectuée par cette enzyme. Comme nous l'avons supposé après le test de coupure de la 8-oxo-dG, le mécanisme d'action de tg1277 est le même que celui de l'Endo III, à savoir la β-élimination. L'hydratation du résidu de sucre qui peut être constatée sur le spectre, ne peut pas être clairement attribuée à l'activité enzymatique et nous supposons qu'elle advient lors de l'analyse de l'échantillon. Pour tg1653 nous ne pouvons pas comparer directement son spectre avec celui de l'Endo III, puisque ces deux enzymes ne sont pas actives sur les mêmes substrats. Néanmoins, le mécanisme de tg1277 ayant été confirmé, il peut désormais servir de contrôle pour tg1653, les deux ayant une activité enzymatique sur la 8-oxo-dG (Figure 77). Nous en concluons alors que le mécanisme de coupure exercé par tg1653 est également la β-élimination, et que la seule différence entre les deux enzymes provenant de *T.gammatolerans* que nous avons étudiées ici consiste en l'étape d'hydratation du sucre qui n'est constatée que pour tg1277.

Les expériences effectuées par spectrométrie de masse MALDI-TOF/MS nous ont permis de déterminer le mécanisme de coupure des deux enzymes isolées de *Thermococcus gammatolerans* et impliquées dans le système BER. En comparant leur mécanisme de coupure avec les mécanismes connus des enzymes bactériennes Fpg et Endo III nous avons ainsi déterminé que les protéines tg1653 et tg1277 exercent leur activité de coupure par un mécanisme de β-élimination.

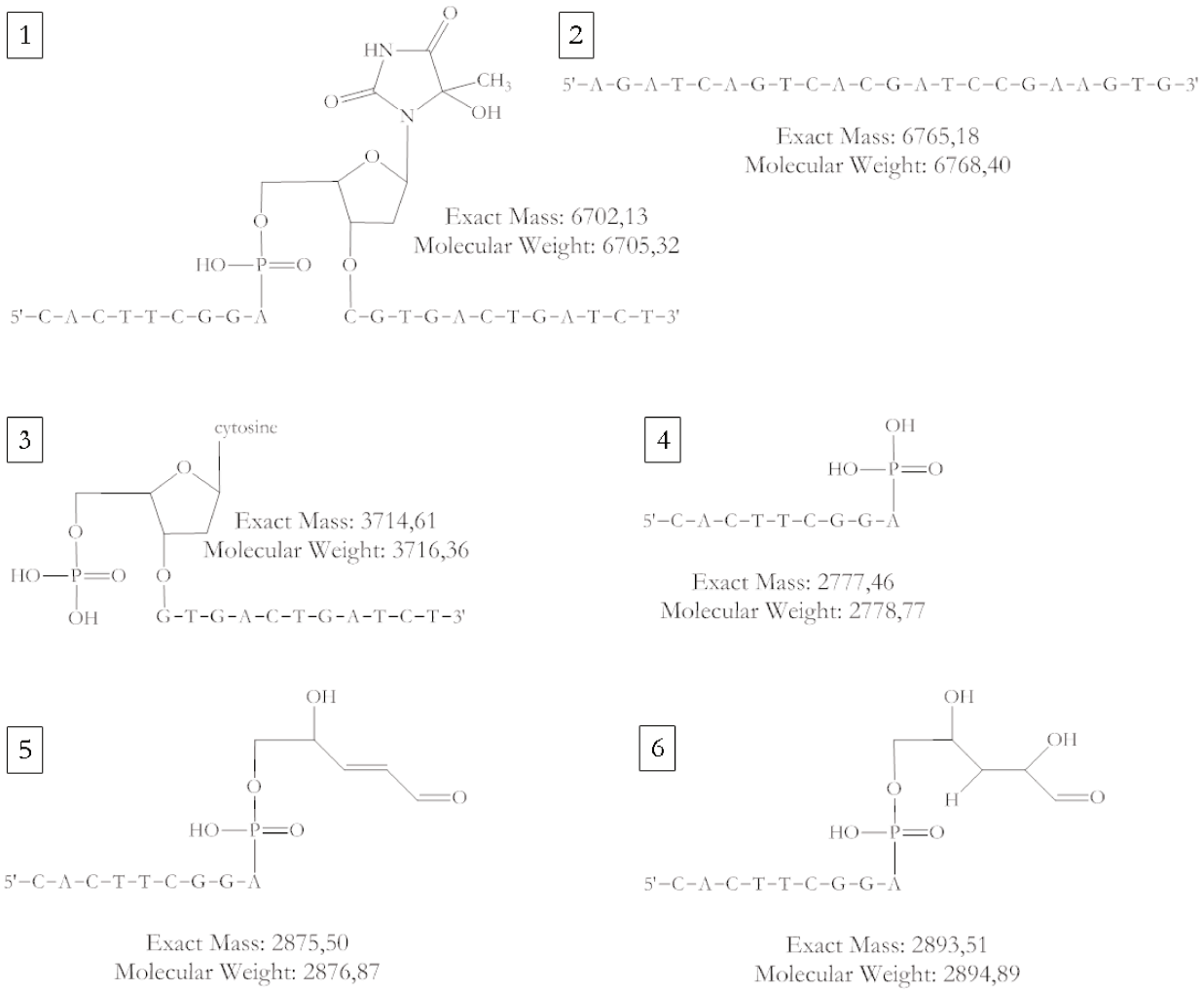


Figure 81 : Structure de l'oligonucléotide portant une 5-OH-5-Me-dHyd et de ses fragments obtenus après la coupure par différentes ADN glycosylases du BER.

2.2.2.3. Activité de coupure en fonction de la base opposée

Après avoir identifié le mécanisme d'action de tg1653 et tg1277, nous avons voulu vérifier si l'activité de ces deux enzymes est influencée par la base qui est opposée à la lésion. Nous avons utilisé pour cela les sondes portant une 8-oxo-dG ou une 8-oxo-dA, et nous les avons hybridées avec l'une des quatre séquences complémentaires, variant d'une base au niveau de la base complémentaire à la lésion (voir : Tableau 7 dans le chapitre « Matériel et méthodes »). La 8-oxo-dA est une lésion moins connue que la 8-oxo-dG, c'est pourquoi nous avons voulu vérifier si les enzymes originaux de *T.gammatolerans* possédaient une activité vis-à-vis d'elle. Dans un premier temps nous avons vérifié si ces mésappariements n'avaient pas d'influence sur l'hybridation de nos sondes, en les faisant migrer sur un gel natif d'acrylamide et en les comparant au contrôle simple-brin. Nous pouvons constater que, quelle que soit la base située en face de la

lésion, ceci n'a pas d'influence sur l'hybridation de nos sondes, puisque toutes les huit sont sous forme double-brin (Figure 82).

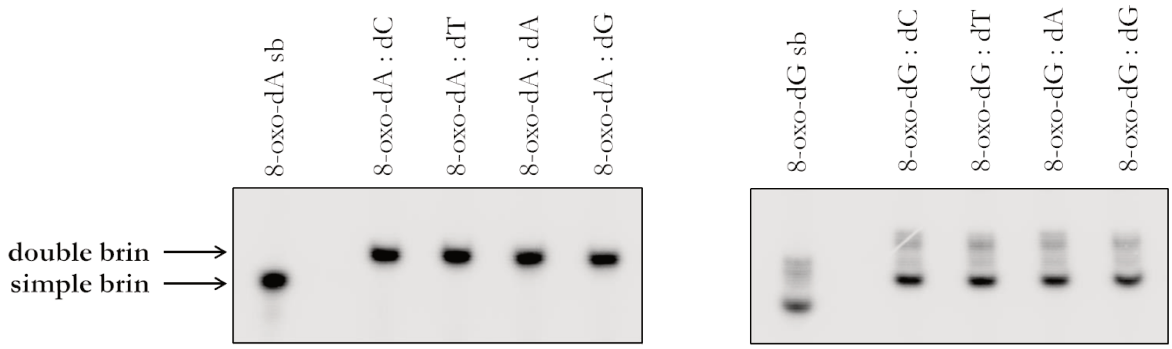


Figure 82 : Vérification de l'hybridation de l'oligonucléotide contenant une 8-oxo-dA ou une 8-oxo-dG avec divers brins complémentaires.

Nous avons ensuite incubé ces sondes duplexes avec trois enzymes : Fpg, qui nous a servi à nouveau de contrôle, tg1653 et tg1277. Comme nous pouvons le voir sur le gel présenté sur la Figure 83, aucune de ces enzymes n'est active sur la 8-oxo-dA.

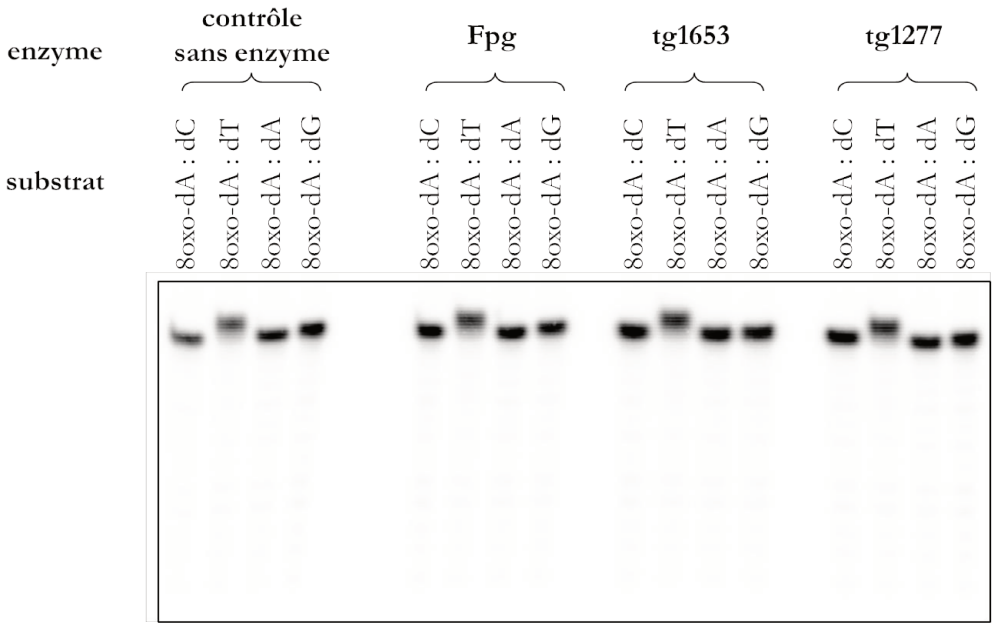


Figure 83 : Activité de coupure des enzymes isolées de *T.gammatolerans* sur les oligonucléotides contenant une 8-oxo-désoxyadénosine (8-oxo-dA) appariée avec une des quatre bases complémentaires.

La même expérience a ensuite été effectuée avec les sondes 8-oxo-dG (Figure 84). Sur le gel, on voit le même profil que celui obtenu lors des premiers tests de coupure : les fragments d'ADN obtenus après la coupure par la Fpg migrent plus loin que ceux générés suite à la coupure par tg1653 ou tg1277. Ceci est indépendant de la base complémentaire, qui a de l'influence uniquement sur l'efficacité de la coupure, et non pas sur son mécanisme.

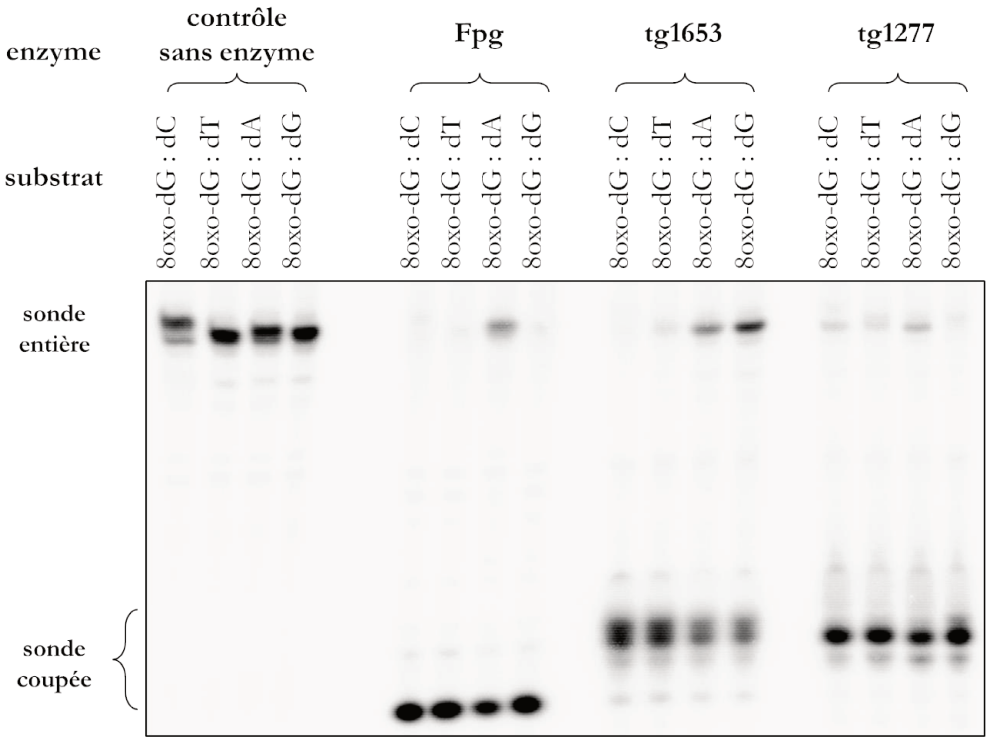


Figure 84 : Activité de coupure des enzymes isolées de *T.gammatolerans* sur les oligonucléotides contenant une 8-oxo-désoxyguanosine (8-oxo-dG) appariée avec une des quatre bases complémentaires.

L'activité de la Fpg n'est presque pas influencée par la base opposée à la lésion, à l'exception du mésappariement avec une adénine, qui affaiblit légèrement l'efficacité de la coupure : nous pouvons remarquer une faible bande résiduelle de l'oligonucléotide de départ présente uniquement dans cette condition. Ce phénomène a déjà été mis en évidence par Tchou et *al.* qui ont montré que l'efficacité de la coupure effectuée par la Fpg sur la 8-oxo-dG était plus faible quand une adénine était localisée sur le brin complémentaire au lieu d'une cytosine. Pour les mésappariements avec une thymine ou une guanine, au contraire, la coupure est même plus efficace, mais nous n'avons pas pu observer cela, puisque dans nos conditions expérimentales utilisées, la coupure était déjà complète pour le duplexe 8-oxo-dG :C [320].

En ce qui concerne l'activité de tg1653, cette protéine est plus sensible à la présence des mésappariements dans la séquence que la Fpg. La coupure au niveau de la sonde lésée est complète si une pyrimidine est localisée en face de la 8-oxo-dG, tandis que la présence d'une purine va avoir un effet inhibiteur sur son activité enzymatique. Quand une adénine ou, plus encore, une guanine est présente en face de la lésion, nous voyons alors la persistance d'une bande intense correspondant au produit de départ, bien que la coupure soit toujours exercée, mais à un degré moindre.

Pour tg1277 un effet similaire à celui vu pour la Fpg, mais moins prononcé, est à noter – des traces de la sonde non-coupée sont visibles dans les trois cas : une cytosine, une thymine ou une adénine située en face de la lésion, et cette bande semble être plus prononcée quand c'est une adénine qui est mésappariée avec la 8-oxo-dG. Cet effet inhibiteur reste cependant très léger, du moins pour la quantité de protéine utilisée dans notre expérience. Pour confirmer cet effet inhibiteur des mésappariements, des expériences complémentaires avec une moindre concentration de l'enzyme devraient être effectuées.

La protéine tg1653 est sensible aux mésappariements en face de la 8-oxo-dG et son activité de coupure est affaiblie quand une purine est située en face de la lésion. Au contraire, l'activité de tg1277 ne semble être que très légèrement perturbée par la présence d'un mésappariement sur le brin complémentaire.

2.2.2.4. Test d'activité des protéines de *T.gammatolerans* sur les lésions complexes d'oxydation

La dernière étape de l'analyse de l'activité enzymatique des enzymes isolées de *T.gammatolerans* a consisté, comme nous l'avons fait précédemment pour la Fpg, à vérifier si elles étaient capables d'exercer une coupure sur les oligonucléotides contenant un cyclonucléoside. Pour tester ceci, nous avons incubé les quatre sondes : 5'S-cyclo-dA, 5'R-cyclo-dA, 5'S-cyclo-dG et 5'S-cyclo-dHT avec les protéines isolées : tg1653 ou tg1277. Un mélange de ces deux enzymes a également été testé pour voir si leurs activités pouvaient se compléter. Comme nous pouvons le voir sur le gel (Figure 85), aucune activité de coupure n'est visible au niveau de ces lésions.

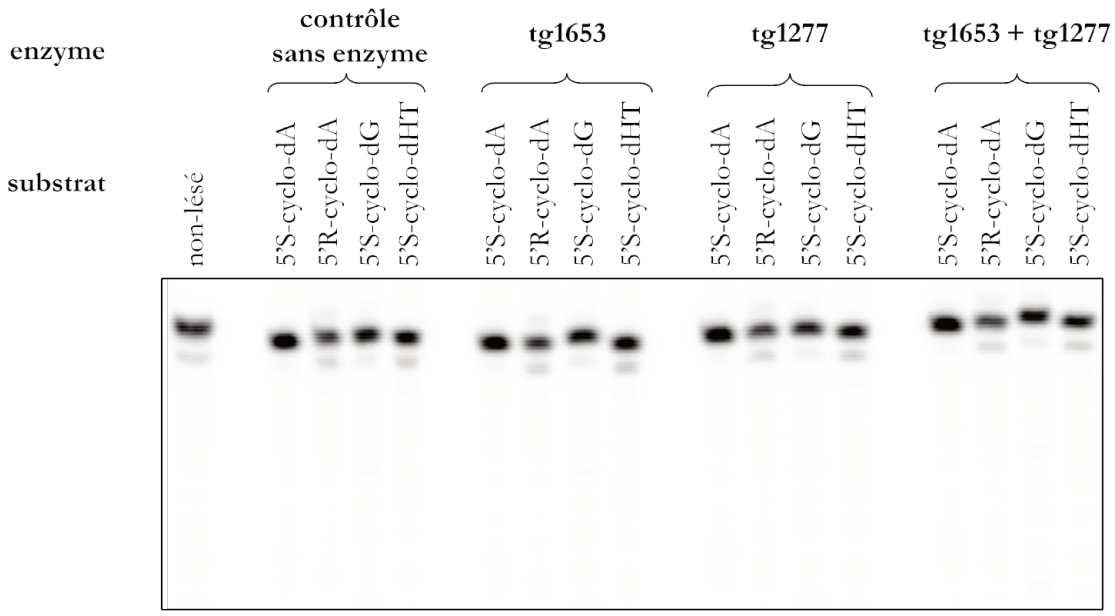


Figure 85 : Activité de coupure des enzymes isolées de *T.gammatolerans* sur les oligonucléotides contenant un cyclonucléoside.

2.2.3. Analyse des interactions entre les cyclonucléosides et les protéines de *Thermococcus gammatolerans*

Afin de déterminer l'affinité des enzymes isolées de *Thermococcus gammatolerans* pour différentes lésions complexes de l'ADN, nous avons procédé de la même manière que pour les tests d'affinité de la glycosylase Fpg. La première étape a consisté à optimiser les conditions d'incubation, qui sont propres à chaque protéine. La principale difficulté envisagée concernait les exigences thermiques de ces protéines provenant d'un organisme thermophile. Nous nous en sommes déjà affranchis pendant les tests d'activité en augmentant la température d'incubation progressivement de 37°C à 65°C, en espérant que cela augmenterait l'activité de la protéine avant de dénaturer son substrat double-brin. Pour les tests d'interaction type « gel shift » cette approche n'est pourtant pas possible, puisque même si une température élevée d'incubation augmentait l'affinité de la protéine pour son substrat, une longue migration à température ambiante pourrait les déstabiliser. De plus, la présence des oligonucléotides sous forme de duplex en tant que contrôle est indispensable, et nous ne pouvons alors pas augmenter la température d'incubation au-delà de leur température de fusion.

Dans un premier temps, nous avons testé l'influence de la température d'incubation sur la formation des complexes. Pour cela, nous avons incubé les sondes non-lésées ou les sondes THF avec les deux enzymes de *T.gammatolerans* pendant 20 minutes à température ambiante ou 37°C,

mais nous n'avons pas constaté de différences, en terme d'affinité, entre ces deux températures (résultats non montrés). Ainsi, nous avons décidé de continuer les expériences à température ambiante, comme cela avait été fait pour l'enzyme Fpg.

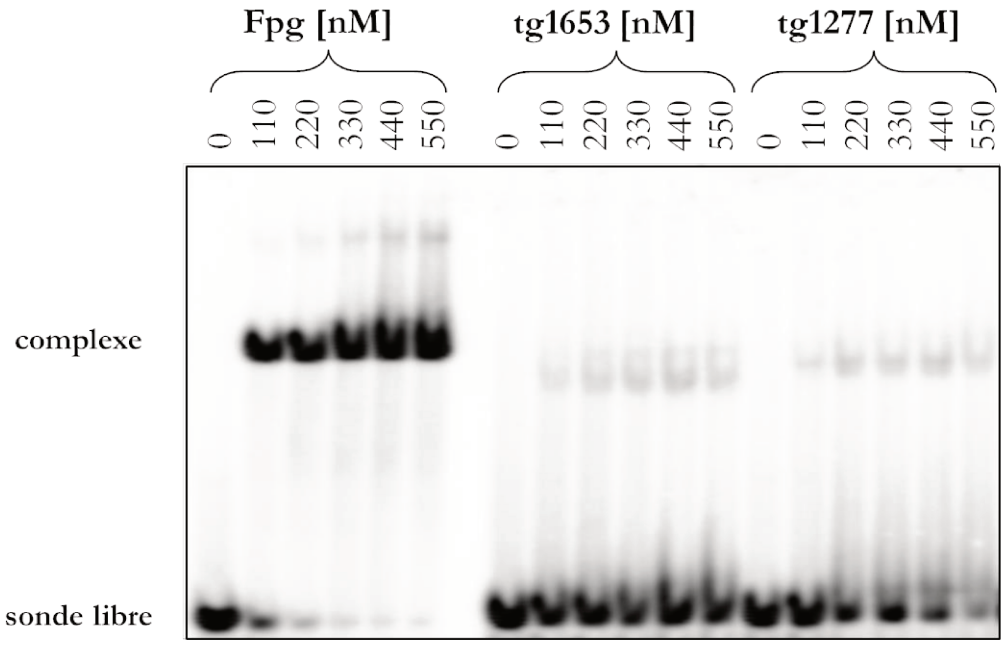


Figure 86 : Interactions entre un oligonucléotide double brin contenant sur un brin un site abasique modifié (THF) et les enzymes : Fpg, tg1653 et tg1277, dans les conditions optimisées initialement pour la protéine Fpg.

Pour comparer l'affinité des enzymes de *T.gammatolerans* à celle de la Fpg, nous avons effectué un test de gel shift avec une sonde THF (« contrôle positif ») pour ces 3 protéines dans les mêmes conditions – celles, que nous avons initialement mis au point pour la Fpg. Comme nous pouvons le constater sur la Figure 86, il n'y a pratiquement pas d'interactions entre cette sonde et les enzymes tg1653 et tg1277, même à une très haute concentration de l'enzyme (550 nM). Cela suggère que soit ces protéines ont une faible affinité pour les sites abasiques et leur concentration serait alors insuffisante, soit les conditions de la réaction ne sont pas optimales.

Le tampon utilisé pour la Fpg est très stringent – il contient 150 mM de KCl – puisque c'était une condition nécessaire pour éliminer les interactions non-spécifiques de la Fpg avec nos sondes. Cette salinité pourrait être trop élevée pour les deux autres protéines étudiées ici, au point d'empêcher toute interaction entre elles et leurs cibles. Pour vérifier cette hypothèse, nous avons incubé les oligonucléotides dans les mêmes conditions, avec comme seule différence une moindre concentration en sel : 50 mM de KCl au lieu de 150 mM.

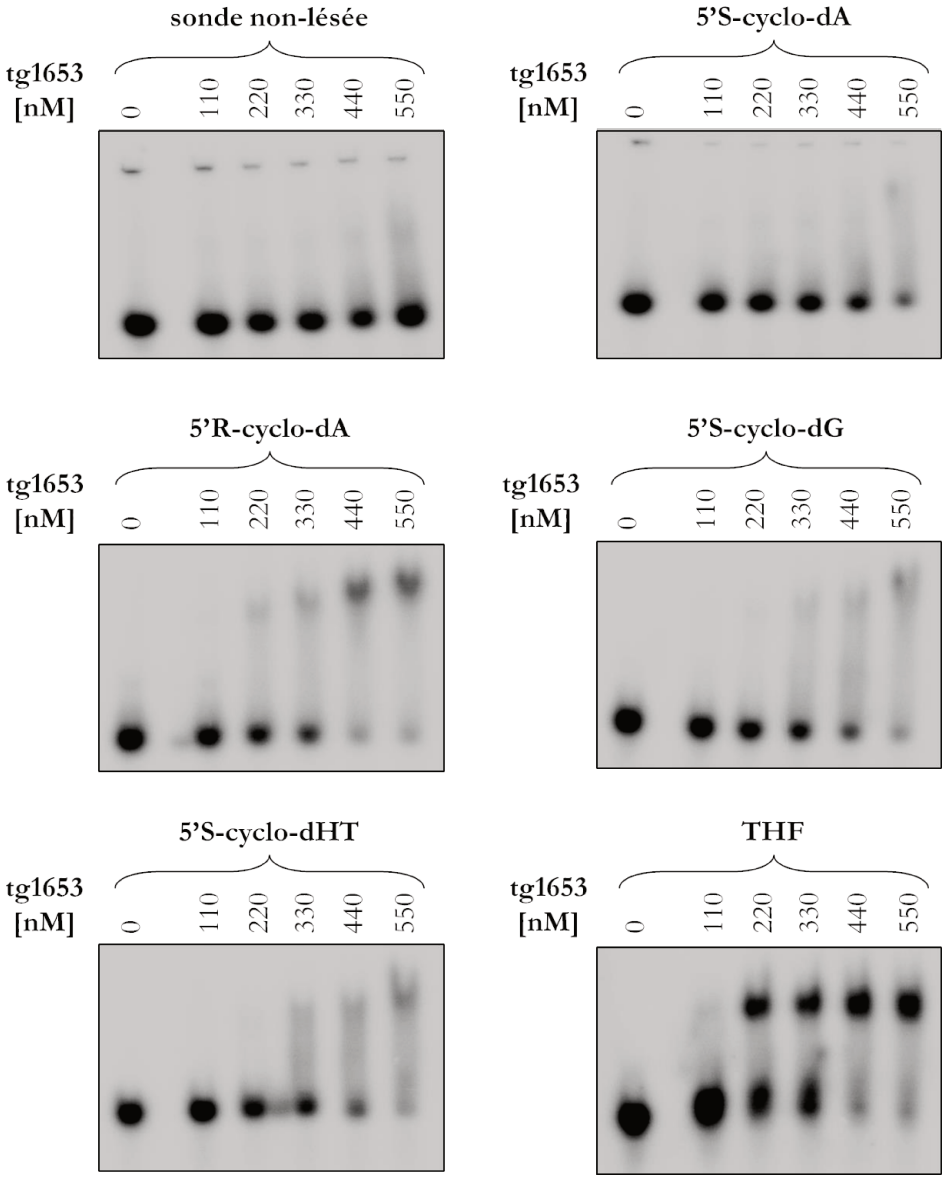


Figure 87 : Interactions entre des oligonucléotides double-brin contenant sur un brin un cyclonucléoside, et l'enzyme tg1653 isolée de *T.gammatolerans*.

Comme nous pouvons le voir sur les Figure 87 et 88, dans les conditions moins salines, les deux protéines de *T.gammatolerans* montrent une affinité pour le THF, ce qui valide le choix du tampon. Ce changement de conditions d'incubation par rapport à la Fpg rend impossible toute comparaison des affinités des protéines tg1653 et tg1277 avec celle de la Fpg pour les lésions testées, mais les interactions de la même protéine avec différentes lésions peuvent néanmoins être analysées.

Pour la protéine tg1653 (Figure 87), nous ne voyons presque pas d'interactions avec les sondes non-lésées, et uniquement une faible interaction est visible pour les plus grandes doses de

la protéine. En ce qui concerne les quatre sondes portant des cyclonucléosides, leur reconnaissance par la tg1653 n'est pas égale : la formation des complexes avec les sondes 5'R-cyclo-dA et, à un moindre degré 5'S-cyclo-dHT, peut être remarquée, tandis que les deux autres sondes : 5'S-cyclo-dA et 5'S-cyclo-dG semblent ne pas être reconnues par l'enzyme.

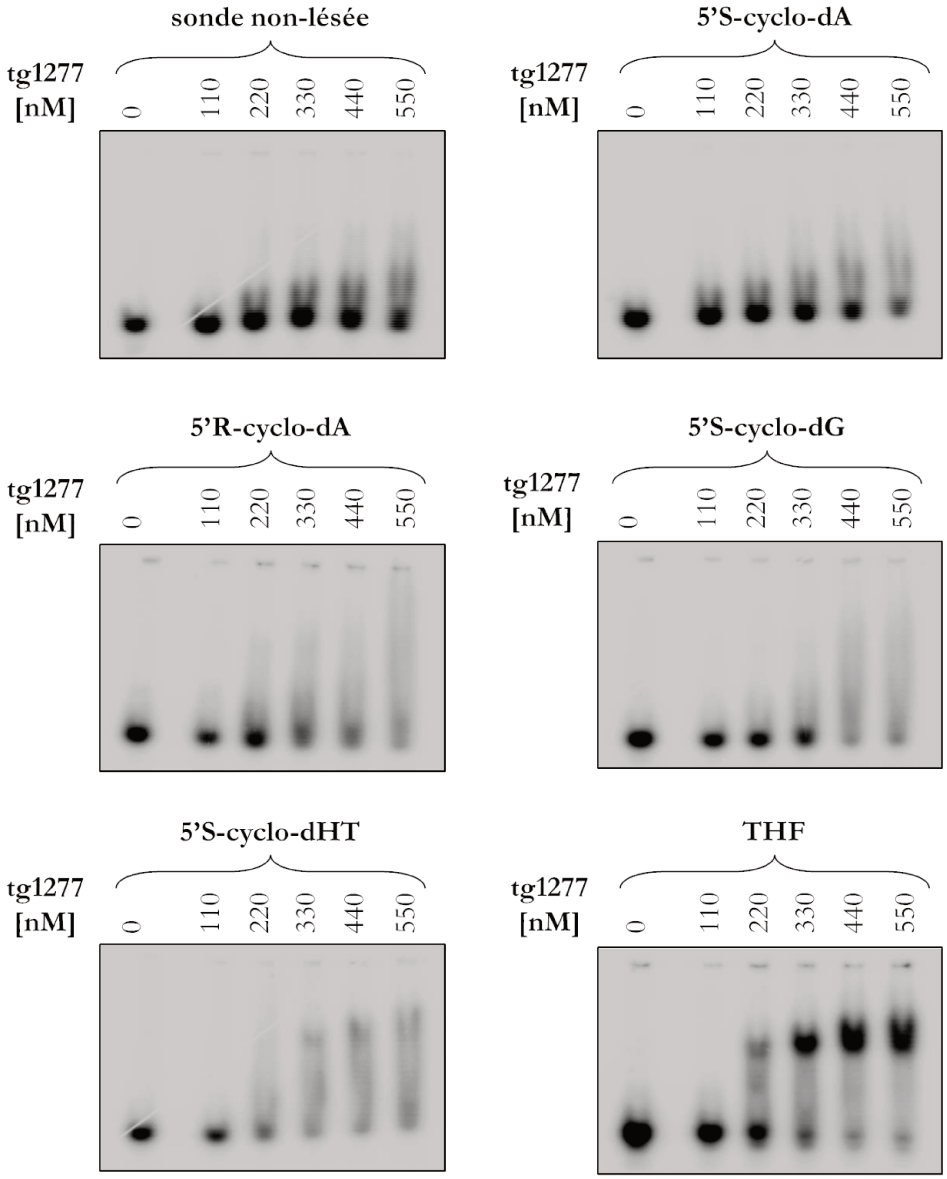


Figure 88 : Interactions entre des oligonucléotides double-brin contenant sur un brin un cyclonucléoside, et l'enzyme tg1277 isolée de *T.gammatolerans*.

En ce qui concerne les interactions entre les mêmes sondes et la protéine tg1277, nous pouvons constater que celles-ci sont beaucoup plus faibles par rapport aux celles observées avec tg1653. Il n'y a pas de complexes ni avec la sonde non-lésée si avec les sondes lésées, à

l'exception de la sonde 5'S-cyclo-dHT, pour laquelle la formation de complexes se décomposant lors de la migration est visible (Figure 88). Ces observations sont cohérentes avec les résultats obtenus auparavant avec la protéine Fpg, dont l'affinité était plus forte pour 5'S-cyclo-dHT et 5'R-cyclo-dA, que pour les deux autres cyclonucléosides.

Les tests effectués sur les protéines tg1653 et tg1277 montrent que ces deux enzymes possèdent une affinité pour le résidu THF. Cette affinité ne peut cependant pas être comparée à celle de la Fpg, car les protéines de *Thermococcus gammatolerans* nécessitent les conditions d'incubation différentes de celles de la Fpg.

Tg1653 possède, en plus, une affinité pour les lésions 5'R-cyclo-dA, 5'S-cyclo-dG et 5'S-cyclo-dHT, tandis que les interactions des cyclonucléosides avec tg1277 sont beaucoup plus faibles – seuls des complexes avec la lésion 5'S-cyclo-dHT ont été observés.

V Discussion

1. Screening des extraits protéiques

Une grande partie de ce travail de thèse a été consacrée à l'optimisation du piège à protéines destiné à identifier des protéines interagissant avec les cyclonucléosides. Différentes sondes ont été conçues et testées, ainsi que différentes méthodes de récupération des protéines. Le piège a été appliqué à deux sources d'extraits protéiques : extraits humains ou bactériens. Nous allons maintenant présenter nos observations et commenter les résultats obtenus concernant tout d'abord cette mise au point technique, puis le sens biologique des interactions identifiées.

1.1. Considérations techniques

1.1.1. Utilité du bras photoclivable

Au cours de l'optimisation du piège, nous avons testé deux variantes pour libérer les protéines. Dans la technique dite « classique », le chauffage du piège dans un tampon dénaturant vise à rompre toutes les interactions entre les protéines et l'ADN, ou n'importe quel élément du piège. Cette stratégie permet de libérer la totalité des protéines, mais son inconvénient majeur consiste en la libération simultanée de toutes les protéines adsorbées à la surface des billes et générant un bruit de fond considérable. C'est pour cette raison que nous avons essayé d'introduire dans la construction de nos pièges un bras photoclivable, situé entre la surface de la bille et la séquence de la sonde. Cette stratégie a déjà été testée par l'équipe de Danièle Praseuth : les chercheurs ont réussi à piéger des protéines participant à la réparation des cassures double-brin : Ku70, Ku80, PARP1 et DNA-PKcs (DNA-dependent protein kinase catalytic subunit). Ils ont ensuite démontré que celles-ci conservaient leur fonctionnalité une fois décrochées de leur support, ce qui indique que le piégeage et le relargage des protéines par irradiation n'ont pas d'influence sur leur activité. Ils ont également prouvé l'utilité d'un bras photoclivable, qui leur a permis d'enrichir les éluats en protéines d'intérêt, et d'éliminer une grande partie des protéines fixées de manière non-spécifique sur la surface des billes couvertes de streptavidine. Néanmoins, ces chercheurs ont également remarqué, que même les protéines d'intérêt possèdaient une

certaine affinité pour les billes, car elles ont été retrouvées dans l'éluat et dans la fraction récupérée sur les billes après l'irradiation [300].

La comparaison que nous avons effectuée entre ces deux stratégies nous a permis de voir les différences à la fois quantitatives et qualitatives entre les ensembles des protéines récupérées. En théorie, l'utilisation d'un bras photoclivable devrait permettre la récupération « propre » de toutes les protéines qui ont reconnu les sondes, que celles-ci soient lésées ou non, sans libérer les protéines attachées sur la surface des billes. Comme nous avons pu le constater, ce cas théorique est loin de la réalité pratique, car beaucoup de protéines sont retrouvées dans l'échantillon provenant des billes seules. Au vu des faibles avantages apportés par cette technique de relargage des protéines par rapport à la stratégie « classique » de chauffage et de dénaturation, nous n'avons pas retenu cette modification du système pour nos pièges à protéines.

1.1.2. Exhaustivité des protéines piégées

Le système de ligand fishing a déjà été exploité au sein du laboratoire pour étudier l'interactome des lésions générées par les dérivés du platine [268] et des modifications épigénétiques de l'ADN [publication en cours]. La problématique de l'exhaustivité de la liste des protéines identifiées a été discutée dans la thèse de Christophe Bounaix du Puch [269]. Nous allons néanmoins rappeler les avantages et les contraintes majeurs de cette technique, que nous avons également rencontrés lors de cette thèse.

Le principal avantage du système ligand fishing est sa grande adaptabilité : en fonction de la lésion que l'on souhaite étudier, l'utilisation des sondes plasmidiques, oligonucléotidiques ou encore synthétisées par la PCR peut être envisagée. Cette méthode permet également une modification facile des conditions de l'incubation et des rinçages, car les sondes immobilisées sur la surface des billes peuvent être reprises dans différents tampons, grâce à leur séparation facile de l'éluat à l'aide d'un simple aimant. Ainsi, la stratégie du ligand fishing permet d'effectuer un screening des protéines au sein des extraits, en testant en parallèle de nombreuses lésions et paramètres d'incubation.

Des inconvénients liés à la technique du ligand fishing existent également. Le système de piège n'est sûrement pas exhaustif, puisque les interactions faibles entre les protéines et leurs cibles sont déterminées par plusieurs facteurs, tels que la composition du tampon, le temps et la température de l'incubation. Il est alors impossible de trouver les conditions universelles dans lesquelles toutes les protéines ayant une affinité pour la lésion ciblée seraient piégées. Egalement,

lorsqu'il s'agit des protéines de réparation, elles agissent le plus souvent de manière concertée et dans un ordre prédéfini, où une protéine ne peut pas reconnaître la lésion tant qu'une autre ne se dissocie pas de celle-ci. C'est aussi pour cela que le piégeage de toutes les protéines agissant en plusieurs étapes est impossible et la liste des protéines identifiées sera toujours incomplète. Les avantages et les inconvénients de cette stratégie, ainsi que d'autres techniques de screening de protéines, ont été récemment détaillés dans une revue rédigée lors de la thèse [324].

Au cours de nos expériences nous n'avons pas réussi à piéger les protéines dont les interactions avec les cyclonucléosides sont connues, comme par exemple la Fpg. Nous avons démontré son affinité pour ces lésions par retard sur gel, mais sans réussir à confirmer cette reconnaissance par ligand fishing, ce qui montre les limites de cette stratégie. Au sein des extraits HeLa, nous avons pu identifier des protéines dont l'interaction avec l'ADN était inhibée par les cyclonucléosides, mais pas de protéines interagissant avec ces derniers. Ainsi, la liste des protéines identifiées sur un type de piège ne peut jamais être considérée comme exhaustive, et plusieurs analyses, faisant varier le temps de l'incubation ainsi que la composition du tampon utilisé, devraient être effectuées pour chaque lésion, afin d'obtenir une vision complète de son interactome.

1.1.3. Application à différents types d'extraits protéiques

Au cours de nos travaux, nous avons testé l'application de notre piège à protéines à différents type d'extraits : extraits nucléaires eucaryotes provenant des cellules HeLa et extraits procaryotes provenant d'*E.coli*. L'utilisation de ces deux sources de protéines n'a pas donné le même type de résultats. Les premiers essais avec les extraits eucaryotes ont montré que beaucoup de protéines non-spécifiques se fixaient sur nos pièges, dont la majorité sur la surface des billes. Les améliorations ont porté ensuite sur la diminution de ce bruit de fond dans le but de favoriser le piégeage des protéines pertinentes.

Au contraire, la tentative du piégeage des protéines bactériennes s'est avérée infructueuse, car peu de protéines ont pu être identifiées. Ainsi, très peu de protéines ont été retrouvées même sur les billes, ce qui signifie que même les interactions non-spécifiques, qui devraient être, faute d'optimisation, les plus courantes, sont inexistantes. Il s'avère donc que notre piège n'est apparemment pas adapté aux extraits protéiques de type bactérien. La raison de cette incompatibilité du piège avec les extraits procaryotes pourrait être due à la nature de ces derniers.

Il existe des études d'interactions entre les protéines bactériennes et les lésions de l'ADN, qui utilisent également le principe de pièges. L'équipe de Zharkov a notamment travaillé sur la reconnaissance des sites abasiques par les enzymes d'*E.coli*. Ces chercheurs ont conçu des pièges à protéines interagissant avec les sites abasiques. Ces dommages sont pris en charge par les protéines du système BER, notamment par les AP lyases ou par les glycosylases bifonctionnelles, mais ils peuvent également interagir avec d'autres protéines, par exemple les polymérases β ou γ, perturbant ainsi certains processus cellulaires [325]. Lors de ces interactions, un intermédiaire est formé entre le site actif de la protéine et le carbone C-1' du désoxyribose, appelé la base de Schiff. Cette base de Schiff peut être ensuite réduite avec le borohydrure de sodium, l'enzyme se trouvant ainsi fixée de manière covalente à la lésion. C'est cette technique qu'ont utilisée les chercheurs de l'équipe de Zharkov afin d'analyser l'interactome des protéines interagissant avec les sites abasiques. Contrairement à notre stratégie, ces chercheurs ont effectué dans un premier temps le piégeage des protéines en solution, et ensuite ils ont fixé les sondes avec les protéines sur la surface des billes fonctionnalisées. Les rinçages stringents ont ensuite été réalisés pour éliminer les protéines non-spécifiques, les sondes ont alors été digérées par la DNAse I et l'éluat contenant les enzymes ayant reconnu les sites abasiques a été récupéré. D'après les résultats des auteurs, plusieurs rinçages avec un tampon à très haute salinité ont été nécessaires pour que les protéines aspécifiques soient éliminées de la surface des billes. Ceci indique que, contrairement à nos expériences, les extraits bactériens se sont avérés, dans leurs conditions expérimentales, compatibles avec leurs pièges [326].

Des tests plus approfondis visant à optimiser les conditions de préparation du piège et de son exposition aux extraits devraient être réalisés pour identifier les paramètres permettant le piégeage des protéines d'*E.coli*. Il est très probable que ces conditions seraient très différentes de celles utilisées pour les extraits protéiques humains, notamment parce que les extraits humains sont de nature nucléaire, tandis que les extraits bactériens sont « totaux ». Ce changement de composition des extraits et la présence des protéines cytosoliques est susceptible de modifier le comportement des enzymes.

1.1.4. Utilisation des pièges de type concatémères

Lors de l'optimisation de notre piège à protéines, nous avons conçu des concatémères afin de s'affranchir du problème des sondes oligonucléotidiques, trop courtes et ne contenant pas suffisamment de lésions. Les concatémères sont composés d'une séquence endommagée ou pas,

répétée plusieurs fois au sein de la même sonde. Ces séquences sont hybridées avec leurs brins complémentaires de manière à ce que les constituants de la sonde se chevauchent, et sont ensuite reliées à l'aide d'une ligase. Ainsi, une longue sonde composée d'une dizaine de séquences répétées est formée. Cette approche nous a permis d'améliorer la construction des pièges utilisés en y intégrant un plus grand nombre des dommages et en augmentant leur distance de la surface du support.

Différents types de sondes ont déjà été utilisés au sein de notre laboratoire ainsi que par d'autres chercheurs, qui ont permis de piéger les protéines reconnaissant des modifications de l'ADN de natures variées. Des exemples des sondes utilisées ainsi que des protéines qui ont été identifiées sont présentés dans le Tableau 17. Le choix de la sonde est souvent guidé par le type de dommage à étudier et par la facilité de son obtention. Ainsi, pour les lésions synthétisées chimiquement, les sondes oligonucléotidiques s'imposent, tandis que des sondes plasmidiques peuvent être utilisées pour étudier des panoplies de dommages générées par l'exposition à un agent génotoxique.

Tableau 17 : Exemples des sondes utilisées pour le piégeage des protéines reconnaissant différentes modifications de l'ADN, et les protéines identifiées avec ces approches. 5-Me-dC = 5-méthyl-désoxycytosine, 5-OH-Me-dC = 5-hydroxyméthyl-désoxycytosine. Pour une revue complète : [269]

Type de sondes	Modification de l'ADN	Exemples des protéines piégées	Référence
plasmide linéarisé de 3 kpb	adduits du cisplatine, de l'oxaliplatine et du satraplatine	HMGB1, UBF, TOX4, TRF2, PNUTS	[268,269]
oligonucléotides de 25 ou 90 pb	adduits du cisplatine et de l'oxaliplatine	protéines à boîte HMG, DDB1, RPA, PARP1, Ku70/Ku80, DNA-PKcs, XRCC1	[327–330]
oligonucléotides de 58 pb contenant un bras photoclivable	cassures double-brin	DNA-PKcs, Ku70/Ku80, PARP1	[300]
sondes synthétisées par PCR de 161 pb	modification épigénétiques de la cytosine : 5-Me-dC et 5-OH-Me-dC	ZBTB2	[publication en cours]

L'utilisation des plasmides linéarisés est la plus facile à mettre en œuvre et elle permet de concevoir les sondes contenant des dizaines de lésions par sonde. Cette approche n'est pourtant pas applicable à l'analyse de l'interactome des cyclonucléosides. Ces dommages complexes ne sont pas les seuls générés par le rayonnement ionisant et l'exposition d'un plasmide à ce stress

formerait une panoplie des lésions, comprenant des lésions simples d'oxydation, des cassures de l'ADN, ainsi que des lésions complexes. Le système des concatémères permet alors d'obtenir les avantages des sondes plasmidiques, soit un grand nombre de lésions, tout en préservant le format unique des dommages étudiés.

1.2. Pertinence et analyse des protéines sélectionnées

Malgré l'utilisation des différentes conditions de piégeage, notamment en termes de temps d'incubation et de composition du tampon, nous n'avons pas réussi à trouver de protéines qui reconnaissent les cyclonucléosides. De nombreuses explications à ce phénomène peuvent alors être avancées.

La première est de nature technique : les paramètres choisis pour effectuer le piégeage ne seraient pas favorables à ces protéines. Les interactions entre les éventuelles protéines de reconnaissance et les cyclonucléosides pourraient être de nature transitoire, et c'est pour ce fait que leur piégeage n'a pas été réalisé. Ainsi, une nouvelle mise au point du piège serait nécessaire, concernant notamment les conditions de l'incubation, afin que d'autres protéines puissent être piégées et identifiées. Il existe cependant d'autres explications possibles de cette absence apparente des protéines capables de reconnaitre les cyclonucléosides. Comme cela a été présenté dans la partie « Analyse bibliographique », il n'y a pas d'unanimité parmi les chercheurs en ce qui concerne l'abondance de ces dommages dans les cellules, que ce soit au niveau basal ou après l'induction d'un stress oxydant. Ainsi, si les cyclonucléosides ne sont pas des lésions courantes dans les cellules eucaryotes, il est possible qu'aucun système permettant leur prise en charge ne soit présent, ce qui expliquerait l'inexistence des protéines interagissant avec elles.

