

Redesign of calcium-regulated protein aequorin towards the
development of a novel ion bioreporter

A Thesis submitted to
University College London for the degree of
Doctor of Philosophy

by
Evlampia-Kyriaki Dimitriadou

Department of Biochemical Engineering
University College London

UCL
April 2014

I, Evlampia-Kyriaki Dimitriadou confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

I dedicate this work to my loving parents,
Sophia and Panagiotis

Abstract

This thesis aimed to design novel sensor proteins that can identify and measure various metal ions *in vivo* and *in situ*. Metal ions play key role in the metabolism of the cell, and monitoring of calcium has helped interrogate cellular processes such as fertilisation, contraction and apoptosis. Real-time monitoring of more divalent metal ions like zinc and copper is required to gain much needed insight into brain function and associated disorders, such as Alzheimer's and Parkinson's disease.

Aequorin is a calcium-regulated photoprotein originally isolated from the jellyfish *Aequorea victoria*. Due to its high sensitivity to calcium and its non-invasive nature, aequorin has been used as a real-time indicator of calcium ions in biological systems for more than forty years. The protein complex consists of the polypeptide chain apoaequorin and a tightly bound chromophore (coelenterazine). Trace amounts of calcium ions trigger conformational changes in the protein, which in turn facilitate the intermolecular oxidation of coelenterazine and concomitant production of CO₂ and a flash of blue light.

Aequorin's light emitting reaction can also be triggered by a range of other divalent and trivalent cations, leading however to significantly lower light yields. Based on aequorin's promiscuity towards other ions, this project tested the hypothesis that aequorin's preference for certain cations could be manipulated through mutations engineered in one or more of the three calcium-binding loops (EF-hands).

In order to test the hypothesis, the following six stages were performed: cloning of the apoaequorin gene for expression in *E. coli*; development of a high-throughput assay for expression and measurement of bioluminescent activity in microwells; design of a library containing forty eight mutant variants of aequorin; screening of the library against seven metal ions; protein purification of wild-type aequorin and one selected mutant; analysis of activity and kinetics of purified wild type and one chosen mutant against all seven ions.

This work produced mutants with shifted selectivity towards new metal ions at the cost of luminescence yield. The impact of mutations is analysed and it is suggested that one of the EF-hands (EF-I) is likely to serve as a gatekeeper to aequorin's selectivity. It was also shown that at least one mutant utilised zinc ions (that wild type failed to utilise) to achieve low levels of bioluminescent activity.

Acknowledgements

I would like to thank my supervisor, Professor Paul Dalby for his valuable input throughout the duration of my project. This work would not have been possible without his continued support. I would also like to thank my advisor, Professor Gary Lye for his advice and encouragement.

Special thanks go to Dr Rigini Papi, Lazaros Batsilas and Dr Jean Aucamp for providing their insights and feedback of my work. I am also grateful to Dr Martina Micheletti for training on the robotic platform I used for all the screening work described in this thesis. I would also like to thank the PhD students of Foster Court and UB18, who made the labs feel like a second home.

Finally, I am most grateful to and for my parents, Sophia and Panagiotis and my dear friends Issy, Maria, Stephanos, Rigini, Lazaros, Spyros, Erato, Jean, Zehra, Rosie and Maria W. for genuine love and support. Thank you for being patient with all my absences and for waiting on the other side.

Contents

Abstract	4
Acknowledgements	5
Contents	6
List of tables	15
List of figures	17
Abbreviations	21
1 Introduction	24
1.1 Bioluminescence	24
1.1.1 Definitions.....	24
1.1.2 Bioluminescence in nature	26
1.1.3 <i>Aequorea victoria</i> , the bioluminescent jellyfish	27
1.1.4 Flash- and glow-type luminescence	28
1.2 Bioreporters.....	29
1.2.1 Definitions.....	29
1.2.2 Need for ion bioreporters	30
1.2.3 Desirable characteristics for bioreporters.....	31
1.2.4 Signal transduction methods	32
1.2.5 Current methods for intracellular ion monitoring	34
1.3 Aequorin: a calcium-sensitive photoprotein	36
1.3.1 Introduction	36
1.3.2 Uses of aequorin.....	40
1.3.3 Properties of aequorin	41
1.3.4 EF-hand motif: general features.....	44
1.3.4.1 EF-hands and cooperativity	46
1.3.5 Proposed mechanism for the bioluminescence reaction	47
1.3.6 Regeneration of aequorin	47
1.3.7 Ion selectivity of aequorin.....	49
1.3.8 Other photoproteins.....	50
1.4 Coordination of metals in proteins.....	51
1.5 Protein engineering	56
1.5.1 Rational design.....	56