L'interactome des cyclonucléosides est censé contenir les acteurs du NER [208, 211]. Le piégeage des protéines du NER sur différentes sondes lésées a déjà été effectué par d'autres chercheurs, qui ont conçu des stratégies particulières permettant l'identification de ces interactions. Une trentaine de protéines sont impliquées dans la prise en charge des lésions volumineuses, leurs interactions avec le dommage étant transitoires et arrivant les unes à la suite des autres [331–335]. Afin de les immobiliser sur les sondes, une technique de photocrosslinking a été utilisée. Elle nécessite le remplacement des substrats classiques du NER par leurs équivalents photoactifs, tels que les dérivés arylazides, dans le but de fixer ensuite la protéine reconnaissant cette modification [328,331]. Cette stratégie possède néanmoins des inconvénients, car le photocrosslinking d'une protéine n'est possible que quand celle-ci rentre en proche contact

avec son substrat. Les enzymes du NER reconnaissent souvent une déformation de la structure générée par la présence de la lésion [336], il est alors possible que la protéine d'intérêt soit localisée trop loin du groupement photoactif pour que son immobilisation puisse avoir lieu.

Même la stratégie de photocrosslinking ne semble pas permettre le piégeage de toutes les protéines du système NER, car aucune des équipes citées ci-dessus n'ont réussi à piéger l'enzyme XPC, qui est indispensable à la reconnaissance des dommages volumineux [328,331]. Ceci montre la difficulté à piéger les protéines du NER et pourrait expliquer que notre système, dépourvu des groupements impliqués dans le crosslinking, ne permette pas non plus l'identification de ces enzymes.

Bien que nous n'ayons identifié aucune protéine capable de reconnaître et éventuellement participer à la prise en charge des cyclonucléosides, nous avons déterminé deux protéines dont l'association à l'ADN est empêchée par la présence de ces dommages.

1.2.1. PARP1

PARP1 est une protéine multifonctionnelle, impliquée dans de nombreux processus cellulaires, tels que la régulation de l'expression des gènes, les modifications de la chromatine ou la signalisation dans la mort cellulaire [309–311]. Son implication dans le système BER est connue depuis longtemps, et il a également été démontré que cette protéine jouait un rôle dans la prise en charge des cassures simple- ou double-brin [309].

Avec le système de pièges, nous avons retrouvé PARP1 majoritairement sur les sondes non-lésées, en constatant sa moindre abondance sur les sondes portant un cyclonucléoside. Nous avons, dans un premier temps, voulu comprendre la raison de la reconnaissance des sondes non-endommagées par cette protéine. Il est connu que PARP1 reconnaît les transitions entre l'ADN simple-brin et double-brin, comme il a été démontré par Jorgenen et *al.* Dans leur étude, ces auteurs ont comparé l'affinité de PARP1 pour différentes structures de l'ADN imitant les extrémités non-protégées des cassures de l'ADN ou les transitions simple brin/double brin. Ainsi, PARP1 semble avoir une meilleure affinité pour les extrémités non-protégées imitant les cassures double-brin (K_d = 0.49 nM) que pour les structures du type « épingle à cheveux » ou boucle (K_d = 4.31 nM et 3.50 nM) [337]. Une autre étude récente indique l'affinité de PARP1 pour d'autres structures de l'ADN, notamment pour les cassures simple-brin, qui seraient reconnues 3 à 5 fois plus efficacement que les cassures double-brin. Une affinité de PARP1 pour les séquence de l'ADN possédant des répétitions A :T a aussi été démontrée [309]. D'après les

auteurs, des fragments poly(A :T) de plus de quatre motifs pourraient conférer à l'ADN une flexibilité accrue, comparable à celle qui est due à la présence d'une cassure simple-brin.

A ce jour, aucune indication d'une reconnaissance de simples séquences d'ADN double-brin par PARP1 n'a été publiée. La possibilité de la reconnaissance de séquences double-brin non-endommagées a même été exclue par certains chercheurs [338], ce qui nous fait nous interroger sur le mécanisme par lequel nous avons réussi à piéger cette enzyme. La séquence que nous avons utilisée pour concevoir nos sondes ne contient pas de répétitions A :T. Nous pourrions supposer que la ligation des concatémères qui forment nos sondes ne soit pas effectuée à 100%, ce qui expliquerait que des cassures simple-brin y seraient présentes. Ainsi, PARP1 pourrait reconnaître celles-ci et non pas la séquence non-lésée. Bien que nos protocoles de formation des concatémères aient été optimisés au mieux, nous ne pouvons pas écarter l'hypothèse de la ligation imparfaite des sondes. De plus, toutes nos sondes, qu'elles portent ou non une lésion, possèdent des extrémités simple-brin plus ou moins longues (6 ou 16 bases) en fonction de l'oligonucléotide qui termine leur séquence. Ce modèle a été testé par Kramer en tant qu'imitation d'une cassure double-brin différente des extrémités franches, et il a été observé que l'affinité de PARP1 serait similaire pour ces deux types d'extrémités [309]. Nous supposons alors que, dans notre cas, PARP1 pourrait reconnaître soit les cassures au niveau des sondes, si elles existent, soit les extrémités non-appariées des sondes.

Bien que ces imperfections des sondes décrites ci-dessus doivent exister également dans les sondes endommagées, sur ces deuxièmes nous ne détectons plus PARP1 après 20 minutes d'interactions. La présence des cyclonucléosides semble être aussi un obstacle important à leur reconnaissance par PARP1. Les interactions de PARP1 avec l'ADN sont effectuées par l'intermédiaire de deux domaines à doigt de zinc qui s'insèrent dans le sillon majeur de l'ADN [339]. Il a été démontré que les cyclonucléosides généraient une distorsion importante dans l'hélice de l'ADN [197,206]. Ainsi, notre première hypothèse était que cette distorsion pourrait rendre impossible les interactions entre PARP1 et l'ADN.

Afin d'élucider la manière dont PARP1 reconnaît nos sondes, nous avons effectué les analyses supplémentaires. Le piégeage dans les extraits protéiques à différents temps d'incubation : 5, 10 ou 20 minutes, nous a permis d'établir une cinétique d'interactions entre PARP1 et les sondes lésées ou non-lésées. Nous avons vu que, dans un premier temps, les sondes portant une lésion étaient reconnues par PARP1 au même degré que les sondes non-lésées. Ces interactions, visibles à 5 minutes, deviennent plus faibles à 10 minutes d'incubation et ne sont plus détectables à 20 minutes. Ainsi, nous pensons que la présence des cyclonucléosides

n'empêche pas les interactions initiales entre PARP1 et nos sondes, mais qu'elle favorise plutôt le départ de l'enzyme après cette reconnaissance.

La fixation de PARP1 sur les dommages de l'ADN est transitoire, car son activation et son auto-modification sont des évènements rapides, survenant après son interaction avec la lésion. L'auto-modification de PARP1 consiste en sa poly(ADP-ribos)ylation (PARylation), c'est-à-dire en l'attachement à sa surface de longs polymères d'ADP-ribose généré à partir de molécules de NAD^+ [309]. La protéine PAR-ylée possède une charge négative importante qui fait qu'elle se dissocie de l'ADN étant lui aussi chargé négativement [340–342]. L'activation de PARP1 sert aussi à attirer les protéines qui vont poursuivre le processus de la réparation de la lésion reconnue. De plus, PARP1 modifie également d'autres protéines, telles que les histones, auxquels la PARylation confère une charge négative, affaiblissant leurs interactions avec l'ADN et rendant ainsi les lésions plus disponibles aux enzymes de réparation [343,344]. Si PARP1 est attiré sur nos sondes par les cassures, nous devrions alors observer son départ des sondes non-lésées dès que d'autres protéines de réparation ont pris en charge ces dommages. Néanmoins, nous n'avons pas de preuve que ce sont ces dommages qui sont effectivement reconnus par PARP1, et nous ne pouvons pas expliquer le manque d'activation de cette protéine senseur de dommages sur les sondes ne contenant pas de cyclonucléosides.

Nous pouvons supposer que, après la reconnaissance initiale des sondes lésées et non-lésées par PARP1, la présence des cyclonucléosides pourrait stimuler son activation et favoriser ainsi sa dissociation de l'ADN. Il a été démontré pour des sondes endommagées par le cisplatine, que leur reconnaissance par PARP1 était possible quand celui-ci était modérément modifié, mais que lorsqu'il était activé par la présence de grandes quantités d'ADP-ribose, sa fixation à la lésion n'était plus possible [330]. Ainsi, le schéma expliquant notre hypothèse est présenté sur la Figure 89.

Nous avons essayé de vérifier l'association de PARP1 aux sondes lésées ou non-lésées par la technique de retard sur gel. Ces expériences n'ont pas confirmé les résultats du piégeage, car les deux sondes sont reconnues par cette enzyme avec une affinité similaire. Au vu du mécanisme de l'association et de la dissociation de PARP1 de ses substrats, il est possible que cette technique ne soit pas adaptée à PARP1. Il a déjà été noté par Zhu et *al.* que PARP1 était en équilibre avec la protéine PARG, ces deux partenaires régulant le degré de modification de PARP1 [330]. Le manque de PARG dans les expériences de retard sur gel, ainsi que l'absence de NAD^+ nécessaire pour la formation des polymères d'ADP-ribose, expliqueraient ainsi que les différences de la reconnaissance des sondes lésées et non-lésées ne soient pas observées sur gel.

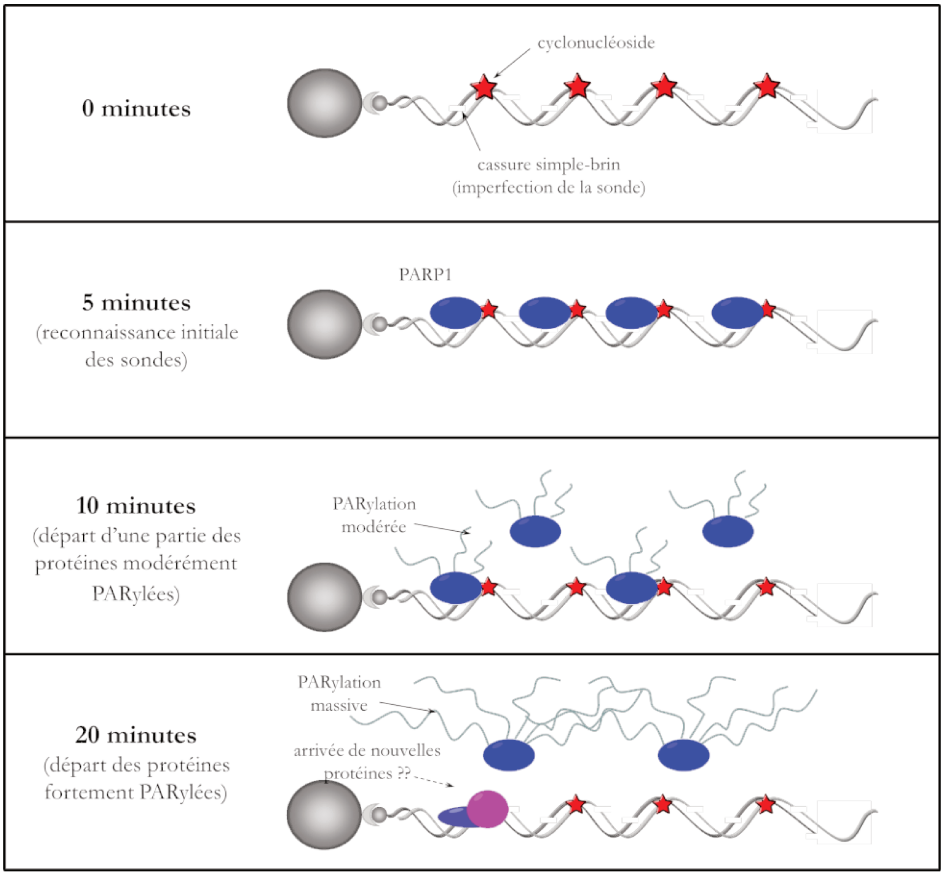


Figure 89 : Schéma présentant l'hypothèse de la reconnaissance des sondes par PARP1 et son départ consécutif dans le temps.

1.2.2. DREF

DREF est une deuxième protéine que nous avons piégée sur nos sondes non-lésées, mais pas sur les sondes portant un cyclonucléoside. Contrairement à PARP1, nous pouvons expliquer la raison de la fixation de ce facteur de transcription sur nos pièges : DREF reconnaît une séquence consensus qui est présente dans nos sondes. Ses interactions sont alors séquence-dépendantes, ce que nous avons prouvé par des expériences en gel retard, qui ont montré une absence des interactions entre DREF et une séquence « scramble » de la même composition en nucléosides.

Un effet inhibiteur des cyclonucléosides sur la fixation des facteurs de transcriptio a déjà été démontré en 2002 par l'équipe de Brooks. Ils ont étudié l'influence d'une cyclo-dA sur la fixation de la protéine TBP (TATA Binding Protein) à sa séquence cible qui est la boîte TATA. Lors d'un premier test *in vitro*, ils ont effectué des expériences de shift et de supershift assay qui ont démontré que la présence d'une cyclo-dA, que ce soit sur le brin transcrit ou non-transcrit, diminuait la fixation de TBP de plus de 95%. Dans des expériences effectuées *in vivo*, ces

chercheurs ont introduit la même séquence TATA contenant une lésion, dans le promoteur CMV qui régissait l'expression du gène de luciférase, et le tout a été transfecté dans les fibroblastes XPA, incapables de réparer les cyclonucléosides. La réduction de l'expression de la luciférase a été de 70% environ à 18 et 48 heures après la transfection, ce qui indique l'effet relativement stable de cette inhibition. L'explication possible de cet effet pourrait venir d'une étude cristallographique, qui indique que la fixation de TBP à la boîte TATA s'effectue dans le sillon mineur et qu'elle engendre une déformation de l'hélice [345] – cette déformation pourrait être rendue impossible par la présence d'un cyclonucléoside. L'activité résiduelle de 20-25% qui a été constatée, était probablement due à la transcription indépendante de la boîte TATA [346]. Dans une étude ultérieure, les chercheurs du même laboratoire ont également démontré que la présence d'une cyclo-dA dans la séquence cible diminuait la fixation *in vitro* des facteurs de transcription séquence-spécifiques : HSF1, CREB et NF-kappaB de plus de 90%, et du facteur HMGA d'approximativement 30% [347]. Ces travaux sont importants du fait que les cyclonucléosides, comme toute autre lésion d'oxydation, seraient plus concentrés dans les régions promotrices, suite à la formation augmentée du peroxyde d'hydrogène par l'histone déméthylase LSD1, qui est recrutée lors de l'activation de l'expression de gène [347–349].

Les résultats que nous avons obtenus sur la protéine DREF sont cohérents avec les études citées ci-dessus. L'inhibition de sa fixation sur sa séquence cible pourrait ainsi avoir un effet inhibiteur sur son rôle dans les cellules. Les cyclonucléosides qui sont introduits dans nos sondes ne sont pas situés dans la séquence consensus. La distorsion de l'hélice qu'ils génèrent à leur proximité serait néanmoins suffisamment importante pour empêcher la reconnaissance de sa séquence cible par DREF.

L'homologue de DREF chez *Drosophila* est un facteur nécessaire pour l'expression des gènes liés à la réplication, et il est positivement régulé dans les processus de croissance et de prolifération cellulaires [350–356]. HDREF pourrait jouer le même rôle dans les cellules humaines, car sa séquence cible a été identifiée dans les régions promotrices des gènes liés à la prolifération et à la régulation du cycle cellulaire [350]. Si cette fonction de hDREF est confirmée, la régulation de cycle cellulaire par hDREF pourrait être perturbée en présence des cyclonucléosides. Alternativement, ceci pourrait être considéré comme un système de protection contre les lésions irréparables.

Les cyclonucléosides sont capables de bloquer l'action de nombreuses enzymes, telles que les DNAses et les polymérases [209]. De plus, la transcription des séquences contenant un cyclonucléoside, par les polymérases qui ne sont pas inhibées, résulte en la formation de transcrits mutés [213]. De ce fait, les cyclonucléosides sont des lésions de l'ADN hautement mutagènes.

L'arrêt du cycle cellulaire et le blocage de la prolifération pourraient ainsi être considérés comme bénéfiques pour les cellules endommagées, car ils permettraient à l'organisme de limiter les dégâts causés par ces dommages et éviter l'apparition des mutations, menant à long terme à la cancérogénèse.

Les dommages de l'ADN n'empêchent pas toujours la fixation des facteurs de transcription à leurs séquences cibles. Par exemple, l'équipe de Lippard a réussi à piéger sur les sondes modifiées par le cisplatine le facteur de transcription YB-1. Ce facteur interagit avec la protéine PCNA (proliferating-cell nuclear antigen), qui est impliquée dans le NER, et sa fixation sur les dommages générés par le platine pourrait ainsi avoir une influence sur leur réparation [328]. Un autre facteur de transcription reconnaissant les lésions formées par les dérivés du platine est la TBP (TATA binding protein). Il a été montré par retard sur gel que son affinité pour les dommages platinylés était supérieure à celle pour la boîte TATA qui est sa cible naturelle. Ainsi, le détournement de la TBP de son rôle habituel par les lésions platinylées serait impliqué dans le mécanisme de toxicité du cisplatine [357].

Ces résultats montrent que chaque lésion de l'ADN peut avoir un effet distinct et spécifique sur le fonctionnement de la cellule. Bien que les cyclonucléosides et les lésions générées par le cisplatine soient pris en charge par le même système de réparation NER, leur reconnaissance par les protéines et leur implication dans les processus cellulaires s'avèrent très différentes. Les premiers empêcheraient la fixation des facteurs de transcription à leur cible, tandis que les deuxièmes détourneraient certains de leur fonction initiale.

2. Travail sur les protéines isolées

2.1. Modèle bactérien : glycosylase Fpg d'*Escherichia coli*

Comme nous l'avons présenté dans le chapitre « Stratégie et objectifs », nous avons choisi la protéine Fpg pour nos travaux, car c'est un modèle couramment étudié et très bien caractérisé pour ses interactions avec les lésions simples d'oxydation. Sa reconnaissance des lésions complexes d'oxydation n'a pas été élucidée, mais des résultats préliminaires intéressants concernant ses interactions avec les cyclonucléosides ont été obtenus il y a quelques années par les membres de notre équipe.

2.1.1. Interactions de la Fpg avec l'ADN non-lésé

Avant d'analyser les interactions qui ont lieu entre la Fpg et les cyclonucléosides, nous allons nous pencher sur ses interactions avec l'ADN non-lésé. Les expériences réalisées avec les sondes non-endommagées visaient à trouver une explication biologique aux premiers résultats obtenus par EMSA. Le coefficient de Hill, qui est un des paramètres de l'équation dont nous nous sommes servis pour calculer la constante de dissociation de la Fpg, est représentatif de la coopérativité entre les protéines. Il peut être également lié au nombre de complexes qui sont formés. La valeur que nous avons obtenue suite à nos expériences suggérait l'existence d'une telle coopérativité.

Nous avons déterminé le nombre de complexes non-spécifiques qui sont formés entre la protéine Fpg et une séquence de longueur variable. Nous avons constaté qu'une molécule de Fpg est capable de se fixer sur l'ADN toutes les dix bases environ. Nous avons observé également que la valeur du coéfficient de Hill « n » était supérieure au nombre de complexes visibles sur le gel : elle correspond approximativement à 2^x (où x = le nombre de complexes visibles sur le gel), une relation que nous ne pouvons pas expliquer sur le plan biologique.

Nous avons aussi constaté que plus les sondes étaient longues, plus la pente de la courbe de formation des complexes était importante. Ainsi, pour les sondes longues, même si plus de complexes sont formés au final, leur formation est plus rapide et les sondes sont consommées plus rapidement que dans le cas des sondes courtes. Il semblerait alors qu'il existe un effet de coopérativité entre les molécules de Fpg et que la reconnaissance initiale effectuée par une protéine faciliterait les interactions avec les molécules suivantes. De nombreux mécanismes pourraient jouer un rôle dans cette coopérativité. La première Fpg pourrait modifier la structure de la sonde, ce qui faciliterait par la suite sa reconnaissance par d'autres molécules de Fpg. Alternativement, les enzymes pourraient interagir non seulement avec l'ADN, mais également avec la protéine située à proximité, et qui est impliquée dans la reconnaissance initiale.

L'équation de Hill est majoritairement utilisée pour décrire les phénomènes de coopérativité et d'interactions allostériques dans le cas de la reconnaissance d'un ligand par une protéine. Le mécanisme de la coopérativité a été décrit pour la première fois pour le cas de l'hémoglobine, qui peut fixer sur elle jusqu'à 4 molécules d'oxygène. Sa valeur du coefficient de Hill est pourtant calculée à 3, ce qui suggérerait plutôt la fixation de 3 molécules d'oxygène sur une molécule d'hémoglobine. Ceci montre les limites de l'interprétation du coefficient de Hill, qui ne peut pas décrire précisément le mécanisme des interactions entre un ligand et son récepteur, ce

coefficient reste cependant un indice très utile de l'existence des phénomènes de coopérativité [358].

Nous avons constaté également qu'une molécule de Fpg se fixait de manière non-spécifique sur l'ADN toutes les 10 bases environ. Ceci est en accord avec des observations faites par d'autres équipes. Notamment, les oligonucléotides de longueur inférieure à 8 bases ne semblent pas capables de former les complexes avec la Fpg, et des changements d'affinité de la Fpg pour une séquence d'ADN seraient observés pour les longueurs des sondes comprises entre 8 et 13 bases [359]. Les études effectuées par la technique de footprinting ont montré qu'une molécule de Fpg protège 9 à 10 nucléotides d'une séquence simple-brin de l'ADN [360,361]. Ainsi, nos résultats confirment que 10 bases environ seraient à la fois nécessaires et suffisantes pour la fixation d'une molécule de Fpg.

2.1.2. Interactions de la Fpg avec les cyclonucléosides

Une série d'expériences effectuées avec la protéine Fpg purifiée d'*E.coli* et les oligonucléotides contenant des cyclonucléosides, nous a permis d'observer les interactions qui existent entre ces deux molécules. Nous avons montré que l'affinité de la Fpg pour les quatre cyclonucléosides : 5'S-cyclo-dA, 5'R-cyclo-dA, 5'S-cyclo-dG et 5'S-cyclo-dHT, était supérieure à celle pour l'ADN non-lésé, mais inférieure à celle pour le THF.

Plusieurs études ont été effectuées antérieurement au sein de notre laboratoire, dans le but d'analyser les interactions entre cette glycosylase et différentes lésions de l'ADN. L'une d'elles, réalisée avec la technique du retard sur gel, a montré que la Fpg d'*Escherichia coli* n'était pas capable de reconnaitre les cyclonucléosides, contrairement à son homologue provenant de *Lactococcus lactis*. Ce dernier formait des complexes non-covalents avec les oligonucléotides contenant une cyclo-dA, une cyclo-dG ou une cyclo-dHT (les deux isoformes de chaque lésion ont été étudiés), mais sans reconnaître la séquence non-endommagée. Pour la Fpg d'*E.coli*, aucune formation des complexes n'a été observée [362]. Néanmoins, des analyses ultérieures ont démontré que les deux protéines Fpg originaires d'*E.coli* ou *L.lactis*, pouvaient former les complexes avec les cyclopurines ou les cyclopyrimidines [résultats non-publiés].

Une étude récente effectuée par l'équipe de Basu démontre, à son tour, l'absence d'interactions entre les cyclopurines et des enzymes du BER bactériennes : Fpg, Endo III, Endo V et Endo VIII ou humaines : OGG1, NEIL1 et NEIL2 [363]. Tous ces résultats ont été

obtenus grâce à l'utilisation de la même technique de retard sur gel et avec de courts oligonucléotides synthétiques. La longueur exacte des sondes et le placement précis de la lésion au sein de la séquence varient entre les études, ce qui pourrait avoir une influence sur les interactions observées, si celles-ci étaient séquence-dépendantes. Cette dépendance n'a pas à ce jour été démontrée, mais nous ne pouvons pas éliminer cette hypothèse qui permettrait d'expliquer les divergences constatées. Une autre explication possible est basée sur les différentes conditions d'incubation, qui peuvent, comme nous l'avons remarqué lors de la mise au point des conditions expérimentales, avoir un rôle primordial sur la reconnaissance de la sonde par une protéine donnée. Toutes les études citées ci-dessus sont uniquement qualitatives et ont été réalisées avec une seule concentration de protéine, nous ne pouvons pas exclure un choix inadapté des conditions, ce qui aurait conduit à des conclusions erronées. Notre étude a été effectuée avec une large gamme de concentrations de Fpg, ce qui nous a permis de voir l'évolution de la formation des complexes et ce qui a enlevé le biais lié à la quantité inadaptée des sondes et des protéines. Aussi, grâce à la mise au point réalisée au préalable, nous avons pu choisir les conditions les plus favorables pour la reconnaissance des cyclonucléosides par la Fpg. Nous mettons ainsi en évidence l'importance de l'optimisation effectuée en amont de chaque expérience visant à analyser les interactions protéines−ADN, et le rôle crucial que jouent les paramètres choisis dans les résultats obtenus et dans les conclusions.

Une autre étude basée sur la technique SPRi, effectuée également au sein de notre laboratoire et dont les résultats ont été publiés en 2008, a montré que la Fpg d'*E.coli* était capable de reconnaître une 5'S-cyclo-dA insérée dans une sonde oligonucléotidique, sans effectuer de coupure. Ainsi, cette lésion a été comparée avec une lésion simple d'oxydation, la 8-oxo-dG. Sur cette dernière, l'activité de coupure exercée par la Fpg était incontestable. Il a été postulé que les cyclonucléosides pourraient agir en tant que pièges à protéines, en détournant la Fpg de son activité habituelle qui est la réparation des lésions simples d'oxydation, et qu'ils seraient néfastes pour le fonctionnement correct de tout le système BER [281]. Nos expériences étendent cette hypothèse aux autres cyclonucléosides, de type cyclopurines et même les cyclopyrimidines, dont nous avons testé un représentant : 5'S-cyclo-dHT. Ainsi, en plus d'être difficilement réparables par les systèmes cellulaires et de bloquer les processus tels que la transcription de l'ADN, les cyclonucléosides pourraient exercer leur toxicité via une inhibition partielle du système de réparation BER.

Un détournement de glycosylase similaire a été observé il y a quelques années pour une autre lésion : N^7-benzyl-Fapy-dG (Bz-Fapy-dG). Ce dommage volumineux est reconnu par la

Fpg avec la même affinité que la lésion Fapy-dG, mais, contrairement à ce deuxième qui est substrat pour la Fpg, il forme un complexe improductif avec cette enzyme. Ainsi, Bz-Fapy-dG agit en tant que piège qui détourne la Fpg de sa cible biologique. Une autre conséquence de ce phénomène a également été proposée. La formation d'un complexe abortif avec une glycosylase pourrait cacher le dommage et empêcher sa réparation par d'autres systèmes, comme par exemple le NER. D'après les auteurs, cette interférence des complexes abortifs impliquant les enzymes du BER avec le système NER pourrait provoquer une accumulation des dommages volumineux *in vivo* et mener, à long terme, à la cancérogénèse [364]. Nous ne pouvons pas exclure l'hypothèse qu'un mécanisme identique soit impliqué dans la prise en charge des cyclonucléosides. Une réparation difficile de ces lésions a été démontrée par le passé [208], et leur protection par des enzymes avec lesquelles elles formeraient des complexes improductifs pourrait ainsi expliquer ce phénomène. L'efficacité de la réparation des cyclonucléosides pourrait, dans le futur, être testée dans les cellules déficientes en système BER, afin de vérifier si l'absence des glycosylases serait susceptible d'améliorer la prise en charge de ces dommages complexes.

Les expériences de gel retard que nous avons effectuées nous ont permis non seulement d'obtenir une réponse qualitative sur la présence ou l'absence de complexes entre la Fpg d'*E.coli* et les cyclonucléosides, mais également de caractériser ces interactions et les comparer entre elles pour différentes lésions. La reconnaissance de nombreuses lésions simples de l'ADN par la Fpg a déjà été étudiée par différentes équipes de chercheurs, et une abondante littérature à ce sujet est disponible (Tableau 18). Néanmoins, il reste toujours difficile de confronter les résultats obtenus dans différentes conditions expérimentales, et surtout de comparer les valeurs de la constante de dissociation K_d calculées pour des interactions observées. La valeur de cette constante est dépendante de multiples facteurs, différents résultats peuvent ainsi être obtenus pour les mêmes interactions, en fonction des conditions expérimentales sélectionnées ou de la longueur et de la séquence de la sonde. Les valeurs de K_d obtenues par d'autres équipes sont globalement plus basses que les notres, ce qui peut s'expliquer par l'utilisation de différentes concentrations de sondes et, par la suite, des protéines. Aussi, nous pouvons constater que les variations de la base opposée à la lésion peuvent jouer un rôle important dans la reconnaissance de celle-ci. Dans nos expériences, la base située en face d'un cyclonucléoside a toujours été sa base complémentaire, tandis que le site abasique était localisé en face d'une cytosine. Il semblerait que l'affinité de Fpg pour THF : dC serait plus forte que celle pour THF :dA [320]. Il serait intéressant de vérifier l'influence des mésappariements sur l'affinité de la Fpg pour les cyclonucléosides.

Tableau 18 : Valeurs de la constante de dissociation (K_d) de la Fpg pour différentes lésions, rapportées dans la littérature, obtenues avec la technique EMSA. 8-oxo-dI = 8-oxo-désoxyinosine, 8-oxo-dN = 8-oxo-désoxygébularine, redAP = site abasique réduit.

lésion	**K_d [nM]**	**référence**
Fapy-dG : dC	6.3 ± 1.5	[365]
Fapy-dG : dA	128.5 ± 27.3	
8-oxo-dG	0.4 – 3.1 (en fonction de la localisation de la lésion)	[366]
redAP	0.3	[319]
1,3-propanediol	2.9	
cyclopentanol	3.0	
THF	25.0	
8-oxo-dG : dC	8.9 ± 2.1	[320]
8-oxo-dG : dT	29 ± 6	
8-oxo-dG : dG	21 ± 6	
8-oxo-dG : dA	340 ± 29	
8-oxo-dI : dC	10 ± 3	
8-oxo-dI : dA	310 ±36	
8-oxo-dN : dC	6.7 ± 1.3	
8-oxo-dN : dA	130 ± 13	
THF : dC	9.5 ± 2.8	
THF : dA	90 ± 20	
THF	190.32 ± 6.93	nos résultats
5'S-cyclo-dA	357.47 ± 11.17	
5'R-cyclo-dA	323.68 ± 9.21	
5'S-cyclo-dG	400.27 ± 10.83	
5'S-cyclo-dHT	299.18 ± 10.80	

Une des lésions les plus étudiées en termes de formation des complexes avec la Fpg est le site abasique. Les sites abasiques sont des substrats pour cette glycosylase bifonctionnelle, qui est capable de les inciser. Pour étudier l'affinité de la Fpg pour ces lésions sans en observer la coupure, deux stratégies peuvent être utilisées : soit le complexe transitoire formé entre le site abasique et la Fpg est figé d'une manière covalente à l'aide de borohydrure de sodium, soit un analogue inactif du site abasique, par exemple le THF, est utilisé à sa place. Pour nos expériences nous avons utilisé le THF en tant que témoin positif, afin de pouvoir comparer l'affinité de la Fpg pour cette lésion avec celle déterminée pour les cyclonucléosides. Nous avons pu constater que l'affinité de la Fpg pour les cyclonucléosides était moindre que pour les sites abasiques, mais qu'elle restait néanmoins plus forte que son affinité pour l'ADN non-lésé. Ainsi, bien qu'une partie de la Fpg puisse être détournée de sa fonction par les cyclonucléosides, une autre fraction

pourrait toujours reconnaître ses substrats principaux, pour lesquels elle a une plus grande affinité.

Nous avons également observé des différences d'affinité de la Fpg pour les quatre cyclonucléosides utilisés au cours de notre étude. Parmi les quatre, nous retrouvons trois cyclopurines et une seule cyclopyrimidine, et c'est cette dernière qui semble interagir avec la Fpg de manière la plus forte. Cette lésion, beaucoup moins étudiée que les cyclopurines, pourrait s'avérer ainsi encore plus cytotoxique que les cyclodésoxyadénosines et les cyclodésoxyguanosines. Néanmoins, aucune étude n'existe à ce jour qui indiquerait une présence de cyclopyrimidines dans l'ADN cellulaire soit au niveau basal, soit après l'induction d'un stress. Les résultats des dosages des cyclopurines ne font pas l'unanimité dans la communauté des chercheurs, et aucune extrapolation ne peut être effectuée sur la quantité des cyclopyrimidines dans les cellules. Si, cependant, ces lésions se révélaient présentes dans l'ADN *in vivo*, elles pourraient être plus toxiques que les cyclopurines.

Parmi les trois cyclopurines que nous avons utilisées pour nos études : 5'S- et 5'R-cyclo-dA, et 5'S-cyclo-dG, l'affinité de la Fpg pour la 5'R-cyclo-dA est la plus forte, suivie, dans l'ordre décroissant, par 5'S-cyclo-dA et 5'S-cyclo-dG. D'après d'autres études, réalisées avec des protéines humaines, le stéréoisomère 5'R semble le moins toxique, car il serait réparé plus facilement que le 5'S, et il bloquerait l'activité des autres enzymes à un moindre degré [209]. Nos résultats vont à l'encontre de ces études, car ils montrent que les stéréoisomères 5'S détourneraient la Fpg de son activité principale avec une moindre efficacité par rapport aux stéréoisomères 5'R. Ce ne sont néanmoins que des résultats obtenus *in vitro*, avec un modèle bactérien, et qui ne nous permettent pas de tirer les conclusions sur l'effet précis de ces dommages *in vivo*.

2.2. *Thermococcus gammatolerans*

Le dosage des lésions d'oxydation, à la fois simples et complexes, dans l'ADN de *Thermococcus gammatolerans* a été effectué afin de mieux connaître cet organisme, pour lequel peu de données sur les dommages de l'ADN sont disponibles. Ce dosage a ainsi été le premier réalisé pour ces organismes originaux. Nous avons également réalisé des tests d'activité des enzymes purifiées : tg1653 et tg1277, dont le but a été de vérifier les hypothèses sur leurs homologies avec

des enzymes bactériennes, avant d'étendre nos études aux tests d'affinité pour les cyclonucléosides.

2.2.1. Dosage des lésions d'oxydation

2.2.1.1. Dosage de la 8-oxo-dG

La lésion « phare » représentant les dommages d'oxydation formés dans l'ADN, est la 8-oxo-dG. Dans un premier temps nous avons effectué un dosage de cette lésion dans les organismes à l'état basal. Nous avons constaté que chez les organismes n'ayant pas subi de stress oxydant, le taux de ces lésions avoisinait 10 pour 1 million de nucléosides normaux, ce qui est supérieur aux niveaux retrouvés dans les cellules eucaryotes (Tableau 19). Ceci pourrait être dû aux conditions de vie de *Thermococcus gammatolerans*, notamment à sa température optimale de 85°C, qui pourrait favoriser la formation de ce dommage d'oxydation.

Nous avons également effectué le dosage de la 8-oxo-dG dans les organismes irradiés à 2.5 kGy ou 5 kGy, qui ont été récoltés immédiatement après le traitement ou remis en culture pendant 1 heure. Nous avons constaté que ce traitement induisait des quantités importantes de 8-oxo-dG, notamment à la plus forte dose d'irradiations. Les expériences effectuées par d'autres chercheurs avec les cellules eucaryotes, comme présenté dans le Tableau 19, ont été réalisées avec des doses beaucoup plus faibles de rayonnement, une induction de la 8-oxo-dG par ce stress peut néanmoins être constatée pour certains modèles. Les expériences menées précédemment dans notre laboratoire ont montré qu'une dose de 90 Gy de rayonnement γ multipliait la quantité de la 8-oxo-dG par 3 environ, par rapport à son niveau basal, dans les cellules THP1 [367]. Il est aussi intéressant de savoir que la quantité de la 8-oxo-dG générée varie en fonction de la source de rayonnement utilisée, ce qui a également été démontré par les chercheurs de notre équipe [368].

La 8-oxo-dG a aussi été dosée dans l'ADN de l'Archée *Halobacterium salinarum* exposée à 7.5 kGy de rayonnement ionisant. Cette dose très élevée n'induit pourtant pas de quantités importantes de 8-oxo-dG chez ces organismes, ce que les auteurs expliquent notamment par l'effet protecteur des concentration élevées en sel dans le milieu intracellulaire et par le ratio favorable Mn/Fe [322]. Les systèmes de réparation de ces dommages chez *Thermococcus gammatolerans* semblent être efficaces, car après une heure d'incubation dans les conditions optimales, le niveau de la 8-oxo-dG retrouve son taux normal. Même dans ses conditions de vie optimales, sans exposition à un stress supplémentaire, la présence d'un système BER efficace est

nécessaire pour la survie de *Thermococcus gammatolerans*, car la température élevée favorise la formation des dommages d'oxydation en grande quantité.

Tableau 19 : Taux de 8-oxo-dG formée à niveau basal ou après traitement avec des rayonnements γ dans différentes cellules. Les estimations de la quantité des lésions par million de nucléosides ont été obtenues par calcul, en admettant que la guanine constitue 25 % des bases dans l'ADN (à l'exception des valeurs pour *H.salinarium* qui ont été calculées par les auteurs pour 10^6 nucléosides)

Type de cellules	**Dose de rayonnement [Gy]**	**8-oxo-dG / 10^5 dG**	**8-oxo-dG / 10^6 nucléosides (estimation)**	**Référence**
Cellules de la rate de souris BDF1	0	0.38 ± 0.06	0.95	[369]
	1	0.49 ± 0.08	1.225	
	2	0.59 ± 0.05	1.475	
	10	0.56 ± 0.06	1.4	
Foie de rats Fischer 344/DuCrj	0 (rats de 6 mois)	0.430 ± 0.069	1.075	[370]
	5 (rats de 6 mois)	0.714 ± 0.042	1.785	
	0 (rats de 27 mois)	0.803 ± 0.089	2.0075	
	5 (rats de 27 mois)	2.140 ± 0.547	5.35	
Lymphocytes	0	1.33 ± 0.24	3.325	[371]
	0.05	1.42 ± 0.04	3.55	
	0.1	1.46 ± 0.20	3.65	
	0.2	1.55 ± 0.12	3.875	
	0.5	1.61 ± 0.20	4.025	
	2	1.34 ± 0.15	3.35	
	5	1.49 ± 0.13	3.725	
Halobacterium salinarum NRC-1	0		2.1	[322]
	7500		3.1	
Thermococcus gammatolerans	0		9.18 ± 0.92	nos résultats
	2500		11.60 ± 1.69	
	0 (contrôle de 2500)		5.09 ± 1.19	
	2500 (+ 1 heure d'incubation)		9.52 ± 1.23	
	0 (contrôle de 2500 + 1 heure d'incubation)		8.40 ± 1.43	
	5000		24.23 ± 7.95	
	0 (contrôle de 5000)		4.15 ± 0.51	
	5000 (+ 1 heure d'incubation		9.74 ± 2.92	
	0 (contrôle de 5000 + 1 heure d'incubation)		6.49 ± 1.11	

Il est intéressant de noter que dans les cellules non-irradiées, mais incubées sur la glace pendant le temps équivalent au temps d'irradiation, une diminution de la quantité de la 8-oxo-dG a été constatée. Sa quantité augmente ensuite, lorsque les cellules sont incubées à nouveau dans

leurs conditions optimales. Ce comportement renforce notre hypothèse, d'après laquelle le taux basal élevé de la 8-oxo-dG chez *Thermococcus gammatolerans* serait dû à la température. La diminution de la température pourrait ralentir leur métabolisme et, par la suite, réduire le stress oxydant endogène et la quantité des lésions formées par celui-ci. L'incubation sur la glace permet à cet organisme de réparer une partie des dommages et c'est pourquoi une diminution de leur niveau est observée chez les organismes contrôles. L'incubation à 85°C cause une égalisation du niveau de 8-oxo-dG chez les organismes irradiés et non-irradiés, ce qui indique que le taux retrouvé d'environ 10 lésions pour 1 million de bases serait le taux « naturel » chez *Thermococcus gammatolerans*.

2.2.1.2. Dosage des cyclonucléosides

Il n'y a pas d'unanimité parmi les chercheurs en ce qui concerne l'abondance de ces dommages dans les cellules eucaryotes, que ce soit au niveau basal ou après l'induction d'un stress oxydant. Ainsi, des informations très diverses peuvent être retrouvées dans la littérature, indiquant un taux basal allant des niveau indétectables jusqu'à 2 résidus de 5'S-cyclo-dG/10^6 paires de bases, ce qui est un niveau comparable à celui de la 8-oxo-dG, l'une des lésions d'oxydation la plus présente dans l'ADN [201]. D'autres articles indiquent un niveau basal plus bas, restant néanmoins non-négligeable, autour de 0.02-0.05 résidus de 5'S-cyclo-dA/10^6 nucléosides [250,254]. Ces valeurs semblent augmenter après l'induction d'un stress oxydant : l'équipe d'Evans montre ainsi que dans les lignées HCC 1937 et MCF-7 traitées avec H_2O_2, le niveau de la 5'S-cyclo-dA doublait en passant d'environ 0.02 à 0.04-0.05 lésions/10^6 nucléosides. Ces données restent en contradiction avec les résultats obtenus dans notre laboratoire, qui indiquent que les cyclonucléosides seraient présents à un taux beaucoup plus bas dans l'ADN, que ce soit au niveau basal ou après l'induction d'un stress oxydant, et que leur niveau se situerait en-dessous des seuils de détection, estimé à 0.1 lésion/10^9 nucléosides normaux [194].

La formation des cyclonucléosides dans l'ADN sous l'effet des rayonnements est influencée par la présence de l'oxygène [194]. Ainsi, *Thermococcus gammatolerans* qui est un organisme pouvant être soumis à des doses extrêmes de rayonnement ionisant et qui, de plus, est un anaérobie obligatoire, semblait être un modèle pertinent pour les études de ces dommages. D'après certains chercheurs, les cyclonucléosides peuvent être retrouvés dans l'ADN humain à taux basal, et leur quantité augmenterait suite à un stress oxydant [254], mais ces résultats n'ont

pas été confirmés par les expériences effectuées dans notre laboratoire. Ainsi, nous avons voulu vérifier si nous pouvions identifier et quantifier ces lésions dans les organismes vivant dans les conditions théoriquement idéales pour leur formation. Nous avons ainsi dosé les cyclodésoxyadénosines et les cyclodésoxyguanosines dans l'ADN extrait de *Thermococcus gammatolerans* non-irradiés ou irradiés à 5 kGy. Dans aucune de ces deux conditions nous n'avons pu détecter les cyclonucléosides. Nous excluons la possibilité d'un problème technique au niveau de la préparation des échantillons ou de l'analyse, car un témoin positif a été utilisé, dans lequel les cyclonucléosides ont pu être dosés, et la qualité de l'ADN analysé a été vérifiée par le dosage des nucléosides normaux. Nous en concluons que l'absence des cyclonucléosides dans nos échantillons n'est pas un artéfact, mais qu'elle reflète la réalité biologique.

Le fait que nous n'ayons pas détecté de cyclonucléosides chez *Thermococcus gammatolerans* non-irradiés n'est pas surprenant. Ces lésions complexes de l'ADN ne sont pas formées dans les cellules de façon spontanée aussi facilement que les lésions simples d'oxydation, comme par exemple la 8-oxo-dG. Le niveau basal de la 8-oxo-dG s'explique par l'activité endogène de toutes les cellules vivantes, qui génèrent en continu les ROS étant à l'origine de ces dommages. Une très faible quantité basale des cyclonucléosides, en dessous des limites de détection, pourrait ainsi confirmer le mécanisme plus compliqué de leur formation, et la nécessité d'un stress particulier pour qu'ils soient générés en grande quantité dans l'ADN. *T.gammatolerans*, bien qu'ils soient adaptés à la vie dans des conditions extrêmes, ne vont alors pas développer des dommages complexes de manière spontanée lorsqu'ils sont cultivés sans stress.

L'absence des cyclonucléosides chez les organismes soumis à 5 kGy de rayonnement ionisant pourrait prouver leur caractère extrêmement rare. Si même dans les conditions qui semblent être optimales pour la formation des cyclonucléosides, nous ne pouvons pas les détecter, ceci voudrait dire que leur présence *in vivo* est hautement questionnable, malgré une littérature abondante les concernant.

2.2.2. Analyse des protéines purifiées

2.2.2.1. Activité de coupure

Les deux protéines pour lesquelles nous avons effectué des tests d'activité de coupure, tg1653 et tg1277, ont été initialement identifiées, via la comparaison des séquences, comme enzymes participant au système BER et étant des homologues probables de la Fpg et de

l'Endo III bactériennes [283]. La recherche de lésions sur lesquelles ces deux protéines exercent leur activité a montré que leurs spectres d'activité ne se superposent pas exactement avec ceux des deux enzymes d'*E.coli*. Parmi les 11 substrats testés, tg1653 semble reconnaître uniquement la 8-oxo-dG, ce qui la rend plus restrictive que la Fpg, tandis que tg1277 semble active sur les mêmes substrats que la Fpg. De plus, cette deuxième enzyme est capable d'exercer une activité de coupure sur les lésions qui ne sont pas substrats pour l'Endo III, comme la 8-oxo-dG. Nous pouvons donc confirmer l'implication de ces deux protéines dans le système BER, cependant la désignation de leurs homologues chez les autres organismes ne peut pas être facilement effectuée. L'identification par spectrométrie de masse MALDI-TOF/MS du mécanisme par lequel tg1653 et tg1277 exercent leur activité d'excision, à savoir la β-élimination, les rapproche de l'Endo III, la seule différence étant l'hydratation de la liaison double au niveau du sucre observé pour tg1277. Cette étude a ainsi permis d'élucider et de caractériser ces deux enzymes issues d'un organisme original, participant dans la réparation des lésions simples de l'ADN.

Les tests d'activité de ces deux enzymes en fonction de la base opposée à la lésion ont montré que tg1653 était sensible aux mésappariements, car la coupure qu'elle exerce était considérablement plus faible quand une purine était localisée en face de la 8-oxo-dG. En plus d'être plus restrictive en termes du choix des substrats, cette enzyme présente également une plus grande sensibilité au contexte dans lequel se trouve la 8-oxo-dG. Au vu de la réparation très efficace de cette lésion dans l'ADN des organismes irradiés, il semble que tg1653 ne pourrait pas être la principale enzyme prenant en charge ces dommages d'oxydation. On peut alors supposer que d'autres glycosylases existent, qui pourraient exciser la 8-oxo-dG efficacement et indépendamment de la base opposée.

Nous ne pouvons cependant pas ignorer le fait que toutes nos expériences ont été effectuées à des températures ne dépassant pas les 65°C, à cause de la stabilité des substrats, qui se trouve diminuée à hautes températures. La température optimale de *Thermococcus gammatolerans* avoisine les 90°C et nous ne pouvons pas exclure qu'à ces températures élevées l'activité des glycosylase soit modifiée et que celles-ci deviennent actives sur d'autres substrats que ceux identifiés à basse température.

Des études récentes montrent l'existence d'une famille de 8-oxo-dG ADN glycosylases propres aux archées et distinctes des celles retrouvées chez les eucaryotes et chez les procaryotes, avec lesquelles elles ne présentent pas d'homologies de séquences. De plus, leur mécanisme de reconnaissance et d'interaction avec la 8-oxo-dG ne serait pas le même que dans leurs homologues eucaryotes et procaryotes. Le représentant de cette famille, appelé AGOG, a été

étudié chez *Pyrobaculum aerophilum*, une archée thermophile, mais il aurait également les homologues chez d'autres archées [372,373].

AGOG de *P.aerophilum* est une glycosylase bifonctionnelle, car elle possède en plus une activité AP-lyase. Ces deux activités semblent néanmoins être découplées, car après 5 minutes d'incubation, 80% de 8-oxo-dG mésappariées avec dG sont excisées, mais la coupure de 20% seulement de sites abasiques a été constatée. Cette enzyme est active sur la 8-oxo-dG quelle que soit sa base opposée, avec cependant une préférence pour la 8-oxo-dG :dG par rapport à la 8-oxo-dG :dC. Les auteurs émettent l'hypothèse selon laquelle les mésappariements 8-oxo-dG :dG seraient formés de manière préférentielle à des températures élevées [374]. Le mécanisme de coupure d'AGOG de *P.aerophilum* a également été élucidé : il s'agit d'une β-élimination. Ainsi, cette enzyme agit sur les mêmes substrats que la Fpg d'*E.coli*, c'est-à-dire la 8-oxo-dG (indépendamment de sa base opposée), mais avec le mécanisme de coupure de l'OGG1 humaine [372].

Nous avons obtenu le même type de résultats avec tg1277 de *T.gammatolerans*, qui est active sur la 8-oxo-dG quelle que soit la base localisée en face d'elle, son mécanisme de coupure étant également la β-élimination. Les systèmes de réparation de l'ADN des archées sont plus réduits par rapport aux systèmes eucaryotes et procaryotes, car ils possèdent uniquement les enzymes-clés, et de nombreuses protéines eucaryotes n'ont pas leurs homologues chez les archées [375]. Nous pourrions alors croire que tg1277 serait la glycosylase principale du BER chez *Thermococcus gammatolerans*.

2.2.2.2. Interactions avec les cyclonucléosides

Les interactions observées entre les deux enzymes de *Thermococcus gammatolerans* et les cyclonucléosides ont été confrontées avec celles que nous avons vues précédemment pour la Fpg. Nous avons remarqué que ces enzymes nécessitaient des conditions d'incubation différentes de celles qui ont été initialement optimisées pour la Fpg, ce qui rend impossible une comparaison précise de leurs affinités respectives pour les mêmes sondes. Néanmoins, pour le choix des conditions d'incubation nous avons procédé de la même manière que dans le cas de la Fpg, c'est-à-dire que nous avons retenu les paramètres pour lesquels les tg1653 et tg1277 présentaient des interactions avec la sonde THF, ces interactions étant ainsi démontrées pour la première fois. Les expériences effectuées dans ces conditions avec les sondes contenant les cyclonucléosides ont

montré que l'affinité des deux protéines provenant de *T.gammatolerans* était extrêmement faible pour ces lésions. Nous pouvons avancer deux hypothèses pour expliquer ceci.

D'abord, nous ne pouvons pas exclure la possibilité de conditions inadaptées aux interactions que nous souhaitions observer : celles-ci ont été optimisées pour la lésion THF, mais il est possible que, en fonction de l'activité et de la spécificité de l'enzyme, la reconnaissance des cyclonucléosides soit effectuée dans d'autres conditions. La température pourrait jouer ainsi un rôle particulièrement intéressant dans cette reconnaissance, car les interactions observées à température ambiante ne vont pas être obligatoirement les mêmes que celles qui ont lieu à 85°C. Afin d'analyser l'affinité des enzymes de *Thermococcus gammatolerans* pour les cyclonucléosides, il faudrait alors concevoir des sondes duplexes stables à haute température (plus longues et avec une proportion des cytosines et des guanines plus élevée, ou encore des sondes auto-complémentaires) et réaliser les expériences de reconnaissance à 80-90°C.

Alternativement, il serait possible que les cyclonucléosides n'exercent pas le même effet sur ces organismes que sur les cellules humaines. Ces dommages complexes sont considérés comme difficilement réparables et hautement toxiques pour les cellules eucaryotes, mais il n'est pas impossible que *Thermococcus gammatolerans* ait développé un mécanisme de protection contre leur formation, ou un système efficace de leur réparation.

La réparation efficace des lésions de l'ADN générées par le rayonnement ionisant est une condition nécessaire pour assurer la radiorésistance exceptionnelle qui est observée chez *Thermococcus gammatolerans*. Le mécanisme le plus souvent associé au phénotype radiorésistant est la réparation des cassures double-brin. Ces dommages extrêmement toxiques, dont un seul peut déclencher la mort de la cellule [181,182], doivent être réparés rapidement. Il a été démontré que *Thermococcus gammatolerans* était capable de reconstruire son génome fragmenté par 5 kGy de rayonnement en quelques heures seulement [297]. Néanmoins, les cassures double-brin, bien que prépondérantes, ne sont pas les seuls dommages formés au niveau de la cellule par le rayonnement. Il serait alors intéressant d'étudier comment les autres systèmes de réparation de l'ADN sont sollicités lorsqu'un endommagement massif de l'ADN survient, et s'ils sont plus efficaces chez les archées que chez les eucaryotes.

VI Conclusions et perspectives

Les travaux de thèse qui sont décrits dans ce manuscrit ont été divisés en deux grandes parties :

- la première consistait en un screening des extraits protéiques humains à la recherche de l'interactome des cyclonucléosides
- la deuxième visait à élucider les interactions existantes entre les protéines du BER provenant d'un modèle bactérien (Fpg d'*Escherichia coli*) ou d'une archée radiorésistante *Thermococcus gammatolerans*, avec des lésions simples ou complexes d'oxydation ; les dosages de ces lésions chez *T.gammatolerans* ont également été effectués afin de mieux caractériser le modèle choisi.

Le screening des protéines humaines dans les extraits nucléaires a été effectué à l'aide d'un système de pièges. Son adaptation a été nécessaire afin d'augmenter le nombre de lésions présentes sur les sondes. Pour cela, les sondes appelées concatémères ont été conçues, qui sont composées d'une courte séquence oligonucléotidique répétée une dizaine de fois. Ce nouveau format s'est avéré particulièrement utile pour travailler avec des dommages qui ne peuvent être intégrés dans les sondes que par une approche de synthèse chimique, ce qui est le cas des cyclonucléosides. Il nous a notamment permis d'intégrer une dizaine de ces lésions au sein des sondes de longueur estimée à 100 à 300 paires de bases. Les concatémères peuvent désormais être utilisés pour le piégeage des protéines sur d'autres lésions complexes, qui ne peuvent pas être obtenues simplement par exposition des sondes à un stress génotoxique. Un exemple de telles lésions est dCyd341, une lésion originale découverte au sein de notre laboratoire, qui est composée d'une cassure simple-brin et d'un pontage inter-brin, et dont des niveaux détectables ont été mesurés dans l'ADN [376].

L'approche technique servant à l'identification de l'interactome des lésions complexes, qui a été développée au cours de cette thèse, pourrait également être utilisée en association avec d'autres méthodes. Nous pouvons envisager d'introduire nos sondes sur la surface des puces qui vont être ensuite exposées aux extraits protéiques. Les protéines piégées pourraient être alors identifiées par spectrométrie de masse, directement sur la surface des puces.

En utilisant les concatémères contenant quatre cyclonucléosides : 5'S-cyclo-dA, 5'R-cyclo-dA, 5'S-cyclo-dG ou 5'S-cyclo-dHT, localisés sur des sondes distinctes, nous n'avons pas pu identifier de protéines capables de prendre en charge ces dommages particuliers de l'ADN. Ceci peut être dû à un choix de conditions d'incubation inadaptées aux enzymes de réparation, ce qui fait que leurs interactions n'ont pas pu être observées. Aussi, les enzymes du système NER semblent être difficiles à piéger à cause de leurs interactions souvent transitoires avec les dommages volumineux. C'est pour cette raison que certaines équipes qui souhaitent piéger et identifier ces facteurs ont développé des approches visant à remplacer les lésions par leurs homologues photoactifs et d'effectuer un crosslinking afin de piéger ces facteurs. Ceci pourrait expliquer l'absence des enzymes du NER dans les listes des protéines que nous avons identifiées sur nos pièges endommagés.

En contrepartie, nous avons identifié deux protéines capables de reconnaître nos sondes non-lésées, mais dont les interactions avec celles-ci disparaissent en présence des cyclonucléosides. La première, PARP1, est une protéine impliquée dans de nombreux processus cellulaires, notamment le BER et la réparation des cassures de l'ADN. Nous supposons alors que sa reconnaissance des séquences non-lésées serait due à la présence dans celles-ci des imperfections de structure, notamment de cassures de l'ADN, qui pourraient être générées suite à une action incomplète de la ligase lors de la formation des concatémères. Ces interactions semblaient être abolies par la présence des cyclonucléosides, mais nous avons montré que les sondes lésées sont reconnues par PARP1 au même degré que les sondes non-lésées. La dissociation de PARP1 des sondes endommagées pourrait être due à son auto-activation stimulée par les cyclonucléosides. La confirmation de cette hypothèse nécessite néanmoins des expériences supplémentaires. Ainsi, dans la suite de ces travaux, des tests de reconnaissance des sondes endommagées ou non-endommagées par PARP1, en présence ou absence de NAD^+, pourraient être envisagés. NAD^+ est nécessaire pour la formation des polymères d'ADP-ribose qui confèrent une charge négative à la protéine modifiée et facilitent ainsi sa dissociation de l'ADN. Son absence pourrait alors avoir un effet inhibiteur sur les interactions observées, et la présence de PARP1 dans ces conditions sur les sondes lésées pourrait valider notre hypothèse.

La deuxième protéine que nous avons identifiée sur nos sondes non-endommagées et dont les interactions avec celles-ci sont abolies par la présence des cyclonucléosides, est DREF. Ce facteur de transcription serait impliqué dans la prolifération cellulaire et nous l'avons piégé grâce à la présence de sa séquence cible dans nos sondes. L'inhibition par les cyclonucléosides de cette interaction avec sa séquence cible pourrait alors avoir un impact important sur le

fonctionnement de toute la machinerie cellulaire. Des expériences complémentaires qui pourraient être envisagées pour confirmer cette hypothèse pourraient comprendre l'introduction d'un cyclonucléoside dans la séquence promotrice reconnue par DREF et l'étude de son impact sur l'expression des gènes contrôlés par ce promoteur. Il serait aussi intéressant de vérifier, avec des sondes contenant un cyclonucléoside à différents endroits de la séquence, quelle est l'étendue de l'effet exercé par ces dommages, et si la distance entre leur localisation et la séquence cible joue sur la force de l'inhibition qu'ils exercent.

Le piège que nous avons mis au point a été appliqué aux extraits des cellules HeLa. Nous avons également essayé d'effectuer un piégeage au sein des extraits bactériens provenant d'*Escherichia coli*, mais sans succès. Il serait particulièrement intéressant de tester l'application de notre système aux extraits protéiques de *Thermococcus gammatolerans*. Si des protéines originales interagissant avec les cyclonucléosides ou avec d'autres lésions, telles que dCyd341, étaient identifiées dans cet organisme, ceci pourrait amener des informations importantes concernant les mécanismes de réparation de l'ADN développés par cette archée et expliquer, au moins en partie, sa radiorésistance exceptionnelle.

La deuxième partie des travaux de cette thèse concernait l'étude d'interactions entre enzymes isolées et lésions d'oxydation.

Les expériences de retard sur gel effectuées avec la protéine Fpg isolée d'*Escherichia coli* ont montré que cette glycosylase bifonctionnelle du système BER était capable de reconnaître ces cyclonucléosides, sans pour autant les exciser. Son affinité pour les cyclonucléosides est supérieure à celle pour l'ADN non-lésé, mais inférieure à celle pour le THF – un analogue stable de site abasique. Les cyclonucléosides pourraient être alors considérés comme des pièges à glycosylases, capables de détourner celles-ci de leur fonction habituelle et capables alors d'impacter négativement le fonctionnement du système BER. Afin de confirmer cette hypothèse, des tests visant à vérifier l'impact des cyclonucléosides sur la réparation des lésions simples d'oxydation pourraient être envisagés. L'introduction d'un plasmide lésé dans les cellules et l'étude de sa capacité à réparer la 8-oxo-dG induite par le stress oxydant serait une expérience à réaliser. L'accumulation de la 8-oxo-dG prouverait ainsi l'inhibition du BER par la présence des cyclonucléosides.

En plus d'analyser les interactions entre la Fpg et l'ADN lésé, nous avons également approfondi nos connaissances sur le mécanisme de reconnaissance de l'ADN non-lésé par cette

enzyme. Nous avons notamment observé qu'une molécule de Fpg pouvait se fixer de manière aspécifique toutes les 10 bases environ. En plus, un effet de coopérativité entre les molécules de cette protéine existe, car la fixation de la première Fpg au niveau de la sonde non-endommagée semble faciliter la reconnaissance de celle-ci par les Fpg suivantes.

Des études d'activité et d'affinité ont été effectuées sur deux protéines isolées de *Thermococcus gammatolerans* : tg1653 et tg1277. Ces enzymes participent au système BER, et nous avons analysé leur spectre d'activité vis-à-vis plusieurs substrats, formés de lésions simples et complexes d'oxydation. Tg1653 est plus restrictive, car elle est capable d'exciser un seul substrat parmi le spectre de lésions testées : la 8-oxo-dG, tandis que tg1277 possède un spectre d'activité plus large, étant capable d'exciser les dérivés de l'hydantoïne. En plus, tg1653 est sensible aux mésappariements et son activité de coupure se trouve diminuée lorqu'une purine est située en face de la 8-oxo-dG. Le mécanisme de coupure de ces deux enzymes, à savoir la β-élimination, a été élucidé avec la technique MALDI-TOF/MS.

Les tests d'affinité ont montré des faibles interactions entre les deux protéines de *Thermococcus gammatolerans* et les cyclonucléosides. Il faut néanmoins signaler que ces expériences ont été effectuées à température ambiante, qui n'est pas la température optimale de ces archées. Ainsi, des sondes duplexes résistantes à la chaleur pourraient être conçues, notamment par augmentation de leur taille et du taux de cytosines et de guanines, afin de réaliser les tests d'affinité à 85°C. Ces tests pourraient être effectués en utilisant d'autres techniques que le retard sur gel, par exemple l'étude des interactions en SPRi pourrait être envisagée.

Afin de mieux caractériser le modèle original utilisé dans ces études, à savoir *Thermococcus gammatolerans*, des dosages des lésions simples et complexes d'oxydation de l'ADN ont été réalisés. Nous avons mesuré le taux basal de 8-oxo-dG, la lésion simple d'oxydation la plus étudiée, ainsi que son induction chez les organismes irradiés à 2.5 ou 5 kGy de rayonnement γ. Le niveau de 8-oxo-dG en absence de stress est plus important que chez les eucaryotes (environ 10 lésions par million de nucléosides), et son augmentation suite à l'exposition au stress oxydant a été constatée. De plus, ce dommage semble être efficacement réparé, car son taux diminue pour revenir à un niveau initial au bout d'une heure d'incubation dans les conditions optimales.

Le dosage des cyclonucléosides chez *Thermococcus gammatolerans* n'a pas permis de détecter ces lésions, que ce soit chez les organismes contrôles, ou irradiés à 5 kGy. Ceci soulève des

questions sur la pertinence biologique de ces dommages, qui ne semblent pas être formés en grande quantité même à des doses élevées de rayonnement.

Nous pouvons envisager pour l'avenir le dosage d'autres lésions complexes générés dans l'ADN par le rayonnement ionisant, comme la dCyd341, afin d'identifier le spectre de dommages créés par ce stress et de mieux comprendre les mécanismes de radiorésistance de *Thermococcus gammatolerans*.

Les travaux sur *Thermococcus gammatolerans* s'inscrivent dans le cadre d'un projet ANR « Gammatolerans ». En plus du dosage des lésions dans l'ADN de ces organismes et de la recherche de l'interactome des lésions complexes, d'autres études sont envisagées. Le protéome de ces archées et ses modifications suite à une irradiation vont ainsi être analysés. L'étude de l'oxydation des protéines suite à l'exposition à des fortes doses d'irradiations fait également partie des perspectives de ce projet.

D'une façon plus générale, les études telles que celles effectuées lors de cette thèse, reflètent la complexité de l'interactome environnant l'ADN endommagé par des agents oxydants. L'étude de ces protéines est indispensable pour comprendre comment les cellules sont affectées ou comment elles se protègent vis-à-vis de ce stress.

Bibliographie

[1] W. Dröge, Free Radicals in the Physiological Control of Cell Function, Physiol. Rev. 82 (2002) 47–95.

[2] A. Barzilai, K.-I. Yamamoto, DNA damage responses to oxidative stress, DNA Repair. 3 (2004) 1109–1115.

[3] R.D. Bont, N. van Larebeke, Endogenous DNA damage in humans: a review of quantitative data, Mutagenesis. 19 (2004) 169–185.

[4] A.J. Kowaltowski, N.C. de Souza-Pinto, R.F. Castilho, A.E. Vercesi, Mitochondria and reactive oxygen species, Free Radic. Biol. Med. 47 (2009) 333–343.

[5] P.C. Hinkle, R.A. Butow, E. Racker, B. Chance, Partial resolution of the enzymes catalyzing oxidative phosphorylation. XV. Reverse electron transfer in the flavin-cytochrome beta region of the respiratory chain of beef heart submitochondrial particles, J. Biol. Chem. 242 (1967) 5169–5173.

[6] P.K. Jensen, Antimycin-insensitive oxidation of succinate and reduced nicotinamide-adenine dinucleotide in electron-transport particles. I. pH dependency and hydrogen peroxide formation, Biochim. Biophys. Acta. 122 (1966) 157–166.

[7] G. Loschen, A. Azzi, L. Flohé, Mitochondrial H2O2 formation: relationship with energy conservation, FEBS Lett. 33 (1973) 84–87.

[8] G. Barja, A. Herrero, Localization at complex I and mechanism of the higher free radical production of brain nonsynaptic mitochondria in the short-lived rat than in the longevous pigeon, J. Bioenerg. Biomembr. 30 (1998) 235–243.

[9] G. Barja, Mitochondrial oxygen radical generation and leak: sites of production in states 4 and 3, organ specificity, and relation to aging and longevity, J. Bioenerg. Biomembr. 31 (1999) 347–366.

[10] J.F. Turrens, A. Boveris, Generation of superoxide anion by the NADH dehydrogenase of bovine heart mitochondria, Biochem. J. 191 (1980) 421–427.

[11] J.F. Turrens, B.A. Freeman, J.G. Levitt, J.D. Crapo, The effect of hyperoxia on superoxide production by lung submitochondrial particles, Arch. Biochem. Biophys. 217 (1982) 401–410.

[12] T. Finkel, Signal transduction by reactive oxygen species, J. Cell Biol. 194 (2011) 7–15.

[13] J.F. Turrens, Mitochondrial formation of reactive oxygen species, J. Physiol. 552 (2003) 335–344.

[14] E. Cadenas, K.J.A. Davies, Mitochondrial free radical generation, oxidative stress, and aging, Free Radic. Biol. Med. 29 (2000) 222–230.

[15] A.L. Jackson, L.A. Loeb, The contribution of endogenous sources of DNA damage to the multiple mutations in cancer, Mutat. Res. Mol. Mech. Mutagen. 477 (2001) 7–21.

[16] J. Cadet, J.R. Wagner, DNA Base Damage by Reactive Oxygen Species, Oxidizing Agents, and UV Radiation, Cold Spring Harb. Perspect. Biol. 5 (2013).

[17] F.C. Fang, Antimicrobial reactive oxygen and nitrogen species: concepts and controversies, Nat. Rev. Microbiol. 2 (2004) 820–832.

[18] G.P. Bienert, J.K. Schjoerring, T.P. Jahn, Membrane transport of hydrogen peroxide, Biochim. Biophys. Acta BBA - Biomembr. 1758 (2006) 994–1003.

[19] J. Chandra, A. Samali, S. Orrenius, Triggering and modulation of apoptosis by oxidative stress, Free Radic. Biol. Med. 29 (2000) 323–333.

[20] R.P. Patel, J. McAndrew, H. Sellak, C.R. White, H. Jo, B.A. Freeman, et al., Biological aspects of reactive nitrogen species, Biochim. Biophys. Acta BBA - Bioenerg. 1411 (1999) 385–400.

[21] C. Nathan, Q.W. Xie, Nitric oxide synthases: roles, tolls, and controls, Cell. 78 (1994) 915–918.

[22] S. Moncada, E.A. Higgs, Endogenous nitric oxide: physiology, pathology and clinical relevance, Eur. J. Clin. Invest. 21 (1991) 361–374.
[23] R.G. Knowles, S. Moncada, Nitric oxide as a signal in blood vessels, Trends Biochem. Sci. 17 (1992) 399–402.
[24] D.J. Hawkins, B.O. Meyrick, J.J. Murray, Activation of guanylate cyclase and inhibition of platelet aggregation by endothelium-derived relaxing factor released from cultured cells, Biochim. Biophys. Acta. 969 (1988) 289–296.
[25] L.B. Adams, J.B. Hibbs Jr, R.R. Taintor, J.L. Krahenbuhl, Microbiostatic effect of murine-activated macrophages for Toxoplasma gondii. Role for synthesis of inorganic nitrogen oxides from L-arginine, J. Immunol. Baltim. Md 1950. 144 (1990) 2725–2729.
[26] D.J. Stuehr, C.F. Nathan, Nitric oxide. A macrophage product responsible for cytostasis and respiratory inhibition in tumor target cells, J. Exp. Med. 169 (1989) 1543–1555.
[27] J. Stadler, W.A. Schmalix, J. Doehmer, Inhibition of biotransformation by nitric oxide (NO) overproduction and toxic consequences, Toxicol. Lett. 82-83 (1995) 215–219.
[28] T. Nguyen, D. Brunson, C.L. Crespi, B.W. Penman, J.S. Wishnok, S.R. Tannenbaum, DNA damage and mutation in human cells exposed to nitric oxide in vitro, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 3030–3034.
[29] K. Dobashi, K. Pahan, A. Chahal, I. Singh, Modulation of endogenous antioxidant enzymes by nitric oxide in rat C6 glial cells, J. Neurochem. 68 (1997) 1896–1903.
[30] M. Asahi, J. Fujii, K. Suzuki, H.G. Seo, T. Kuzuya, M. Hori, et al., Inactivation of glutathione peroxidase by nitric oxide. Implication for cytotoxicity, J. Biol. Chem. 270 (1995) 21035–21039.
[31] L. Boscá, S. Hortelano, Mechanisms of nitric oxide-dependent apoptosis: involvement of mitochondrial mediators, Cell. Signal. 11 (1999) 239–244.
[32] K.T. Lin, J.Y. Xue, M. Nomen, B. Spur, P.Y. Wong, Peroxynitrite-induced apoptosis in HL-60 cells, J. Biol. Chem. 270 (1995) 16487–16490.
[33] G. Loschen, L. Flohé, B. Chance, Respiratory chain linked H(2)O(2) production in pigeon heart mitochondria, FEBS Lett. 18 (1971) 261–264.
[34] K.J. Davies, Oxidative stress: the paradox of aerobic life, Biochem. Soc. Symp. 61 (1995) 1–31.
[35] M. Ott, V. Gogvadze, S. Orrenius, B. Zhivotovsky, Mitochondria, oxidative stress and cell death, Apoptosis. 12 (2007) 913–922.
[36] S. Reuter, S.C. Gupta, M.M. Chaturvedi, B.B. Aggarwal, Oxidative stress, inflammation, and cancer: How are they linked?, Free Radic. Biol. Med. 49 (2010) 1603–1616.
[37] A.J. Schetter, N.H.H. Heegaard, C.C. Harris, Inflammation and cancer: interweaving microRNA, free radical, cytokine and p53 pathways, Carcinogenesis. 31 (2010) 37–49.
[38] B.B. Aggarwal, R.V. Vijayalekshmi, B. Sung, Targeting inflammatory pathways for prevention and therapy of cancer: short-term friend, long-term foe, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 15 (2009) 425–430.
[39] W.-W. Lin, M. Karin, A cytokine-mediated link between innate immunity, inflammation, and cancer, J. Clin. Invest. 117 (2007) 1175–1183.
[40] S.P. Hussain, L.J. Hofseth, C.C. Harris, Radical causes of cancer, Nat. Rev. Cancer. 3 (2003) 276–285.
[41] L.M. Coussens, Z. Werb, Inflammation and cancer, Nature. 420 (2002) 860–867.
[42] A. Federico, F. Morgillo, C. Tuccillo, F. Ciardiello, C. Loguercio, Chronic inflammation and oxidative stress in human carcinogenesis, Int. J. Cancer. 121 (2007) 2381–2386.
[43] S. Mena, A. Ortega, J.M. Estrela, Oxidative stress in environmental-induced carcinogenesis, Mutat. Res. Toxicol. Environ. Mutagen. 674 (2009) 36–44.
[44] J.T. Thaiparambil, M.V. Vadhanam, C. Srinivasan, C.G. Gairola, R.C. Gupta, Time-Dependent Formation of 8-Oxo-deoxyguanosine in the Lungs of Mice Exposed to Cigarette Smoke, Chem. Res. Toxicol. 20 (2007) 1737–1740.

[45] H. van der Vaart, D. Postma, W. Timens, H. Ten, Acute effects of cigarette smoke on inflammation and oxidative stress: a review, Thorax. 59 (2004) 713–721.
[46] H. Bartsch, U. Nair, A. Risch, M. Rojas, H. Wikman, K. Alexandrov, Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers, Cancer Epidemiol. Biomarkers Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol. 9 (2000) 3–28.
[47] A. Altieri, W. Garavello, C. Bosetti, S. Gallus, C. La Vecchia, Alcohol consumption and risk of laryngeal cancer, Oral Oncol. 41 (2005) 956–965.
[48] G. Corrao, V. Bagnardi, A. Zambon, S. Arico, Exploring the dose-response relationship between alcohol consumption and the risk of several alcohol-related conditions: a meta-analysis, Addict. Abingdon Engl. 94 (1999) 1551–1573.
[49] S.K. Das, D.M. Vasudevan, Alcohol-induced oxidative stress, Life Sci. 81 (2007) 177–187.
[50] O. Sergent, B. Griffon, P. Cillard, J. Cillard, [Alcohol and oxidative stress], Pathol. Biol. (Paris). 49 (2001) 689–695.
[51] L. Conde de la Rosa, H. Moshage, N. Nieto, [Hepatocyte oxidant stress and alcoholic liver disease], Rev. Española Enfermedades Dig. Organo Of. Soc. Española Patol. Dig. 100 (2008) 156–163.
[52] K. Jung, R. Reszka, Mitochondria as subcellular targets for clinically useful anthracyclines, Adv. Drug Deliv. Rev. 49 (2001) 87–105.
[53] M. Valko, C.J. Rhodes, J. Moncol, M. Izakovic, M. Mazur, Free radicals, metals and antioxidants in oxidative stress-induced cancer, Chem. Biol. Interact. 160 (2006) 1–40.
[54] L.T. Stayner, D.A. Dankovic, R.A. Lemen, Occupational exposure to chrysotile asbestos and cancer risk: a review of the amphibole hypothesis., Am. J. Public Health. 86 (1996) 179–186.
[55] D.J. Price, J.G. Joshi, Ferritin. Binding of beryllium and other divalent metal ions., J. Biol. Chem. 258 (1983) 10873–10880.
[56] E. Casalino, C. Sblano, C. Landriscina, Enzyme Activity Alteration by Cadmium Administration to Rats: The Possibility of Iron Involvement in Lipid Peroxidation, Arch. Biochem. Biophys. 346 (1997) 171–179.
[57] K. Sankaranarayanan, Estimation of the genetic risks of exposure to ionizing radiation in humans: current status and emerging perspectives, J. Radiat. Res. (Tokyo). 47 Suppl B (2006) B57–66.
[58] Rayonnements ionisants et santé, (n.d.).
[59] Comprendre le nucléaire - Les rayonnements proviennent-ils uniquement des centrales nucléaires? - Association nucléaire canadienne, Assoc. Nucléaire Can. (n.d.).
[60] R. Platzman, The physical and chemical basis of mechanisms in radiation biology, in: W Claus Ed Radiat. Biol. Med. Sel. Rev. Life Sci., Addison-Wesley, Reading, MA, 1958: pp. 15–72.
[61] E.I. Azzam, J.-P. Jay-Gerin, D. Pain, Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury, Cancer Lett. 327 (2012) 48–60.
[62] S.D. Varma, S. Kovtun, K.R. Hegde, Role of UV Irradiation and Oxidative Stress in Cataract Formation. Medical Prevention by Nutritional Antioxidants and Metabolic Agonists, Eye Contact Lens. 37 (2011) 233–245.
[63] O.A. Bessey, O.H. Lowry, Factors Influencing the Riboflavin Content of the Cornea, J. Biol. Chem. 155 (1944) 635–643.
[64] L.A. Herrera, L. Benítez-Bribiesca, A. Mohar, P. Ostrosky-Wegman, Role of infectious diseases in human carcinogenesis, Environ. Mol. Mutagen. 45 (2005) 284–303.
[65] H.H. Guo, L.A. Loeb, Tumbling down a different pathway to genetic instability, J. Clin. Invest. 112 (2003) 1793–1795.
[66] A.G. Georgakilas, W.G. Mosley, S. Georgakila, D. Ziech, M.I. Panayiotidis, Viral-induced human carcinogenesis: an oxidative stress perspective, Mol. Biosyst. 6 (2010) 1162–1172.

[67] X.W. Wang, S.P. Hussain, T.-I. Huo, C.-G. Wu, M. Forgues, L.J. Hofseth, et al., Molecular pathogenesis of human hepatocellular carcinoma, Toxicology. 181-182 (2002) 43–47.

[68] J.-H. Wang, C. Yun, S. Kim, J.-H. Lee, G. Yoon, M.-O. Lee, et al., Reactive oxygen species modulates the intracellular level of HBx viral oncoprotein, Biochem. Biophys. Res. Commun. 310 (2003) 32–39.

[69] S. Abel, M. De Kock, D.J. van Schalkwyk, S. Swanevelder, M.C. Kew, W.C.A. Gelderblom, Altered lipid profile, oxidative status and hepatitis B virus interactions in human hepatocellular carcinoma, Prostaglandins Leukot. Essent. Fatty Acids. 81 (2009) 391–399.

[70] R. Shimoda, M. Nagashima, M. Sakamoto, N. Yamaguchi, S. Hirohashi, J. Yokota, et al., Increased formation of oxidative DNA damage, 8-hydroxydeoxyguanosine, in human livers with chronic hepatitis, Cancer Res. 54 (1994) 3171–3172.

[71] F. Farinati, R. Cardin, M. Bortolami, A. Grottola, M. Manno, A. Colantoni, et al., Estrogens receptors and oxidative damage in the liver, Mol. Cell. Endocrinol. 193 (2002) 85–88.

[72] R. Cardin, G. Saccoccio, F. Masutti, S. Bellentani, F. Farinati, C. Tiribelli, DNA oxidative damage in leukocytes correlates with the severity of HCV-related liver disease: validation in an open population study, J. Hepatol. 34 (2001) 587–592.

[73] P. Aukrust, L. Luna, T. Ueland, R.F. Johansen, F. Müller, S.S. Frøland, et al., Impaired base excision repair and accumulation of oxidative base lesions in CD4+ T cells of HIV-infected patients, Blood. 105 (2005) 4730–4735.

[74] M. Husain, L.G. Meggs, H. Vashistha, S. Simoes, K.O. Griffiths, D. Kumar, et al., Inhibition of p66ShcA longevity gene rescues podocytes from HIV-1-induced oxidative stress and apoptosis, J. Biol. Chem. 284 (2009) 16648–16658.

[75] T. Söderdahl, M. Enoksson, M. Lundberg, A. Holmgren, O.P. Ottersen, S. Orrenius, et al., Visualization of the compartmentalization of glutathione and protein-glutathione mixed disulfides in cultured cells, FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. 17 (2003) 124–126.

[76] A.Y. Andreyev, Y.E. Kushnareva, A.A. Starkov, Mitochondrial metabolism of reactive oxygen species, Biochem. Biokhimiia. 70 (2005) 200–214.

[77] A. Wahlländer, S. Soboll, H. Sies, I. Linke, M. Müller, Hepatic mitochondrial and cytosolic glutathione content and the subcellular distribution of GSH-S-transferases, FEBS Lett. 97 (1979) 138–140.

[78] T.R. Pushpa-Rekha, A.L. Burdsall, L.M. Oleksa, G.M. Chisolm, D.M. Driscoll, Rat phospholipid-hydroperoxide glutathione peroxidase. cDNA cloning and identification of multiple transcription and translation start sites, J. Biol. Chem. 270 (1995) 26993–26999.

[79] H. Van Remmen, C. Salvador, H. Yang, T.T. Huang, C.J. Epstein, A. Richardson, Characterization of the antioxidant status of the heterozygous manganese superoxide dismutase knockout mouse, Arch. Biochem. Biophys. 363 (1999) 91–97.

[80] M.O. Pereverzev, T.V. Vygodina, A.A. Konstantinov, V.P. Skulachev, Cytochrome c, an ideal antioxidant, Biochem. Soc. Trans. 31 (2003) 1312–1315.

[81] K. Mailer, Superoxide radical as electron donor for oxidative phosphorylation of ADP, Biochem. Biophys. Res. Commun. 170 (1990) 59–64.

[82] R.M. Pope, Apoptosis as a therapeutic tool in rheumatoid arthritis, Nat. Rev. Immunol. 2 (2002) 527–535.

[83] S. Haupt, M. Berger, Z. Goldberg, Y. Haupt, Apoptosis - the p53 network, J. Cell Sci. 116 (2003) 4077–4085.

[84] M. Mayer, M. Noble, N-acetyl-L-cysteine is a pluripotent protector against cell death and enhancer of trophic factor-mediated cell survival in vitro, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 7496–7500.

[85] N.N. Danial, S.J. Korsmeyer, Cell Death: Critical Control Points, Cell. 116 (2004) 205–219.
[86] K.C. Kregel, H.J. Zhang, An integrated view of oxidative stress in aging: basic mechanisms, functional effects, and pathological considerations, Am. J. Physiol. - Regul. Integr. Comp. Physiol. 292 (2007) R18–R36.
[87] K. Rothkamm, M. Löbrich, Evidence for a lack of DNA double-strand break repair in human cells exposed to very low x-ray doses, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 5057–5062.
[88] W.P. Roos, B. Kaina, DNA damage-induced cell death: from specific DNA lesions to the DNA damage response and apoptosis, Cancer Lett. 332 (2013) 237–248.
[89] G. Melino, F. Bernassola, R.A. Knight, M.T. Corasaniti, G. Nistico, A. Finazzi-Agro, S-nitrosylation regulates apoptosis, Nature. 388 (1997) 432–433.
[90] M. Leist, B. Single, H. Naumann, E. Fava, B. Simon, S. Kühnle, et al., Inhibition of mitochondrial ATP generation by nitric oxide switches apoptosis to necrosis, Exp. Cell Res. 249 (1999) 396–403.
[91] Y. Tsujimoto, S. Shimizu, Y. Eguchi, W. Kamiike, H. Matsuda, Bcl-2 and Bcl-xL block apoptosis as well as necrosis: possible involvement of common mediators in apoptotic and necrotic signal transduction pathways, Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund UK. 11 Suppl 3 (1997) 380–382.
[92] A. Vigneron, K.H. Vousden, p53, ROS and senescence in the control of aging, Aging. 2 (2010) 471–474.
[93] R. Colavitti, T. Finkel, Reactive Oxygen Species as Mediators of Cellular Senescence, IUBMB Life. 57 (2005) 277–281.
[94] L. Hayflick, The limited in vitro lifetime of human diploid cell strains, Exp. Cell Res. 37 (1965) 614–636.
[95] T. Lu, T. Finkel, Free radicals and senescence, Exp. Cell Res. 314 (2008) 1918–1922.
[96] D. Harman, Aging: a theory based on free radical and radiation chemistry., J. Gerontol. 11 (1956).
[97] R.S. Sohal, S. Agarwal, A. Dubey, W.C. Orr, Protein oxidative damage is associated with life expectancy of houseflies., Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 7255–7259.
[98] J. Bejma, L.L. Ji, Aging and acute exercise enhance free radical generation in rat skeletal muscle, J. Appl. Physiol. Bethesda Md 1985. 87 (1999) 465–470.
[99] J. Bejma, P. Ramires, L.L. Ji, Free radical generation and oxidative stress with ageing and exercise: differential effects in the myocardium and liver, Acta Physiol. Scand. 169 (2000) 343–351.
[100] A.S. Driver, P.R. Kodavanti, W.R. Mundy, Age-related changes in reactive oxygen species production in rat brain homogenates, Neurotoxicol. Teratol. 22 (2000) 175–181.
[101] F. Gomi, H. Utsumi, A. Hamada, M. Matsuo, Aging retards spin clearance from mouse brain and food restriction prevents its age-dependent retardation, Life Sci. 52 (1993) 2027–2033.
[102] H.J. Zhang, L. Xu, V.J. Drake, L. Xie, L.W. Oberley, K.C. Kregel, Heat-induced liver injury in old rats is associated with exaggerated oxidative stress and altered transcription factor activation, FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. 17 (2003) 2293–2295.
[103] J.M. McCord, I. Fridovich, Superoxide dismutase: the first twenty years (1968-1988), Free Radic. Biol. Med. 5 (1988) 363–369.
[104] V. Cecarini, J. Gee, E. Fioretti, M. Amici, M. Angeletti, A.M. Eleuteri, et al., Protein oxidation and cellular homeostasis: Emphasis on metabolism, Biochim. Biophys. Acta BBA - Mol. Cell Res. 1773 (2007) 93–104.
[105] E.R. Stadtman, R.L. Levine, Free radical-mediated oxidation of free amino acids and amino acid residues in proteins, Amino Acids. 25 (2003) 207–218.
[106] E.R. Stadtman, R.L. Levine, Protein oxidation, Ann. N. Y. Acad. Sci. 899 (2000) 191–208.

[107] W.M. Garrison, Comparative free radical chemistry in the radiolytic deamination and dephosphorylation of bio-organic molecules. II. Oxygenated solutions, Free Radic. Biol. Med. 7 (1989) 171–175.
[108] L.J. Marnett, Lipid peroxidation—DNA damage by malondialdehyde, Mutat. Res. Mol. Mech. Mutagen. 424 (1999) 83–95.
[109] Z. Radak, Z. Zhao, S. Goto, E. Koltai, Age-associated neurodegeneration and oxidative damage to lipids, proteins and DNA, Mol. Aspects Med. 32 (2011) 305–315.
[110] A. Catalá, Lipid peroxidation of membrane phospholipids generates hydroxy-alkenals and oxidized phospholipids active in physiological and/or pathological conditions, Chem. Phys. Lipids. 157 (2009) 1–11.
[111] G. Ferretti, T. Bacchetti, Peroxidation of lipoproteins in multiple sclerosis, J. Neurol. Sci. 311 (2011) 92–97.
[112] I. Levitan, S. Volkov, P.V. Subbaiah, Oxidized LDL: diversity, patterns of recognition, and pathophysiology, Antioxidants Redox Signal. 13 (2010) 39–75.
[113] H. Itabe, Oxidative modification of LDL: its pathological role in atherosclerosis, Clin. Rev. Allergy Immunol. 37 (2009) 4–11.
[114] G. Ferretti, T. Bacchetti, A. Nègre-Salvayre, R. Salvayre, N. Dousset, G. Curatola, Structural modifications of HDL and functional consequences, Atherosclerosis. 184 (2006) 1–7.
[115] Y. Zhang, O. Marcillat, C. Giulivi, L. Ernster, K.J. Davies, The oxidative inactivation of mitochondrial electron transport chain components and ATPase, J. Biol. Chem. 265 (1990) 16330–16336.
[116] E. Albano, G. Bellomo, M. Parola, R. Carini, M.U. Dianzani, Stimulation of lipid peroxidation increases the intracellular calcium content of isolated hepatocytes, Biochim. Biophys. Acta. 1091 (1991) 310–316.
[117] B.R. Bacon, R. O'Neill, R.S. Britton, Hepatic mitochondrial energy production in rats with chronic iron overload, Gastroenterology. 105 (1993) 1134–1140.
[118] V.E. Kagan, V.A. Tyurin, J. Jiang, Y.Y. Tyurina, V.B. Ritov, A.A. Amoscato, et al., Cytochrome c acts as a cardiolipin oxygenase required for release of proapoptotic factors, Nat. Chem. Biol. 1 (2005) 223–232.
[119] M.D. Evans, M. Dizdaroglu, M.S. Cooke, Oxidative DNA damage and disease: induction, repair and significance, Mutat. Res. 567 (2004) 1–61.
[120] H.E. Poulsen, E. Specht, K. Broedbaek, T. Henriksen, C. Ellervik, T. Mandrup-Poulsen, et al., RNA modifications by oxidation: A novel disease mechanism?, Free Radic. Biol. Med. 52 (2012) 1353–1361.
[121] H.H. Thorp, The importance of being r: greater oxidative stability of RNA compared with DNA, Chem. Biol. 7 (2000) R33–36.
[122] M. Dizdaroglu, Oxidatively induced DNA damage: Mechanisms, repair and disease, Cancer Lett. 327 (2012) 26–47.
[123] J. Cadet, T. Douki, D. Gasparutto, J.-L. Ravanat, Oxidative damage to DNA: formation, measurement and biochemical features, Mutat. Res. Mol. Mech. Mutagen. 531 (2003) 5–23.
[124] M.L. Wood, M. Dizdaroglu, E. Gajewski, J.M. Essigmann, Mechanistic studies of ionizing radiation and oxidative mutagenesis: genetic effects of a single 8-hydroxyguanine (7-hydro-8-oxoguanine) residue inserted at a unique site in a viral genome, Biochemistry (Mosc.). 29 (1990) 7024–7032.
[125] S. Shibutani, M. Takeshita, A.P. Grollman, Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG, Nature. 349 (1991) 431–434.
[126] K.C. Cheng, D.S. Cahill, H. Kasai, S. Nishimura, L.A. Loeb, 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions, J. Biol. Chem. 267 (1992) 166–172.

[127] A.P. Grollman, M. Moriya, Mutagenesis by 8-oxoguanine: an enemy within, Trends Genet. TIG. 9 (1993) 246–249.
[128] B. van Loon, E. Markkanen, U. Hübscher, Oxygen as a friend and enemy: How to combat the mutational potential of 8-oxo-guanine, DNA Repair. 9 (2010) 604–616.
[129] T. Lindahl, D.E. Barnes, Repair of endogenous DNA damage, Cold Spring Harb. Symp. Quant. Biol. 65 (2000) 127–133.
[130] H. Kamiya, H. Miura, N. Murata-Kamiya, H. Ishikawa, T. Sakaguchi, H. Inoue, et al., 8-Hydroxyadenine (7, 8-dihydro-8-Oxoadenine) induces misincorporation in in vitro DNA synthesis and mutations in NIH 3T3 cells, Nucleic Acids Res. 23 (1995) 2893–2899.
[131] X. Tan, A.P. Grollman, S. Shibutani, Comparison of the mutagenic properties of 8-oxo-7,8-dihydro-2'-deoxyadenosine and 8-oxo-7,8-dihydro-2'-deoxyguanosine DNA lesions in mammalian cells, Carcinogenesis. 20 (1999) 2287–2292.
[132] D. Pluskota-Karwatka, Modifications of nucleosides by endogenous mutagens–DNA adducts arising from cellular processes, Bioorganic Chem. 36 (2008) 198–213.
[133] Y.-J. Kim, D.M. Wilson 3rd, Overview of base excision repair biochemistry, Curr. Mol. Pharmacol. 5 (2012) 3–13.
[134] B.R. Berquist, D.M. Wilson III, Pathways for repairing and tolerating the spectrum of oxidative DNA lesions, Cancer Lett. 327 (2012) 61–72.
[135] D.M. Wilson III, V.A. Bohr, The mechanics of base excision repair, and its relationship to aging and disease, DNA Repair. 6 (2007) 544–559.
[136] H.E. Krokan, R. Standal, G. Slupphaug, DNA glycosylases in the base excision repair of DNA, Biochem. J. 325 (Pt 1) (1997) 1–16.
[137] S. Mitra, DNA glycosylases: specificity and mechanisms, Prog. Nucleic Acid Res. Mol. Biol. 68 (2001) 189–192.
[138] M. Bhagwat, J.A. Gerlt, 3'- and 5'-strand cleavage reactions catalyzed by the Fpg protein from Escherichia coli occur via successive beta- and delta-elimination mechanisms, respectively, Biochemistry (Mosc.). 35 (1996) 659–665.
[139] D.M. Wilson III, D. Barsky, The major human abasic endonuclease: formation, consequences and repair of abasic lesions in DNA, Mutat. Res. Repair. 485 (2001) 283–307.
[140] B. Demple, J.-S. Sung, Molecular and biological roles of Ape1 protein in mammalian base excision repair, DNA Repair. 4 (2005) 1442–1449.
[141] K.-M. Chou, Y.-C. Cheng, An exonucleolytic activity of human apurinic/apyrimidinic endonuclease on 3' mispaired DNA, Nature. 415 (2002) 655–659.
[142] M.Z. Hadi, K. Ginalski, L.H. Nguyen, D.M. Wilson 3rd, Determinants in nuclease specificity of Ape1 and Ape2, human homologues of Escherichia coli exonuclease III, J. Mol. Biol. 316 (2002) 853–866.
[143] L.A. Loeb, R.J. Monnat Jr, DNA polymerases and human disease, Nat. Rev. Genet. 9 (2008) 594–604.
[144] W.A. Beard, R. Prasad, S.H. Wilson, Activities and mechanism of DNA polymerase beta, Methods Enzymol. 408 (2006) 91–107.
[145] N.K. Bernstein, R.S. Williams, M.L. Rakovszky, D. Cui, R. Green, F. Karimi-Busheri, et al., The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase, Mol. Cell. 17 (2005) 657–670.
[146] A.B. Robertson, A. Klungland, T. Rognes, I. Leiros, DNA repair in mammalian cells: Base excision repair: the long and short of it, Cell. Mol. Life Sci. CMLS. 66 (2009) 981–993.
[147] G. Frosina, P. Fortini, O. Rossi, F. Carrozzino, G. Raspaglio, L.S. Cox, et al., Two pathways for base excision repair in mammalian cells, J. Biol. Chem. 271 (1996) 9573–9578.