1.5.2	Directed evolution.....	57
1.5.3	Integration of random and rational approaches.....	58
1.5.4	Metalloprotein engineering.....	59
1.6	Overview.....	60
1.6.1	Aequorin mutants as potential bioreporters.....	60
1.6.2	Potential activator ions for aequorin mutants.....	61
1.7	Aims and organisation of thesis.....	62
2	Materials and Methods.....	64
2.1	Preparation of buffers and media.....	64
2.1.1	LB medium.....	64
2.1.2	LB agar.....	64
2.1.3	YT agar.....	65
2.1.4	SOC medium.....	65
2.1.5	Kanamycin.....	65
2.1.6	Ampicillin.....	65
2.1.7	X-gal.....	66
2.1.8	Tris-HCl buffer.....	66
2.1.9	EDTA.....	66
2.2	Cell culture.....	67
2.2.1	<i>E. coli</i> strains.....	67
2.2.1.1	TOP10 electrocompetent.....	67
2.2.1.2	BL21Star™(DE3).....	67
2.2.2	Streaked plates.....	67
2.2.3	Overnight cultures.....	68
2.2.4	Shake flask cultures.....	68
2.2.5	Glycerol stocks.....	68
2.2.6	Sonication.....	68
2.3	DNA.....	69
2.3.1	Preparation of plasmid DNA.....	69
2.3.2	PCR amplification.....	69
2.3.3	QuikChange® site-directed mutagenesis.....	70
2.3.4	Sequencing.....	73
2.3.5	Oligonucleotides for amplification and sequencing.....	73
2.3.6	Restriction digests.....	73

2.3.7	Ligation	74
2.3.7.1	Standard ligation	74
2.3.7.2	Ligation of TOPO [®] open vector	75
2.3.8	Electrocompetent cell transformation	76
2.3.9	DNA agarose gel electrophoresis	76
2.3.10	Absorbance measurement of DNA	77
2.4	Protein	78
2.4.1	Determination of protein	78
2.4.1.1	Determination by UV absorbance	78
2.4.1.2	Determination by Bradford assay	78
2.4.2	Protein size – SDS-PAGE	79
2.4.2.1	Protein quantity by SDS-PAGE	80
2.4.2.2	Protein sample purity by SDS-PAGE	80
2.5	Metal ion concentration	81
2.5.1	Ion solution preparation	81
2.5.2	Atomic absorption	82
2.6	Automation	82
2.6.1	Automated colony picking	82
2.6.2	Automated liquid handling	82
2.7	Luminescence measurements	83
2.7.1	Platereader	83
2.7.1.1	Visualisation of results	84
3	Cloning of apoaquorin gene	85
3.1	Introduction	85
3.1.1	Envisaged expression vector	85
3.1.2	Cloning strategies	86
3.2	Apoaquorin in original capture vector	86
3.2.1	Verification of apoaquorin gene in pHAQ vector	89
3.3	pET cloning route	90
3.3.1	Vector pET26b	91
3.3.2	Schematic of pET cloning route	92
3.3.3	Methods for pET cloning	93
3.3.4	Results and discussion for pET cloning	95
3.4	TOPO [®] cloning route	100

3.4.1	Schematic of TOPO [®] cloning route	102
3.4.2	Methods for TOPO [®] cloning.....	103
3.4.2.1	Design of primers.....	103
3.4.2.2	PCR amplifications	103
3.4.2.3	TOPO [®] ligation.....	106
3.4.3	Results and discussion for TOPO [®] cloning	107
3.5	Conclusions	113
4	High-throughput expression and screening.....	115
4.1	Introduction	115
4.1.1	Steps to developing a high-throughput process	115
4.1.2	Requirements for a high-throughput assay	116
4.2	Materials.....	118
4.2.1	Chemicals.....	118
4.2.2	Consumables	118
4.3	Method development strategy	119
4.3.1	Building the assay protocol.....	119
4.3.2	Experimental setup A.....	120
4.3.3	Experimental setup B	122
4.3.4	Experimental setup C	123
4.4	Results and discussion	124
4.4.1	Measurement of cell density in microplates.....	124
4.4.2	Process parameters	126
4.4.2.1	Choice of vector.....	126
4.4.2.2	Part of culture as screening material.....	127
4.4.2.3	IPTG induction	127
4.4.2.4	Choice of bacterial strain	128
4.4.2.5	Check for background signal	129
4.4.2.6	Activity profile across the microplate.....	129
4.4.2.7	Cell growth and activity versus time	132
4.4.2.8	Optimum coelenterazine concentration	132
4.4.2.9	Injection, mixing.....	133
4.4.3	Check for cross-contamination between wells.....	134
4.4.4	Yield of aequorin in microwells.....	135
4.5	Conclusions	137