[148] P. Fortini, E. Dogliotti, Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways, DNA Repair. 6 (2007) 398–409.
[149] D.O. Zharkov, Base excision DNA repair, Cell. Mol. Life Sci. CMLS. 65 (2008) 1544–1565.
[150] E. Petermann, M. Ziegler, S.L. Oei, ATP-dependent selection between single nucleotide and long patch base excision repair, DNA Repair. 2 (2003) 1101–1114.
[151] J.-S. Sung, B. Demple, Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA, FEBS J. 273 (2006) 1620–1629.
[152] A.E. Tomkinson, S. Vijayakumar, J.M. Pascal, T. Ellenberger, DNA ligases: structure, reaction mechanism, and function, Chem. Rev. 106 (2006) 687–699.
[153] E. Cappelli, R. Taylor, M. Cevasco, A. Abbondandolo, K. Caldecott, G. Frosina, Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair, J. Biol. Chem. 272 (1997) 23970–23975.
[154] K.W. Caldecott, J.D. Tucker, L.H. Stanker, L.H. Thompson, Characterization of the XRCC1-DNA ligase III complex in vitro and its absence from mutant hamster cells, Nucleic Acids Res. 23 (1995) 4836–4843.
[155] M. Dizdaroglu, Base-excision repair of oxidative DNA damage by DNA glycosylases, Mutat. Res. Mol. Mech. Mutagen. 591 (2005) 45–59.
[156] M.L. Hegde, T.K. Hazra, S. Mitra, Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells, Cell Res. 18 (2008) 27–47.
[157] Y. Xie, H. Yang, C. Cunanan, K. Okamoto, D. Shibata, J. Pan, et al., Deficiencies in Mouse Myh and Ogg1 Result in Tumor Predisposition and G to T Mutations in Codon 12 of the K-Ras Oncogene in Lung Tumors, Cancer Res. 64 (2004) 3096–3102.
[158] M.K. Chan, M.T. Ocampo-Hafalla, V. Vartanian, P. Jaruga, G. Kirkali, K.L. Koenig, et al., Targeted deletion of the genes encoding NTH1 and NEIL1 DNA N-glycosylases reveals the existence of novel carcinogenic oxidative damage to DNA, DNA Repair. 8 (2009) 786–794.
[159] S. Baglioni, G. Melean, F. Gensini, M. Santucci, M. Scatizzi, L. Papi, et al., A kindred with MYH-associated polyposis and pilomatricomas, Am. J. Med. Genet. A. 134A (2005) 212–214.
[160] R.W. Sobol, J.K. Horton, R. Kühn, H. Gu, R.K. Singhal, R. Prasad, et al., Requirement of mammalian DNA polymerase-beta in base-excision repair, Nature. 379 (1996) 183–186.
[161] L.B. Meira, S. Devaraj, G.E. Kisby, D.K. Burns, R.L. Daniel, R.E. Hammer, et al., Heterozygosity for the mouse Apex gene results in phenotypes associated with oxidative stress, Cancer Res. 61 (2001) 5552–5557.
[162] R.S. Tebbs, M.L. Flannery, J.J. Meneses, A. Hartmann, J.D. Tucker, L.H. Thompson, et al., Requirement for the Xrcc1 DNA base excision repair gene during early mouse development, Dev. Biol. 208 (1999) 513–529.
[163] D. Bentley, J. Selfridge, J.K. Millar, K. Samuel, N. Hole, J.D. Ansell, et al., DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability, Nat. Genet. 13 (1996) 489–491.
[164] N. Puebla-Osorio, D.B. Lacey, F.W. Alt, C. Zhu, Early embryonic lethality due to targeted inactivation of DNA ligase III, Mol. Cell. Biol. 26 (2006) 3935–3941.
[165] J. Huamani, C.A. McMahan, D.C. Herbert, R. Reddick, J.R. McCarrey, M.I. MacInnes, et al., Spontaneous mutagenesis is enhanced in Apex heterozygous mice, Mol. Cell. Biol. 24 (2004) 8145–8153.
[166] D.C. Cabelof, Z. Guo, J.J. Raffoul, R.W. Sobol, S.H. Wilson, A. Richardson, et al., Base excision repair deficiency caused by polymerase beta haploinsufficiency: accelerated DNA damage and increased mutational response to carcinogens, Cancer Res. 63 (2003) 5799–5807.

[167] D.C. Cabelof, Y. Ikeno, A. Nyska, R.A. Busuttil, N. Anyangwe, J. Vijg, et al., Haploinsufficiency in DNA polymerase beta increases cancer risk with age and alters mortality rate, Cancer Res. 66 (2006) 7460–7465.
[168] D.R. McNeill, P.-C. Lin, M.G. Miller, P.J. Pistell, N.C. de Souza-Pinto, K.W. Fishbein, et al., XRCC1 haploinsufficiency in mice has little effect on aging, but adversely modifies exposure-dependent susceptibility, Nucleic Acids Res. 39 (2011) 7992–8004.
[169] T. Nouspikel, P.C. Hanawalt, DNA repair in terminally differentiated cells, DNA Repair. 1 (2002) 59–75.
[170] I.I. Kruman, E. Schwartz, Y. Kruman, R.G. Cutler, X. Zhu, N.H. Greig, et al., Suppression of uracil-DNA glycosylase induces neuronal apoptosis, J. Biol. Chem. 279 (2004) 43952–43960.
[171] A.G. Georgakilas, Processing of DNA damage clusters in human cells: current status of knowledge, Mol. Biosyst. 4 (2008) 30–35.
[172] E. Gollapalle, R. Wang, R. Adetolu, D. Tsao, D. Francisco, G. Sigounas, et al., Detection of oxidative clustered DNA lesions in X-irradiated mouse skin tissues and human MCF-7 breast cancer cells, Radiat. Res. 167 (2007) 207–216.
[173] V.A. Semenenko, R.D. Stewart, A fast Monte Carlo algorithm to simulate the spectrum of DNA damages formed by ionizing radiation, Radiat. Res. 161 (2004) 451–457.
[174] V. Michalik, Model of DNA damage induced by radiations of various qualities, Int. J. Radiat. Biol. 62 (1992) 9–20.
[175] B.M. Sutherland, P.V. Bennett, J.C. Sutherland, J. Laval, Clustered DNA damages induced by x rays in human cells, Radiat. Res. 157 (2002) 611–616.
[176] M.A. Chaudhry, M. Weinfeld, The action of Escherichia coli endonuclease III on multiply damaged sites in DNA, J. Mol. Biol. 249 (1995) 914–922.
[177] M.H. David-Cordonnier, J. Laval, P. O'Neill, Clustered DNA damage, influence on damage excision by XRS5 nuclear extracts and Escherichia coli Nth and Fpg proteins, J. Biol. Chem. 275 (2000) 11865–11873.
[178] A.G. Georgakilas, P.V. Bennett, D.M. Wilson 3rd, B.M. Sutherland, Processing of bistranded abasic DNA clusters in gamma-irradiated human hematopoietic cells, Nucleic Acids Res. 32 (2004) 5609–5620.
[179] E. Sage, L. Harrison, Clustered DNA lesion repair in eukaryotes: Relevance to mutagenesis and cell survival, Mutat. Res. Mol. Mech. Mutagen. 711 (2011) 123–133.
[180] J.V. Harper, J.A. Anderson, P. O'Neill, Radiation induced DNA DSBs: Contribution from stalled replication forks?, DNA Repair. 9 (2010) 907–913.
[181] T.R. Kasparek, T.C. Humphrey, DNA double-strand break repair pathways, chromosomal rearrangements and cancer, Semin. Cell Dev. Biol. 22 (2011) 886–897.
[182] M.A. Resnick, P. Martin, The repair of double-strand breaks in the nuclear DNA of Saccharomyces cerevisiae and its genetic control, Mol. Gen. Genet. MGG. 143 (1976) 119–129.
[183] B. Pardo, B. Gómez-González, A. Aguilera, DNA repair in mammalian cells: DNA double-strand break repair: how to fix a broken relationship, Cell. Mol. Life Sci. CMLS. 66 (2009) 1039–1056.
[184] A.J. Hartlerode, R. Scully, Mechanisms of double-strand break repair in somatic mammalian cells, Biochem. J. 423 (2009) 157–168.
[185] M. Takata, M.S. Sasaki, E. Sonoda, C. Morrison, M. Hashimoto, H. Utsumi, et al., Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells, EMBO J. 17 (1998) 5497–5508.
[186] F. Delacôte, B.S. Lopez, Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance: the trans-S double-strand break repair model, Cell Cycle Georget. Tex. 7 (2008) 33–38.

[187] R. Holliday, A mechanism for gene conversion in fungi, Genet. Res. 89 (2007) 285–307.
[188] G. Ira, A. Malkova, G. Liberi, M. Foiani, J.E. Haber, Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in yeast, Cell. 115 (2003) 401–411.
[189] T. Robert, D. Dervins, F. Fabre, S. Gangloff, Mrc1 and Srs2 are major actors in the regulation of spontaneous crossover, EMBO J. 25 (2006) 2837–2846.
[190] M. Dizdaroglu, Free-radical-induced formation of an 8,5'-cyclo-2'-deoxyguanosine moiety in deoxyribonucleic acid, Biochem. J. 238 (1986) 247–254.
[191] K. Keck, Formation of cyclonucleotides during irradiation of aqueous solutions of purine nucleotides, Z. Für Naturforschung Teil B Chem. Biochem. Biophys. Biol. 23 (1968) 1034–1043.
[192] N. Mariaggi, J. Cadet, R. Teoule, Cyclisation radicalaire de la desoxy-2'-adenosine en solution aqueuse, sous l'effet du rayonnement gamma, Tetrahedron. 32 (1976) 2385–2387.
[193] P.J. Brooks, The case for 8,5'-cyclopurine-2'-deoxynucleosides as endogenous DNA lesions that cause neurodegeneration in xeroderma pigmentosum, Neuroscience. 145 (2007) 1407–1417.
[194] N. Belmadoui, F. Boussicault, M. Guerra, J.-L. Ravanat, C. Chatgilialoglu, J. Cadet, Radiation-induced formation of purine 5',8-cyclonucleosides in isolated and cellular DNA: high stereospecificity and modulating effect of oxygen, Org. Biomol. Chem. 8 (2010) 3211-3219.
[195] C. Chatgilialoglu, C. Ferreri, M.A. Terzidis, Purine 5',8-cyclonucleoside lesions: chemistry and biology, Chem. Soc. Rev. 40 (2011) 1368.
[196] L.B. Jimenez, S. Encinas, C. Chatgilialoglu, M.A. Miranda, Solar one-way photoisomerisation of 5',8-cyclo-2'-deoxyadenosine, Org. Biomol. Chem. 6 (2008) 1083.
[197] P. Jaruga, M. Dizdaroglu, 8,5'-Cyclopurine-2'-deoxynucleosides in DNA: Mechanisms of formation, measurement, repair and biological effects, DNA Repair (Amst.). 7 (2008) 1413–1425.
[198] K. Randerath, G.-D. Zhou, R.L. Somers, J.H. Robbins, P.J. Brooks, A 32P-Postlabeling Assay for the Oxidative DNA Lesion 8,5'-Cyclo-2'-deoxyadenosine in Mammalian Tissues; Evidence that four type II I-compounds are dinucleotides containing the lesion in the 3' nucleotide, J. Biol. Chem. 276 (2001) 36051–36057.
[199] P. Jaruga, M. Dizdaroglu, Identification and quantification of (5'R)- and (5'S)-8,5'-cyclo-2'-deoxyadenosines in human urine as putative biomarkers of oxidatively induced damage to DNA, Biochem. Biophys. Res. Commun. 397 (2010) 48–52.
[200] P. Jaruga, Y. Xiao, B.C. Nelson, M. Dizdaroglu, Measurement of (5'R)- and (5'S)-8,5'-cyclo-2'-deoxyadenosines in DNA in vivo by liquid chromatography/isotope-dilution tandem mass spectrometry, Biochem. Biophys. Res. Commun. 386 (2009) 656–660.
[201] R.A. Egler, E. Fernandes, K. Rothermund, S. Sereika, N. de Souza-Pinto, P. Jaruga, et al., Regulation of reactive oxygen species, DNA damage, and c-Myc function by peroxiredoxin 1, Oncogene. 24 (2005) 8038–8050.
[202] G. Kirkali, N.C. de Souza-Pinto, P. Jaruga, V.A. Bohr, M. Dizdaroglu, Accumulation of (5'S)-8,5'-cyclo-2'-deoxyadenosine in organs of Cockayne syndrome complementation group B gene knockout mice, DNA Repair (Amst.). 8 (2009) 274–278.
[203] D.C. Malins, K.M. Anderson, J.J. Stegeman, P. Jaruga, V.M. Green, N.K. Gilman, et al., Biomarkers signal contaminant effects on the organs of English sole (Parophrys vetulus) from Puget Sound, Environ. Health Perspect. 114 (2006) 823–829.
[204] K.M. Anderson, P. Jaruga, C.R. Ramsey, N.K. Gilman, V.M. Green, S.W. Rostad, et al., Structural alterations in breast stromal and epithelial DNA: the influence of 8,5'-cyclo-2'-deoxyadenosine, Cell Cycle 5 (2006) 1240–1244.
[205] G. Kirkali, M. Tunca, S. Genc, P. Jaruga, M. Dizdaroglu, Oxidative DNA damage in polymorphonuclear leukocytes of patients with familial Mediterranean fever, Free Radic. Biol. Med. 44 (2008) 386–393.

[206] K. Miaskiewicz, J.H. Miller, A.F. Fuciarelli, Theoretical analysis of DNA intrastrand cross linking by formation of 8,5'-cyclodeoxyadenosine, Nucleic Acids Res. 23 (1995) 515–521.
[207] H. Huang, R.S. Das, A.K. Basu, M.P. Stone, Structure of (5'S)-8,5'-Cyclo-2'-deoxyguanosine in DNA, J. Am. Chem. Soc. 133 (2011) 20357–20368.
[208] I. Kuraoka, Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells, Proc. Natl. Acad. Sci. 97 (2000) 3832–3837.
[209] I. Kuraoka, P. Robins, C. Masutani, F. Hanaoka, D. Gasparutto, J. Cadet, et al., Oxygen Free Radical Damage to DNA, J. Biol. Chem. 276 (2001) 49283–49288.
[210] A.L. Swanson, J. Wang, Y. Wang, Accurate and Efficient Bypass of 8,5'-Cyclopurine-2'-Deoxynucleosides by Human and Yeast DNA Polymerase η, Chem. Res. Toxicol. 25 (2012) 1682–1691.
[211] P.J. Brooks, D.S. Wise, D.A. Berry, J.V. Kosmoski, M.J. Smerdon, R.L. Somers, et al., The Oxidative DNA Lesion 8,5'-(S)-Cyclo-2'-deoxyadenosine Is Repaired by the Nucleotide Excision Repair Pathway and Blocks Gene Expression in Mammalian Cells, J. Biol. Chem. 275 (2000) 22355–22362.
[212] A.N. Kapanidis, E. Margeat, S.O. Ho, E. Kortkhonjia, S. Weiss, R.H. Ebright, Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism, Science. 314 (2006) 1144–1147.
[213] C. Marietta, P.J. Brooks, Transcriptional bypass of bulky DNA lesions causes new mutant RNA transcripts in human cells, EMBO Rep. 8 (2007) 388–393.
[214] C. You, X. Dai, B. Yuan, J. Wang, J. Wang, P.J. Brooks, et al., A Quantitative Assay for Assessing the Effects of DNA Lesions on Transcription, Nat. Chem. Biol. 8 (2012) 817–822.
[215] J.A. Theruvathu, P. Jaruga, M. Dizdaroglu, P.J. Brooks, The oxidatively induced DNA lesions 8,5'-cyclo-2'-deoxyadenosine and 8-hydroxy-2'-deoxyadenosine are strongly resistant to acid-induced hydrolysis of the glycosidic bond, Mech. Ageing Dev. 128 (2007) 494–502.
[216] G.N. Wogan, S.S. Hecht, J.S. Felton, A.H. Conney, L.A. Loeb, Environmental and chemical carcinogenesis, Semin. Cancer Biol. 14 (2004) 473–486.
[217] T. Nouspikel, DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility, Cell. Mol. Life Sci. CMLS. 66 (2009) 994–1009.
[218] G.P. Pfeifer, Formation and processing of UV photoproducts: effects of DNA sequence and chromatin environment, Photochem. Photobiol. 65 (1997) 270–283.
[219] E.E. Trimmer, J.M. Essigmann, Cisplatin, Essays Biochem. 34 (1999) 191–211.
[220] K. Sugasawa, J.M. Ng, C. Masutani, S. Iwai, P.J. van der Spek, A.P. Eker, et al., Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair, Mol. Cell. 2 (1998) 223–232.
[221] K. Sugasawa, Y. Okuda, M. Saijo, R. Nishi, N. Matsuda, G. Chu, et al., UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex, Cell. 121 (2005) 387–400.
[222] W.L. de Laat, N.G.J. Jaspers, J.H.J. Hoeijmakers, Molecular mechanism of nucleotide excision repair, Genes Dev. 13 (1999) 768–785.
[223] M. Araki, C. Masutani, M. Takemura, A. Uchida, K. Sugasawa, J. Kondoh, et al., Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair, J. Biol. Chem. 276 (2001) 18665–18672.
[224] J.K. Kim, D. Patel, B.S. Choi, Contrasting structural impacts induced by cis-syn cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity, Photochem. Photobiol. 62 (1995) 44–50.

[225] J. Tang, G. Chu, Xeroderma pigmentosum complementation group E and UV-damaged DNA-binding protein, DNA Repair. 1 (2002) 601–616.
[226] Y. Hirose, Y. Ohkuma, Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression, J. Biochem. (Tokyo). 141 (2007) 601–608.
[227] P. Schultz, S. Fribourg, A. Poterszman, V. Mallouh, D. Moras, J.M. Egly, Molecular structure of human TFIIH, Cell. 102 (2000) 599–607.
[228] M. Volker, M.J. Moné, P. Karmakar, A. van Hoffen, W. Schul, W. Vermeulen, et al., Sequential assembly of the nucleotide excision repair factors in vivo, Mol. Cell. 8 (2001) 213–224.
[229] T. Hey, G. Lipps, K. Sugasawa, S. Iwai, F. Hanaoka, G. Krauss, The XPC−HR23B Complex Displays High Affinity and Specificity for Damaged DNA in a True-Equilibrium Fluorescence Assay†, Biochemistry (Mosc.). 41 (2002) 6583–6587.
[230] L.J. Blackwell, J.A. Borowiec, Human replication protein A binds single-stranded DNA in two distinct complexes, Mol. Cell. Biol. 14 (1994) 3993–4001.
[231] L.J. Blackwell, J.A. Borowiec, I.A. Mastrangelo, Single-stranded-DNA binding alters human replication protein A structure and facilitates interaction with DNA-dependent protein kinase, Mol. Cell. Biol. 16 (1996) 4798–4807.
[232] C. Kim, R.O. Snyder, M.S. Wold, Binding properties of replication protein A from human and yeast cells, Mol. Cell. Biol. 12 (1992) 3050–3059.
[233] C. Kim, B.F. Paulus, M.S. Wold, Interactions of human replication protein A with oligonucleotides, Biochemistry (Mosc.). 33 (1994) 14197–14206.
[234] W.L. de Laat, E. Appeldoorn, K. Sugasawa, E. Weterings, N.G. Jaspers, J.H. Hoeijmakers, DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair, Genes Dev. 12 (1998) 2598–2609.
[235] A. O'Donovan, A.A. Davies, J.G. Moggs, S.C. West, R.D. Wood, XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair, Nature. 371 (1994) 432–435.
[236] D. Mu, D.S. Hsu, A. Sancar, Reaction mechanism of human DNA repair excision nuclease, J. Biol. Chem. 271 (1996) 8285–8294.
[237] O. Popanda, H.W. Thielmann, The function of DNA polymerases in DNA repair synthesis of ultraviolet-irradiated human fibroblasts, Biochim. Biophys. Acta. 1129 (1992) 155–160.
[238] K.K. Shivji, M.K. Kenny, R.D. Wood, Proliferating cell nuclear antigen is required for DNA excision repair, Cell. 69 (1992) 367–374.
[239] A.F. Nichols, A. Sancar, Purification of PCNA as a nucleotide excision repair protein, Nucleic Acids Res. 20 (1992) 2441–2446.
[240] K. Diderich, M. Alanazi, J.H.J. Hoeijmakers, Premature aging and cancer in nucleotide excision repair-disorders, DNA Repair. 10 (2011) 772–780.
[241] A.R. Lehmann, D. McGibbon, M. Stefanini, Xeroderma pigmentosum, Orphanet J. Rare Dis. 6 (2011) 70.
[242] K.H. Kraemer, N.J. Patronas, R. Schiffmann, B.P. Brooks, D. Tamura, J.J. DiGiovanna, Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: A complex genotype–phenotype relationship, Neuroscience. 145 (2007) 1388–1396.
[243] A.R. Lehmann, DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy, Biochimie. 85 (2003) 1101–1111.
[244] M. de la L. Arenas-Sordo, E. Hernández-Zamora, L.A. Montoya-Pérez, B.C. Aldape-Barrios, Cockayne's syndrome: a case report. Literature review, Med. Oral Patol. Oral Cirugía Bucal. 11 (2006) E236–238.

[245] T. Stevnsner, M. Muftuoglu, M.D. Aamann, V.A. Bohr, The role of Cockayne Syndrome group B (CSB) protein in base excision repair and aging, Mech. Ageing Dev. 129 (2008) 441–448.
[246] K.M. Weidenheim, D.W. Dickson, I. Rapin, Neuropathology of Cockayne syndrome: Evidence for impaired development, premature aging, and neurodegeneration, Mech. Ageing Dev. 130 (2009) 619–636.
[247] M. D'Errico, E. Parlanti, M. Teson, P. Degan, T. Lemma, A. Calcagnile, et al., The role of CSA in the response to oxidative DNA damage in human cells, Oncogene. 26 (2007) 4336–4343.
[248] T. Matsunaga, D. Mu, C.-H. Park, J.T. Reardon, A. Sancar, Human DNA Repair Excision Nuclease; Analysis of the roles of the subunits involved in dual incisions by using anti-XPG and anti-ERCC1 antibodies, J. Biol. Chem. 270 (1995) 20862–20869.
[249] M. Muftuoglu, N.C. de Souza-Pinto, A. Dogan, M. Aamann, T. Stevnsner, I. Rybanska, et al., Cockayne syndrome group B protein stimulates repair of formamidopyrimidines by NEIL1 DNA glycosylase, J. Biol. Chem. 284 (2009) 9270–9279.
[250] P. Jaruga, Y. Xiao, V. Vartanian, R.S. Lloyd, M. Dizdaroglu, Evidence for the Involvement of DNA Repair Enzyme NEIL1 in Nucleotide Excision Repair of (5'R)- and (5'S)-8,5'-Cyclo-2'-deoxyadenosines, Biochemistry (Mosc.). 49 (2010) 1053–1055.
[251] A.D. Andrews, S.F. Barrett, J.H. Robbins, Xeroderma pigmentosum neurological abnormalities correlate with colony-forming ability after ultraviolet radiation., Proc. Natl. Acad. Sci. U. S. A. 75 (1978) 1984–1988.
[252] J.H. Robbins, A childhood neurodegeneration due to defective DNA repair: a novel concept of disease based on studies xeroderma pigmentosum, J. Child Neurol. 4 (1989) 143–146.
[253] P.J. Brooks, The 8,5'-cyclopurine-2'-deoxynucleosides: Candidate neurodegenerative DNA lesions in xeroderma pigmentosum, and unique probes of transcription and nucleotide excision repair, DNA Repair (Amst.). 7 (2008) 1168–1179.
[254] S.G. Nyaga, P. Jaruga, A. Lohani, M. Dizdaroglu, M.K. Evans, Accumulation of Oxidatively Induced DNA Damage in Human Breast Cancer Cell Lines Following Treatment with Hydrogen Peroxide, Cell Cycle. 6 (2007) 1471–1477.
[255] T.P. Szatrowski, C.F. Nathan, Production of large amounts of hydrogen peroxide by human tumor cells, Cancer Res. 51 (1991) 794–798.
[256] R. Scully, J. Chen, A. Plug, Y. Xiao, D. Weaver, J. Feunteun, et al., Association of BRCA1 with Rad51 in mitotic and meiotic cells, Cell. 88 (1997) 265–275.
[257] M.E. Moynahan, J.W. Chiu, B.H. Koller, M. Jasin, Brca1 controls homology-directed DNA repair, Mol. Cell. 4 (1999) 511–518.
[258] A.R. Venkitaraman, Functions of BRCA1 and BRCA2 in the biological response to DNA damage, J. Cell Sci. 114 (2001) 3591–3598.
[259] F. Le Page, V. Randrianarison, D. Marot, J. Cabannes, M. Perricaudet, J. Feunteun, et al., BRCA1 and BRCA2 are necessary for the transcription-coupled repair of the oxidative 8-oxoguanine lesion in human cells, Cancer Res. 60 (2000) 5548–5552.
[260] A.-R. Hartman, J.M. Ford, BRCA1 induces DNA damage recognition factors and enhances nucleotide excision repair, Nat. Genet. 32 (2002) 180–184.
[261] R. Takimoto, T.K. MacLachlan, D.T. Dicker, Y. Niitsu, T. Mori, W.S. el-Deiry, BRCA1 transcriptionally regulates damaged DNA binding protein (DDB2) in the DNA repair response following UV-irradiation, Cancer Biol. Ther. 1 (2002) 177–186.
[262] S.X. Shen, Z. Weaver, X. Xu, C. Li, M. Weinstein, L. Chen, et al., A targeted disruption of the murine Brca1 gene causes gamma-irradiation hypersensitivity and genetic instability, Oncogene. 17 (1998) 3115–3124.
[263] D.W. Abbott, M.E. Thompson, C. Robinson-Benion, G. Tomlinson, R.A. Jensen, J.T. Holt, BRCA1 expression restores radiation resistance in BRCA1-defective cancer cells

through enhancement of transcription-coupled DNA repair, J. Biol. Chem. 274 (1999) 18808–18812.

[264] H. Rodriguez, P. Jaruga, D. Leber, S.G. Nyaga, M.K. Evans, M. Dizdaroglu, Lymphoblasts of Women with BRCA1 Mutations Are Deficient in Cellular Repair of 8,5'-Cyclopurine-2'-deoxynucleosides and 8-Hydroxy-2'-deoxyguanosine†, Biochemistry (Mosc.). 46 (2007) 2488–2496.

[265] T. Sjöblom, S. Jones, L.D. Wood, D.W. Parsons, J. Lin, T.D. Barber, et al., The consensus coding sequences of human breast and colorectal cancers, Science. 314 (2006) 268–274.

[266] L.D. Wood, D.W. Parsons, S. Jones, J. Lin, T. Sjöblom, R.J. Leary, et al., The genomic landscapes of human breast and colorectal cancers, Science. 318 (2007) 1108–1113.

[267] G. Kirkali, D. Keles, A.E. Canda, C. Terzi, P.T. Reddy, P. Jaruga, et al., Evidence for upregulated repair of oxidatively induced DNA damage in human colorectal cancer, DNA Repair (Amst.). 10 (2011) 1114–1120.

[268] C.B.M. du Puch, E. Barbier, A. Kraut, Y. Couté, J. Fuchs, A. Buhot, et al., TOX4 and its binding partners recognize DNA adducts generated by platinum anticancer drugs, Arch. Biochem. Biophys. 507 (2011) 296–303.

[269] C.B.M. Du Puch, Analyse des interactions ADN lésé / protéines : Optimisations méthodologiques et applications aux dommages de l'ADN engendrés par les dérivés du platine, Université Joseph-Fourier - Grenoble I, 2010.

[270] D. Mittelman, J.H. Wilson, The fractured genome of HeLa cells, Genome Biol. 14 (2013) 111.

[271] C.J. Chetsanga, T. Lindahl, Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from Escherichia coli., Nucleic Acids Res. 6 (1979) 3673–3684.

[272] S. Boiteux, J. Laval, Imidazole open ring 7-methylguanine : An inhibitor of DNA synthesis, Biochem. Biophys. Res. Commun. 110 (1983) 552–558.

[273] E. Gajewski, G. Rao, Z. Nackerdien, M. Dizdaroglu, Modification of DNA bases in mammalian chromatin by radiation-generated free radicals, Biochemistry (Mosc.). 29 (1990) 7876–7882.

[274] S. Boiteux, O. Huisman, Isolation of a formamidopyrimidine-DNA glycosylase (fpg) mutant of Escherichia coli K12, Mol. Gen. Genet. MGG. 215 (1989) 300–305.

[275] J. Tchou, H. Kasai, S. Shibutani, M.H. Chung, J. Laval, A.P. Grollman, et al., 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity., Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 4690–4694.

[276] J.C. Fromme, G.L. Verdine, Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM, Nat. Struct. Biol. 9 (2002) 544–552.

[277] S. Boiteux, T.R. O'Connor, J. Laval, Formamidopyrimidine-DNA glycosylase of Escherichia coli: cloning and sequencing of the fpg structural gene and overproduction of the protein., EMBO J. 6 (1987) 3177–3183.

[278] S. Boiteux, T.R. O'Connor, F. Lederer, A. Gouyette, J. Laval, Homogeneous Escherichia coli FPG protein. A DNA glycosylase which excises imidazole ring-opened purines and nicks DNA at apurinic/apyrimidinic sites., J. Biol. Chem. 265 (1990) 3916–3922.

[279] J. Tchou, M.L. Michaels, J.H. Miller, A.P. Grollman, Function of the zinc finger in Escherichia coli Fpg protein., J. Biol. Chem. 268 (1993) 26738–26744.

[280] T.R. O'Connor, R.J. Graves, G. de Murcia, B. Castaing, J. Laval, Fpg protein of Escherichia coli is a zinc finger protein whose cysteine residues have a structural and/or functional role, J. Biol. Chem. 268 (1993) 9063–9070.

[281] C. Corne, J.-B. Fiche, D. Gasparutto, V. Cunin, E. Suraniti, A. Buhot, et al., SPR imaging for label-free multiplexed analyses of DNA N-glycosylase interactions with damaged DNA duplexes, Analyst. 133 (2008) 1036-1045.

[282] Y.-V.L. Bihan, M.A. Izquierdo, F. Coste, P. Aller, F. Culard, T.H. Gehrke, et al., 5-Hydroxy-5-methylhydantoin DNA lesion, a molecular trap for DNA glycosylases, Nucleic Acids Res. 39 (2011) 6277–6290.

[283] Y. Zivanovic, J. Armengaud, A. Lagorce, C. Leplat, P. Guérin, M. Dutertre, et al., Genome analysis and genome-wide proteomics of Thermococcus gammatolerans, the most radioresistant organism known amongst the Archaea, Genome Biol. 10 (2009) R70.

[284] E. Jolivet, S. L'Haridon, E. Corre, P. Forterre, D. Prieur, Thermococcus gammatolerans sp. nov., a hyperthermophilic archaeon from a deep-sea hydrothermal vent that resists ionizing radiation, Int. J. Syst. Evol. Microbiol. 53 (2003) 847–851.

[285] D.W. Grogan, Stability and repair of DNA in hyperthermophilic Archaea, Curr. Issues Mol. Biol. 6 (2004) 137–144.

[286] D.W. Grogan, The question of DNA repair in hyperthermophilic archaea, Trends Microbiol. 8 (2000) 180–185.

[287] T. Lindahl, Instability and decay of the primary structure of DNA, Nature. 362 (1993) 709–715.

[288] Z. Kelman, M.F. White, Archaeal DNA replication and repair, Curr. Opin. Microbiol. 8 (2005) 669–676.

[289] D.W. Grogan, G.T. Carver, J.W. Drake, Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon Sulfolobus acidocaldarius, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 7928–7933.

[290] J. DiRuggiero, J.R. Brown, A.P. Bogert, F.T. Robb, DNA Repair Systems in Archaea: Mementos from the Last Universal Common Ancestor?, J. Mol. Evol. 49 (1999) 474–484.

[291] C.A. Prail, W.D. Taylor, Responses of Halobacterium halobium and Sulfolobus solfataricus to hydrogen peroxide and N-methyl-N'-nitro-N-nitrosoguanidine [correction of N-methyl-N-nitrosoguanidine] exposure, Microbiol. Res. 152 (1997) 257–260.

[292] M. Skorvaga, N.D. Raven, G.P. Margison, Thermostable archaeal O6-alkylguanine-DNA alkyltransferases, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 6711–6715.

[293] A. Koulis, D.A. Cowan, L.H. Pearl, R. Savva, Uracil-DNA glycosylase activities in hyperthermophilic micro-organisms, FEMS Microbiol. Lett. 143 (1996) 267–271.

[294] T. Lindahl, B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic acid, Biochemistry (Mosc.). 13 (1974) 3405–3410.

[295] C. Rouillon, M.F. White, The evolution and mechanisms of nucleotide excision repair proteins, Res. Microbiol. 162 (2011) 19–26.

[296] E.R. Wood, F. Ghané, D.W. Grogan, Genetic responses of the thermophilic archaeon Sulfolobus acidocaldarius to short-wavelength UV light, J. Bacteriol. 179 (1997) 5693–5698.

[297] A. Tapias, C. Leplat, F. Confalonieri, Recovery of ionizing-radiation damage after high doses of gamma ray in the hyperthermophilic archaeon Thermococcus gammatolerans, Extremophiles 13 (2009) 333–343.

[298] A. Romieu, D. Gasparutto, J. Cadet, Synthesis and Characterization of Oligonucleotides Containing 5',8-Cyclopurine 2'-Deoxyribonucleosides: (5'R)-5',8-Cyclo-2'-deoxyadenosine, (5'S)-5',8-Cyclo-2'-deoxyguanosine, and (5'R)-5',8-Cyclo-2'-deoxyguanosine, Chem. Res. Toxicol. 12 (1999) 412–421.

[299] O. Bogdanovic, G.J.C. Veenstra, Affinity-based enrichment strategies to assay methyl-CpG binding activity and DNA methylation in early Xenopus embryos, BMC Res. Notes. 4 (2011) 300.

[300] N. Hégarat, G.M. Cardoso, F. Rusconi, J.-C. François, D. Praseuth, Analytical biochemistry of DNA–protein assemblies from crude cell extracts, Nucleic Acids Res. 35 (2007) e92.

[301] B. Bernay, M.-C. Gaillard, V. Guryča, A. Emadali, L. Kuhn, A. Bertrand, et al., Discovering New Bioactive Neuropeptides in the Striatum Secretome Using in Vivo Microdialysis and Versatile Proteomics, Mol. Cell. Proteomics. 8 (2009) 946–958.
[302] V. Dupierris, C. Masselon, M. Court, S. Kieffer-Jaquinod, C. Bruley, A toolbox for validation of mass spectrometry peptides identification and generation of database: IRMa, Bioinformatics. 25 (2009) 1980–1981.
[303] S. Li, W.W. Kuhne, A. Kulharya, F.Z. Hudson, K. Ha, Z. Cao, et al., Involvement of p54(nrb), a PSF partner protein, in DNA double-strand break repair and radioresistance, Nucleic Acids Res. 37 (2009) 6746–6753.
[304] Y. Shav-Tal, D. Zipori, PSF and p54nrb/NonO – multi-functional nuclear proteins, FEBS Lett. 531 (2002) 109–114.
[305] M. Salton, Y. Lerenthal, S.-Y. Wang, D.J. Chen, Y. Shiloh, Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response, Cell Cycle. 9 (2010) 1568–1576.
[306] K. Ha, Y. Takeda, W.S. Dynan, Sequences in PSF/SFPQ mediate radioresistance and recruitment of PSF/SFPQ-containing complexes to DNA damage sites in human cells, DNA Repair. 10 (2011) 252–259.
[307] C. Rajesh, D.K. Baker, A.J. Pierce, D.L. Pittman, The splicing-factor related protein SFPQ/PSF interacts with RAD51D and is necessary for homology-directed repair and sister chromatid cohesion, Nucleic Acids Res. 39 (2011) 132–145.
[308] Y. Zhao, H.D. Thomas, M.A. Batey, I.G. Cowell, C.J. Richardson, R.J. Griffin, et al., Preclinical evaluation of a potent novel DNA-dependent protein kinase inhibitor NU7441, Cancer Res. 66 (2006) 5354–5362.
[309] M.A. Kramer, Poly (ADP-ribose) polymerase 1 (PARP1) and its DNA-binding characteristics, COLORADO STATE UNIVERSITY, 2011.
[310] J.-C. Amé, C. Spenlehauer, G. de Murcia, The PARP superfamily, BioEssays News Rev. Mol. Cell. Dev. Biol. 26 (2004) 882–893.
[311] D. D'Amours, S. Desnoyers, I. D'Silva, G.G. Poirier, Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions, Biochem. J. 342 (Pt 2) (1999) 249–268.
[312] N. Ohshima, M. Takahashi, F. Hirose, Identification of a Human Homologue of the DREF Transcription Factor with a Potential Role in Regulation of the Histone H1 Gene, J. Biol. Chem. 278 (2003) 22928–22938.
[313] F. Hirose, M. Yamaguchi, K. Kuroda, A. Omori, T. Hachiya, M. Ikeda, et al., Isolation and characterization of cDNA for DREF, a promoter-activating factor for Drosophila DNA replication-related genes, J. Biol. Chem. 271 (1996) 3930–3937.
[314] Y. Takahashi, F. Hirose, A. Matsukage, M. Yamaguchi, Identification of three conserved regions in the DREF transcription factors from Drosophila melanogaster and Drosophila virilis, Nucleic Acids Res. 27 (1999) 510–516.
[315] T. Nagase, K. Ishikawa, M. Suyama, R. Kikuno, N. Miyajima, A. Tanaka, et al., Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro, DNA Res. Int. J. Rapid Publ. Reports Genes Genomes. 5 (1998) 277–286.
[316] D. Yamashita, Y. Sano, Y. Adachi, Y. Okamoto, H. Osada, T. Takahashi, et al., hDREF Regulates Cell Proliferation and Expression of Ribosomal Protein Genes, Mol. Cell. Biol. 27 (2007) 2003–2013.
[317] D. Gasparutto, E. Muller, S. Boiteux, J. Cadet, Excision of the oxidatively formed 5-hydroxyhydantoin and 5-hydroxy-5-methylhydantoin pyrimidine lesions by Escherichia coli and Saccharomyces cerevisiae DNA N-glycosylases, Biochim. Biophys. Acta BBA - Gen. Subj. 1790 (2009) 16–24.
[318] M. Redrejo-Rodríguez, C. Saint-Pierre, S. Couve, A. Mazouzi, A.A. Ishchenko, D. Gasparutto, et al., New insights in the removal of the hydantoins, oxidation product of

pyrimidines, via the base excision and nucleotide incision repair pathways, PloS One. 6 (2011) e21039.
[319] B. Castaing, J.L. Fourrey, N. Hervouet, M. Thomas, S. Boiteux, C. Zelwer, AP site structural determinants for Fpg specific recognition., Nucleic Acids Res. 27 (1999) 608–615.
[320] J. Tchou, V. Bodepudi, S. Shibutani, I. Antoshechkin, J. Miller, A.P. Grollman, et al., Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA., J. Biol. Chem. 269 (1994) 15318–15324.
[321] B. Castaing, S. Boiteux, C. Zelwer, DNA containing a chemically reduced apurinic site is a high affinity ligand for the E. coli formamidopyrimidine-DNA glycosylase., Nucleic Acids Res. 20 (1992) 389–394.
[322] A. Kish, G. Kirkali, C. Robinson, R. Rosenblatt, P. Jaruga, M. Dizdaroglu, et al., Salt shield: intracellular salts provide cellular protection against ionizing radiation in the halophilic archaeon, Halobacterium salinarum NRC-1, Environ. Microbiol. 11 (2009) 1066–1078.
[323] A. Palud, G. Villani, S. L'Haridon, J. Querellou, J.-P. Raffin, G. Henneke, Intrinsic properties of the two replicative DNA polymerases of Pyrococcus abyssi in replicating abasic sites: possible role in DNA damage tolerance?, Mol. Microbiol. 70 (2008) 746–761.
[324] C. Bounaix Morand du Puch, E. Barbier, S. Sauvaigo, D. Gasparutto, J. Breton, Tools and strategies for DNA damage interactome analysis, Mutat. Res. 752 (2013) 72–83.
[325] C.E. Piersen, R. Prasad, S.H. Wilson, R.S. Lloyd, Evidence for an imino intermediate in the DNA polymerase beta deoxyribose phosphate excision reaction, J. Biol. Chem. 271 (1996) 17811–17815.
[326] R.A. Rieger, E.I. Zaika, W. Xie, F. Johnson, A.P. Grollman, C.R. Iden, et al., Proteomic Approach to Identification of Proteins Reactive for Abasic Sites in DNA, Mol. Cell. Proteomics. 5 (2006) 858–867.
[327] C.X. Zhang, P.V. Chang, S.J. Lippard, Identification of Nuclear Proteins that Interact with Platinum-Modified DNA by Photoaffinity Labeling, J. Am. Chem. Soc. 126 (2004) 6536–6537.
[328] E.R. Guggenheim, D. Xu, C.X. Zhang, P.V. Chang, S.J. Lippard, Photo-Affinity Isolation and Identification of Proteins in Cancer Cell Extracts that Bind to Platinum-Modified DNA, Chembiochem Eur. J. Chem. Biol. 10 (2009) 141–157.
[329] G. Zhu, S.J. Lippard, Photoaffinity Labeling Reveals Nuclear Proteins that Uniquely Recognize Cisplatin-DNA Interstrand Cross-Links, Biochemistry (Mosc.). 48 (2009) 4916–4925.
[330] G. Zhu, P. Chang, S.J. Lippard, Recognition of Platinum-DNA Damage by Poly(ADP-Ribose) Polymerase-1, Biochemistry (Mosc.). 49 (2010) 6177–6183.
[331] E.A. Maltseva, N.I. Rechkunova, L.C. Gillet, I.O. Petruseva, O.D. Schärer, O.I. Lavrik, Crosslinking of the NER damage recognition proteins XPC-HR23B, XPA and RPA to photoreactive probes that mimic DNA damages, Biochim. Biophys. Acta BBA - Gen. Subj. 1770 (2007) 781–789.
[332] R.D. Wood, DNA repair in eukaryotes, Annu. Rev. Biochem. 65 (1996) 135–167.
[333] A. Sancar, DNA excision repair, Annu. Rev. Biochem. 65 (1996) 43–81.
[334] O.D. Schärer, Chemistry and biology of DNA repair, Angew. Chem. Int. Ed Engl. 42 (2003) 2946–2974.
[335] L.C.J. Gillet, O.D. Schärer, Molecular mechanisms of mammalian global genome nucleotide excision repair, Chem. Rev. 106 (2006) 253–276.
[336] T. Buterin, C. Meyer, B. Giese, H. Naegeli, DNA quality control by conformational readout on the undamaged strand of the double helix, Chem. Biol. 12 (2005) 913–922.

[337] T.J. Jorgensen, K. Chen, S. Chasovskikh, R. Roy, A. Dritschilo, A. Uren, Binding kinetics and activity of human poly(ADP-ribose) polymerase-1 on oligo-deoxyribonucleotide substrates, J. Mol. Recognit. JMR. 22 (2009) 446–452.
[338] I. Lonskaya, V.N. Potaman, L.S. Shlyakhtenko, E.A. Oussatcheva, Y.L. Lyubchenko, V.A. Soldatenkov, Regulation of Poly(ADP-ribose) Polymerase-1 by DNA Structure-specific Binding, J. Biol. Chem. 280 (2005) 17076–17083.
[339] M.-F. Langelier, J.L. Planck, S. Roy, J.M. Pascal, Crystal structures of poly(ADP-ribose) polymerase-1 (PARP-1) zinc fingers bound to DNA: structural and functional insights into DNA-dependent PARP-1 activity, J. Biol. Chem. 286 (2011) 10690–10701.
[340] A.E.O. Fisher, H. Hochegger, S. Takeda, K.W. Caldecott, Poly(ADP-Ribose) Polymerase 1 Accelerates Single-Strand Break Repair in Concert with Poly(ADP-Ribose) Glycohydrolase, Mol. Cell. Biol. 27 (2007) 5597–5605.
[341] A.M. Ferro, B.M. Olivera, Poly(ADP-ribosylation) in vitro. Reaction parameters and enzyme mechanism, J. Biol. Chem. 257 (1982) 7808–7813.
[342] P. Zahradka, K. Ebisuzaki, A shuttle mechanism for DNA-protein interactions. The regulation of poly(ADP-ribose) polymerase, Eur. J. Biochem. FEBS. 127 (1982) 579–585.
[343] E.R. Guggenheim, A.E. Ondrus, M. Movassaghi, S.J. Lippard, Poly(ADP-ribose) polymerase-1 activity facilitates the dissociation of nuclear proteins from platinum-modified DNA, Bioorg. Med. Chem. 16 (2008) 10121–10128.
[344] V. Schreiber, F. Dantzer, J.-C. Ame, G. de Murcia, Poly(ADP-ribose): novel functions for an old molecule, Nat. Rev. Mol. Cell Biol. 7 (2006) 517–528.
[345] Z.S. Juo, T.K. Chiu, P.M. Leiberman, I. Baikalov, A.J. Berk, R.E. Dickerson, How proteins recognize the TATA box, J. Mol. Biol. 261 (1996) 239–254.
[346] C. Marietta, H. Gulam, P.J. Brooks, A single 8,5'-cyclo-2'-deoxyadenosine lesion in a TATA box prevents binding of the TATA binding protein and strongly reduces transcription in vivo, DNA Repair (Amst.). 1 (2002) 967–975.
[347] J. Abraham, P.J. Brooks, Divergent effects of oxidatively induced modification to the C8 of 2'- deoxyadenosine on transcription factor binding: 8,5'(S)- cyclo- 2'- deoxyadenosine inhibits the binding of multiple sequence specific transcription factors, while 8- oxo- 2'- deoxyadenosine increases binding of CREB and NF- kappa B to DNA, Environ. Mol. Mutagen. 52 (2011) 287–295.
[348] F. Forneris, C. Binda, M.A. Vanoni, A. Mattevi, E. Battaglioli, Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process, FEBS Lett. 579 (2005) 2203–2207.
[349] B. Perillo, M.N. Ombra, A. Bertoni, C. Cuozzo, S. Sacchetti, A. Sasso, et al., DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression, Science. 319 (2008) 202–206.
[350] D. Yamashita, H. Komori, Y. Higuchi, T. Yamaguchi, T. Osumi, F. Hirose, Human DNA replication-related element binding factor (hDREF) self-association via hATC domain is necessary for its nuclear accumulation and DNA binding, J. Biol. Chem. 282 (2007) 7563–7575.
[351] F. Hirose, M. Yamaguchi, H. Handa, Y. Inomata, A. Matsukage, Novel 8-base pair sequence (Drosophila DNA replication-related element) and specific binding factor involved in the expression of Drosophila genes for DNA polymerase alpha and proliferating cell nuclear antigen, J. Biol. Chem. 268 (1993) 2092–2099.
[352] M. Yamaguchi, Y. Hayashi, Y. Nishimoto, F. Hirose, A. Matsukage, A nucleotide sequence essential for the function of DRE, a common promoter element for Drosophila DNa replication-related genes, J. Biol. Chem. 270 (1995) 15808–15814.
[353] Y. Takahashi, M. Yamaguchi, F. Hirose, S. Cotterill, J. Kobayashi, S. Miyajima, et al., DNA replication-related elements cooperate to enhance promoter activity of the

drosophila DNA polymerase alpha 73-kDa subunit gene, J. Biol. Chem. 271 (1996) 14541–14547.

[354] T. Sawado, F. Hirose, Y. Takahashi, T. Sasaki, T. Shinomiya, K. Sakaguchi, et al., The DNA replication-related element (DRE)/DRE-binding factor system is a transcriptional regulator of the Drosophila E2F gene, J. Biol. Chem. 273 (1998) 26042–26051.

[355] F. Hirose, N. Ohshima, M. Shiraki, Y.H. Inoue, O. Taguchi, Y. Nishi, et al., Ectopic expression of DREF induces DNA synthesis, apoptosis, and unusual morphogenesis in the Drosophila eye imaginal disc: possible interaction with Polycomb and trithorax group proteins, Mol. Cell. Biol. 21 (2001) 7231–7242.

[356] J. Hyun, H. Jasper, D. Bohmann, DREF is required for efficient growth and cell cycle progression in Drosophila imaginal discs, Mol. Cell. Biol. 25 (2005) 5590–5598.

[357] Y. Jung, Y. Mikata, S.J. Lippard, Kinetic Studies of the TATA-binding Protein Interaction with Cisplatin-modified DNA „ J. Biol. Chem. 276 (2001) 43589–43596.

[358] S.J. Edelstein, N. Le Novère, Cooperativity of allosteric receptors, J. Mol. Biol. 425 (2013) 1424–1432.

[359] D.O. Zharkov, G. Shoham, A.P. Grollman, Structural characterization of the Fpg family of DNA glycosylases, DNA Repair. 2 (2003) 839–862.

[360] A.A. Ishchenko, N.V. Bulychev, G.A. Maksakova, F. Johnson, G.A. Nevinskiĭ, [Interaction of Escherichia coli 8-oxoguanine DNA glycosylase with single-stranded oligodeoxyribonucleotides and their complexes], Mol. Biol. (Mosk.). 32 (1998) 549–558.

[361] A.A. Ishchenko, N.L. Vasilenko, O.I. Sinitsina, V.I. Yamkovoy, O.S. Fedorova, K.T. Douglas, et al., Thermodynamic, Kinetic, and Structural Basis for Recognition and Repair of 8-Oxoguanine in DNA by Fpg Protein from Escherichia coli†, Biochemistry (Mosc.). 41 (2002) 7540–7548.

[362] E. Muller, D. Gasparutto, B. Castaing, A. Favier, J. Cadet, Recognition of Cyclonucleoside Lesions by the Lactococcus lactis FPG Protein, Nucleosides Nucleotides Nucleic Acids. 22 (2003) 1563–1565.

[363] P. Pande, R.S. Das, C. Sheppard, Y.W. Kow, A.K. Basu, Repair efficiency of (5'S)-8,5'-cyclo-2'-deoxyguanosine and (5'S)-8,5'-cyclo-2'-deoxyadenosine depends on the complementary base, DNA Repair. (n.d.).

[364] F. Coste, M. Ober, Y.-V. Le Bihan, M.A. Izquierdo, N. Hervouet, H. Mueller, et al., Bacterial Base Excision Repair Enzyme Fpg Recognizes Bulky N7-Substituted-FapydG Lesion via Unproductive Binding Mode, Chem. Biol. 15 (2008) 706–717.

[365] C.J. Wiederholt, M.O. Delaney, M.A. Pope, S.S. David, M.M. Greenberg, Repair of DNA Containing Fapy·dG and Its β-C-Nucleoside Analogue by Formamidopyrimidine DNA Glycosylase and MutY†, Biochemistry (Mosc.). 42 (2003) 9755–9760.

[366] M.V. Rogacheva, M.K. Saparbaev, I.M. Afanasov, S.A. Kuznetsova, Two sequential phosphates 3' adjacent to the 8-oxoguanosine are crucial for lesion excision by E. coli Fpg protein and human 8-oxoguanine-DNA glycosylase, Biochimie. 87 (2005) 1079–1088.

[367] S. Frelon, T. Douki, J.L. Ravanat, J.P. Pouget, C. Tornabene, J. Cadet, High-performance liquid chromatography--tandem mass spectrometry measurement of radiation-induced base damage to isolated and cellular DNA, Chem. Res. Toxicol. 13 (2000) 1002–1010.

[368] T. Douki, J.-L. Ravanat, J.-P. Pouget, I. Testard, J. Cadet, Minor contribution of direct ionization to DNA base damage inducedby heavy ions, Int. J. Radiat. Biol. 82 (2006) 119–127.

[369] S. Tahara, T. Kaneko, Susceptibility of Mouse Splenic Cells to Oxidative DNA Damage by X-Ray Irradiation, Biol. Pharm. Bull. 27 (2004) 105–108.

[370] T. Kaneko, S. Tahara, M. Tanno, T. Taguchi, Effect of age on the induction of 8-oxo-2'-deoxyguanosine-releasing enzyme in rat liver by γ-ray irradiation, Arch. Gerontol. Geriatr. 36 (2003) 23–35.

[371] W. Sudprasert, P. Navasumrit, M. Ruchirawat, Effects of low-dose gamma radiation on DNA damage, chromosomal aberration and expression of repair genes in human blood cells, Int. J. Hyg. Environ. Health. 209 (2006) 503–511.
[372] A.A. Sartori, G.M. Lingaraju, P. Hunziker, F.K. Winkler, J. Jiricny, Pa-AGOG, the founding member of a new family of archaeal 8-oxoguanine DNA-glycosylases, Nucleic Acids Res. 32 (2004) 6531–6539.
[373] F. Faucher, S. Doublié, Z. Jia, 8-oxoguanine DNA glycosylases: one lesion, three subfamilies, Int. J. Mol. Sci. 13 (2012) 6711–6729.
[374] V. Thiviyanathan, A. Somasunderam, T.K. Hazra, S. Mitra, D.G. Gorenstein, Solution structure of a DNA duplex containing 8-hydroxy-2'-deoxyguanosine opposite deoxyguanosine, J. Mol. Biol. 325 (2003) 433–442.
[375] R. Fujikane, S. Ishino, Y. Ishino, P. Forterre, Genetic analysis of DNA repair in the hyperthermophilic archaeon, Thermococcus kodakaraensis, Genes Genet. Syst. 85 (2010) 243–257.
[376] P. Regulus, B. Duroux, P.-A. Bayle, A. Favier, J. Cadet, J.-L. Ravanat, Oxidation of the sugar moiety of DNA by ionizing radiation or bleomycin could induce the formation of a cluster DNA lesion, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 14032–14037.

Annexes

Tableau A 1 : Protéines retenues uniquement sur les sondes lésées ou non-lésées, dans les trois analyses protéomiques. Si plusieurs sondes lésées ont été utilisées dans une analyse, seules les protéines ayant été retenues sur toutes les sondes lésées en même temps ont été prises en compte. Pour chaque protéine, le nombre de peptides ayant servi à son identification est donné, suivi de son score Mascot. Pour les protéines présentes sur toutes les sondes lésées d'une analyse, ces chiffres sont donnés dans l'ordre des sondes indiqué dans le tableau. Les protéines identifiées avec moins de 2 peptides ou un score Mascot inférieur à 70 ont été écartées de l'analyse. Pour toutes les protéines présentées, les résultats de leur identification dans les deux autres analyses sont également indiqués avec les données de leur identification sur toutes les sondes où elles sont présentes. (*) = protéine non-retrouvée dans la base de données Uniprot, son numéro d'accession correspond à une autre base de données

entrée (Uniprot)	nom de la protéine	nom du gène	1ère analyse		2ème analyse		3ème analyse	
			5'S-cyclo-dG	non-lésé	5'R-cyclo-dA ET 5'S-cyclo-dG	non-lésé	5'S-cyclo-dA, 5'R-cyclo-dA, 5'S-cyclo-dG ET 5'S-cyclo-dHT	non-lésé
K1C18_HUMAN	Keratin, type I cytoskeletal 18	KRT18	28/845.50	---	5'S-cyclo-dG : 28/889.42 non-lésé : 29/886.70 billes : 28/863.05		absente de l'analyse	
NONO_HUMAN	Non-POU domain-containing octamer-binding protein	NONO	16/400.75	---	5'R-cyclo-dA : 9/254.55 5'S-cyclo-dG : 11/361.55 non-lésé : 8/210.43 billes : 20/500.74		5'S-cyclo-dA : 1/38.66 5'R-cyclo-dA : 1/41.28 5'S-cyclo-dHT : 2/82.88	
K1C14_HUMAN	Keratin, type I cytoskeletal 14	KRT14	7/274.08	---	billes : 12/373.65		5'S-cyclo-dG : 13/417.04 5'S-cyclo-dHT : 13/470.19 non-lésé : 26/748.77 billes : 27/759.80	
COF1_HUMAN	Cofilin-1	CFL1	2/93.75	---	absente de l'analyse		absente de l'analyse	
DDX3X_HUMAN	ATP-dependent RNA helicase DDX3X	DDX3X	2/77.79	---	absente de l'analyse		absente de l'analyse	
MYO6_HUMAN	Unconventional myosin-VI	MYO6	5/201.32	---	5'R-cyclo-dA : 5/202.12 5'S-cyclo-dG : 4/166.65 non-lésé : 2/41.84 billes : 4/197.26		1/43.02 ; 2/63.77 ; 2/57.12 ; 2/74.79	---
G3P_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	4/118.61	---	5'S-cyclo-dG : 8/320.07 non-lésé : 6/283.33 billes : 2/124.38		5'S-cyclo-dA : 3/115.46 5'S-cyclo-dG : 1/43.18 billes : 3/119.14	
LIMA1_HUMAN	LIM domain and actin-binding protein 1	LIMA1	3/109.85	---	5'R-cyclo-dA : 7/215.89 5'S-cyclo-dG : 11/360.30 non-lése : 6/170.11 billes : 10/336.99		5'S-cyclo-dA : 6/142.73 5'R-cyclo-dA : 9/315.46 5'S-cyclo-dG : 4/116.64 5'S-cyclo-dHT : 11/318.88 non-lésé : 3/140.35	
NEST_HUMAN	Nestin	NES	2/72.33	---	5'S-cyclo-dG : 2/60.30 billes : 1/55.92		5'S-cyclo-dA : 5/191.98 5'R-cyclo-dA : 8/224.27 5'S-cyclo-dG : 1/51.32	
K1C10_HUMAN	Keratin, type I cytoskeletal 10	KRT10	---	26/749.31	5'R-cyclo-dA : 21/811.33		5'S-cyclo-dG : 40/1200.93	

					5'S-cyclo-dG : 29/947.79 non-lésé : 35/1179.46 billes : 39/1156.34		non-lésé : 42/1174.81	
Q96DE1_HUMAN	Putative uncharacterized protein	---	---	50/699.06	absente de l'analyse		absente de l'analyse	
PARP1_HUMAN	Poly [ADP-ribose] polymerase 1	PARP1	---	15/477.19	5'S-cyclo-dG : 2/73.07 non-lésé : 10/380.47		---	8/257.61
TBB3_HUMAN	Tubulin, beta 3	TUBB3	---	15/405.86	---	9/302.89	absente de l'analyse	
COBL_HUMAN	Protein cordon-bleu	COBL	---	3/118.03	5'S-cyclo-dG : 2/76.15 non-lésé : 2/63.14 billes : 3/121.82		absente de l'analyse	
ZBED1_HUMAN	Zinc finger BED domain-containing protein 1	ZBED1 (DREF)	---	6/244.82	---	11/386.37	---	8/287.48
HNRPC_HUMAN	Heterogeneous nuclear ribonucleoproteins C1/C2	HNRNPC	---	2/73.32	5'S-cyclo-dG : 1/74.39		absente de l'analyse	
FLNA_HUMAN	Filamin-A	FLNA	7/271.47	10/323.40	2/91.68 ; 7/608.19	---	absente de l'analyse	
K22E_HUMAN	Keratin, type II cytoskeletal 2 epidermal	KRT2	5'S-cyclo-dG : 22/667.97 non-lésé : 17/660.22 billes : 22/697.91		9/307.26 ; 11/408.87	---	5'R-cyclo-dA : 44/1301.51 5'S-cyclo-dG : 37/1204.74 5'S-cyclo-dHT: 30/1165.68	
K1C13_HUMAN	Keratin, type I cytoskeletal 13	KRT13	absente de l'analyse		2/115.35 ; 6/208.48	---	absente de l'analyse	
TBA1A_HUMAN	Tubulin alpha-1A chain	TUBA1A	absente de l'analyse		---	15/559.6	absente de l'analyse	
LAP2B_HUMAN	Lamina-associated polypeptide 2, isoforms beta/gamma	TMPO	absente de l'analyse		---	11/401.48	5'S-cyclo-dG : 7/256.44 non-lésé : 34/1058.09	
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich	SFPQ	5'S-cyclo-dG : 21/521.54 non-lésé : 14/363.06 billes : 11/296.57		---	10/327.41	5'S-cyclo-dA : 2/62.78 5'R-cyclo-dA : 2/97.61 5'S-cyclo-dG: 1/51.65 5'S-cyclo-dHT : 2/39.49 billes : 2/46.48	
IQGA1_HUMAN	Ras GTPase-activating-like protein IQGAP1	IQGAP1	5'S-cyclo-dG : 9/338.70 non-lésé : 15/461.25 billes : 2/104.16		5'R-cyclo-dA : 12/384.07 5'S-cyclo-dG : 18/676.80 non-lésé: 16/495.24 billes: 16/437.74		7/221.26 ; 11/382.12 ; 9/300.84 ; 12/430.47	---
TBA4A_HUMAN	Tubulin alpha-4A chain	TUBA4A	absente de l'analyse		absente de l'analyse		7/279.04 ; 9/225.17 ; 11/305.61 ; 9/317.94	---
B4DW52_HUMAN	cDNA FLJ55253, highly similar to Actin, cytoplasmic 1		absente de l'analyse		absente de l'analyse		---	34/653.40
TOP1_HUMAN	DNA topoisomerase 1	TOP1	---	2/54.64	absente de l'analyse		---	12/402.69
dbj|BAJ20861.1 (*)	Histone cluster 1, H1c		absente de l'analyse		absente de l'analyse		---	8/219.57
prf||1208320A (*)	Histone H1b		absente de l'analyse		absente de l'analyse		---	8/219.57
RRBP1_HUMAN	Ribosome-binding protein 1	RRBP1	absente de l'analyse		absente de l'analyse		---	2/72.66

Liste des communications et des publications :

Posters :

Ewa Barbier, Christophe Bounaix Morand du Puch, Stéphane Mouret, Thierry Douki, Coralie Philippe, Anne Forestier, Edwige Guais, Alexandra Kraut, Yohann Couté, Didier Gasparutto, Philippe Becuwe, Jean Breton
« Insight into bulky DNA lesions recognition and repair by DDB2. »
CLARA 6th Scientific Forum, mars 2011 – Lyon, France

Ewa Barbier, Christine Saint-Pierre, Didier Gasparutto, Jean Breton
« Study of protein / cyclonucleosides interactions. Preliminary results obtained with the Fpg enzyme. »
CERLIB 2012, mars 2012 – les Houches, France

Communications orales :

Ewa Barbier, Christophe Bounaix Morand du Puch, Lucie Morand, Stéphane Mouret, Alexandra Kraut, Yohann Couté, Thierry Livache, Didier Gasparutto, Thierry Douki, Philippe Becuwe, Jean Breton
« Contribution of Damaged-DNA Binding Protein 2 (DDB2) to recognition and repair of bulky DNA adducts generated by UV-B and platinating agents. »
41st EEMS Meeting, juillet 2011 – Barcelone, Espagne

Ewa Barbier, Christine Saint-Pierre, Didier Gasparutto, Jean Breton
« Etude des interactions protéines / cyclonucléosides. Résultats préliminaires obtenus à partir de l'enzyme Fpg. »
Forum de la SFTG, mai 2012 – Rennes, France

Publications :

Bounaix Morand du Puch C, Barbier E, Kraut A, Couté Y, Fuchs J, Buhot A, Livache T, Sève M, Favier A, Douki T, Gasparutto D, Sauvaigo S, Breton J.
TOX4 and its binding partners recognize DNA adducts generated by platinum anticancer drugs.
Arch Biochem Biophys. 2011 Mar 15;507(2):296-303

Bounaix Morand du Puch C, Barbier E, Sauvaigo S, Gasparutto D, Breton J.
Tools and strategies for DNA damage interactome analysis.
Mutat Res. 2013 Apr-Jun;752(2):72-83

Céline Lafaye, Ewa Barbier, Audrey Miscioscia, Christine Saint-Pierre, Alexandra Kraut, Yoann Couté, Isabelle Plo, Didier Gasparutto, Jean-Luc Ravanat, Jean Breton
A DNA pull-down assay coupled to proteomics demonstrates that DNA binding of the p21 repressor ZBTB2 is inhibited by cytosine hydroxymethylation.
(en cours de publication)

Archives of Biochemistry and Biophysics 507 (2011) 296–303

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

TOX4 and its binding partners recognize DNA adducts generated by platinum anticancer drugs

Christophe Bounaix Morand du Puch [a], Ewa Barbier [a], Alexandra Kraut [b], Yohann Couté [b], Julia Fuchs [c], Arnaud Buhot [c], Thierry Livache [c], Michel Sève [d], Alain Favier [a], Thierry Douki [a], Didier Gasparutto [a], Sylvie Sauvaigo [a], Jean Breton [a,*]

[a] CEA Grenoble, INAC, SCIB (UMR E_3 CEA-Université Joseph Fourier, CNRS FRE3200) – Laboratoire Lésions des Acides Nucléiques, 17 rue des Martyrs, 38054 Grenoble cedex 09, France
[b] CEA Grenoble, iRTSV, Laboratoire d'Étude de la Dynamique des Protéomes (INSERM U880/Université Joseph Fourier), 17 rue des Martyrs, 38054 Grenoble cedex 09, France
[c] CEA Grenoble, INAC, SPrAM (UMR 5819 CNRS-CEA-UJF), 17 rue des Martyrs, 38054 Grenoble cedex 09, France
[d] INSERM U823, Institut A. Bonniot, BP 170, 38 042 Grenoble cedex 9, France

ARTICLE INFO

Article history:
Received 20 September 2010
and in revised form 12 December 2010
Available online 22 December 2010

Keywords:
Platinum anticancer drugs
PNUTS
TOX4
WDR82
SPRi
DNA lesions

ABSTRACT

Platinating agents are commonly prescribed anticancer drugs damaging DNA. Induced lesions are recognized by a wide range of proteins. These are involved in cellular mechanisms such as DNA repair, mediation of cytotoxicity or chromatin remodeling. They therefore constitute crucial actors to understand pharmacology of these drugs. To expand our knowledge about this subproteome, we developed a ligand fishing trap coupled to high throughput proteomic tools. This trap is made of damaged plasmids attached to magnetic beads, and was exposed to cell nuclear extracts. Retained proteins were identified by nano-HPLC coupled to tandem mass spectrometry. This approach allowed us to establish a list of 38 proteins interacting with DNA adducts generated by cisplatin, oxaliplatin and satraplatin. Some of them were already known interactome members like high mobility group protein 1 (HMGB1) or the human upstream binding factor (hUBF), but we also succeeded in identifying unexpected proteins such as TOX HMG box family member 4 (TOX4), phosphatase 1 nuclear targeting subunit (PNUTS), and WD repeat-containing protein 82 (WDR82), members of a recently discovered complex. Interaction between TOX4 and platinated DNA was subsequently validated by surface plasmon resonance imaging (SPRi). These interactions highlight new cellular responses to DNA damage induced by chemotherapeutic agents.
© 2010 Elsevier Inc. All rights reserved.

Introduction

Cisplatin was the first member of platinum-based compounds approved as an anticancer drug in 1978 [1]. This molecule is still widely used to treat patients suffering from tumoral diseases such as testicular, upper aero-digestive tract, lung or ovarian cancers. In order to circumvent adverse side effects or to enlarge the panel of targeted tumors, further research led to the development and approbation of carboplatin and oxaliplatin, the latter being often used in chemotherapeutic regimen to treat colorectal tumors. Hundreds of other platinating agents are still in the pipelines at research or clinical trial stages, including compounds such as satraplatin or picoplatin [2]. Following intracellular activation, the main mechanism of action of these molecules consists in forming DNA lesions such as intra- and interstrand crosslinks (ICL)[1], the most common being 1,2-d(GpG) and 1,2-d(ApG) intrastrand adducts [3]. These damages impair transcription and replication, and generally cause a cell-cycle arrest to provide a time frame for DNA repair. When the damage extent exceeds repair capacities, apoptosis is induced. These cellular and molecular events involve a wide range of proteins that interact with DNA lesions or with subsequent DNA distortions and constitute the "platinated DNA interactome". These include damage-response proteins (DRP) but also other factors whose ordinary targets are not DNA lesions.

Xeroderma pigmentosum complementation group C (XPC) is an important member of this subproteome as it binds to platinum adducts during the recognition step of the global genome repair

[1] *Abbreviations used:* CPD, cyclobutane pyrimidine dimers; DRP, damage-response protein; DSB, double-strand breaks; HMG, high mobility group; HMGB1, high mobility group protein B1; ICL, interstrand crosslinks; LCP1, epidermal Langerhans cell protein 1; NER, nucleotide excision repair; PNUTS, phosphatase 1 nuclear targeting subunit; SPRi, surface plasmon resonance imaging; TOX4, TOX HMG box family member 4; TRF2, telomeric repeat factor 2; hUBF, human upstream binding factor; WDR82, WD repeat-containing protein 82; (6-4)-PP, (6-4)-photoproducts.

* Corresponding author. Fax: +33 438785090.
E-mail address: jean.breton@cea.fr (J. Breton).

0003-9861/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2010.12.021

subpathway of nucleotide excision repair (NER). Among the other well-studied members also stand the high mobility group (HMG) proteins, whose affinity for platinum adducts has been known for years [4]. Indeed, HMGB1 was one of the first proteins identified as being able to interact with platinated DNA [5]. The biological repercussions of HMGB1 binding are still controversial, as some published works demonstrated a subsequent facilitation of DNA repair whereas others suggest a shielding effect towards NER factors [6]. A second example is hUBF. This ribosomal RNA transcription factor is hijacked from its natural binding site by cisplatin lesions, for which it demonstrates a strong affinity [7]. Lists of other proteins already characterized as capable of recognizing DNA platinum adducts are accessible in dedicated reviews [8]. Discovery of these protein–DNA interactions already brought crucial data for the understanding of platinating agents pharmacology. In order to establish the cartography of damaged DNA-interacting proteins, protocols using proteomic tools were recently designed for DNA lesions such as double-strand breaks [9], abasic sites [10], oxidative damage [11] and DNA crosslinks generated by platinating agents [12,13]. The aim of our work was to optimize one of these approaches combining a ligand fishing system and high throughput proteomics, to allow us to trap and identify new proteins interacting with DNA alterations. Our tool is made of plasmids damaged by one of three different platinum drugs (cisplatin, oxaliplatin or the satraplatin metabolite JM118) and immobilized onto magnetic beads. This system was exposed to HeLa (cervical cancer cell line) and MDA-MB231 (breast cancer cell line) nuclear extracts, and retained proteins were identified by proteomics. This strategy allowed selection of relevant candidates that were validated by immunoassays and SPRi. Identified proteins may improve our understanding of molecular and cellular responses to this particular type of anticancer drugs.

Materials and methods

Protein sources

Two different protein sources were used in our study. HeLa nuclear extracts were purchased from CIL Biotech (Mons, Belgium). MDA-MB231 extracts were prepared in our laboratory. This cell line was a gift from Prof. Philippe Becuwe (University of Nancy, France). Cells were grown in RPMI 1640 medium supplemented with 2 mM L-glutamine, 10% (v/v) fetal calf serum and 800 μg/ml geneticin (Invitrogen) at 37 °C under 5% CO_2. Cells were passaged or harvested at ~80% confluence. Nuclear extracts were prepared as previously described [14].

Ligand fishing

Probe synthesis

Principle for this step was adapted from [15]. Plasmid pBlue-Script II SK (−) (Stratagene) was linearized using restriction enzymes ClaI and BamHI (NEB). The larger fragment obtained was purified with Quick Spin G-50 Sephadex columns (Roche), precipitated and resuspended in water.

Cohesive ends obtained through ClaI and BamHI restrictions allowed 3′ labeling of the plasmids by biotinylated dATP using Klenow polymerase (NEB), biotin-14-dATP (Invitrogen). Alpha-thio 2′dNTPs (Trilink biotechnologies) or normal dNTPs (GE Healthcare) were added to fill in the remaining part of the overhangs. Biotinylated DNA was purified by isopropanol/sodium acetate precipitation.

Biotinylated plasmids (100 μg) were exposed at 37 °C in the dark to either 67 μM cisplatin (Bayer Pharma) for 2 h, 200 μM oxaliplatin (Sanofi Aventis) for 4 h, or 67 μM JM118 (kindly provided by Agennix) for 2 h. Probes were then precipitated, resuspended in water and stored at −20 °C. These conditions generated homogenous levels of adducts: 7 1,2-d(ApG) and 35 1,2-d(GpG) crosslinks/plasmid for cisplatin, 2 1,2-d(ApG) and 22 1,2-d(GpG) crosslinks/plasmid for oxaliplatin, 5 1,2-d(ApG) and 31 1,2-d(GpG) crosslinks/plasmid for JM118. UVC-damaged plasmid probes were also employed in our study. These probes were prepared with biotinylated plasmids following a method previously described [16], which allowed the generation of 12 (6–4)-photoproducts ((6–4)-PP) and 26 cyclobutane pyrimide dimers (CPD) per plasmid. Adduct levels were measured by HPLC-MS/MS (1100 Agilent chromatographic system coupled to SCIEX API 3000 electrospray triple quadrupolar spectrometer, multiple reaction monitoring analysis) after plasmid digestion to obtain a mix of normal deoxynucleosides and monophosphate adducts. The amount of each lesion was determined by using an external calibration obtained by the injection of known amounts of the authentic compound. Normal nucleosides were quantified by HPLC-UV with a detector set at 260 nm placed at the outlet of the column. Representative spectra for both UV and mass detections are available in Supplementary Fig. 1.

Probes were finally attached to polystyrene magnetic beads (Dynabeads kilobaseBINDER, Invitrogen) according to the manufacturer instructions, using 500 μg beads for 10 μg damaged plasmids. Functionalised beads were finally rinsed twice with water, and then stored in water at 4 °C.

Ligand fishing

Buffer A (final concentrations: 40 mM Hepes-KOH pH7.8, 50 mM KCl, 5 mM $MgCl_2$, 0.5 mM DTT) was used to equilibrate functionalised beads. In a 100 μL final volume, 1.2 μg of DNA probes, 120 μg of fish sperm DNA (Roche), and 100 μg of nuclear extracts were mixed in buffer A. Samples were incubated under shaking for 45 min at 37 °C. Supernatant was eliminated using a magnetic tray, and beads were rinsed twice with 100 μL of rinsing buffer: 1 M (high stringency), 150 mM (intermediate stringency) or 50 mM (low stringency) NaCl, 20 mM Hepes-KOH pH 7.8, 5 mM $MgCl_2$, 0.025% Nonidet-P-40. Proteins were finally released in 20 μL of Laemmli buffer (Tris–HCl pH 6.8 25 mM, SDS 2%, glycerol 4%, β-mercaptoethanol 5%, bromophenol blue 0.0008%) by heating at 90 °C for 10 min.

Proteomic analyses

Proteins contained in ligand fishing eluates were stacked by SDS–PAGE in a 2 mm-high band and stained with Coomassie blue. Protein bands were manually excised from the gels, proteins digested in-gel with trypsin, and resulting peptides extracted using a sample preparation robot (EVO150, Tecan) as previously described [17]. Extracted peptides were analyzed by nano-liquid chromatography coupled to tandem mass spectrometry (Ultimate 3000, Dionex and LTQ-Orbitrap Discovery, Thermo Fischer) and identified using Mascot software (Matrix Science) and the Uniprot database (*Homo sapiens* taxonomy) as described in [18]. Proteins identified with a minimum score of 35 were automatically validated using IRMa software [19]. Proteins showing a score $\geqslant 70$ and a number of identified peptides $\geqslant 2$ are provided in Supplementary Table 1. Functions of identified proteins were assessed using Generic GO Term Mapper (http://go.princeton.edu).

Western blot

Proteins obtained by ligand fishing were separated by SDS–PAGE, transferred to nitrocellulose and analyzed by Western blot using the following monoclonal antibodies: rabbit anti-HMGB1 (Abcam), rabbit anti-PNUTS (Abcam), rabbit anti-TOX4 (Abcam) and mouse anti-β actin (Sigma). Secondary antibodies were monkey anti-rabbit and anti-mouse coupled to horseradish-peroxydase

Table 1
Sequences of oligonucleotides utilized to construct the SPRi probes.

Target		5′-CCT CTC TGG ACC TTC C-3′
Zip probes	Zip4	5′-Py-(5T)-TGC GGG TAC AGC ACC TAC CTT GCG-3′
	Zip6	5′-Py-(5T)-GAC CGG TAT GCG ACC TGG TAT GCG-3′
	Zip7	5′-Py-(5T)-TGC GAT CGC AGC GGT AAC CTG ACC-3′
	Zip9	5′-Py-(5T)-GAC CAT CGT GCG GGT AGG TAG ACC-3′
Complementary sequences	Zip4c	5′-GGA AGG TCC AGA GAG GCG CAA GGT AGG TGC TGT ACC CGC A-3′
	Zip6c	5′-GGA AGG TCC AGA GAG GCG CAT ACC AGG TCG CAT ACC GGT C-3′
	Zip7c	5′-GGA AGG TCC AGA GAG GGG TCA GGT TAC CGC TGC GAT CGC A-3′
	Zip9c	5′-GGA AGG TCC AGA GAG GGG TCT ACC TAC CCG CAC GAT GGT C-3′

(GE Healthcare). ECL Plus and Hyperfilm ECL (GE Healthcare) were used to detect the bands.

SPRi

DNA probe synthesis and control

The design of SPRi ODN probes is illustrated on Fig. 3. Sequences employed are detailed in Table 1. Four Zip ODNs were grafted on the chip, allowing localization of each type of short damaged or control ODNs on specific areas of the gold layer (DNA damaged with cisplatin, oxaliplatin and JM118, and undamaged DNA). Long 40-base sequences complementary to Zip and damaged probes were used to obtain the double-stranded structure. The short sequences bearing DNA crosslinks were adapted from [20]. Short damaged ODNs were synthesized using the phosphoramidite chemistry carried out on an Applied Biosystems 392 DNA synthesizer. Target ODNs (20 nmol) were site-specifically damaged by overnight incubation at 37 °C with cisplatin, oxaliplatin or JM118 (150 nmol), then purified by RP-HPLC. Correct masses were measured by MALDI-TOF mass spectrometry (Bruker) measurements. 5′-pyrrole-modified-Zip and long complementary sequences were obtained from Eurogentec.

Grafting of DNA probes

SPRi biochips were made of glass prisms covered with a 50 nm gold layer (GenOptics), onto which DNA probes were grafted. Electro-copolymerization of these probes using pyrrole chemistry was performed according to [21], except that the electric pulse voltage was 2.4 V. This voltage was applied by a potentiostat galvanostat PG 263 (Princeton Applied Research). Quality of spotting was assessed by controlling electrochemical charges of each spot. All sequences were grafted in triplicate.

Target/probe hybridization and interaction analysis

Duplex of damaged targets ODNs (275 nM) was obtained by mixing damaged/undamaged short ODN (275 nM) with long complementary ODNs (250 nM) in buffer A already described in the ligand fishing section. Solution was heated for 5 min at 95 °C, and cooled at room temperature to ensure hybridization. SPRi experiments were performed on a homemade system running under LabView software (GenOptics) for real-time data collection. Prism was placed in a heated cell allowing temperature control for target hybridization (37 °C) and protein interaction (25 °C). All steps were performed at a constant flow rate of 2.5 mL/h in buffer A. SPRi chip surface was first blocked with 1% BSA in buffer A. Damaged ODNs were subsequently injected for final hybridization with zip sequences. Tested proteins were HMGB1 and TOX4 full length recombinant proteins (Abnova) and ultrapure BSA (USB), injected at 20 nM in buffer A. Proteins were separated from the probes with a 1.5 M NaCl solution. At the end of each experiment the biochip was regenerated with a 0.1 M NaOH solution, rinsed with water, dried with argon and stored at 4 °C.

Results

Ligand fishing

The list of proteins isolated from HeLa nuclear extracts by our ligand fishing approach was obtained after proper optimizations of crucial parameters including (i) levels of platinum DNA damage within grafted probes (1,2-(d(GpG) and 1,2-d(ApG) crosslinks levels were respectively of ≈22–35 and ≈2–7 lesions/plasmid depending on the platinating agent), (ii) plasmidic probes concentration/beads quantity ratio (iii) beads blocking and (iv) methods and buffers used to wash the trap and to release proteins before identification. The sharpest criterion was the releasing protocol. Methods aiming at reducing the proportion of non specific proteins binding (mainly those adsorbed on magnetic beads), including the use of endonucleases to digest the DNA fishing probes, led to poorly reproducible results (data not shown). Actually, the best results were obtained by releasing the proteins directly within Laemmli buffer. Beads without DNA and beads grafted with undamaged plasmids were used as controls to discriminate between interesting candidates and non specific proteins identified in the proteomic experiments. After optimizations, each experiment was performed twice, both using low and intermediate stringency rinsing buffers to extend the range of potentially retained proteins (we observed that the highest salt concentration was a too stringent condition and caused loss of proteins of interest). Four different analyses per plasmid probe were therefore performed. The extensive list of identified proteins obtained by proteomics for each experiment and condition (including controls) is available as Supplementary Table 1. A summary of these results is described in Fig. 1 and Table 2. Taking only into account reliably identified proteins (i.e. with sufficient proteomic scores and peptide numbers in two independent experiments), 38 of them were specifically retained by platinated probes, 47% of which possess a known DNA binding activity according to the Gene Ontology

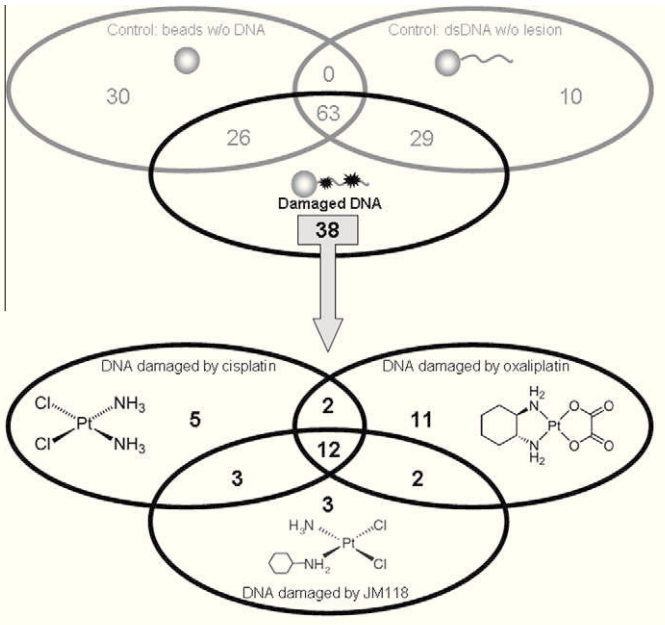


Fig. 1. Proteomics results: numbers of proteins reliably identified in the different ligand fishing assays. Indicated numbers are limited to proteins identified in two distinct experiments with Mascot scores ⩾70 and numbers of identified peptides ⩾2. Results obtained with low and intermediate stringency buffers are mixed.

Table 2
Proteins specifically identified with plasmidic probes damaged with cisplatin (CDDP), oxaliplatin (OxaliPt) and JM118. The list is restricted to proteins reliably identified in two distinct ligand-fishing/proteomics experiments (for more detailed lists, see Supplementary Table 1). Candidates whose affinity for platinum adducts had not been shown previously are highlighted in bold. S-: low stringency rinsing, S+: intermediate stringency rinsing.

Protein name (Uniprot)	Entry name (Uniprot)	Number of identified peptides (min–max)	Mascot scores (min–max)	Cisplatin S+	Cisplatin S–	Oxaliplatin S+	Oxaliplatin S–	JM118 S+	JM118 S–	Protein functions
Apoptotic chromatin condensation inducer in the nucleus ACIN1	**ACINU_HUMAN**	**2–3**	**103–156**	**X**	**–**	**–**	**–**	**–**	**–**	**Component of a complex responsible for exon junction during RNA maturation**
ADP-ribosylation factor 6 ARF6	**ARF6_HUMAN**	**2–3**	**97–181**	**X**	–	–	–	–	–	**Involved in cell motility; essential for cytokinesis**
60 kDa heat shock protein, mitochondrial HSPD1	**CH60_HUMAN**	**3–4**	**144–264**	–	–	**X**	–	–	–	**Involved in protein folding and import inside the mitochondrion**
Probable ATP-dependent RNA helicase DDX46	**DDX46_HUMAN**	**2**	**77–121**	**X**	**X**	**–**	**–**	**–**	**–**	**Responsible for mRNA processing (splicing)**
DNA ligase 3	DNLI3_HUMAN	2–7	96–436	–	–	X	-	X	X	DNA repair (base excision repair and double-strand breaks repair)
HEATR1	**HEAT1_HUMAN**	**2–12**	**75–710**	**X**	**–**	**–**	**–**	**–**	**X**	**–**
High mobility group box 1 HMGB1	HMGB1_HUMAN	11–16	472–1168	X	X	X	X	X	X	Regulation of chromatin structure and transcription, inflammation mediator
HMGB2	HMGB2_HUMAN	3–11	137–611	X	X	X	X	X	X	Regulation of chromatin structure and transcription
HMGB3	HMGB3_HUMAN	2–6	92–389	X	X	X	–	–	–	DNA binding and unwinding
Eukaryotic translation initiation factor 2 subunit 1 IF2a	**IF2A_HUMAN**	**2–3**	**83–120**	**–**	**–**	**–**	**X**	**X**	**-**	**Translation initiation**
Eukaryotic translation initiation factor 2 subunit 2 IF2b	**IF2B_HUMAN**	**2–3**	**96–178**	**–**	**–**	**X**	**–**	**–**	**–**	
Ras GTPase-activating-like protein IQGA1	**IQGA1_HUMAN**	**2–4**	**95–246**	**X**	**–**	**–**	**–**	**–**	**–**	**Potentially involved in cytoskeleton reorganization**
Leucine-rich repeat-containing protein 40 LRRC40	**LRC40_HUMAN**	**2**	**87–95**	**–**	**–**	**X**	**–**	**–**	**–**	**-**
Mediator of DNA damage checkpoint protein 1 MDC1	**MDC1_HUMAN**	**2–4**	**99–208**	**–**	**–**	**X**	**–**	**–**	**–**	**Recruited at site of double-strand breaks (DSB); required for DNA damage checkpoint (S phase and G_2/M transition)**
Metastasis-associated protein MTA2	**MTA2_HUMAN**	**2**	**80–136**	**–**	**–**	**X**	**–**	**–**	**–**	**Transcription repression via chromatin remodeling and histone deacetylation**
Nucleolar protein 11 NOL11	**NOL11_HUMAN**	**2–4**	**99–176**	**-**	**-**	**X**	**–**	**–**	**–**	**–**
Putative ribosomal RNA methyltransferase NOP2/NOL1	**NOP2_HUMAN**	**2–3**	**90–226**	**–**	**–**	**–**	**–**	**X**	**–**	**–**
Phosphatidylinositol-binding clathrin assembly protein PICALM	**PICAL_HUMAN**	**2–5**	**93–253**	**X**	**–**	**X**	**–**	**X**	**–**	**Mediates clathrin-vesicle assembly**
Pinin	**PININ_HUMAN**	**2**	**80–92**	**–**	**–**	**X**	**–**	**–**	**–**	**Transcriptional activator of the E-cadherin gene**
Serine/threonine-protein phosphatase 1 subunit 10 PNUTS	**PP1RA_HUMAN**	**2–12**	**90–653**	**X**	**X**	**X**	**X**	**X**	**X**	**Phosphatases inhibition, regulation of cell proliferation and apoptosis**
PC4 and SFRS1-interacting protein PSIP1	**PSIP1_HUMAN**	**2–3**	**120–162**	**–**	**–**	**X**	**–**	**–**	**–**	**Transcriptional coactivator involved in neuroepithelial stem cell differentiation and neurogenesis**
Telomere-associated protein RIF1	**RIF1_HUMAN**	**2–8**	**103–583**	**X**	**–**	**–**	**–**	**X**	**X**	**Negative regulator of telomere elongation; prevents replication in presence of DSB**
U2 small nuclear ribonucleoprotein A′ SNRPA1	**RU2A_HUMAN**	**2**	**89–96**	**–**	**–**	**X**	**–**	**–**	**–**	**–**

(*continued on next page*)

Table 2 (*continued*)

Protein name (Uniprot)	Entry name (Uniprot)	Number of identified peptides (min–max)	Mascot scores (min–max)	Cisplatin S+	Cisplatin S−	Oxaliplatin S+	Oxaliplatin S−	JM118 S+	JM118 S−	Protein functions
RuvB-like 2 RUVBL2 (reptin)	**RUVB2_HUMAN**	**2–4**	**83–260**	**–**	**–**	**–**	**–**	**X**	**–**	**Activator of ATM-mediated DNA damage checkpoint; chromatin remodeling (member of the TIP60/NuA4 complex)**
Septin-2	**SEPT2_HUMAN**	**2–5**	**98–281**	**–**	**X**	**–**	**X**	**–**	**X**	**Cytoskeleton organization**
Structural maintenance of chromosomes protein 3 SMC3	**SMC3_HUMAN**	**2–7**	**92–397**	**X**	**–**	**–**	**X**	**X**	**X**	**Involved in chromosome cohesion during cell cycle and in DNA repair**
Transcription activator BRG1	SMCA4_HUMAN	2	76–157	X	–	–	–	X	–	Transcriptional activator
SMARCA5	**SMCA5_HUMAN**	**2–3**	**106–140**	**–**	**–**	**–**	**–**	**X**	**–**	**Helicase with nucleosome-remodeling activity**
Chromatin-specific transcription elongation factor SPT16	SP16H_HUMAN	6–18	403–1292	X	X	X	X	X	X	Nucleosome reorganization
Signal recognition particle 68 kDa protein SRP68	**SRP68_HUMAN**	**2–5**	**85–278**	**X**	**–**	**X**	**–**	**–**	**–**	**Targets secretory proteins to the rough endoplasmic reticulum membrane**
Structure-specific recognition protein 1 SSRP1	SSRP1_HUMAN	5–12	254–625	X	X	X	X	X	X	Nucleosome reorganization
Telomeric repeat-binding factor 2 TRF2	**TERF2_HUMAN**	**2–5**	**129–254**	**X**	**X**	**X**	**X**	**X**	**–**	**Responsible for protection of telomeres**
Serine/threonine-protein kinase tousled-like 1 TLK1	**TLK1_HUMAN**	**2**	**92–113**	**–**	**–**	**X**	**–**	**–**	**–**	**Involved in nucleosome sliding, chromatin remodeling and DNA repair**
TOX high mobility group box family member 4 TOX4	**TOX4_HUMAN**	**2–6**	**111–381**	**X**	**X**	**X**	**X**	**X**	**–**	**Transcription activator; involved in thymocytes differentiation**
Human upstream binding factor hUBF	UBF1_HUMAN	6–22	340–1264	X	X	X	X	X	X	Ribosomal RNA transcription factor
WD repeat-containing protein 82 WDR82	**WDR82_HUMAN**	**2–7**	**140–430**	**X**	**–**	**X**	**–**	**–**	**X**	**Histone methylation**
X-ray repair cross-complementing protein 6/Ku70	XRCC6_HUMAN	2	73–92	X	–	–	–	–	–	Helicase activity ; involved in chromosome translocation and double-strand break repair
Zinc finger protein 638 ZN638	**ZN638_HUMAN**	**2–3**	**104–153**	**–**	**–**	**X**	**–**	**–**	**–**	**Binds to cytidine clusters in double-stranded DNA**

annotation. Nine out of the 38 proteins are already known members of the platinated DNA interactome (namely HMGB1, HMGB2, HMGB3, hUBF, SSRP1, SPT16, Ku70, SMCA4 and DNA ligase III). To our knowledge, the 29 other candidates had never been described before as associated with platinum-damaged DNA. Twelve out of the 38 proteins were recovered from ligand fishing reactions with the 3 different damaging agents, while 26 proteins were retained by only one or two types of damaged probes. Oxaliplatin adducts were those that retained the greatest number of unique protein (11), as opposed to cisplatin (5) and JM118 (3).

Our proteomics results were tested and confirmed by Western blot analyses of the ligand fishing eluate for 3 identified candidates (Fig. 2): HMGB1 as a well-known platinum adduct-interacting protein, but also TOX4 and PNUTS as original candidates linked to each others in recently published works [22,23]. Also, capture of TOX4 (6 peptides, score 273), PNUTS (7 peptides, score 319) and WDR82 (7 peptides, score 269) onto the cisplatin-damaged probes was observed using MDA-MB231 nuclear extracts as a second source of proteins. No additional original candidate was identified with these extracts. However, we noticed the presence of TFAM (18 peptides, score 543), another well-known platinum adduct-recognizing factor [4] that was absent from the protein list obtained with HeLa extracts.

The affinity of HMGB1 for another type of bulky adducts, namely UV-induced pyrimidine dimers, has already been reported [24]. We thus evaluated by proteomics the binding of TOX4 and PNUTS from HeLa nuclear extracts onto a UVC-damaged plasmid probe. As a positive control, we also monitored the binding of DNA Damage Binding protein 2 (DDB2) on these probes. DDB2 is a member of the UV-DDB complex responsible for UV-damage recognition and triggers the recruitment of NER machinery. Neither TOX4 nor PNUTS were captured with this particular trap, while DDB2 was (9 peptides, score 345). The MDA-MB231 nuclear extracts (in which DDB2 is absent) were also employed with the UVC-damaged probe. Once again, TOX4 and PNUTS were absent, demonstrating that their interaction with these particular lesions is not impaired by a shielding mechanism of DDB2.

We introduced thio-dNTPs at the extremities of our probes to avoid their degradation by nucleases present in the nuclear extracts [25]. In order to check that the identification of TOX4 and PNUTS was not an artifact originating from the formation of adducts on the internucleotide thiophosphate groups, we carried out an experiment with normal dNTPs to fill in the 5′ and 3′ ends of the plasmids. Western blot analyses confirmed the trapping of TOX4 and PNUTS (Supplementary Fig. 2A).

Also, high stringency washes (salt concentration $\geqslant 1$ M) are generally employed when using probes that capture proteins thanks to the formation of a covalent bond. It is not the case for the protocol we developed, and high stringency buffers revealed less efficient: HMGB1 signal was reduced (Supplementary

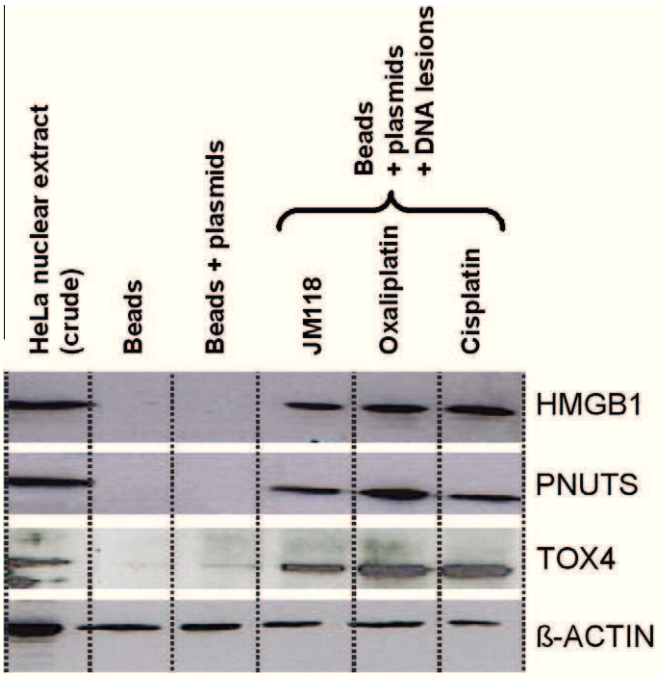


Fig. 2. Confirmation of proteomics identification by Western-blotting performed on ligand fishing eluates obtained from HeLa nuclear extracts. β-actin is used as a loading control. Nuclear extracts contain β-actin originating from actin nuclear pool and/or cytoplasmic contamination. This highly abundant protein is known to adsorb onto beads presenting different surface chemistries including polystyrene [44]. It is one of the proteins non-specifically adsorbed on magnetic beads and is therefore released in all eluates (Table S1).

Fig. 2B), while that of TOX4 and PNUTS was abolished. This is the reason why we relied on intermediate and low stringencies, thereby preserving these interactions, but also the presence of non-specific proteins. This explains the large number of total proteins found in each sample (between 92 and 194, Supplementary Table 1). Appropriate controls (undamaged DNA grafted onto the beads, and beads alone) are therefore indispensable to sort relevant candidates from all the identified proteins.

SPRi

SPRi is a versatile tool allowing the simultaneous, label-free and real-time study of biomolecular interactions [26]. We used this approach to assess in parallel interactions between recombinant purified proteins and four different double-stranded DNA probes: oligodeoxynucleotides (ODNs) bearing 1,2-d(GpG) cisplatin, oxaliplatin and JM118 crosslinks and ODNs without lesion. These damaged ODNs were synthesized, purified, controlled and grafted onto SPRi biochips. The tested proteins were successively HMGB1 (positive control), TOX4 (original candidate selected from proteomic results) and BSA (negative control). These experiments were performed at least twice with two different SPRi biochips for each protein. TOX4 showed a significant but lower capacity to bind platinated probes in comparison with HMGB1 (Fig. 3). It also showed a stronger affinity for all 1,2-d(GpG) crosslinks tested in comparison with non-damaged ODNs. Our repeated experiments did not demonstrate clear differences in binding affinity for TOX4 among the 3 platinated lesions tested. We nonetheless noticed that in our experimental conditions, HMGB1 recognizes cisplatin and JM118 crosslinks with a stronger affinity than oxaliplatin lesions.

Discussion

Proteomics methods applied to the identification of proteins recognizing DNA lesions constitute a recent research area. Various approaches have been developed to study these groups of proteins. Regarding lesions induced by platinating agents, Stansfield et al. [27] used protein microarrays incubated with cisplatin-modified fluorescent ODNs. Using this tool, the authors identified new interactions such as an association between cisplatin-modified DNA and Aurora kinase A. The most common strategy consists in trapping proteins using site-specifically damaged DNA immobilized on a solid phase such as magnetic beads. Extensive lists of proteins interacting with 1,2-instrastrand crosslinks, 1,3-instrastrand crosslinks and ICL were thus recently obtained by photocrosslinking proteins onto the photoactive cisplatin analogue Pt-BP6 inserted in duplex or dumbbell probes [12,13,28]. These original methods allowed the highlighting of new proteins interacting with platinum adducts, such as PARP-1. To carry out our experiments, we chose a different type of DNA probe consisting in 3 kb linearized plasmids exposed to three platinating agents: cisplatin, oxaliplatin or JM118. This strategy offers the following advantages: (i) a longer DNA sequence potentially allowing association with large protein complexes (ii) presence of several lesions/plasmid enhancing the quantity of trapped proteins and therefore the sensitivity of the method (iii) presence of the whole spectrum of lesions generated by the real pharmacological agent. Dealing with this last point, it must be emphasized that our probes contain a mixture of different crosslinks distributed all

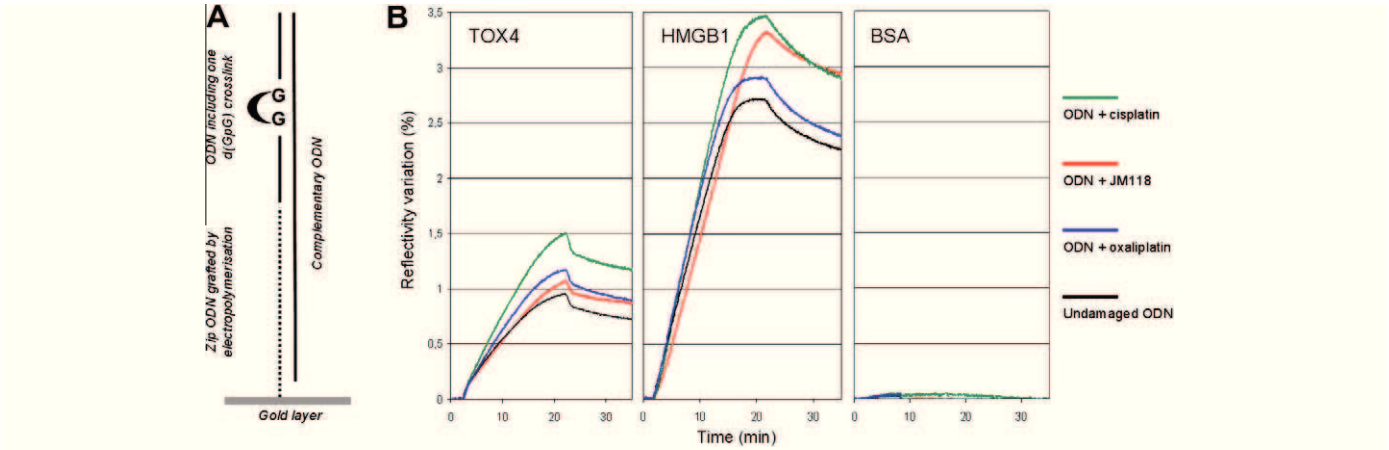


Fig. 3. SPR imaging. (A) Probes design (B) SPRi sensorgrams obtained with 20 nM TOX4 (tested candidate), HMGB1 (protein already known to interact with platinated DNA) and BSA (negative control).

along the sequence rather than one well characterized alteration at a specific site. These differences with other ligand fishing probes may explain some discrepancies observed with the abovementioned studies [12,13]. Indeed, when comparing the lists of proteins identified in our study with previously published works, we can observe the presence of common elements (HMGB1, HMGB2, HMGB3, hUBF, DNA ligase III, Ku70) but also of proteins specific to our study. For instance, we succeeded in trapping both SSRP1 (not found by other global methods but already known to interact with platinated DNA) and SPT16, but also more original candidates such as WDR82, PNUTS or TOX4. In addition to the identification of novel proteins involved in platinum adducts recognition, we have also demonstrated the binding of already known candidates to satraplatin lesions. Limited clinical development work is currently being done with satraplatin, which may be ultimately used as an oral chemotherapeutic drug [29]. Binding of HMGB1 to JM118-related DNA lesions was already demonstrated [30]. In this work, we showed that DNA ligase III, hUBF, SSRP1, SPT16 and HMGB2 also interact with JM118 lesions.

The 3 platinating compounds that were tested present different carrier ligands (those which remain bound to platinum after DNA adduction). They comprise a small ammonia group for cisplatin, a bulky 1,2-diaminocyclohexane for oxaliplatin, and asymmetrical ammonia and cyclohexylamine groups for JM118. These nonleaving groups are the cause of structural differences between adducts and have distinct repercussions on DNA bending and unwinding [8,31]. It has been demonstrated that these differences are affecting the affinity of proteins for DNA lesions. As an example, HMGB1 has a less important affinity for DNA crosslinks generated by oxaliplatin than for cisplatin adducts [32]. These structural differences may also explain the different panels of proteins that have been trapped with the 3 agents.

Among proteins listed in Table 2, we must underline the absence of known repair factors from the NER pathway. Nevertheless, we noticed the presence of Ku70 and MDC1, proteins recruited to DSB [33], and ARF6, a factor implicated in cell motility but also capable of interacting with Ku70 [34]. In addition, a large portion of our list comprises chromatin remodeling factors, including HMGB factors but also other proteins involved in this process: RuvBL2, SMARCA5, TLK1, MTA2, SMC3. Interactions between these candidates and platinum DNA damage require further validation using complementary methods.

For our study, we selected TOX4, also called epidermal Langerhans Cell Protein 1 (LCP1). Indeed, this protein has never been directly linked to processes associated with damaged DNA. TOX4 belongs to the HMG domain proteins family [35]. Several HMG-box proteins have an established capacity to bind DNA damaged by platinating agents through their HMG domain, a DNA binding motif made of 3 alpha helices [3,36]. Depending on the HMG protein, this binding has been associated with a wide range of molecular mechanisms such as shielding toward DNA repair factors, recruitment of DNA repair factors or hijacking of transcription factors. The TOX subfamily of HMG-box proteins has been recently discovered and the biological role of TOX4 is far from being elucidated. Its N-terminus domain was characterized as a strong transcription activator [23]. Its identification as a platinated-DNA interacting protein suggests it may be hijacked from its main role by adducts as it is the case for hUBF. Additionally, we identified TOX4 using the same ligand fishing strategy in nuclear extracts of another cell line (MDA-MB231, breast cancer). To bring complementary clues of TOX4/damaged DNA interaction, we confirmed proteomic data using Western blot and took advantage of the SPRi method. This approach had already been used in our team to analyse interactions between oxidative DNA lesions and base excision repair proteins [37]. Here, it allowed the assessment of recombinant TOX4 binding on ODNs bearing 1,2-d(GpG) crosslinks, and showed that platinated drugs induced lesions are recognized with a higher affinity than ODNs without DNA damage. These parallel interactions studies also showed that binding of TOX4 was weaker than that of HMGB1 for the 3 lesions and non-damaged DNA.

Using our ligand fishing strategy on HeLa nuclear extracts but also on MDA-MB231 nuclear extracts, we identified PNUTS (also called PPP1R10) and WDR82. PNUTS inhibits protein phosphatases PPP1CA and PPP1CC, binds to nucleic acids and plays a role in mechanisms such as cell proliferation and apoptosis by regulating key actors such as pRb [38] and p53 [39]. Our bait also trapped the WDR82 protein, involved in histone methylation and initiation of transcription. Interestingly, an association between PNUTS and TOX4 has recently been demonstrated [23], followed by another study showing the formation of a PNUTS/TOX4/WDR82 complex possibly involved in chromatin structuration during the mitosis/interphase transition [22]. PNUTS/TOX4 interaction suppresses the transcriptional activity of TOX4. It would be of great interest to find out if the interaction between damaged DNA and PNUTS is affecting its biological properties, and if the biological role of the PNUTS/TOX4/WDR82 complex is modified by platinating agents. Moreover, it would be interesting to test if PNUTS and WDR82 directly bind platinated crosslinks or only through their interaction with TOX4.

In addition, two telomeric proteins were identified by our assay: telomeric repeat-binding factor 2 (TRF2) and RIF1. A recent publication by Kim et al. [40] suggests that TRF2 directly interacts with PNUTS, this association playing a role in telomere length regulation. TRF2 is a member of the shelterin complex in charge of telomere protection from end-to-end fusion process or DNA damage sensors. Using telomere mimicking ODNs, Ourliac-Garnier et al. [41] found a decreased affinity of TRF2 towards platinated sequences in comparison with intact ODNs. In addition, telomere chromatin immunoprecipitation indicated a 60% decrease of TRF2 binding when cells were treated by cisplatin. This discrepancy with our results may be explained by differences of DNA probes sequences or DNA adducts patterns. The decrease in telomere/TRF2 association in cisplatin-treated cells may also be explained by a reduction of the available pool of TRF2 being once again hijacked by cisplatin adducts in non-telomeric DNA. Demonstration of these assumptions and a better understanding of relationships between DNA lesions and TRF2 have to be further assessed.

Our study brought results that partially overlap those of other teams, but also comprise several new candidates. This emphasizes the usefulness of carrying out complementary methods rather than one single approach to tend toward a complete cartography of the factors capable of interacting with damaged DNA. Following validation, these new members of the protein network surrounding platinated DNA lesions may be of great interest to go further in the understanding of the pharmacology of these therapeutic drugs and improve it in terms of efficiency, resistance overcoming and side effects limitations. These proteins might also constitute future targets for their potentialisation. These interactions could even find therapeutic applications in non cancerous diseases: low dose cisplatin regimens are for example tested as HMGB1 intracellular sequestration methods to prevent the release of this pro-inflammatory mediator [42]. Future directions include the adaptation of the presented tools to allow the screening of platinated DNA-interacting proteins within a chromatin context [43], which would permit to take into account the complex environment that might interfere with other proteins/damaged DNA interactions. More generally, both the ligand fishing system and the SPRi biochip developed for this study may be applied to screen interactomes associated with other types of DNA lesions generated by other anticancer drugs, radiations or environmental mutagenic compounds.

Acknowledgments

We thank R Bombera for his help in the optimization of SPRi, C Saint-Pierre for ODN controls and JL Ravanat for his expertise in mass spectrometry. We are also thankful to Agennix for providing JM118. This work was supported by a grant from "La Ligue contre le Cancer, comité de la Drôme".

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.abb.2010.12.021.

References

[1] L. Kelland, Nat. Rev. Cancer 7 (2007) 573–584.
[2] B.A. Teicher, Clin. Cancer Res. 14 (2008) 1610–1617.
[3] D. Wang, S.J. Lippard, Nat. Rev. Drug Discov. 4 (2005) 307–320.
[4] C.S. Chow, J.P. Whitehead, S.J. Lippard, Biochemistry 33 (1994) 15124–15130.
[5] E.N. Hughes, B.N. Engelsberg, P.C. Billings, J. Biol. Chem. 267 (1992) 13520–13527.
[6] S.S. Lange, K.M. Vasquez, Mol. Carcinog. 48 (2009) 571–580.
[7] D.K. Treiber, X. Zhai, H.M. Jantzen, J.M. Essigmann, Proc. Natl. Acad. Sci. USA 91 (1994) 5672–5676.
[8] Y. Jung, S.J. Lippard, Chem. Rev. 107 (2007) 1387–1407.
[9] N. Hegarat, G.M. Cardoso, F. Rusconi, J.C. Francois, D. Praseuth, Nucleic Acids Res. 35 (2007) e92.
[10] R.A. Rieger, E.I. Zaika, W. Xie, F. Johnson, A.P. Grollman, C.R. Iden, D.O. Zharkov, Mol. Cell Proteomics 5 (2006) 858–867.
[11] M. Winnacker, V. Welzmiller, R. Strasser, T. Carell, Chembiochem 11 (2010) 1345–1349.
[12] E.R. Guggenheim, D. Xu, C.X. Zhang, P.V. Chang, S.J. Lippard, Chembiochem 10 (2009) 141–157.
[13] G. Zhu, S.J. Lippard, Biochemistry 48 (2009) 4916–4925.
[14] M.I. Olnes, R.N. Kurl, Biotechniques 17 (1994) 828–829.
[15] A. Gérard, S.E. Polo, D. Roche, G. Almouzni, L.C. Judith, M. Paul, Methods in Enzymology, Academic Press, 2006. pp. 358–374.
[16] J.F. Millau, A.L. Raffin, S. Caillat, C. Claudet, G. Arras, N. Ugolin, T. Douki, J.L. Ravanat, J. Breton, T. Oddos, C. Dumontet, A. Sarasin, S. Chevillard, A. Favier, S. Sauvaigo, Lab Chip 8 (2008) 1713–1722.
[17] M. Jaquinod, F. Villiers, S. Kieffer-Jaquinod, V. Hugouvieux, C. Bruley, J. Garin, J. Bourguignon, Mol. Cell Proteomics 6 (2007) 394–412.
[18] B. Bernay, M.C. Gaillard, V. Guryca, A. Emadali, L. Kuhn, A. Bertrand, I. Detraz, C. Carcenac, M. Savasta, E. Brouillet, J. Garin, J.M. Elalouf, Mol. Cell Proteomics 8 (2009) 946–958.
[19] V. Dupierris, C. Masselon, M. Court, S. Kieffer-Jaquinod, C. Bruley, Bioinformatics 25 (2009) 1980–1981.
[20] Q. He, U.M. Ohndorf, S.J. Lippard, Biochemistry 39 (2000) 14426–14435.
[21] J.B. Fiche, J. Fuchs, A. Buhot, R. Calemczuk, T. Livache, Anal. Chem. 80 (2008) 1049–1057.
[22] J.H. Lee, J. You, E. Dobrota, D.G. Skalnik, J. Biol. Chem. 285 (2010) 24466–24476.
[23] S.J. Lee, J.K. Lee, Y.S. Maeng, Y.M. Kim, Y.G. Kwon, Exp. Mol. Med. 41 (2009) 189–200.
[24] E.A. Pasheva, I.G. Pashev, A. Favre, J. Biol. Chem. 273 (1998) 24730–24736.
[25] J.A. Mello, J.G. Moggs, G. Almouzni, Methods Mol. Biol. 314 (2006) 477–487.
[26] S. Scarano, M. Mascini, A.P. Turner, M. Minunni, Biosens. Bioelectron. 25 (2010) 957–966.
[27] H.E. Stansfield, B.P. Kulczewski, K.E. Lybrand, E.R. Jamieson, J. Biol. Inorg. Chem. 14 (2009) 193–199.
[28] C.X. Zhang, P.V. Chang, S.J. Lippard, J. Am. Chem. Soc. 126 (2004) 6536–6537.
[29] H. Choy, C. Park, M. Yao, Clin. Cancer Res. 14 (2008) 1633–1638.
[30] A. Vaisman, S.E. Lim, S.M. Patrick, W.C. Copeland, D.C. Hinkle, J.J. Turchi, S.G. Chaney, Biochemistry 38 (1999) 11026–11039.
[31] H. Kostrhunova, O. Vrana, T. Suchankova, D. Gibson, J. Kasparkova, V. Brabec, Chem. Res. Toxicol. 23 (2010) 1833–1842.
[32] M. Wei, S.M. Cohen, A.P. Silverman, S.J. Lippard, J. Biol. Chem. 276 (2001) 38774–38780.
[33] T.A. Mochan, M. Venere, R.A. DiTullio Jr., T.D. Halazonetis, Cancer Res. 63 (2003) 8586–8591.
[34] J.K. Schweitzer, C. D'Souza-Schorey, Exp. Cell. Res. 311 (2005) 74–83.
[35] E. O'Flaherty, J. Kaye, BMC Genomics 4 (2003) 13.
[36] U.M. Ohndorf, M.A. Rould, Q. He, C.O. Pabo, S.J. Lippard, Nature 399 (1999) 708–712.
[37] C. Corne, J.B. Fiche, D. Gasparutto, V. Cunin, E. Suraniti, A. Buhot, J. Fuchs, R. Calemczuk, T. Livache, A. Favier, Analyst 133 (2008) 1036–1045.
[38] G. De Leon, T.C. Sherry, N.A. Krucher, Cancer Biol. Ther. 7 (2008) 833–841.
[39] S.J. Lee, C.J. Lim, J.K. Min, J.K. Lee, Y.M. Kim, J.Y. Lee, M.H. Won, Y.G. Kwon, Cell Death Differ. 14 (2007) 1106–1116.
[40] H. Kim, O.H. Lee, H. Xin, L.Y. Chen, J. Qin, H.K. Chae, S.Y. Lin, A. Safari, D. Liu, Z. Songyang, Nat. Struct. Mol. Biol. 16 (2009) 372–379.
[41] I. Ourliac-Garnier, A. Poulet, R. Charif, S. Amiard, F. Magdinier, K. Rezai, E. Gilson, M.J. Giraud-Panis, S. Bombard, J. Biol. Inorg. Chem. 15 (2010) 641–654.
[42] P. Pan, J. Cardinal, R. Dhupar, M.R. Rosengart, M.T. Lotze, D.A. Geller, T.R. Billiar, A. Tsung, J. Leukoc. Biol. 86 (2009) 625–632.
[43] B. Wu, C.A. Davey, Chem. Biol. 15 (2008) 1023–1028.
[44] M. Ehrenberg, J.L. MacGrath, Acta Biomater. 1 (2005) 305–315.

Mutation Research 752 (2013) 72–83

Contents lists available at SciVerse ScienceDirect

Mutation Research/Reviews in Mutation Research

journal homepage: www.elsevier.com/locate/reviewsmr
Community address: www.elsevier.com/locate/mutres

Review

Tools and strategies for DNA damage interactome analysis

Christophe Bounaix Morand du Puch [a], Ewa Barbier [a], Sylvie Sauvaigo [a], Didier Gasparutto [a], Jean Breton [a,b,*]

[a] Laboratoire Lésions des Acides Nucléiques, SCIB, UMR-E3 CEA/UJF-Grenoble 1, INAC, 17 rue des Martyrs, Grenoble, F-38054, France
[b] UFR de Pharmacie, Université Joseph Fourier-Grenoble 1, Domaine de la Merci, La Tronche, F-38706, France

ARTICLE INFO

Article history:
Received 19 June 2012
Received in revised form 1 November 2012
Accepted 22 November 2012
Available online 7 December 2012

Keywords:
DNA damage
Protein–DNA interactions
Proteomics

ABSTRACT

DNA is the target of multiple endogenous and exogenous agents generating chemical lesions on the double helix. Cellular DNA damage response pathways rely on a myriad of proteins interacting with DNA alterations. The cartography of this interactome currently includes well known actors of chromatin remodelling, DNA repair or proteins hijacked from their natural functions such as transcription factors. In order to go further into the characterisation of these protein networks, proteomics-based methods began to be used in the early 2000s. The strategies are diverse and include mainly (i) damaged DNA molecules used as targets on protein microarrays, (ii) damaged DNA probes used to trap within complex cellular extracts proteins that are then separated and identified by proteomics, (iii) identification of chromatin- bound proteins after a genotoxic stress, or (iv) identification of proteins associated with other proteins already known to be part of DNA damage interactome. All these approaches have already been performed to find new proteins recognizing oxidised bases, abasic sites, strand breaks or crosslinks generated by anticancer drugs such as nitrogen mustards and platinating agents. Identified interactions are generally confirmed using complementary methods such as electromobility shift assays or surface plasmon resonance. These strategies allowed, for example, demonstration of interactions between cisplatin–DNA crosslinks and PARP-1 or the protein complex PTW/PP. The next challenging step will be to understand the biological repercussions of these newly identified interactions which may help to unravel new mechanisms involved in genetic toxicology, discover new cellular responses to anticancer drugs or identify new biomarkers and therapeutic targets.

© 2012 Elsevier B.V. All rights reserved.

Contents

1. Introduction . . . 73
2. Main tools and strategies . . . 73
 2.1. Overall strategy . . . 73
 2.2. Mid- to high-throughput methods for the identification of new DNA damage/protein interactions . . . 74
  2.2.1. Protein microarrays . . . 74
  2.2.2. DNA pull down coupled to proteomics . . . 75
  2.2.3. Other approaches . . . 78
 2.3. Confirmation methods . . . 79
  2.3.1. Filter binding assay . . . 79
  2.3.2. DNA footprinting . . . 79
  2.3.3. Electrophoretic mobility shift assay (EMSA) . . . 79

Abbreviations: DDR, DNA damage response; DIGE, differential in-gel electrophoresis; DPC, DNA-protein crosslink; EMSA, electrophoretic mobility shift assay; ESI-MS/MS, electrospray ionisation tandem mass spectrometry; FRET, fluorescence resonance energy transfer; IRIF, ionizing radiation-induced foci; ITC, isothermal titration calorimetry; MALDI-TOF, matrix-assisted laser desorption/ionisation coupled to time-of-flight mass spectrometry; NER, nucleotide excision repair; PTM, post-translational modification; SPR, surface plasmon resonance; SPRi, surface plasmon resonance imaging; TF, transcription factor.

* Corresponding author at: Laboratoire Lésions des Acides Nucléiques, SCIB, UMR-E3 CEA/UJF-Grenoble 1, INAC, 17 rue des Martyrs, Grenoble, F-38054, France. Tel.: +33 4 38 78 56 01; fax: +33 4 38 78 50 90.
E-mail address: jean.breton@cea.fr (J. Breton).

1383-5742/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.mrrev.2012.11.002

2.3.4. Surface plasmon resonance (SPR) . . . 79
2.3.5. Fluorescence anisotropy . . . 80
2.3.6. Fluorescence resonance energy transfer (FRET) . . . 80
2.3.7. Isothermal titration calorimetry (ITC) . . . 80
2.3.8. *In cellulo* approaches . . . 80
3. Comments and perspectives . . . 80
4. Concluding remarks . . . 81
Acknowledgements . . . 82
References . . . 82

1. Introduction

DNA is subjected to endogenous or exogenous factors that produce a wide range of chemical lesions [1–3]. Sources of DNA damage can be divided into three classes. The first class of lesions are due to spontaneous causes, which include mechanisms such as deamination of cytosine to uracile, or 5-methylcytosine leading to thymine. Another example is provided by mismatches or insertions/deletions erroneously introduced by polymerases during DNA replication. The second class of lesions are due to reactive species endogenously generated by physiological or pathological processes (such as mitochondrial respiration, inflammation or infectious diseases). These processes can alter DNA primary structure, for instance through DNA oxidation or formation of abasic sites. Reactive oxygen species (ROS), reactive nitrogen species (RNS) or lipid peroxidation products are classic examples of these factors. The third class of lesions are due to exogenous physical or chemical agents that are able to generate a large variety of lesions including strand breaks, DNA bulky adducts or DNA crosslinks. It is worth noting that some of these agents are voluntarily administered to cancer patients in order to treat tumours, using strategies such as radiotherapy or alkylating drug-based chemotherapies [4].

The occurrence of lesions in DNA initiates a complex series of molecular and cellular reactions, comprising deleterious events leading to cell death or mutagenesis and defensive mechanisms requiring cell cycle arrest, DNA remodelling and DNA repair. The main components in these reactions are a myriad of proteins, including those that are in the forefront, *i.e.* in close contact with DNA lesions. To illustrate this notion of "DNA damage interactome", we can cite examples of proteins associated with DNA lesions generated by platinum compounds [5]. Cisplatin, carboplatin and oxaliplatin are the leading molecules of this class of anticancer drugs. These agents generate DNA damage comprising mainly intrastrand crosslinks and, to a lesser extent, interstrand and DNA-protein crosslinks. By screening the biomedical literature, we observe that more than 40 proteins have been characterised as being able to interact individually or within a complex with DNA damaged by these agents (Table 1). For instance, the heterodimeric complex between XPC and HR23B is a sensor of distorting adducts such as cisplatin lesions, playing a major role in the initiation of nucleotide excision repair (NER) process [6]. But the DNA damage interactome is far from being limited to DNA repair factors. Indeed, portions of DNA altered by platinum-based drugs can also constitute a trap for transcription factors (TF) hijacked from their consensus sequence. This phenomenon was described for TF such as TATA binding protein (TBP) [7], Y-box binding protein 1 (YB-1) [8] or human upstream binding factor (hUBF, TF involved in rRNA genes transcription) [9]. A last example of proteins known to be associated with platinum damaged DNA is provided by chromatin remodelling factors high mobility group box 1 (HMGB1) or structure specific recognition protein 1 (SSRP1) [10]. These proteins may hinder (HMGB1) [11], or on the contrary enhance, DNA repair by shielding or facilitating lesion accessibility to other proteins. Checkpoints and DNA repair proteins assembled at double-strand breaks (DSBs) induced by ionizing radiations represent another widely studied interactome [12]. Ionizing radiation-induced foci (IRIF) are dynamic and complex structures including for instance the phosphorylated histone variant γH2AX which is often used as a marker to assess the extent of DSBs in cells.

Proteins present at the vicinity of DNA lesions are therefore essential elements for the understanding of genetic toxicology mechanisms or cellular response to anticancer treatments. Until recently, these factors were essentially identified and characterised by low-throughput methods which enabled the discovery of specific interactions of one DNA lesion with one, or very few, proteins. Scientists specialised in DNA damage response (DDR) took advantage of the development of proteomics approaches such as protein microarrays and mass spectrometry techniques and began to apply these tools, mainly after 2005, in order to complete the cartography of proteins recognizing DNA alterations. This review will first focus on proteomics strategies used in the last few years to screen DNA damage interactomes. Methods applied to validate and further characterise the identified interactions will also be described, which will highlight the original biological interactions discovered using those types of protocols. In the second part of the review, a potential means to improve these approaches in terms of methodology and biological relevance will be discussed.

2. Main tools and strategies

2.1. Overall strategy

A three-stage scheme can be drawn to describe scientific works that have been published in the field of this review so far (Fig. 1). The first step consists of developing tools able to generate a list of candidate proteins interacting with a given DNA lesion. From a technical point of view, this stage can be considered as the most challenging, since each family of DNA lesions (*e.g.* oxidised bases, bulky adducts, strand breaks, *etc.*) needs to be included in different molecular baits interacting with proteins. These proteins then have to be identified. The development of these approaches requires various competencies including DNA engineering, protein biochemistry, large scale biology approaches and bioinformatics. Once the first step has provided a list of candidates that are potentially associated with DNA damage, this assumption has to be validated by complementary methods which are usually simpler (*e.g.* gel shift approaches, immunodetection methods, *etc.*). This stage can either simply provide a "yes or no" answer, or bring new information about interaction parameters such as specificity or kinetic data. Ideally, the final step should consist in understanding the biological relevance of the newly identified interaction. For example, relevant questions include: in which cellular process is this interaction involved? Or does it play a role in the toxicity of the lesion? Each of these three steps will be detailed in the following chapters.

Table 1
Illustration of a well-studied but still-to-expand DNA damage interactome: proteins associated with DNA altered by anticancer platinating agents. At least 44 proteins can be considered as members of this interactome to date.

Biological function	Protein	Reference
Chromatin remodelling/ chromatin structure	SSRP1	[10]
	HMGB1	[79]
	HMGB2	[79]
	HMGB3	[22]
	SMARCA2	[22]
	SMARCA3	[22]
	SMARCC4	[22]
	PB-1	[22]
	LIM domain kinase 2	[16]
	Histone H1	[80]
Transcription factors	hUBF	[9]
	ZNF143	[81]
	LEF-1	[82]
	mTFA	[82]
	SRY	[82]
	TBP	[7]
	p53	[83]
	YB-1	[84]
	Zinc finger protein 684	[16]
DNA repair	XPC-HR23B	[6]
	XPA	[85]
	RPA	[86]
	PARP-1	[8]
	MSH2	[87]
	AAG	[88]
	PNK	[25]
	MutSβ	[25]
	XRCC1	[25]
	Ku70/Ku80	[89]
	RFC1	[22]
	RFC2	[22]
	DNA ligase III	[22]
	DDB1	[22]
	DDB2	[90]
Relaxation of supercoiled DNA structures	DNA topoisomerase I	[91]
Structural role during mitosis	NUMA1	[25]
Cell cycle regulation	Aurora kinase A	[16]
	Aurora kinase B	[16]
Metabolism of amino-acids, lipids and sugars	Propionyl coenzyme A carboxylase	[16]
Alternative splicing	Heterogeneous nuclear ribonucleoprotein C	[16]
RNA edition	Apolipoprotein B messenger RNA editing enzyme	[16]
	Ankyrin repeat and zinc finger domain containing 1	[16]
	Melanoma antigen family B, 4	[16]
Transcription regulation	TOX4	[20]

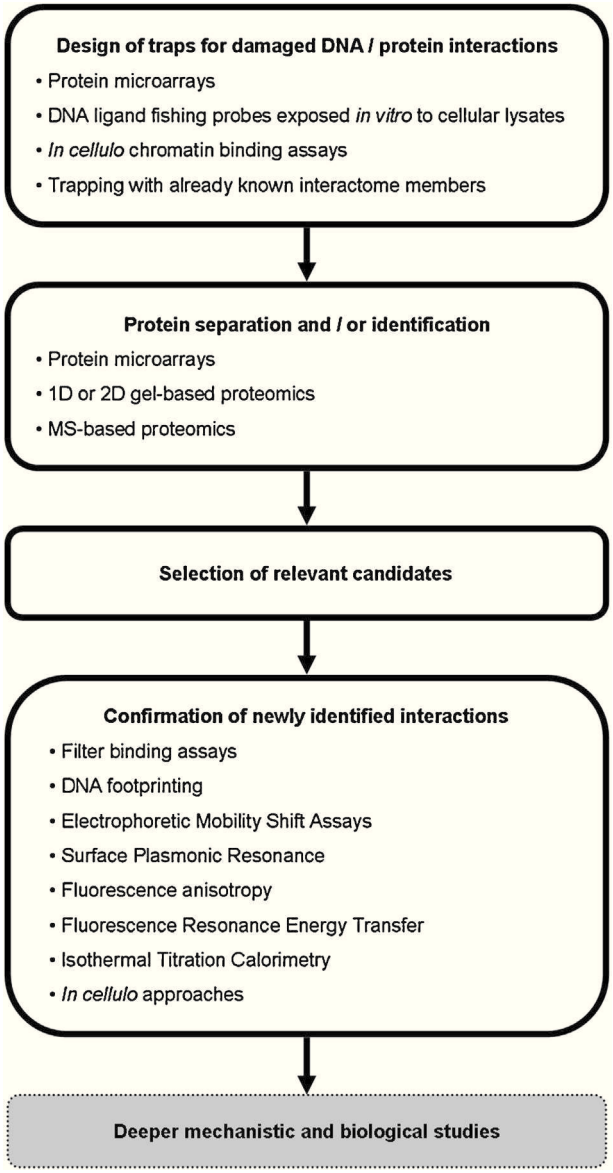


Fig. 1. Overall strategy used to identify new interactions between damaged DNA and proteins. Proteome-wide methods are often generating numerous candidates that have to be confirmed by techniques allowing the characterisation of interactions between defined DNA damage and a specific protein. Based on different techniques, all these approaches are not targeted towards the same panels of proteins and are therefore revealing different but complementary sets of interactions.

2.2. Mid- to high-throughput methods for the identification of new DNA damage/protein interactions

2.2.1. Protein microarrays

The traditional design of protein microarrays relies on a planar surface on which is immobilised a range of individualised proteins. These miniaturised systems allow the parallel analysis of multiple biological interactions between proteins and limited quantities of targets injected in, or deposited onto, the chip [13]. The main application field of these devices has been the screening of protein-protein interactions, but its usefulness to study protein–DNA interactions has also been demonstrated. By arraying TFs on a chip and using labelled oligodeoxynucleotides (ODNs) as targets, protein microarrays can be used to find new TFs/specific DNA sequence interactions [14].

In the field of damaged DNA/protein interactions, the protein microarray strategy has been used to detect proteins recognizing mismatches or abasic sites contained in ODNs labelled with the fluorescent dye Cy3 [15]. Grafted proteins were produced from the *Escherichia coli* genome and purified using a high-throughput protocol to obtain 4256 different proteins. Among the interesting biological results to come from this study, the authors identified

YbcN (previously of unknown function) as being able to interact with G:T mismatches and the YbaZ protein as being able to bind abasic sites. By using complementary biochemical experiments, these binding activities were confirmed and the base flipping properties of YbcN and YbaZ demonstrated. This base flipping step corresponds to the capacity of some DNA repair enzymes to swing the damaged base out of the helix into the active site of the enzyme.

A second relevant example is described in the work of Stansfield et al. [16] who tested two versions of a commercially available protein microarray (ProtoArray™, Invitrogen) to identify proteins bound to fluorescent ODNs modified by cisplatin or chromium. About 1800 and 3000 different human proteins were arrayed on the two versions of this chip. Of those, there are between 25 and 29 proteins identified as able to recognise platinated DNA. Within this list, the authors chose the protein aurora kinase A for further confirmation using a "macroarray" approach that utilised a home-made array containing only the candidate to be validated and a control protein (GST). Aurora enzymes are kinases involved in mitosis regulation and have been found overexpressed in cancerous tissues. If validated by further studies, the aurora/platinated DNA interaction may therefore constitute an interesting phenomenon to understand the response of tumours to chemotherapy.

In addition to the classically cited advantages of protein microarrays (*e.g.* parallelised and miniaturised system, small amount of proteins and targets), both latter studies have the benefit of being well-suited for detection of interactions with proteins that are weakly expressed in the cells. In these arrays, proteins are indeed all deposited at similar levels instead of being presented in cellular extracts with a wide dynamic range of protein concentrations, the most concentrated possibly hindering relevant interactions with those present at low copy numbers. Another potentially interesting point is the opportunity to use the same chip for protein-protein interactions. For example, Chen et al. [15] took advantage of their microarray to find protein partners of YbaZ shown previously to bind damaged DNA. In spite of these favourable elements, the drawbacks that may be highlighted are (i) the difficulty to get protein immobilisation procedures preserving key structural parameters necessary to establish DNA-protein interactions, (ii) the difficulty to identify binding with the diversity of post-translationally modified protein forms and protein complexes, and (iii) the difficulty to obtain a set of relevant proteins spotted onto the chip. Indeed, commercial arrays do not always contain the sufficient range of proteins fitting with DNA damage studies. The development of tools allowing high-throughput production and purification of all the proteins expressed in a model organism is very promising but still limited to specialised laboratories.

2.2.2. DNA pull down coupled to proteomics

The principle of the DNA pull down approach, also called "ligand fishing", resides in the use of DNA probes containing the chemical lesion of interest. These DNA fragments are mixed with crude cellular or nuclear extracts and serve as traps to retain proteins that interact with DNA damage. It is of interest to note that this type of altered DNA probe was already used in the 1980s for a series of protein replica filters derived from cDNA expression libraries [17]. This method led to the identification of SSRP1 as a protein associated with lesions generated by cisplatin. However, as mentioned for protein microarrays, it does not allow the study of interactions that may occur with proteins containing specific post-translational modifications (PTMs) or proteins interacting indirectly as members of a complex.

The main experimental points to keep in mind when optimizing a DNA pull down strategy are summarised in Fig. 2. To compare the different approaches that can be classified as "DNA pull down coupled to proteomics", we will focus our discussion on a selection of four parameters: (i) nature of DNA probes, (ii) modes of protein retention on DNA probes, (iii) methods to release trapped proteins and (iv) proteomics methods used to identify proteins.

(i) Nature of DNA probe

It would be an oversimplification to limit the design of protein traps to ODNs containing the lesion of interest. While optimising the structure of these baits, the main points to keep in mind are: methods used to synthesise the probe and to introduce DNA damage, adaptations to avoid probe degradation by nucleases contained in the cellular extract, methods used to immobilise probes on a solid support and finally appropriate controls to avoid the identification of non-specific candidates. These technical criteria are discussed in the following paragraphs.

DNA probes can, for example, be obtained by solid phase chemical assembly using a DNA synthesiser. In this case, the lesion can be introduced by providing the corresponding functionalised monomer to the synthesiser before deprotection, purification and control of produced ODNs. The advantage of such approach is the possibility to add one or several lesions at a specific position in the DNA sequence [18]. The drawback is the availability of the

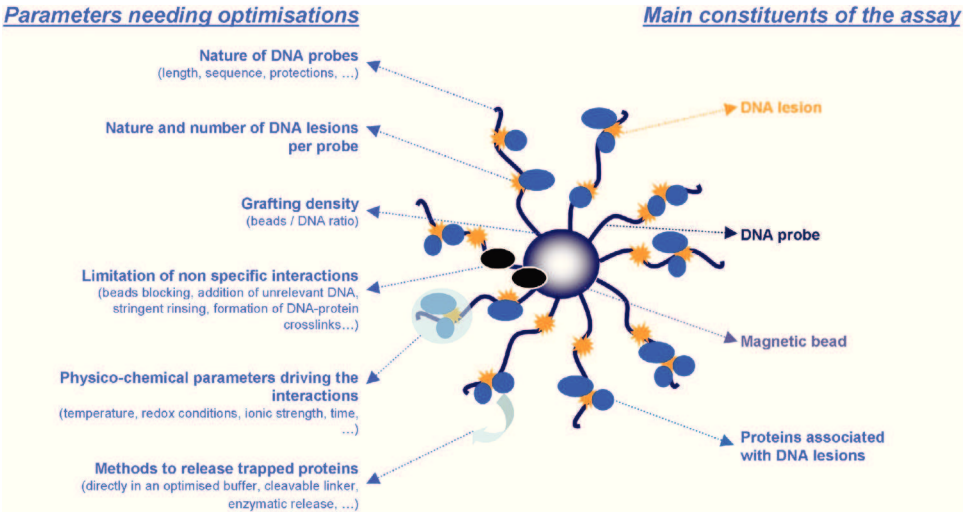


Fig. 2. DNA pull down approaches: main optimisations. The design of pull down strategies requires the optimisation of numerous experimental conditions needing the cooperation of specialists in DNA chemistry (synthesis of modified DNA), protein biochemistry and proteomics methods that are subsequently used to study trapped factors.

modified synthon, which can be commercially provided for commonly studied DNA alterations but reserved to laboratories specialised in nucleic acids chemistry for original damage. Lesions in synthesised ODNs can also be introduced by exposing the nucleic acid fragment to the desired physical or chemical damaging agent. In comparison to the earlier method described in this section, it does not offer the possibility to perfectly control the position and the chemical nature of the alteration. The desired modified ODNs have therefore to be separated from contaminating ODNs and subsequently quality-controlled, using for example high performance liquid chromatography (HPLC) for the first step and matrix-assisted laser desorption/ionisation (MALDI) coupled to time-of-flight (TOF) mass spectrometry for the latter.

Another way to obtain DNA baits is to use Polymerase Chain Reaction (PCR). This option offers the possibility to produce longer probes in comparison to those coming from a DNA synthesiser whose length is often limited to a maximum of a hundred nucleotides. This strategy has not as yet been applied to the screening of DNA damage interactome. The main reasons are the lack of modified dNTPs to be introduced into the PCR mix and the necessity to study DNA modifications with a corresponding polymerase able to introduce the modified dNTPs in the synthesised strand and bypass them with a sufficient fidelity during the next cycles.

'Finally, some teams use plasmid probes [19,20]. This offers the opportunity to obtain traps with numerous DNA lesions and therefore to increase the amount of captured proteins and, by doing so, the sensitivity of the method. These baits are altered by exposure to physical or chemical genotoxic agents, providing a spectrum of different lesions that have to be controlled (identification, quantification) using for example mass spectrometry. It has to be highlighted that this approach provides a mixture of various alterations instead of a pre-determined and well-defined DNA lesion.

Another important point when optimising the design of these lesion-containing probes is to counteract their potential degradation by nucleases contained in cell extracts. Although protection from endonucleases is difficult to obtain, various adaptations have been applied to protect the end of DNA fragments from exonucleases. For example, Mello et al. [19] have incorporated α-thio-dNTPs at the 3′ end of their plasmidic probes. Phosphorothioate bonds are indeed known as being able to block digestion by exonucleases [21]. Another solution would be to close DNA ends by designing DNA autocomplementary hairpin probes terminated by a single-stranded loop instead of hybridising two separated ODNs. Guggenheim et al. [22] proposed an ODN structure closed at both extremities by two single stranded loops forming a dumbbell probe. That bait allowed the trapping of proteins recognizing platinated DNA. These closed ends also offered the advantage of limiting potential contamination by proteins associated with DNA breaks which are mimicked by usual blunt ends. It should also be mentioned here that, for some DNA lesions such as abasic sites, damaged probes can also be rapidly cleaved by DNA repair proteins contained in the tested biological extract. DNA pull down is therefore not applicable to all DNA lesions.

For DNA pull down purposes, probes are mainly maintained on a solid support. A first strategy compatible with proteomics methods is to graft DNA fragments onto a solid support (such as sepharose) packed in a column. After injection of cell extract and washing steps, proteins are analysed in elution fractions. But the

Table 2
Comparative overview of protein microarrays and *in vitro* DNA pull down approaches used for DNA damage interactome screening.

Type of lesion	Approach used (main characteristics)	Examples of proteins retained by the trap	Confirmation methods	References
Cisplatin adducts (1,2- and 1,3-d(GpG), interstrand crosslinks) and oxaliplatin adducts (1,2-d(GpG))	ODNs containing a modified (benzophenone group) platinum adduct immobilised on a solid support; photocrosslinking of proteins	HMG box proteins, DDB1, RPA, PARP-1, Ku70/Ku80, DNA-PK$_{cs}$, hMutSβ (*ICL*), XRCC1, ... (human)	EMSA (hMutSβ)	[8,22,25,92]
Platinum adducts from cisplatin, oxaliplatin and satraplatin	Damaged linearised plasmids grafted onto magnetic beads	TOX4 (human)	SPRi	[20]
Abasic sites	ODNs containing abasic sites immobilised on beads; $NaBH_4$ crosslinking of proteins	AroF, DnaK, MutM, PolA, TnA, TufA, UvrA (*E. coli*); ARC1,Ygl245wp (*S. cerevisiae*)	Crosslinking of purified protein to radiolabelled ODN (Yg1245wp)	[24]
Double-strand breaks	Photocleavable ODNs fixed on magnetic beads	DNA-PK$_{cs}$, Ku70/Ku80, PARP-1 (human)	-	[28]
Mechlorethamine adducts (nitrogen mustard)	Joint incubation of ODNs, cellular extracts and alkylating agent (chemical crosslinking); immobilisation of probes on beads	AGT, XRCC1, PARP-1, Ku70/86, Ref-1, ... (human)	-	[23]
1,2-d(GpG) adduct of cisplatin	Commercial human protein chip; fluorescent ODNs containing the lesion used as targets	Kinases Aurora A et B, zinc finger protein 684, ankyrin repeat and zinc finger domain containing 1, apolipoprotein B messenger RNA editing enzyme, LIM domain kinase 2, ... (human)	Protein "macro-array" format (Aurora kinase A)	[16]
Mismatch (A:C and G:T); abasic sites	*E. coli* proteome chip; fluorescent ODNs containing the lesion used as targets	cspE, ybcN, ybaZ, dgt (*E. coli*)	EMSA (ybaZ)	[15]
8-oxo-dG	ODNs containing a lesion with diazirine group in front of (photocrosslinking of proteins); fixation on agarose beads	OGG1, PARP-1, HMGB1, uracile DNA-glycosylase, APEX (human)	N.C.	[26]

most popular approach these last few years has been to immobilise DNA probes onto magnetic beads (Table 2), mainly through biotin/streptavidin interactions. As for many other biological experiments, the handling of these beads facilitates washing and elution stages by avoiding columns or centrifugation steps.

Finally, when optimising the set of DNA pull down baits, it is of course important to include a control DNA fragment without DNA damage. Most of the time, this control is very helpful to eliminate candidates from the list of identified proteins that actually interact with the non-damaged part of the probe, recognise DNA ends or are adsorbed and released from the solid support (beads).

(ii) Modes of protein retention on DNA probes

Along with the subsequent releasing stage, this parameter plays a critical role in the specificity and sensitivity of DNA pull down assays. It may be observed that two main strategies have been developed regarding the retention of proteins on damaged DNA probes: with or without formation of a covalent crosslink between DNA and proteins.

- The first strategy consists of forming a covalent crosslink between DNA bait and the trapped protein.

This stable capture can be obtained naturally when the studied genotoxic agent is already known for its capacity to form DNA-protein crosslinks (DPCs). For example, Loeber et al. [23] identified proteins crosslinked to DNA lesions generated by the alkylating drug mechlorethamine. In addition, the authors analysed individual amino-acids after DPCs formation to determine the nature of amino acid – nucleobase conjugates. This was achieved by incubating an ODN with a HeLa nuclear extract and the alkylating agent. DPCs were released by thermal hydrolysis, digested using trypsin, carboxypeptidase Y and proteinase K to obtain the individual amino acid – nucleobase conjugates, analysed by MS. This work confirmed the alkylation of guanine N7 position by mechlorethamine and the preferential crosslink of this DNA adduct with cystein thiols of proteins.

An alternative approach relies on the generation of protein – damaged DNA crosslinks by using an external chemical or physical agent. Rieger et al. [24] took advantage of $NaBH_4$ as a chemical agent that allows the crosslinking of bifunctional glycosylases with the Schiff base intermediate they generate during the base excision repair (BER) process. This approach identified proteins from *E. coli* and *Saccharomyces cerevisiae* reacting with abasic sites. Among their list of candidates, the authors produced a recombinant Ygl245wp and used that purified yeast protein to confirm the interaction. The formation of a covalent link between damaged DNA and proteins can also be generated by photo-crosslinking. This technique requires the replacement of the DNA lesion by a photoreactive analogue. This photoaffinity trapping method was applied to platinated DNA by using a cisplatin analogue comprising a benzophenone photoreactive group [8,22,25], and to oxidative DNA damage with a 8-oxodG analogue containing a diazirine moiety [26]. This last trap gave relevant results, as exemplied by the identification of hOGG1 which is a BER glycosylase well-known to play an important role in 8-oxodG excision. The photoreactive arylazido analogues mimicking bulky adducts developed by Maltseva et al. [27] may also be employed as they have demonstrated their capacity to retain NER factors such as XPC-HR23B or XPA.

Crosslinking with DNA probes allows the application of stringent washing steps before recovering trapped proteins and therefore has the distinct advantage to eliminate unspecific proteins retained on undamaged DNA or solid support. Moreover, these approaches can be exploited to understand which protein site is interacting with the DNA lesion of interest [23]. However, the following points have to be kept in mind when establishing protocols and interpreting the results: (a) these methods are not well-suited to study protein complexes interacting with DNA lesions as factors which are not in direct contact with DNA are eliminated. (b) DNA lesion analogues are not always easily available and require competencies in synthesis chemistry. (c) Additional moieties such as those used for photoaffinity may not exactly reflect interactions occurring with the "real" DNA alteration.

- The second strategy therefore consists in the non-covalent retention of proteins. This strategy has already been applied to ODNs mimicking DNA strand breaks [28] or plasmids containing lesions generated by platinating drugs [20]. The advantages and drawbacks are the opposite of those described in the previous paragraph. The advantages include the possibility to identify protein complexes and the inclusion into the probe of biologically relevant DNA lesions. This approach also gives the opportunity to release functional proteins, as demonstrated by Hégarat et al. [28] who succeeded in obtaining phosphorylase and DNA ligation activities after DNA pull down. The disadvantage is that this strategy is much less specific. Optimisations of releasing methods have been tested to reduce contamination by irrelevant proteins.

(iii) Methods to release trapped proteins

The choice of releasing method is tightly linked to that of previously described trapping stages. We can establish a short list of protocols already tested:

- Direct release in a dedicated buffer

Protein elution can be achieved by a buffer that has increased salt and detergent concentrations compared to washing buffers. Some protocols simply consist in eluting proteins into the loading buffer (for example Laemmli buffer), which is heated and used in the subsequent electrophoretic steps [20]. In the case of proteins covalently bound to DNA probes, the retrieval of unspecific proteins is limited by the preceding washing steps in stringent conditions [22]. The use of an increased temperature is particularly favourable to the release of some proteins crosslinked to DNA lesions such as the thermally labile N7-alkylguanine adducts, by cleavage of the N-glycosidic bond [23].

However, in the case of non-covalent binding strategies, this method brings a high quantity of non-specific proteins not eliminated by washing steps and adsorbed on beads or interacting with non-damaged DNA. Control conditions should therefore include analysis of proteins retained on beads and on non-damaged DNA in order to exclude them from the list of proteins coming from the "beads + damaged DNA" condition [20].

- Non-specific digestion of the nucleic probe

Another way to release trapped proteins is to degrade the DNA probe using nucleases such as DNAse I [24]. The quantity of this additional factor should be as low as possible as it is likely to bring an additional source of contaminating peptides during proteomics analysis.

- Use of a cleavable linker

This smart strategy employs a cleaving agent specifically acting on the link between the solid support (bead) and the nucleic probe [29]. This allows a limited use of buffers releasing non-specific proteins. The cleavage reaction can be either chemical (linker with disulfide bond uncoupled by DTT), physical or enzymatic in nature. In this last case, the linker contains a restriction site digested by the corresponding enzyme [30], which is also susceptible to introduce contaminating peptides. The most promising cleavage strategy is probably of physical nature with

photocleavable linkers (commercially available) such as those using 1-(2 nitrophenyl)ethyl groups. These elements have already been successfully tested by Hégarat et al. [28] in the context of protein/damaged DNA interactions. Briefly, ODN baits containing the linker bound to a biotin moiety are immobilised on streptavidin coated beads. After incubation of this trap with the cellular extract and washing stages, the mixture is UV irradiated, beads are pelleted and DNA-protein complexes released in the supernatant. A filtration-concentration step precedes incubation in Laemmli buffer followed by proteins separation using 1D or 2D electrophoresis.

(iv) Proteomics methods used to identify proteins

The various protocols used by the different authors are roughly similar. Briefly, trapped proteins are separated by 1D SDS-PAGE or sometimes by 2D electrophoresis [28]. After staining, spots or bands are excised from the gel and submitted to trypsin digestion. Peptides and subsequently proteins are identified (with or without preliminary HPLC or nanoHPLC separation) using a MALDI-TOF mass spectrometer [24] or electrospray ionisation tandem mass spectrometry (ESI-MS/MS) tools, refer to these citations for current literature: [23,26]. Modern quantitative proteomics approaches have not yet been applied to protein mixtures coming from damaged-DNA pull down experiments. If the aim was to quantitatively compare proteomes of lesions A and B, one of the main technical challenges would be to control the quantity of proteins coming from trapping stages and subjected to proteomics analyses. This point will be further discussed later.

2.2.3. Other approaches

(i) Identification of chromatin-binding proteins

A complementary approach to identify new proteins associated with DNA damage consists of isolating chromatin-binding proteins after having exposed cells to a genotoxic stress. Kim et al. [31] established by differential in-gel electrophoresis (DIGE) maps of chromatin proteins from budding yeast with or without exposure to the alkylating agent methyl methanesulfonate (MMS). Among proteins identified as upregulated in genotoxic conditions, the authors identified subunits of the replication factor A complex, which is a key player in DNA replication and repair process. Already known yeast DNA damage associated proteins such as BMH1 were also captured along with other candidates for which roles in DNA damage response still need to be confirmed. In human cells, Lee et al. [32] used γ-radiations as a source of genotoxic stress, and isolated chromatin-binding proteins through sequential salt extraction before separating and identifying candidates by 2D electrophoresis and MALDI-TOF mass spectrometry. Nucleophosmin, already cited in DDR publications, and the less expected heterogeneous nuclear ribonucleoproteins (hnRNP) were the main factors highlighted by this work. A last and more recent illustration is given by Chou et al. [33] who took advantage of the stable isotope labelling in cell culture (SILAC) proteomics approach to compare chromatin-binding proteins in cells exposed or not to high doses of UV light. This strategy led to the identification of transcriptional repressors from the polycomb group proteins. Further experiments showed their recruitment to UV laser-induced DNA lesions by PARP.

This family of methods is therefore a fruitful source of information and is more amenable to comparative and quantitative studies than previously described strategies. The main difference is that the identified proteins are far from being limited to players in direct vicinity of DNA damage. Identified candidates can indeed be bound to other chromatin-associated proteins such as histones, to non-damaged DNA, or even be considered as "false positive" proteins since fractionated chromatin is potentially contaminated by other cellular fractions. As an example, nucleophosmin and hnRNP selected by Lee et al. [32] have been confirmed as able to bind chromatin after DNA damage, but they do not accumulate at the specific site of DNA lesion. Moreover, and as it is the case for a lot of approaches described in this review, these techniques are more reliable for highly expressed proteins than for low abundant factors in spite of the fractionation provided by chromatin extraction [34].

(ii) Identification of proteins interacting with already characterised damaged DNA binding factors.

Instead of using damaged DNA as a bait, this strategy consists of isolating partners of proteins known to be in close contact to DNA lesions. With a GST pull-down assay coupled to LC–MS/MS, Lévy et al. [35] showed that XRCC1 (the bait in this work) was interacting with the p58 subunit of DNA polymerase α-primase. This result, along with complementary experiments, suggests that XRCC1 may be involved in the regulation of S phase DNA replication in the context of a single-strand DNA break. Vago et al. [36] used a different protocol to screen DNA ligase I partners, which entailed extracting proteins from HeLa cells in conditions maintaining protein-protein interactions. This can be achieved by using extraction buffers containing non-ionic mild detergents such as Tween and low salt concentrations (50 mM NaCl in this study). The lysate was then deposited onto a commercial antibody array dedicated to cell cycle regulator studies. After incubation with a labelled anti-ligase I antibody, the authors observed an interaction of ligase I with Nbs 1, a factor for repair of DNA double strand breaks. This may indicate a new role for ligase I in this process.

Finally, the team of Chen [37,38] has associated the production of FLAG-tagged H2AX with SILAC to allow the identification of H2AX binding partners in different cell lines and/or after exposure to ionizing radiations. One of the most interesting applications developed is the acquisition of dynamic data, *i.e.* the interactome study at different time points after exposure to stress. Specifically, they noticed dynamic changes for PARP-1/H2AX association, which was detectable 30 min following irradiation but disappeared after 2 h and 30 min.

(iii) *In cellulo* protein crosslinking

We previously mentioned a DNA pull down approach which consisted in artificially crosslinking proteins from nuclear extracts onto modified DNA probes. Another strategy is to create DPCs by exposing cells to chemical agents and then to identify proteins bound to DNA extracted from these cells. On the one hand, Qiu and Wang [39] have considered this approach suitable to trap a range of proteins possessing a wide panel of functions in DNA metabolism (*i.e.* transcription, replication, repair, *etc.*). In HL-60 human leukaemia cells, they used formaldehyde as a crosslinking agent. On the other hand, this tool can also be considered to address toxicological issues by using known toxic agents as crosslinking compounds. Here, the aim is to characterise proteins involved in genotoxic DPCs such as those generated by camptothecin, ionizing radiations [40,41] or 6-thioguanine/UVA combination [42]. In this last example, the authors determined proteins crosslinked to DNA by focusing their attention on those that showed, using DIGE, decreased levels in the nucleus after crosslinking. Trapping of identified factors such as MSH2 was confirmed by immunoblotting of proteins bound on DNA fractions. It has to be noted that Loeber et al. [23] applied a slightly different approach that was closer to previously described for *in vitro* DNA pull down, whereby they co-incubated non-damaged DNA probes with nuclear proteins and mechlorethamine before proteins recovery and identification.

This short list of methods is of course likely to be expanded in the years to come by other approaches. It is of interest to note that, to our knowledge, chromatin immunoprecipitation (ChIP) using antibodies recognizing DNA lesions has not been tested yet to screen the DNA damage interactome.

2.3. Confirmation methods

The methods described in Section 2.2 generate lists containing sometimes hundreds of candidates identified through mass spectrometry or protein micro-arrays. The next step involves selecting those which are biologically relevant and validating newly identified interactions. First, a preliminary approach can be to confirm the results using the same trapping strategy, followed by specific detection of the selected candidate instead of proteomics methods. As an example, the chosen protein can be analysed by Western-blotting in the mixture of proteins obtained after the DNA pull down technique [20]. But this type of redundant validation is not sufficient and has to be strengthened by different and complementary assays.

This stage is essential and useful because it allows the demonstration that the association between damaged DNA and the protein was not the result of a technical artefact. Moreover, these complementary tools provide further information such as kinetic data or parallel testing of the interaction with other types of DNA damage. The following paragraphs describe methods that have been or may be used to verify and further characterise these damaged DNA/protein interactions.

2.3.1. Filter binding assay

In the 1970s, one of the first methods used for studying the specific interactions between DNA and proteins was a filter binding assay. This technique is now rather historical, but its principle of selective retention of proteins, but not of small DNA fragments, on a nitrocellulose membrane, made it a simple and rapid means for analysing specific interactions. The protein of interest is incubated with an ODN previously labelled with ^{32}P, and filtered through a membrane, onto which the protein–DNA complexes may be subsequently visualised. At the same time, the radioactivity measured in the filtrate represents the non-bound DNA, which is a value that enables determination of kinetic parameters of the studied interaction. Rapidity and simplicity of the filter binding assay are highly advantageous, especially when optimisation of the interaction conditions and parameters is required. However, the protein used in the assay must be of high purity, otherwise a multitude of non-specific complexes will be observed and will make any analysis impossible. This technique, initially used for studies of repressor and operator interactions, also permitted the determination of the ability of DDB protein to bind UV-damaged DNA [43].

2.3.2. DNA footprinting

Another rather pioneering method for analysing protein–DNA interactions is DNA footprinting. Described for the first time in 1978 for the study of protein interactions with Lac operon repressors [44], this technique was also used to analyse the binding of proteins to DNA lesions [45]. DNA is incubated with a non-specific endonuclease, DNAse I, which cuts it randomly and generates fragments of different lengths. Since each phosphodiester bond may be cut with equal probability, the action of DNAse I results in a mix of DNA fractions differing by one nucleotide. However, in the presence of a protein, the portion of DNA that participates in its binding will be protected from the endonuclease attack which will therefore provide a different digestion profile after electrophoresis. This way, the localisation of protein binding site on the DNA sequence can also be determined.

A variant of DNA footprinting, taking advantage of the hydroxyl radical instead of DNase I, was also conceived. Its benefit is a better precision in the determination of the sequence necessary for protein binding, which is due to the smaller size of radicals compared to enzymes. The radical, generated by Fenton reaction from H_2O_2, is able to cleave phosphodiester bonds that would be inaccessible to the endonuclease, and therefore provides the information about the nucleotides actually participating in the protein recognition. Thanks to its extreme precision, the hydroxyl footprinting enabled for example the identification of the exact localisation of DNA platinum crosslinks in nucleosomes [46].

2.3.3. Electrophoretic mobility shift assay (EMSA)

EMSA, also called ''gel shift'', is a common and widely used technique for verification of the interactions between damaged DNA and specific proteins. It is based on the principle that, during electrophoretic separation, fragments of damaged DNA that are complexed with a protein migrate slower than DNA alone. Therefore, the migration of fluorescent or radioactively labelled DNA incubated with the protein of interest will result in a band corresponding to free DNA and another one to the complex. If more complexes are to be created, more bands will be visible. In such a way, not only the interaction of interest can be confirmed, but also the range of protein concentration may be tested in order to find the dissociation constant of the complex. As this approach is the most popular, many examples can be found in the literature. We can for example highlight the results of Zhu and Lippard [25] who identified by a DNA pull down assay the mismatch repair factor hMutSβ as being able to bind cisplatin–DNA interstrand crosslinks. This result was confirmed by the authors using EMSA.

Apart from determining the properties of a complex of known identity, EMSA enables a way to control the identity of the protein interacting with a given lesion, if the proper antibody is available. In such case, a band corresponding to the complex DNA/protein/antibody is observed. This method was reported as supershift [47].

It has to be noticed that coupling EMSA to proteomics is possible, but in that case the radioactive labelling has to be replaced by a fluorescent one. DNA is incubated with crude extracts and separated on a 1D or 2D gel. The desired bands or spots are subsequently cut from the gel and analysed by mass spectrometry to identify proteins that generate the complex. Such a combination of these two methods was employed for the study of TFs [48], but we can imagine its application in the study of the damaged DNA interactome as well.

2.3.4. Surface plasmon resonance (SPR)

Based on the principle of an ''optical scale'', SPR relies on the measurement of biomolecular interactions between targets injected in a mobile phase that will interact with bioreceptors immobilised onto the optical sensing surface. This surface is a gold film deposited on a glass slide or prism. The mass variation generated by the target-bioreceptor interaction induces a shift of the refractive index close to the gold surface. This reaction results in a variation of intensity of the reflected light at a given incidence angle. Without the need of target labelling, these optical changes are converted to a real-time signal. Quantitative data regarding kinetics and strength of interactions can be obtained. One drawback of this approach is the need for a large quantity of purified proteins. Regarding the field of damaged DNA-protein interactions, this technique has been performed with modified ODNs as immobilised probes and purified proteins as injected targets. It allowed, for example, the analysis of interactions between the N-methylpurine-DNA glycosylase with ODNs containing hypoxanthine or $1,N^6$-ethenoadenine [49].

A more advanced form of SPR is surface plasmon resonance imaging (SPRi) [50], which enables the real-time visualisation of the interactions and also the parallel analysis of multiple probes on the same SPR surface. Our team used this array configuration to simultaneously validate the interaction between DNA altered by three different platinating agents and the TOX HMG box family member 4 (TOX4) protein previously identified through DNA pull

down and mass spectrometry [20]. We also performed SPRi experiments that allowed the study of Fpg binding and excision activity towards DNA containing oxidative DNA lesions [51].

2.3.5. Fluorescence anisotropy

Fluorescence anisotropy is based on the principle of depolarisation of light by fluorescent molecules. The light emitted by some fluorophores is polarised and it was observed that the degree of the polarisation changes with temperature, viscosity or molecule size. That last property forming the basis of the fluorescence anisotropy technique is that the bigger is the molecule, the smaller is its rotational mobility, and therefore the polarisation of the emitted light is greater.

In the experiments where this method was applied, the fluorescently labelled ODN contained the damage of interest. When the ODN is free in solution, its mobility is elevated, which results in a low light polarisation. Upon binding of a protein to the ODN, its mobility decreases and an increase in emitted light polarisation is observed. The anisotropy changes can be measured by a spectrofluorometer. The binding of XPC-HR23B complex to UV-damaged DNA as well as interactions between XPA, RPA, and cisplatin-damaged DNA, were studied thanks to this method [52,53].

2.3.6. Fluorescence resonance energy transfer (FRET)

FRET is another technique that takes advantage of the fluorescent properties of molecules. Very useful for the study of DNA-protein interactions, FRET necessitates the fluorescent labelling of both interacting partners with a donor molecule and an acceptor one. The principle of this method is based on the transfer of energy between the two fluorescent partners, provided that the emission spectrum of the donor and the absorption spectrum of the acceptor overlap and that acceptor and donor are close to each other. If DNA and protein do interact, the excited donor will transfer energy to the acceptor, which will emit fluorescence. In such way the recognition of deoxyinosine by endonuclease V was investigated [54]. Here, since the recognition step was followed by an excision of damaged DNA by the enzyme, the system also provided data regarding enzymatic cleavage. Lopez-Crapez et al. took advantage of a long-lived emission fluorescent europium cryptate as a donor to study interactions between MutS protein and mismatched DNA duplex or platinated DNA [55].

2.3.7. Isothermal titration calorimetry (ITC)

This method relies on the heat uptake or release associated with binding of biomolecules [56]. In the case of protein/DNA interactions monitoring, equal-volume protein solutions are injected in a DNA solution contained in the "sample cell" of an ITC instrument. This device measures the small heat variation and adjusts it to maintain the temperature in the "sample cell" at the same level as in a "reference cell". The power needed for this operation is accurately recorded at each protein injection. This system thus allows the calculation of quantitative data such as enthalpy, entropy, free energy and also binding constants or number of binding sites. This approach has already been applied to binding between (i) Fpg protein and abasic sites or 8-oxo-dG analogues [57], (ii) thymine DNA glycosylase and G:T mismatches [58], (iii) AlkB and methylated DNA [59] or (iv) DNA photolyase and cyclobutane pyrimidine dimers [60].

2.3.8. In cellulo *approaches*

All the techniques described above study interactions *in vitro*, without taking into consideration the cellular environment. They are extremely useful for confirmation of the recognition of the lesion by a given protein. However the reaction conditions, which are highly specific, must be adapted each time to the protein and the damage analysed. *In cellulo* methods have the considerable advantage over *in vitro* ones that the interactions take place in real conditions. To fit with the aim of detecting proteins in close contact to DNA damage, protocols that have been used are (i) to damage DNA at localised areas using micro-irradiation, (ii) to localise generated lesions by immunofluorescence and (iii) to detect a colocalisation with the protein of interest. As an illustration, Chou et al. [33] induced DSBs in HeLa cells by a high dose of UV delivered by laser microirradiation, indirectly visualised these lesions by detecting γH2AX by immunofluorescence, and then observed a colocalisation with polycomb group members previously identified in a proteomic screen. As another example, Gérard et al. [61] locally generated UV DNA damage, detected these lesions using immunofluorescence with anti-cyclobutane pyrimidine dimer (CPD) antibodies, and demonstrated the colocalisation of altered DNA and the chromatin assembly factor p60 CAF-1. Fluorescence microscopy experiments using antibodies directed against DNA lesions and the simultaneous detection of their associated proteins offer interesting perspectives. For example, the possibility to genetically tag proteins with molecules such as the green fluorescent protein (GFP) or mCherry provides relevant tools already used to study interactions between UV-damaged DNA and NER proteins [62,63]. In addition, these approaches can be optimised to collect data about the temporal and spatial organisations of complex DNA repair processes using fluorescent recovery after photobleaching (FRAP) experiments [62].

Wang and Li [64] designed an original assay (episomal chromatin immunoprecipitation) enabling the detection of proteins recruited at DNA damage sites. A plasmid containing the EBV replication origin and an interstrand crosslink was introduced into human cells in culture. Proteins stalled by the lesion during the EBV episomal replication were detected by ChIP. This strategy identified, for example, the recruitment of Fanconi anemia proteins on these blocking lesions [65].

3. Comments and perspectives

The application of proteomics methods to the study of DNA damage interactome should evolve in the coming years through the development of new techniques. These improvements should allow access to more and more relevant and accurate biological data. This third section provides a non-exhaustive list of technical advances that could be applied to the study of proteins associated to DNA damage. We will also highlight biological issues that have not yet been fully taken into account in abovementioned reference works.

One of the most interesting technical developments in MS-based proteomics methods is the continuous improvement in their quantitative dimension. Some of these approaches are based on the assumption that highly abundant proteins generate more MS-MS fragmentation spectra than less frequent ones. Other versions rely on differential isotope labelling of proteins coming from different biological samples. Even if already mentioned in techniques identifying chromatin-bound proteins, quantitative proteomics has probably not been used as far as it could in the field of DNA damage interactome. Interestingly, some examples can be found in studies dealing with proteins associated with DNA epigenetic modifications [66]. The SILAC strategy was used by Mittler et al. [30] and Bartels et al. [67] to select proteins bound to sequences containing 5-methyl cytosine. In SILAC, cells are grown with or without isotope-labelled amino-acids. In these cases, unlabelled extracts are mixed with DNA pull down probes without 5-methyl cytosine, and heavy-labelled lysates are incubated with baits containing the epigenetic modifications. The MS allows the differentiation of proteins coming from the two types of ligands, which is why the technique can decrease the risk of identifying non-specific proteins (in addition to its relative quantification

capacity). This strategy demonstrated the ability of TF RBP-J to bind DNA in a methylation-dependent way [67]. It would be of interest to apply this kind of approach to baits containing DNA lesions instead of epigenetic modifications. Other recent advances in proteomics not been employed in DNA damage interactome studies yet, is their growing capacity in the determination of proteins PTMs. Indeed PTMs such as phosphorylations, acetylations or ubiquitinations are playing important roles in cellular responses to DNA lesions [12]. It would therefore be of interest to use specific PTMs enrichment strategies, advanced MS-based approaches or databases dedicated to these modifications to study proteins mentioned in this review [68].

Among the other technical improvements that could be fruitfully tested are methods permitting the simultaneous study of several lesions. The optimisation of protein-binding microarrays [14] where different modified ODNs are grafted on chips and exposed to protein mixtures has not been published yet to our knowledge. One could also envision the application of the promising coupling of SPR assays to MS-based proteomics [69,70]. In such approach, damaged DNA probes would be grafted on a SPR chip, allowing the detection of an interaction after injection of a crude whole cellular (or nuclear) extract. This stage would be followed by an on-chip mass spectrometry identification of trapped proteins.

Further technical improvements will be necessary to answer biological issues that are not always easily accessible using techniques we expose in this review. Indeed, even if presenting a high-throughput capacity to identify proteins, these approaches are not always well-suited to screen a sufficiently wide range of important criteria in the study of DNA/protein interactions. The ideal tool would allow the implementation of several conditions for each of these parameters:

- sequence context

It is known that some damaged DNA-associated factors, such as several BER enzymes, are binding DNA according to a sequence context [71]. The screening of proteins interacting with DNA lesions localised within a range of specific sequences, such as various positions within telomeric repeats, is also biologically relevant [72]. For a better mechanistic understanding of DNA repair, it is also interesting to determine on DNA the precise site of protein interaction compared to the localisation of the processed lesion [73]. Tools to detect easily proteins bound to DNA lesions in different sequence contexts would therefore provide valuable information.

- nucleosomal context

With the exception of "*in cellulo*" methods, many of the previously described approaches use naked damaged DNA as a bait to trap proteins. This does therefore not take into account the fact that DNA-protein interactions are often occurring in a chromatin context. The binding of HMGB1 on cisplatin-treated DNA for example, is dependant of nucleosome positioning [74]. *De novo* nucleosome assembly methods by histone deposition on DNA probes have already been evaluated with DNA damaged by UV-C or DNase I [61] to study proteins – damaged DNA interactions in this chromatin context.

- Redox, temperature and other physico-chemical conditions

The experimenter should be able to alter these parameters that may influence the panel of proteins interacting with DNA. Considering again the example of cisplatin-treated DNA and the protein HMGB1, Park and Lippard [75] demonstrated that redox parameters were conditioning their interaction.

- Interaction dynamics

Other valuable information, notably regarding DNA repair processes, are the timing and order of arrival of the different interactome members. Live cell imaging techniques or approaches based on the detection of proteins at locally induced damage can be used to collect these data, but only for individual or very few proteins at the same time [76]. If applied to –omics methods described in this review, access to these dynamic data would require that an interactome cartography be obtained at several time points. This strategy is currently already possible but difficult to perform, as each experiment is often heavy, time-consuming (including data analysis) and expensive.

To add another point to these perspectives, it has to be highlighted that the range of DNA lesions studied with abovementioned interactome screening techniques are rather limited (Table 2). Bulky DNA adducts generated by food or environmental contaminants such as mycotoxins or polycyclic aromatic hydrocarbons would constitute interesting targets to be added on this list. The constantly growing family of complex DNA lesions and/or clusters of DNA damage induced by ionizing radiations also constitute relevant candidates. The number of biological models used to collect DNA damage interactome members, and to validate these interactions, is however still limited, with a predominance of experiments performed on the classic HeLa cell line in cancer related studies [8,20,22,23,25,28,32,33]. Using more diversified cellular and *in vivo* models at tissue or organ levels would probably provide interesting supplementary data since the network of proteins dealing with DNA lesions can vary depending on cell types. The Cdx2 protein interacts with the DDR protein Ku and inhibits DNA repair in colon cancer cells but not in leukaemia cells [77]. As another example, the FACT chromatin remodelling complex, known to interact with DNA lesions such as cisplatin crosslinks, is not expressed in most cells from mouse and human adult tissues but shows higher levels in stem cells [78]. Another way to get original results from DNA damage interactome screening approaches would therefore be to use them in order to trap and identify proteins involved in DDRs expressed within human cells, tissues, or original organisms for which these pathways are not characterised yet.

4. Concluding remarks

Over the past few years, several approaches have been tested to screen with a mid- to high-throughput proteins directly and/or indirectly bound to DNA lesions. We must keep in mind that these techniques are quite different from each other and are therefore not expected to deliver the same picture of DNA damage interactome. None of them is exhaustive, and obtaining a complete cartography of these subproteomes using only one of these assays is of course not achievable. Nevertheless, published works described in this review have demonstrated that even a partial view can highlight new DNA-protein interactions. Recent and future improvements in proteomics methods should enhance the quantity and quality of generated data. For this purpose, specialists in DNA damage and DNA repair biochemistry have to create or maintain fruitful exchanges with proteomics teams. The next step is to go deeper into the understanding of mechanisms and cellular consequences of these newly identified interactions. As one would expect, the combination of the abovementioned competencies should lead to the discovery of new DNA repair pathways in various organisms,

new cellular responses to genotoxic stress or new biomarkers related to DNA damaging cancer treatments.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Acknowledgement

The authors thank Pr Donald Martin (Université Joseph Fourier) for providing language help and comments on the manuscript.

References

[1] K.S. Gates, An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals, Chem. Res. Toxicol. 22 (2009) 1747–1760.
[2] J. Cadet, T. Douki, J.-L. Ravanat, Oxidatively generated base damage to cellular DNA, Free Radic. Biol. Med. 49 (2010) 9–21.
[3] M.P. Stone, H. Huang, K.L. Brown, G. Shanmugam, Chemistry and structural biology of DNA damage and biological consequences, Chem. Biodivers. 8 (2011) 1571–1615.
[4] C.J. Lord, A. Ashworth, The DNA damage response and cancer therapy, Nature 481 (2012) 287–294.
[5] Y. Jung, S.J. Lippard, Direct cellular responses to platinum-induced DNA damage, Chem. Rev. 107 (2007) 1387–1407.
[6] K. Sugasawa, J.M. Ng, C. Masutani, S. Iwai, P.J. van der Spek, A.P. Eker, et al., Xeroderma Pigmentosum Group C protein complex is the initiator of global genome nucleotide excision repair, Mol. Cell. 2 (1998) 223–232.
[7] P. Vichi, F. Coin, J.P. Renaud, W. Vermeulen, J.H. Hoeijmakers, D. Moras, et al., Cisplatin- and UV-damaged DNA lure the basal transcription factor TFIID/TBP, EMBOJ 16 (1997) 7444–7456.
[8] C.X. Zhang, P.V. Chang, S.J. Lippard, Identification of nuclear proteins that interact with platinum-modified DNA by photoaffinity labeling, J. Am. Chem. Soc. 126 (2004) 6536–6537.
[9] D.K. Treiber, X. Zhai, H.M. Jantzen, J.M. Essigmann, Cisplatin–DNA adducts are molecular decoys for the ribosomal RNA transcription factor hUBF (human upstream binding factor), Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 5672–5676.
[10] S.L. Bruhn, P.M. Pil, J.M. Essigmann, D.E. Housman, S.J. Lippard, Isolation and characterization of human cDNA clones encoding a high mobility group box protein that recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 2307–2311.
[11] J.J. Turchi, K.M. Henkels, I.L. Hermanson, S.M. Patrick, Interactions of mammalian proteins with cisplatin-damaged DNA, J. Inorg. Biochem. 77 (1999) 83–87.
[12] M.S.Y. Huen, J. Chen, Assembly of checkpoint and repair machineries at DNA damage sites, Trends Biochem. Sci. 35 (2010) 101–108.
[13] L. Berrade, A. Garcia, J. Camarero, Protein microarrays: novel developments and applications, Pharm. Res. 28 (2011) 1480–1499.
[14] M.L. Bulyk, DNA microarray technologies for measuring protein–DNA interactions, Curr. Opin. Biotechnol. 17 (2006) 422–430.
[15] C.-S. Chen, E. Korobkova, H. Chen, J. Zhu, X. Jian, S.-C. Tao, et al., A proteome chip approach reveals new DNA damage recognition activities in *Escherichia coli*, Nat. Methods. 5 (2008) 69–74.
[16] H.E. Stansfield, B.P. Kulczewski, K.E. Lybrand, E.R. Jamieson, Identifying protein interactions with metal-modified DNA using microarray technology, J. Biol. Inorg. Chem. 14 (2009) 193–199.
[17] J.H. Toney, B.A. Donahue, P.J. Kellett, S.L. Bruhn, J.M. Essigmann, S.J. Lippard, Isolation of cDNAs encoding a human protein that binds selectively to DNA modified by the anticancer drug cis-diamminedichloroplatinum(II), Proc. Natl. Acad. Sci. U.S.A. 86 (1989) 8328–8332.
[18] D. Gasparutto, A.-G. Bourdat, C. D'Ham, V. Duarte, A. Romieu, J. Cadet, Repair and replication of oxidized DNA bases using modified oligodeoxyribonucleotides, Biochimie 82 (2000) 19–24.
[19] J.A. Mello, J.G. Moggs, G. Almouzni, Analysis of DNA repair and chromatin assembly in vitro using immobilized damaged DNA substrates, in: A.H. Schönthal, J.M. Walker (Eds.), Checkpoint Controls and Cancer, Humana Press, 2004, pp. 271–281.
[20] C.B.M. du Puch, E. Barbier, A. Kraut, Y. Couté, J. Fuchs, A. Buhot, et al., TOX4 and its binding partners recognize DNA adducts generated by platinum anticancer drugs, Arch. Biochem. Biophys. 507 (2011) 296–303.
[21] S.D. Putney, S.J. Benkovic, P.R. Schimmel, A DNA fragment with an alpha-phosphorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated in vivo, Proc. Natl. Acad. Sci. U.S.A. 78 (1981) 7350–7354.
[22] E.R. Guggenheim, D. Xu, C.X. Zhang, P.V. Chang, S.J. Lippard, Photo-affinity isolation and identification of proteins in cancer cell extracts that bind to platinum-modified DNA, Chembiochem 10 (2009) 141–157.
[23] R.L. Loeber, E.D. Michaelson-Richie, S.G. Codreanu, D.C. Liebler, C.R. Campbell, N.Y. Tretyakova, Proteomic analysis of DNA-protein cross-linking by antitumor nitrogen mustards, Chem. Res. Toxicol. 22 (2009) 1151–1162.
[24] R.A. Rieger, E.I. Zaika, W. Xie, F. Johnson, A.P. Grollman, C.R. Iden, et al., Proteomic approach to identification of proteins reactive for abasic sites in DNA, Mol. Cell Proteomics 5 (2006) 858–867.
[25] G. Zhu, S.J. Lippard, Photoaffinity labeling reveals nuclear proteins that uniquely recognize cisplatin–DNA interstrand cross-links, Biochemistry 48 (2009) 4916–4925.
[26] M. Winnacker, V. Welzmiller, R. Strasser, T. Carell, Development of a DNA photoaffinity probe for the analysis of 8-OxodG-binding proteins in a human proteome, Chembiochem 11 (2010) 1345–1349.
[27] E.A. Maltseva, N.I. Rechkunova, L.C. Gillet, I.O. Petruseva, O.D. Schärer, O.I. Lavrik, Crosslinking of the NER damage recognition proteins XPC-HR23B, XPA and RPA to photoreactive probes that mimic DNA damages, Biochim. Biophys. Acta 1770 (2007) 781–789.
[28] N. Hégarat, G.M. Cardoso, F. Rusconi, J.-C. François, D. Praseuth, Analytical biochemistry of DNA–protein assemblies from crude cell extracts, Nucleic Acids Res. 35 (2007) e92.
[29] N. Hégarat, J.-C. François, D. Praseuth, Modern tools for identification of nucleic acid-binding proteins, Biochimie 90 (2008) 1265–1272.
[30] G. Mittler, F. Butter, M. Mann, A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements, Genome Res. 19 (2009) 284–293.
[31] D.R. Kim, R.D. Gidvani, B.P. Ingalls, B.P. Duncker, B.J. McConkey, Differential chromatin proteomics of the MMS-induced DNA damage response in yeast, Proteome Sci. 9 (2011) 62.
[32] S.Y. Lee, J.-H. Park, S. Kim, E.-J. Park, Y. Yun, J. Kwon, A proteomics approach for the identification of nucleophosmin and heterogeneous nuclear ribonucleoprotein C1/C2 as chromatin-binding proteins in response to DNA double-strand breaks, Biochem. J. 388 (2005) 7–15.
[33] D.M. Chou, B. Adamson, N.E. Dephoure, X. Tan, A.C. Nottke, K.E. Hurov, et al., A chromatin localization screen reveals poly (ADP ribose)-regulated recruitment of the repressive polycomb and NuRD complexes to sites of DNA damage, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 18475–18480.
[34] B.N. Wardleworth, J.A. Downs, Spotting new DNA damage-responsive chromatin-binding proteins, Biochem. J. 388 (2005).
[35] N. Lévy, M. Oehlmann, F. Delalande, H.P. Nasheuer, A. Van Dorsselaer, V. Schreiber, et al., XRCC1 interacts with the p58 subunit of DNA pol α-primase and may coordinate DNA repair and replication during S phase, Nucleic Acids Res. 37 (2009) 3177–3188.
[36] R. Vago, V. Leva, G. Biamonti, A. Montecucco, DNA ligase I and Nbs1 proteins associate in a complex and colocalize at replication factories, Cell Cycle 8 (2009) 2600–2607.
[37] Y.-C. Du, S. Gu, J. Zhou, T. Wang, H. Cai, M.A. MacInnes, et al., The dynamic alterations of H2AX complex during DNA repair detected by a proteomic approach reveal the critical roles of Ca^{2+}/calmodulin in the ionizing radiation-induced cell cycle arrest, Mol. Cell Proteomics 5 (2006) 1033–1044.
[38] X. Yang, P. Zou, J. Yao, D. Yun, H. Bao, R. Du, et al., Proteomic dissection of cell type-specific H2AX-interacting protein complex associated with hepatocellular carcinoma, J. Proteome Res. 9 (2010) 1402–1415.
[39] H. Qiu, Y. Wang, Exploring DNA-binding proteins with in vivo chemical cross-linking and mass spectrometry, J. Proteome Res. 8 (2009) 1983–1991.
[40] S. Barker, D. Murray, J. Zheng, L. Li, M. Weinfeld, A method for the isolation of covalent DNA–protein crosslinks suitable for proteomics analysis, Anal. Biochem. 344 (2005) 204–215.
[41] S. Barker, M. Weinfeld, J. Zheng, L. Li, D. Murray, Identification of mammalian proteins cross-linked to DNA by ionizing radiation, J. Biol. Chem. 280 (2005) 33826–33838.
[42] Q. Gueranger, A. Kia, D. Frith, P. Karran, Crosslinking of DNA repair and replication proteins to DNA in cells treated with 6-thioguanine and UVA, Nucleic Acids Res. 39 (2011) 5057–5066.
[43] R.S. Feldberg, L. Grossman, A DNA binding protein from human placenta specific for ultraviolet damaged DNA, Biochemistry 15 (1976) 2402–2408.
[44] D.J. Galas, A. Schmitz, DNAse footprinting: a simple method for the detection of protein–DNA binding specificity, Nucleic Acids Res. 5 (1978) 3157–3170.
[45] R. Visse, M. De Ruijter, J. Brouwer, J.A. Brandsma, P. Van De Putte, UVR excision repair protein complex of *Escherichia coli* binds to the convex side of a cisplatin-induced kink in the DNA, J. Biol. Chem. 266 (1991) 7609–7617.
[46] A.J. Danford, D. Wang, Q. Wang, T.D. Tullius, S.J. Lippard, Platinum anticancer drug damage enforces a particular rotational setting of DNA in nucleosomes, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 12311–12316.
[47] S.M. Patrick, J.J. Turchi, Xeroderma Pigmentosum Complementation Group A Protein (XPA) modulates RPA–DNA interactions via enhanced complex stability and inhibition of strand separation activity, J. Biol. Chem. 277 (2002) 16096–16101.
[48] J.A. Stead, J.N. Keen, K.J. McDowall, The identification of nucleic acid-interacting proteins using a simple proteomics-based approach that directly incorporates the electrophoretic mobility shift assay, Mol. Cell Proteomics 5 (2006) 1697–1702.
[49] S. Adhikari, A. Üren, R. Roy, Excised damaged base determines the turnover of human N-methylpurine-DNA glycosylase, DNA Repair (Amst.) 8 (2009) 1201–1206.
[50] S. Scarano, M. Mascini, A.P.F. Turner, M. Minunni, Surface plasmon resonance imaging for affinity-based biosensors, Biosens. Bioelectron. 25 (2010) 957–966.
[51] C. Corne, J.-B. Fiche, D. Gasparutto, V. Cunin, E. Suraniti, A. Buhot, et al., SPR imaging for label-free multiplexed analyses of DNA N-glycosylase interactions with damaged DNA duplexes, Analyst 133 (2008) 1036–1045.
[52] T. Hey, G. Lipps, G. Krauss, Binding of XPA and RPA to damaged DNA investigated by fluorescence anisotropy, Biochemistry 40 (2001) 2901–2910.
[53] T. Hey, G. Lipps, K. Sugasawa, S. Iwai, F. Hanaoka, G. Krauss, The XPC–HR23B complex displays high affinity and specificity for damaged dna in a true-equilibrium fluorescence assay, Biochemistry 41 (2002) 6583–6587.

[54] J. Lin, H. Gao, K.A. Schallhorn, R.M. Harris, W. Cao, P.C. Ke, Lesion recognition and cleavage by endonuclease V, Biochemistry 46 (2007) 7132–7137.
[55] E. Lopez-Crapez, J.-M. Malinge, F. Gatchitch, L. Casano, T. Langlois, M. Pugnière, et al., A homogeneous resonance energy transfer-based assay to monitor MutS/DNA interactions, Anal. Biochem. 383 (2008) 301–306.
[56] J.E. Ladbury, Application of isothermal titration calorimetry in the biological sciences: things are heating up! Biotechniques 37 (2004) 885–887.
[57] C.A.S. Minetti, D.P. Remeta, D.O. Zharkov, G. Eric Plum, F. Johnson, A.P. Grollman, et al., Energetics of lesion recognition by a DNA repair protein: thermodynamic characterization of formamidopyrimidine-glycosylase (Fpg) interactions with damaged DNA duplexes, J. Mol. Biol. 328 (2003) 1047–1060.
[58] A. Maiti, M.T. Morgan, E. Pozharski, A.C. Drohat, Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 8890–8895.
[59] B. Bleijlevens, T. Shivarattan, E. Flashman, Y. Yang, P.J. Simpson, P. Koivisto, et al., Dynamic states of the DNA repair enzyme AlkB regulate product release, EMBO Rep. 9 (2008) 872–877.
[60] T.J. Wilson, M.A. Crystal, M.C. Rohrbaugh, K.P. Sokolowsky, Y.M. Gindt, Evidence from thermodynamics that DNA photolyase recognizes a solvent-exposed CPD lesion, J. Phys. Chem. B 115 (2011) 13746–13754.
[61] A. Gérard, S.E. Polo, D. Roche, G. Almouzni, Methods for studying chromatin assembly coupled to DNA repair, DNA Rep., Part B (2006) 358–374 (Academic Press).
[62] W. Vermeulen, Dynamics of mammalian NER proteins, DNA Repair (Amst.) 10 (2011) 760–771.
[63] M.S. Luijsterburg, M. Lindh, K. Acs, M.G. Vrouwe, A. Pines, H. van Attikum, et al., DDB2 promotes chromatin decondensation at UV-induced DNA damage, J. Cell Biol. 197 (2012) 267–281.
[64] Y. Wang, L. Li, Detecting recruitment of DNA damage response factors through the eChIP approach, in: W.X. Li, J.M. Walker (Eds.), Cell Cycle Checkpoints, Humana Press, 2011, pp. 245–255.
[65] X. Shen, H. Do, Y. Li, W.-H. Chung, M. Tomasz, J.P. de Winter, et al., Recruitment of Fanconi anemia and breast cancer proteins to DNA damage sites is differentially governed by replication, Mol. Cell. 35 (2009) 716–723.
[66] H.C. Eberl, M. Mann, M. Vermeulen, Quantitative proteomics for epigenetics, Chembiochem 12 (2011) 224–234.
[67] S.J.J. Bartels, C.G. Spruijt, A.B. Brinkman, P.W.T.C. Jansen, M. Vermeulen, H.G. Stunnenberg, A SILAC-based screen for methyl-CpG binding proteins identifies RBP-J as a DNA methylation and sequence-specific binding protein, PLoS One 6 (2011).
[68] K.S. Kamath, M.S. Vasavada, S. Srivastava, Proteomic databases and tools to decipher post-translational modifications, J. Proteomics 75 (2011) 127–144.
[69] J. Buijs, G.C. Franklin, SPR-MS in functional proteomics, Brief Funct. Genomics Proteomic 4 (2005) 39–47.
[70] S. Bellon, W. Buchmann, F. Gonnet, N. Jarroux, M. Anger-Leroy, F. Guillonneau, et al., Hyphenation of surface plasmon resonance imaging to matrix-assisted laser desorption ionization mass spectrometry by on-chip mass spectrometry and tandem mass spectrometry analysis, Anal. Chem. 81 (2009) 7695–7702.
[71] K. Donigan, J.B. Sweasy, Sequence context specific mutagenesis and base excision repair, Mol. Carcinog. 48 (2009) 362–368.
[72] D.B. Rhee, A. Ghosh, J. Lu, V.A. Bohr, Y. Liu, Factors that influence telomeric oxidative base damage and repair by DNA glycosylase OGG1, DNA Repair (Amst.) 10 (2011) 34–44.
[73] Y.S. Krasikova, N.I. Rechkunova, E.A. Maltseva, I.O. Petruseva, O.I. Lavrik, Localization of xeroderma pigmentosum group A protein and replication protein A on damaged DNA in nucleotide excision repair, Nucleic Acids Res. 38 (2010) 8083–8094.
[74] I. Ugrinova, S. Zlateva, I.G. Pashev, E.A. Pasheva, Native HMGB1 protein inhibits repair of cisplatin-damaged nucleosomes in vitro, Int. J. Biochem. Cell Biol. 41 (2009) 1556–1562.
[75] S. Park, S.J. Lippard, Redox state-dependent interaction of HMGB1 and cisplatin-modified DNA, Biochemistry 50 (2011) 2567–2574.
[76] S. Feuerhahn, J.-M. Egly, Tools to study DNA repair: what's in the box? Trends Genet. 24 (2008) 467–474.
[77] B. Renouf, C. Soret, T. Saandi, F. Delalande, E. Martin, M. Vanier, et al., Cdx2 homeoprotein inhibits non-homologous end joining in colon cancer but not in leukemia cells, Nucleic Acids Res. 40 (2012) 3456–3469.
[78] H. Garcia, D. Fleyshman, K. Kolesnikova, A. Safina, M. Commane, G. Paszkiewicz, et al., Expression of FACT in mammalian tissues suggests its role in maintaining of undifferentiated state of cells, Oncotarget 2 (2011) 783–796.
[79] E.N. Hughes, B.N. Engelsberg, P.C. Billings, Purification of nuclear proteins that bind to cisplatin-damaged DNA. Identity with high mobility group proteins 1 and 2, J. Biol. Chem. 267 (1992) 13520–13527.
[80] J. Yaneva, S.H. Leuba, K. van Holde, J. Zlatanova, The major chromatin protein histone H1 binds preferentially to *cis*-platinum-damaged DNA, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 13448–13451.
[81] H. Ishiguchi, H. Izumi, T. Torigoe, Y. Yoshida, H. Kubota, S. Tsuji, et al., ZNF143 activates gene expression in response to DNA damage and binds to cisplatin-modified DNA, Int. J. Cancer 111 (2004) 900–909.
[82] C.S. Chow, J.P. Whitehead, S.J. Lippard, HMG domain proteins induce sharp bends in cisplatin-modified DNA, Biochemistry 33 (1994) 15124–15130.
[83] C.C. Wetzel, S.J. Berberich, DNA binding activities of p53 protein following cisplatin damage of ovarian cells, Oncol. Res. 10 (1998) 151–161.
[84] T. Ise, G. Nagatani, T. Imamura, K. Kato, H. Takano, M. Nomoto, et al., Transcription factor Y-box binding protein 1 binds preferentially to cisplatin-modified DNA and interacts with proliferating cell nuclear antigen, Cancer Res. 59 (1999) 342–346.
[85] C.J. Jones, R.D. Wood, Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA, Biochemistry 32 (1993) 12096–12104.
[86] C.K. Clugston, K. McLaughlin, M.K. Kenny, R. Brown, Binding of human single-stranded DNA binding protein to DNA damaged by the anticancer drug cis-diamminedichloroplatinum (II), Cancer Res. 52 (1992) 6375–6379.
[87] J.A. Mello, S. Acharya, R. Fishel, J.M. Essigmann, The mismatch-repair protein hMSH2 binds selectively to DNA adducts of the anticancer drug cisplatin, Chem. Biol. 3 (1996) 579–589.
[88] M. Kartalou, L.D. Samson, J.M. Essigmann, Cisplatin adducts inhibit 1,N^6-ethenoadenine repair by interacting with the human 3-methyladenine DNA glycosylase, Biochemistry 39 (2000) 8032–8038.
[89] J.J. Turchi, K. Henkels, Human Ku autoantigen binds cisplatin-damaged DNA but fails to stimulate human DNA-activated protein kinase, J. Biol. Chem. 271 (1996) 13861–13867.
[90] A. Payne, G. Chu, Xeroderma pigmentosum group E binding factor recognizes a broad spectrum of DNA damage, Mutat. Res. 310 (1994) 89–102.
[91] S. Kobayashi, M. Furukawa, C. Dohi, H. Hamashima, T. Arai, A. Tanaka, Topology effect for DNA structure of cisplatin: topological transformation of cisplatin-closed circular DNA adducts by DNA topoisomerase I, Chem. Pharm. Bull. 47 (1999) 783–790.
[92] G. Zhu, P. Chang, S.J. Lippard, Recognition of platinum-DNA damage by poly(ADP-ribose) polymerase-1, Biochemistry 49 (2010) 6177–6183.

Publication en cours de rédaction

A DNA pull-down assay coupled to proteomics demonstrates that DNA binding of the p21 repressor ZBTB2 is inhibited by cytosine hydroxymethylation.

Céline Lafaye[1], Ewa Barbier[1], Audrey Miscioscia[1], Christine Saint-Pierre[1], Alexandra Kraut[2], Yoann Couté[2], Isabelle Plo[3], Didier Gasparutto[1], Jean-Luc Ravanat[1], Jean Breton[1,4]

[1] Laboratoire Lésions des Acides Nucléiques, Service de Chimie Inorganique et Biologique, UMR E_3 CEA/UJF-Grenoble 1, INAC, 17 rue des Martyrs, Grenoble, F-38054, France

[2] Etude de la Dynamique des Protéomes, Biologie à Grande Echelle, UMR S_1038 CEA/INSERM/UJF-Grenoble 1, iRTSV, 17 rue des Martyrs, Grenoble, F-38054, France

[3] Hématopoièse normale et pathologique, UMR 1009 INSERM/IGR/Université Paris 11, Institut Gustave Roussy, 114 rue Edouard Vaillant, Villejuif, F-94805, France

[4] UFR de Pharmacie, Université Joseph Fourier – Grenoble 1, Domaine de la Merci, La Tronche, F-38706, France

ABSTRACT

Recent studies have demonstrated that the modified base 5-hydroxymethylcytosine (5-hmC) is detectable at various rates in DNA extracted from human tissues or cellular and animal experimental models. This oxidative product of 5-methylcytosine (5-mC) constitutes a new and important actor of epigenetic mechanisms. We designed a DNA pull down assay to trap and identify through proteomic tools nuclear proteins bound to 5-hmC and/or 5-mC. We applied this strategy to three cancerous cell lines (HeLa, SH-SY5Y and UT7-MPL) in which we also measured 5-mC and 5-hmC levels by HPLC-MS-MS. We found that the putative oncoprotein Zinc finger and BTB domain-containing protein 2 (ZBTB2) is associated with methylated DNA sequences and that this interaction is inhibited by the presence of 5-hmC replacing 5-mC. As published data mention ZBTB2 recognition of p21 regulating sequences, we verified by electromobility shift assays that this sequence specific binding was also alleviated by 5-hmC. ZBTB2 being considered as a multifunctional cell proliferation activator, notably through p21 repression, this work points out new epigenetic processes potentially involved in carcinogenesis.

Methylation of DNA cytosine at position C5 (5-methylcytosine, 5-mC) is one of the most extensively studied epigenetic mark[1]. One of its oxidation products, 5-hydroxymethylcytosine (5-hmC), has been identified in the 1950s in bacteriophages[2] and twenty years later in rodent DNA[3]. This modified base led to a few publications until 2009 when two groups simultaneously reported the presence of 5-hmC in mouse nuclear DNA extracted from embryonic stem cells[4] and brain cells[5]. Both studies opened the way for an impressive amount of original researches demonstrating that 5-hmC levels regulation plays a role in stem cell pluripotency modulation, embryo development or carcinogenesis[6,7].

Some of the key regulation mechanisms involving DNA elements containing 5-mC and 5-hmC are governed by their interactions with proteins[1,8]. A first type of interaction is of enzymatic nature. As an example, the TET2 (Ten Eleven Translocation 2) oxygenase recognizes 5-mC and converts it to 5-hmC. Secondly, it is for exemple now well known that hypermethylation of CpG islands located in genes promoter regions are associated with transcription abrogation. This can be explained by 5-mC directly inhibiting DNA binding of transcription factors such as USF[9]. Another demonstrated process is the intervention of methyl-CpG binding proteins (MBPs) which recruit chromatin modifying proteins such as histone deacetylases also blocking the transcription process. The first identified MBP was MeCP2[10], one of the mechanisms explaining its capacity to repress transcription consists in recruiting histone deacetylases[11]. A wide range of data about protein interactions with methylated DNA is now available[8]. On the contrary, elements about protein interactions with 5-hmC are still scarce. It has been demonstrated that DNA binding of some MBPs such as MeCP2[12] or ZFP57[13] is inhibited by the presence of 5-hmC unlike Uhrf1. This DNA methylation maintenance factor is indeed able to recognize DNA sequences containing 5-mC as well as those containing 5-hmC[14].

In this context, the aim of this work is to identify new DNA interacting proteins sensible to the hydroxymethylation status of CpG rich sequences. We begin by controlling the presence of 5-hmC in the selected cellular models, its level is measured by a highly sensible and specific protocol using HPLC coupled to tandem mass spectrometry (HPLC-MS/MS). We next develop a DNA pull down assay to trap proteins associated to 5-mC and/or 5-hmC which are then identified by proteomic tools. Proteomics results are verified by western-blotting using antibodies against the identified candidates. Finally, the selected protein candidate is produced, purified and used along with modified oligodeoxynucleotides (ODNs) of biologically relevant sequences in order to check the interaction by ElectroMobility Shift Assay (EMSA).

Measurement of 5-mC and 5-hmC in cellular models by HPLC-MS/MS.

Our screening for DNA binding proteins was realized in nuclear extracts isolated from 3 cancerous cell lines: HeLa (derived from cervical cancer), SH-SY5Y (derived from neuroblastoma) and UT7-MPL (derived from megakaryoblastic leukemia). Taking into consideration that 5-hmC rates are not known in some of these models, we first verified its presence. The level of cytosine, 5-mC and 5-hmC (actually their nucleosides) was quantified by HPLC-MS/MS after extraction and enzymatic hydrolysis of DNA.

DNA pull down of proteins interacting with 5-mC and 5-hmC containing probes.

We next optimize a DNA ligand fishing tool to pull down proteins bound to 5-mC or 5-hmC sequences. Associated to a subsequent proteomic analysis, this strategy has already been applied to find new proteins recognizing DNA lesions such as oxidative DNA damage or

cisplatin induced crosslinks[15-17]. Regarding epigenetic modifications, at least two groups designed methyl-CpG baits coupled to the SILAC (Stable Isotope Labeling by Amino acids in Cell culture) proteomic approach[18,19]. To our knowledge, traps including 5-hmC rather than 5-mC have never been tested. The principle of our system is depicted in figure 2. The first step consists in synthetizing modified DNA probes. We produced them using PCR (Polymerase Chain Reaction) by introducing dCTP (2'-deoxycytidine 5'-triphosphate), 5-mdCTP (5-methyl-2'-deoxycytidine 5'-triphosphate) or 5-hmdCTP (5-hydroxymethyl-2'-deoxycytidine 5'-triphosphate) to obtain baits containing respectively cytosine, 5-mC or 5-hmC[20]. These PCR products are amplified using biotinylated primers allowing their grafting on streptavidin coated magnetic beads. We used PCR products in order to include several DNA modifications within a reasonably long probe. Indeed, shorter ODN baits containing few modified cytosines would retain lower quantities of proteins hence affecting the sensibility of our method. This system was then exposed to nuclear extracts. Two incubation buffers and conditions (including different salts concentrations, temperatures, incubation times) were tested in order to enlarge the panel of retained proteins. Four distinct subproteomes are obtained for each experiment: proteins released from beads without DNA probe (to exclude proteins non specifically adsorbed) and proteins released from C, 5-mC and 5-hmC containing baits. Candidates are identified through proteomics protocols (nano-liquid chromatography coupled to tandem mass spectrometry). The most significant result was the remarkably reproducible identification of ZBTB2 (Zinc finger and BTB domain-containing protein 2) trapped on C and 5-mC containing probes, while systematically missing on 5-hmC baits (figure 3a). This observation was done for the three cell lines tested and did not vary with the different conditions of baits exposition to nuclear extracts. We confirmed this proteomics identification of ZBTB2 by analyzing proteins isolated from the same traps using western-blotting (figure 3b). Other MBPs were observed using this DNA pull down strategy although they were not

Figure 2

Figure 3

identified in all tested cell lines and experimental conditions as it was the case for ZBTB2. MeCP2 was trapped on the 5-mC bait in the UT7-MPL cell line and we confirmed in the HeLa cell line that uhrf1 is able to bind 5-mC as well as 5-hmC containing DNA (figure 3a). Differences in the identity of trapped proteins throughout the different cell extracts can, at least partly, be explained by differences in protein expressions in our cellular models. The expression of MeCP2 is for example low in the HeLa cell line[21].

Study of ZBTB2 interactions with 5-mC and 5-hmC by EMSA

The aim of these gel shift assays was to further validate results previously obtained by DNA pull down, proteomics and western-blot using a complementary approach. ZBTB2 is a potential cell cycle modulator through a negative regulation of the ARF-HDM2-p53-p21 pathway[22]. One of the mechanisms may involve *p21* transcription repression by ZBTB2 binding on the proximal promoter GC-box 5/6 which should normally be recognized by the transcription factor Sp1. We therefore designed two types of EMSA probes: ODNs mimicking one portion of DNA baits used in the DNA pull down assay and ODNs containing a sequence taken from the GC-box 5/6 found in the p21 promoter region (later designated in this paper as "*p21* ODNs"). Each ^{32}P labeled probe was synthetized into 5 versions: ODNs containing only C, ODNs containing 1 or 3 5-mC and ODNs containing 1 or 3 5-hmC. The protein that was produced and purified to perform EMSA was the N-terminal zinc finger domain of ZBTB2. In order to compare ZBTB2 EMSA results with those of a known MBP, the methyl-CpG binding domain (MBD) of MeCP2 was also produced and submitted to the same gel shift assays.

EMSA confirmed that the zinc finger domain of ZBTB2 is able to bind C and 5-mC containing ODNs, this interaction being inhibited by 5-hmC introduction (Figure 4).

Figure 4

Electrophoretic profiles were slightly different according to the sequence context: binding of ZBTB2 was demonstrated on the *p21* ODNs and on random ODNs whereas interaction with sequences containing only C was only visible in *p21* ODNs. We can make the hypothesis that the affinity of ZBTB2 for the *p21* sequence is stronger than for the random sequence, trapping on this latter DNA fragment (in EMSA and DNA pool down assays) being explained by its high content of CpG dinucleotides. Parallel control experiments performed with MeCP2 MBD showed as previously published that this protein recognizes DNA substrates containing 5-mC but not those containing 5-hmC[12]. We also observed that the shift using MeCP2 MBD was obtained using lower quantities of DNA and protein in comparison with ZBTB2 zinc finger domain. These data showed that the affinity of ZBTB2 for the tested DNA sequences is weaker than the affinity demonstrated by MeCP2.

The goal of this study was to identify DNA binding proteins sensible to cytosine hydroxymethylation. Our tool consisted in associating proteomics to a DNA pull down assay based on magnetic beads grafted with PCR products containing C, 5-mC or 5-hmC. When interpreting data extracted from these ligand fishing strategies, one has to keep in mind that they can not provide an exhaustive list of all proteins in close contact to DNA modifications. Indeed, the identity of trapped proteins can vary according to parameters such as buffers stringency, crosslinking of proteins on DNA probes, DNA sequence context, temperature, duration of DNA/extract contact or nature of the biological model[17]. However, even if not exhaustive, these approaches demonstrated their usefulness by highlighting new interactions between proteins and DNA modifications. As an illustration, these assays demonstrated that PARP-1[23] or the protein complex PTW/PP[16] were able to bind cisplatin induced DNA crosslinks. In the field of epigenetic modifications, a similar test identified the TF RBP-J as a new methyl-CpG binding protein[19]. In our hands, screening experiments pointed out ZBTB2.

We decided to focus our attention on this candidate because: (i) this protein was reliably identified in all condition tested as being able to bind 5-mC rich sequences but not their 5-hmC rich counterparts, (ii) it has recently been characterized as a regulator of *p21* transcription that may imply an interesting role in tumor biology[22].

ZBTB2 belongs to the BTB/POZ (Broad complex, Tramtrak, Bric à brac / Pox virus and zinc finger) family of zinc finger transcription factors. To date, 3 members of this family have already been recognized as MBPs able to block transcription: Kaiso/ZBTB33[24], ZBTB4 and ZBTB38[25]. In addition to their N-terminal BTB/POZ domain involved in protein/protein interactions, these 3 proteins contain in their N terminal parts Kruppel-like zinc fingers known to bind a consensus Kaiso binding site and methylated DNA. Published data about ZBTB2 are limited to a few studies including [22] indicating its capacity to repress the ARF-HDM2-p53-p21 pathway. One of the identified mechanisms is a Sp1 / ZBTB2 binding competition for GC-boxes 5/6 contained in p21 regulating sequences. ZBTB2 is also able to compete with p53 for p53 binding elements found in p21 regulating sequences. ZBTB2 also affects cell cycle control by directly interacting with Sp1 and p53 proteins. It was recently postulated that *ZBTB2* acts as an oncogene in human gastric cancers and that its repressor microRNA miR-149 may be considered as a tumor suppressor[26]. Finally, *ZBTB2* may play a role in the development of chronic myeloid leukemia as a genome-wide association study highlighted a chromosomal locus corresponding to this gene[27].

This report is to our knowledge the first indicating that ZBTB2 DNA binding is sensible to cytosine epigenetic modifications. ZBTB2 association with DNA is not affected by cytosine methylation but by its further oxidation into 5-hmC. It is worth to note that ZBTB2 binding mechanisms to methylated DNA are probably different from those of Kaiso, ZBTB4 and ZBTB38. Indeed, ZBTB2 N-terminal zinc finger domain sequence does not match with

Kaiso-like zinc fingers. Additionally, ZBTB2 recognizes un-methylated GC-box 5/6 which sequence is not containing Kaiso binding sites.

Already published experiments seem to indicate that ZBTB2 can be considered as an oncoprotein, for exemple through cell proliferation stimulation by p21 repression. If cytosine hydroxymethylation inhibits this ZBTB2 deleterious property, one may assume that 5-hmC protects the cell from this carcinogenic process. This hypothesis needs of course further validations and studies exploring for exemple 5-hmC impact on DNA binding proteins competing with ZBTB2 such as Sp1. This exemple highlights that 5-hmC and not only 5-mC has now to be taken into account when studying epigenetic regulation of key cellular mechanisms involved in carcinogenesis. The improvement of tools aimed at trapping and identifying proteins associated with DNA modifications should help to unravel these yet unknown processes.

METHODS

Cellular models

SH-SY5Y, UT7-MPL and HeLa extracts were prepared in our laboratory. UT7-MPL cell line was a gift from I. Plo (Institut Gustave Roussy, France). Cells were grown at 37°c under 5% CO_2 in MEM Eagle medium supplemented with 2 mM L-glutamine, 10% (v/v) fetal calf serum, 50 µg/ml penicillin-streptomycin (Invitrogen) and 5 ng/ml GM-CSF (Sigma) for UT7-mpl cells; RPMI 1640 medium supplemented with 2 mM L-glutamine, 10% (v/v) fetal calf serum and 800 µg/ml geneticin (Invitrogen) for HeLa cells and RPMI 1640 medium supplemented with 2 mM L-glutamine, 10% (v/v) fetal calf serum and 800 µg/ml geneticin

(Invitrogen) for UT7-mpl cells. Cells were passaged or harvested at 80% confluence. Commercial HeLa nuclear protein extracts from CIL Biotech (Mons, Belgium) were also used.

5-mC and 5-hmC quantification

DNA extraction

HPLC-MS-MS

Trapping of proteins bound to 5-mC and/or 5-hmC

Probe synthesis - The procedure for the production of oligonucleotides containing modified cytosines was derived from [20]. It was accomplished using polymerase chain reaction (PCR) amplification through the use of modified deoxycytidine triphosphates, 5-methyldeoxycytidinetriphosphate (5mdCTP) (Fermentas) and 5-hydroxymethyldeoxycytidinetriphosphate (5hmdCTP) (Bioline). A starting amount of 0.5 ng of single-stranded 161-mer DNA sequence (sequence 5'-CCTCACCATCTCAACCAATATATATTATCGATATTACGTTATCGTATACGTAATCGTTATATTAACGTTACGTATACGTTTATACGCGTTTTAACGTATTACGATTATATAAAACGTACGTAATATTACGTCGATATTACGTATCGTTAATTATTGAGGGAGAAGTGGTGA- 3') containing one BstUI restriction sites (5'-CGCG) was used to generate DNA amplicons by PCR reactions with reaction buffer containing 0.1mM of each dNTP (or 5mdCTP or 5hmdCTP in place of dCTP), and Taq DNA polymerase (Fermentas). PCR cycling conditions in 25 µl reaction volumes were as follows: 94°C for 2 min and then 22 cycles of PCR at 94°C for 20 s, 55°C for 25 s and 72°C for 30 s, followed by a final extension

step at 72°C for 2 min, using the forward biotinylated primer 5'-CCTCACCATCTCAACCAATA-3' and the reverse primer 5'-TCACCACTTCTCCCTCAAT-3'. In order to effectively remove unmodified DNA templates from the final products, another 30 cycles of subsequent PCR amplifications were performed using 0.5 µl of first round PCR products in 50 µl of reaction volume under the same reaction conditions. PCR products were then purified using PCR purification kits (Invitrogen). These three oligonucleotides containing C, 5-mC or 5-hmC are referred to as ODN-C, ODN-5-mC and ODN-5-hmC respectively. In order to confirm the presence of modified cytosine, the purified PCR products were digested with the methylation-sensitive restriction enzyme BstUI (Fermentas) and the digested products were separated by electrophoresis on a 3% agarose gel containing ethidium bromide and were visualized by UV light.

Biotinylated probes were finally attached to magnetic beads (Dynabeads kilobaseBINDER, Invitrogen) according to the manufacturer instructions, using 500 µg beads for 1.5 µg DNA probe. Functionalised beads were then rinsed twice with water, and then stored in water at 4°C.

DNA pull down - Principle for this step was adapted from [16]. Nuclear extracts were prepared as previously described [28]. Two experimental conditions named DPD1 and DPD2 were used. Protein binding buffer (buffer A-DPD1 (40 mM Hepes-KOH pH7.8, 50 mM KCl, 5 mM $MgCl_2$, 0.5 mM DTT) or buffer A-DPD2 (50 mM NaCl, 50 mM Tris-Hcl pH 8.0, 0.25% NP40, 0.5 mM DTT and complete protease inhibitors-EDTA (Roche)) was used to equilibrate functionalised beads. Nuclear extracts (in exception of the commercial one) were dialysed against protein binding buffer using amicon-Ultra-0.5, 3kDa (Merck Millipore). In a 100 µL final volume, 80 µg of functionalised beads, 120 µg of fish sperm DNA (Roche), and 100 µg of nuclear extracts were mixed in protein binding buffer. Samples were incubated under

shaking for 45 min at 37 °C (DPD1) or 2 h at 4 °C (DPD2). Supernatant was eliminated using a magnetic tray, and beads were rinsed twice with 100 µL of buffer A-DPD2 or buffer B-DPD1 (150 mM NaCl, 20 mM Hepes-KOH pH 7.8, 5 mM $MgCl_2$, 0.025% Nonidet-P-40). Proteins were finally released in 20 µL of Laemmli buffer (Tris–HCl pH 6.8 25 mM, SDS 2%, glycerol 4%, β-mercaptoethanol 5%, bromophenol blue 0.0008%) by heating at 90 °C for 10 min.

Proteomic

Proteins contained in ligand fishing eluates were stacked by SDS–PAGE in a 2 mm-high band and stained with Coomassie blue. Protein bands were manually excised from the gels, proteins digested in-gel with trypsin, and resulting peptides extracted using a sample preparation robot (EVO150, Tecan) as previously described [29]. Extracted peptides were analyzed by nano-liquid chromatography coupled to tandem mass spectrometry (Ultimate 3000, Dionex and LTQ-Orbitrap Discovery, Thermo Fischer) and identified using Mascot software (Matrix Science) and the Uniprot database (*Homo sapiens* taxonomy) as described in [30]. Proteins identified with a minimum score of 35 were automatically validated using IRMa software [31]. Functions of identified proteins were assessed using Generic GO Term Mapper (http://go.princeton.edu).

Western blot

Proteins obtained by DNA pull down were separated by SDS–PAGE, transferred to nitrocellulose and analyzed by Western blot using the monoclonal rabbit polyclonal antibody anti-ZBTB2 (Abnova PAB16678). Secondary antibody was a monkey anti-rabbit coupled to

horseradish-peroxydase (GE Healthcare). ECL Plus and Hyperfilm ECL (GE Healthcare) were used to detect the bands.

Production of recombinant proteins

ZF-ZBTB2 recombinant protein – To prepare recombinant HIS-ZF-ZBTB2, the plasmid PUC-HISZFZBTB2 containing the cDNA fragment encoding the zinc fingers (amino acids 254-468) domain of ZBTB2 in fusion with an N-terminal MAHHHHHHGHHHQLENLYFQG tag containing a tobacco etch virus (TEV) protease cleavage site was synthesised by Eurogentec. The gene was then cloned into the plasmid pet-30b between the restriction sites Nde1 and EcoRI (pet30b-His-ZFZBTB2). The pet30b-His-ZF-ZBTB2 plasmid was transformed into *E. coli* BL21(DE3). 500 mL LB broth medium containing 1 mM kanamycine, 0.1% glucose and 1 mM $ZnSO_4$ was inoculated with 10 mL of an overnight culture at 37 °c. At an OD_{600} of approximately 0.6, expression of the recombinant ZF-ZBTB2 was induced by the addition of 1 mM IPTG and cells were grown for an additional 16 h at 16°c. Cells were pelleted by centrifugation (4000g, 4°c, 30 min), resuspended in buffer A (50 mM Napi pH 8.0, 300 mM NaCl, and complete protease inhibitors-EDTA (Roche)) and disrupted using a sonicator. The soluble fraction was recovered by centrifugation (21 000g, 4°c, 30 min) and mixed with 1 mL of a suspension of Ni-NTA resin (Qiagen) pre-equilibrated with buffer A at 4°c. The resin was then loaded onto a column. To remove unbound proteins, the resin was extensively washed with 10 mL buffer A and 10 mL buffer A containing 10 mM imidazol. The protein was eluted with 10 mL buffer A containing 300 mM imidazole. The purified protein were then concentrated to 2 mL using Centricon devices (Amicon, 3 000 Da cutoff) and loaded onto a size-exclusion chromatography column (Hiload Superdex 75, 16/60, GE Healthcare) for the final step of the

purification procedure. The column was equilibrated with buffer A and the pooled peak fractions were concentrated to 60 µM (0,94 mg/ml). The purity of the protein solutions was confirmed by SDS-PAGE. The final concentration was determined by UV spectroscopy assuming molar absorption coefficients at 280 nm of 15 190 M^{-1} cm^{-1}.

MeCP2 recombinant protein - The plasmid pet30bhMECP2 (76-167) used to purify the MBD domain of MeCP2 was a gift from Bird A.P. [32]. It encodes residues 76 to 167 of MeCP2 with the heterologous N-terminal sequence MHHHHHHAM. The pet30b-hMeCP2 plasmid was transformed into *E. coli* BL21(DE3). 1 L LB broth medium containing 1 mM kanamycine, was inoculated with 10 mL of an overnight culture at 37 °C. At an OD_{600} of approximately 0.6, expression of the recombinant MeCP2 was induced by the addition of 1 mM IPTG and cells were grown for an additional 4 h at 37°C. Cells were pelleted by centrifugation (4000g, 4°c, 30 min), resuspended in buffer A (50 mM Napi pH 8.0, 300 mM NaCl, 10% glycerol and 10 mM β-mercaptoethanol) and disrupted using a sonicator. The soluble fraction was recovered by ultracentrifugation (100 000g, 4°c, 30 min) and loaded on a 1 mL HisTrap HP column (GE Healthcare) pre-equilibrated with 15 mL of buffer A 4°C. To remove unbound proteins, the column was extensively washed with 25 mL buffer A. The proteins were eluted using a linear gradient of buffer A containing 0–250 mM imidazole. The pooled peak fractions was then concentrated to 2 mL using Centricon devices (Amicon 3 000 Da cutoff) and loaded onto a size-exclusion chromatography column (Hiload Superdex 75, 16/60, GE Healthcar) for the final step of the purification procedure. The column was equilibrated with buffer A and the pooled peak fractions were concentrated to 200 µM (2.05 mg/ml). The purity of the protein solutions was confirmed by SDS-PAGE (Fig. X). The final concentration was determined by UV spectroscopy assuming molar absorption coefficients at 280 nm of 11 460 M^{-1} cm^{-1}.

EMSA

ODNs used in EMSA are listed in Table SX. Two probes were used: the "p21 probe" (as described previously [22]) (5'-GAT**C**GGG**C**GGGGG**C**GGTTGTATATCA-3') which corresponds to Sp1 response element on the p21 gene proximal promoter and the "DPD probe" (5'-TATATTAACGTTACGTATACGTTTA-3') which corresponds to a part of the probe used in the DNA pull down assay. Each probe contains either Cytosine (p21-C and DPD-C), 5-mC (p21-mC and DPD-mC) or 5-hmC (p21-hmC and DPD-hmC). p21 probes were synthesised by Eurogentec and DPD probes were synthesised in our laboratory as described previously.

Probes were radiolabeled using T4 polynucleotide kinase (Fermentas) and [γ-32P]ATP (GE-Healthcare). Binding reaction between recombinant protein (ZF-ZBTB2 or MeCP2) and radiolabeled probes were carried out for 30 min at 25°C in a final volume of 20µL. The binding buffer was 20 mM HEPES pH 7.9, 1 mM EDTA, 10 mM $(NH_4)_2SO_4$, 30 mM KCl, 0.5 mM DTT, 0.01% Tween 20 and 10 µg/ml polydIdC (only for the complex with ZF-ZBTB2). Complexes were resolved by non-denaturing electrophoresis on a 4% (ZFZBT2) or 12% (MeCP2) polyacrylamide-Tris-acetate-EDTA gel.

ACKNOWLEDGMENTS

This work was supported by the Agence Nationale de la Recherche (EpigenomeANR project, coordinator: Murat Saparbaev, Institut Gustave Roussy). We thank the Membrane Protein Purification Platform, from the Institut de Biologie Structurale in Grenoble (PSB/IBS), for the purification of the protein MeCP2

REFERENCES

(1) Nabel, C. S.; Manning, S. A.; Kohli, R. M. *ACS Chem Biol*, *7*, 20-30.

(2) Wyatt, G. R.; Cohen, S. S. *Biochem J* **1953**, *55*, 774-82.

(3) Penn, N. W.; Suwalski, R.; O'Riley, C.; Bojanowski, K.; Yura, R. *Biochem J* **1972**, *126*, 781-90.

(4) Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A. *Science (New York, N Y)* **2009**, *324*, 930-5.

(5) Kriaucionis, S.; Heintz, N. *Science* **2009**, *324*, 929-30.

(6) Dahl, C.; Gronbaek, K.; Guldberg, P. *Clin Chim Acta*, *412*, 831-6.

(7) Munzel, M.; Globisch, D.; Carell, T. *Angew Chem Int Ed Engl* **2011**, *50*, 6460-8.

(8) Buck-Koehntop, B. A.; Defossez, P. A. *Epigenetics*, *8*.

(9) Watt, F.; Molloy, P. L. *Genes Dev* **1988**, *2*, 1136-43.

(10) Meehan, R. R.; Lewis, J. D.; Bird, A. P. *Nucleic Acids Res* **1992**, *20*, 5085-92.

(11) Jones, P. L.; Veenstra, G. J.; Wade, P. A.; Vermaak, D.; Kass, S. U.; Landsberger, N.; Strouboulis, J.; Wolffe, A. P. *Nat Genet* **1998**, *19*, 187-91.

(12) Valinluck, V.; Tsai, H.-H.; Rogstad, D. K.; Burdzy, A.; Bird, A.; Sowers, L. C. *Nucleic acids research* **2004**, *32*, 4100-8.

(13) Liu, Y.; Toh, H.; Sasaki, H.; Zhang, X.; Cheng, X. *Genes Dev*, *26*, 2374-9.

(14) Frauer, C.; Hoffmann, T.; Bultmann, S.; Casa, V.; Cardoso, M. C.; Antes, I.; Leonhardt, H. *PLoS One*, *6*, e21306.

(15) Winnacker, M.; Welzmiller, V.; Strasser, R.; Carell, T. *Chembiochem*, *11*, 1345-9.

(16) Bounaix Morand du Puch, C.; Barbier, E.; Kraut, A.; Coute, Y.; Fuchs, J.; Buhot, A.; Livache, T.; Seve, M.; Favier, A.; Douki, T.; Gasparutto, D.; Sauvaigo, S.; Breton, J. *Arch Biochem Biophys* **2011**, *507*, 296-303.

(17) du Puch, C. B.; Barbier, E.; Sauvaigo, S.; Gasparutto, D.; Breton, J. *Mutat Res*.

(18) Mittler, G.; Butter, F.; Mann, M. *Genome Res* **2009**, *19*, 284-93.

(19) Bartels, S. J.; Spruijt, C. G.; Brinkman, A. B.; Jansen, P. W.; Vermeulen, M.; Stunnenberg, H. G. *PLoS One* **2011**, *6*, e25884.

(20) Jin, S.-G.; Kadam, S.; Pfeifer, G. P. *Nucleic acids research* **2010**, *38*, e125.

(21) Ng, H. H.; Zhang, Y.; Hendrich, B.; Johnson, C. A.; Turner, B. M.; Erdjument-Bromage, H.; Tempst, P.; Reinberg, D.; Bird, A. *Nature genetics* **1999**, *23*, 58-61.

(22) Jeon, B. N.; Choi, W. I.; Yu, M. Y.; Yoon, A. R.; Kim, M. H.; Yun, C. O.; Hur, M. W. *J Biol Chem* **2009**, *284*, 17935-46.

(23) Zhang, C. X.; Chang, P. V.; Lippard, S. J. *J Am Chem Soc* **2004**, *126*, 6536-7.

(24) Prokhortchouk, A.; Hendrich, B.; Jorgensen, H.; Ruzov, A.; Wilm, M.; Georgiev, G.; Bird, A.; Prokhortchouk, E. *Genes Dev* **2001**, *15*, 1613-8.

(25) Filion, G. J.; Zhenilo, S.; Salozhin, S.; Yamada, D.; Prokhortchouk, E.; Defossez, P. A. *Mol Cell Biol* **2006**, *26*, 169-81.

(26) Wang, Y.; Zheng, X.; Zhang, Z.; Zhou, J.; Zhao, G.; Yang, J.; Xia, L.; Wang, R.; Cai, X.; Hu, H.; Zhu, C.; Nie, Y.; Wu, K.; Zhang, D.; Fan, D. *PLoS One*, *7*, e41693.

(27) Kim, D. H.; Lee, S. T.; Won, H. H.; Kim, S.; Kim, M. J.; Kim, H. J.; Kim, S. H.; Kim, J. W.; Kim, Y. K.; Sohn, S. K.; Moon, J. H.; Jung, C. W.; Lipton, J. H. *Blood*, *117*, 6906-11.

(28) Olnes, M. I.; Kurl, R. N. *Biotechniques* **1994**, *17*, 828-9.

(29) Jaquinod, M.; Villiers, F.; Kieffer-Jaquinod, S.; Hugouvieux, V.; Bruley, C.; Garin, J.; Bourguignon, J. *Mol Cell Proteomics* **2007**, *6*, 394-412.

(30) Bernay, B.; Gaillard, M. C.; Guryca, V.; Emadali, A.; Kuhn, L.; Bertrand, A.; Detraz, I.; Carcenac, C.; Savasta, M.; Brouillet, E.; Garin, J.; Elalouf, J. M. *Mol Cell Proteomics* **2009**, *8*, 946-58.

(31) Dupierris, V.; Masselon, C.; Court, M.; Kieffer-Jaquinod, S.; Bruley, C. *Bioinformatics* **2009**, *25*, 1980-1.

(32) Klose, R. J.; Bird, A. P. *J Biol Chem* **2004**, *279*, 46490-6.

FIGURE CAPTIONS

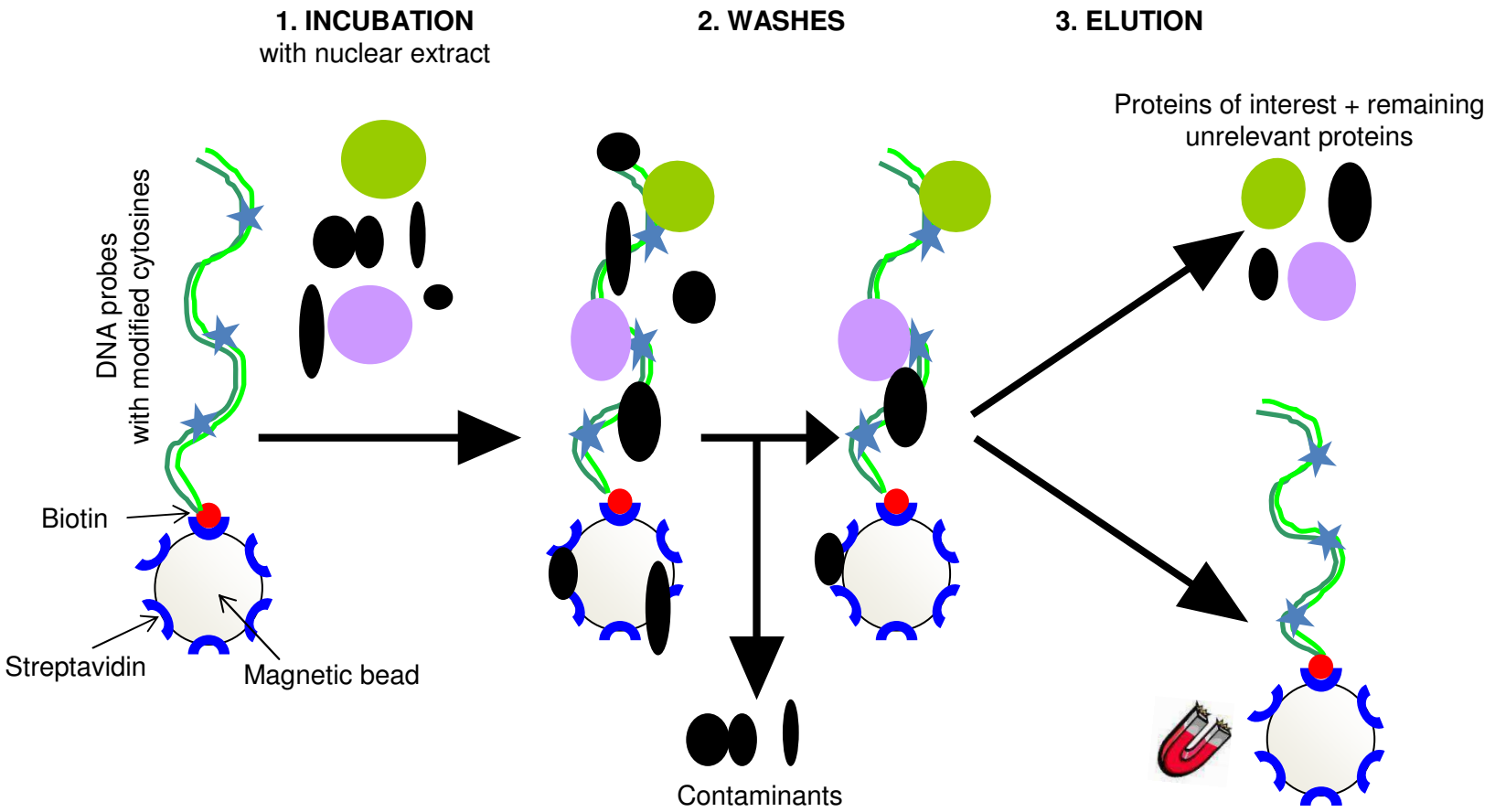
Figure 1. Measurement of modified cytosines by HPLC-MS-MS

Figure 2. DNA pull down principle. DNA probes obtained by PCR are containing C, 5-mC or 5-hmC and are immobilized on magnetic beads through biotin / streptavidin interactions. After incubation with nuclear extracts and washing steps, proteins are eluted directly in Laemmli buffer before analysis by proteomics or western-blotting. The systematic use of control conditions using DNA without epigenetic modifications and beads without probes limits the identification of unrelevant proteins.

Figure 3. Analysis of proteins obtained through DNA pull down. (a) Selection of proteins identified by nanoHPLC-MS-MS. ZBTB2 has systematically been found on C and 5-mC baits but absent from 5-hmC containing DNA. (b) Western-blotting confirmation of proteomics results regarding ZBTB2: inhibition of DNA binding related to the presence of 5-hmC in DNA probes.

Figure 4. Interaction of ZBTB2ZF and MeCP2 with ODNs containing C, 5-mC or 5-hmC. Cytosines in red are replaced by 5-mC or 5-hmC in EMSA probes containing 3 methylated or hydoxymethylated positions. Underscored cytosines are replaced by 5-mC or 5-hmC in probes containing only one methylated or hydoxymethylated position. (a) Results obtained with an ODN containing a sequence taken from the GC-box 5/6 found in the *p21* promoter region. (b) Results obtained with an ODN containing one portion of DNA probes used in the DNA pull down assays.

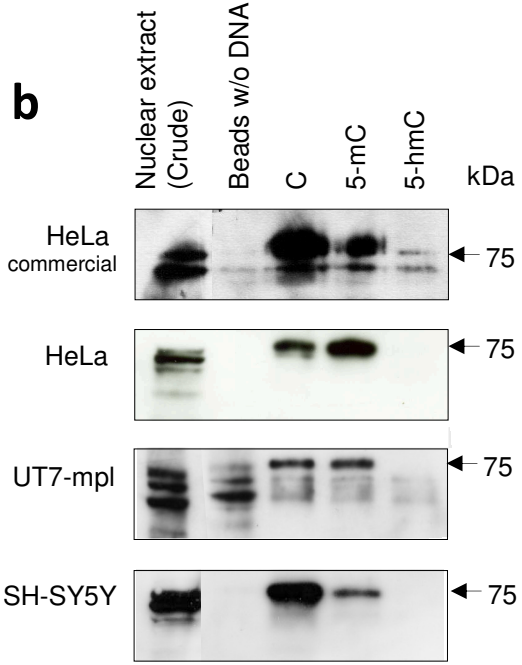
Figure 2

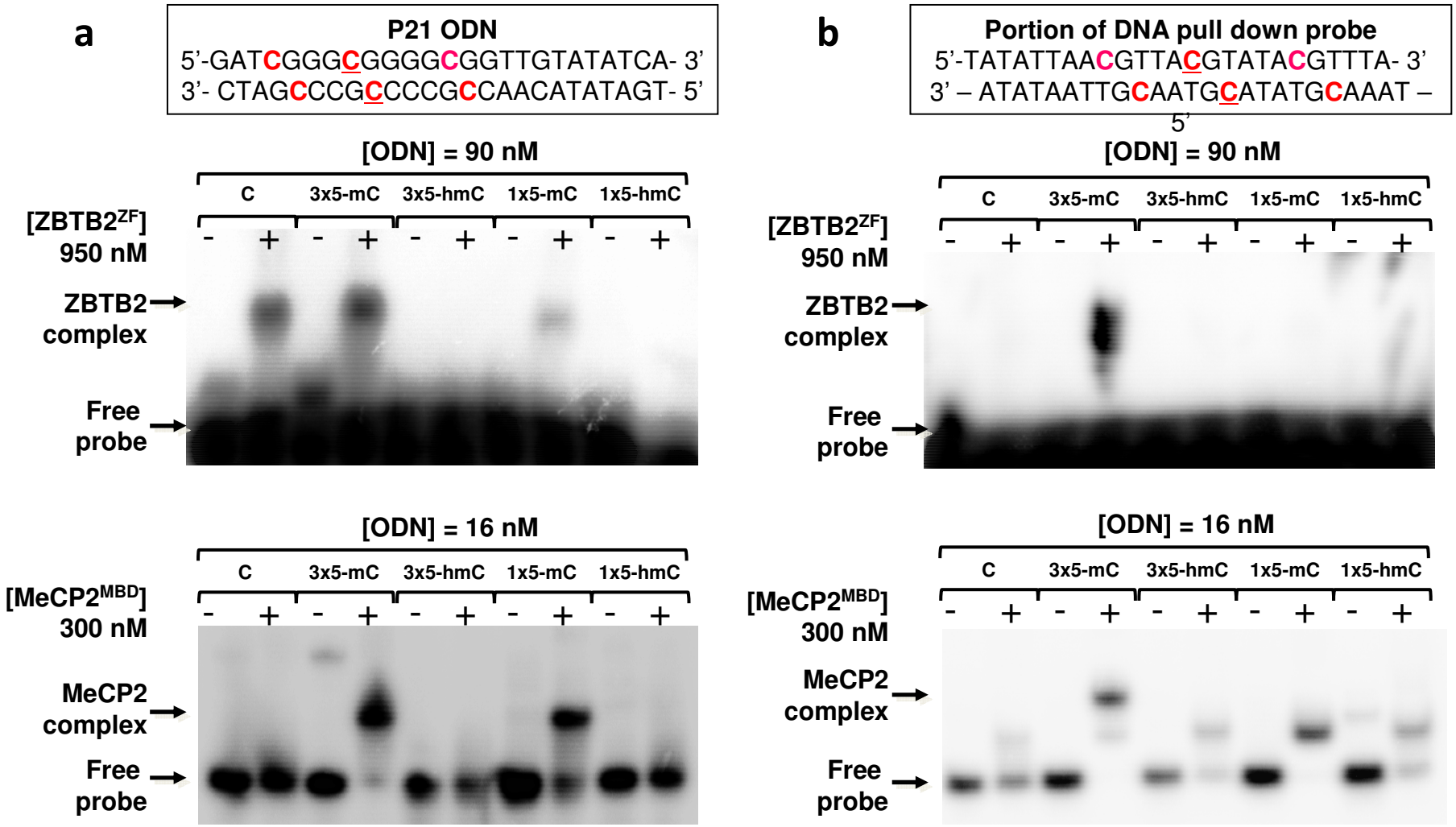

1. INCUBATION
with nuclear extract
2. WASHES
3. ELUTION
DNA probes
with modified cytosines
Biotin
Streptavidin
Magnetic bead
Contaminants
Proteins of interest + remaining
unrelevant proteins

a

	DNA pull down probes							
	Beads w/o DNA		**C**		**5-mC**		**5-hmC**	
Protein name	Mascot score	Peptides number	Mascot score	Peptides number	Mascot score	Peptides number	Mascot score	Peptides number
Commercial HeLa nuclear extract; experimental conditions DPD1								
ZBTB2	-	-	249-186	5-4	102-215	3-4	-	-
Commercial HeLa nuclear extract; experimental conditions DPD2								
ZBTB2	-	-	390-732	7-11	309-513	6-7	-	-
MBD2	-	-	-	-	183-188	3-3	-	-
UHRF1	-	-	-		236	4	672	11
HeLa nuclear extract; experimental conditions DPD2								
ZBTB2	Not analysed		Not analysed		573	9	-	-
SH-SY5Y nuclear extract; experimental conditions DPD2								
ZBTB2	-	-	527	8	88	2	-	-
UT7-mpl nuclear extract; experimental conditions DPD2								
ZBTB2	-	-	750	11	932	12	-	-
MeCP2	-	-	-	-	143	3		

b




a
P21 ODN
5'-GATCGGGCGGGGCGGTTGTATATCA- 3'
3'- CTAGCCCGCCCCGCCAACATATAGT- 5'
[ODN] = 90 nM
C
3x5-mC
3x5-hmC
1x5-mC
1x5-hmC
[ZBTB2ZF] 950 nM
ZBTB2 complex
Free probe
[ODN] = 16 nM
[MeCP2MBD] 300 nM
MeCP2 complex
Free probe
b
Portion of DNA pull down probe
5'-TATATTAACGTTACGTATACGTTTA- 3'
3' – ATATAATTGCAATGCATATGCAAAT – 5'
[ODN] = 90 nM
[ZBTB2ZF] 950 nM
ZBTB2 complex
Free probe
[ODN] = 16 nM
[MeCP2MBD] 300 nM
MeCP2 complex
Free probe