4.5.1	Final process	137
4.5.1.1	Cloning and microwell cultures.....	138
4.5.1.2	Charging of active protein	139
4.5.1.3	96-well microplate library	139
4.5.1.4	High-throughput screening assay.....	139
4.5.2	Further improvements to the method.....	140
5	Aequorin structure and mutant library design	141
5.1	Aequorin: detailed structure.....	142
5.1.1	Coelenterazine-binding cavity and calcium-binding EF-hands .	142
5.1.2	H-bond coupling of EF-hands.....	146
5.1.3	Theories on signal transduction following calcium binding	149
5.2	Previous mutagenesis on aequorin.....	151
5.2.1	Calcium sensitivity.....	151
5.2.2	Impaired EF-hands / central Gly→Arg mutations	152
5.2.3	Spectral shift.....	154
5.2.4	Intensity.....	154
5.2.5	Thermostability	155
5.3	Mutation of other calcium-binding structures.....	158
5.3.1	Short peptides with calcium-binding ability	158
5.3.2	EF-hand loops in model proteins	159
5.4	Structural analysis methods	161
5.4.1	Crystallographic structure examination	161
5.4.2	Interatomic overlap calculations	161
5.5	Results and discussion	163
5.5.1	Calcium-free versus calcium-bound aequorin structure	163
5.5.2	Movement of residues to coordinate calcium	165
5.5.3	Overlap of atomic radii in the ion-binding loop	169
5.6	Conclusions.....	173
5.6.1	Conclusions from structural analysis	173
5.6.2	Conclusions from previous mutational studies	174
5.6.3	Choice of mutant library	174
6	Library screening and mutant selection	180
6.1	Introduction.....	180
6.1.1	Creation of mutant library.....	180

6.1.2	Previous work in methods of aequorin mutagenesis.....	180
6.1.3	Screening of mutant library.....	182
6.2	Methods.....	184
6.2.1	Mutant library construction.....	184
6.2.2	Mutant library layout and screening	186
6.2.2.1	Metal ion concentrations for library injection	189
6.2.2.2	Gain settings	189
6.2.3	Data processing of high-throughput screening	191
6.3	Results.....	194
6.3.1	Library creation results	194
6.3.2	Library screening results.....	196
6.4	Discussion	210
6.4.1	Overall effect on yield.....	210
6.4.2	Selectivity (D) versus retained activity (C).....	211
6.4.3	EF-I versus EF-III	211
6.4.4	Ala-scan mutants and importance of side chains	212
6.4.5	Importance of loop positions.....	212
6.4.6	Type of substitution	213
6.4.7	Impairment of EF-hands: central Gly mutants.....	215
6.4.8	Double mutations	217
6.4.9	Software for prediction of mutation effects	217
6.5	Conclusions.....	218
6.5.1	Choice of mutant for further study.....	220
7	Protein purification.....	221
7.1	Introduction.....	221
7.2	Materials.....	222
7.3	Methods.....	223
7.3.1	Culture.....	223
7.3.1.1	Host strain and plasmid.....	223
7.3.1.2	Cell growth	223
7.3.2	Regeneration and extraction of aequorin	223
7.3.2.1	One-step extraction and charging of active aequorin	223
7.3.2.2	Periplasmic fraction	224
7.3.3	Concentration of aequorin.....	225

7.3.3.1	Spin-filtration.....	225
7.3.4	Luminescence assay	226
7.3.5	Size exclusion chromatography	226
7.3.5.1	Protein size by mass spectrometry.....	227
7.3.6	Protein handling considerations	228
7.4	Results and Discussion.....	229
7.4.1	Bacterial cell culture	229
7.4.1.1	OD and activity versus incubation time.....	229
7.4.2	Localisation of protein	231
7.4.2.1	Localisation of protein in the cell	231
7.4.2.2	Localisation of protein after coelenterazine incubation.....	232
7.4.2.3	Periplasmic fraction	234
7.4.3	Clarification of protein sample.....	236
7.4.4	Enrichment of aequorin prior to chromatography.....	237
7.4.4.1	Spin-filtration.....	237
7.4.4.2	Mass spectrometry of spin-concentrated sample	238
7.4.5	Storage of purified fractions.....	238
7.4.6	Size exclusion chromatography of aequorin: wild-type, mutant and commercial aequorin.....	238
7.4.7	SDS-PAGE of purified samples.....	241
7.4.8	Mass spectrometry of purified samples	242
7.5	Conclusions	244
7.5.1	Improvements to the process.....	246
8	Response of wild-type aequorin and mutant Asn28Cys/Ser32His to seven metal ions	247
8.1	Introduction	248
8.1.1	Binding studies.....	248
8.1.2	Affinity and apparent affinity.....	249
8.1.3	Allostery and cooperativity	252
8.1.4	Concentration-response curves	254
8.1.5	The aequorin system	256
8.1.5.1	Fundamental work on the kinetics of aequorin.....	257
8.1.5.2	A model of double exponential decay	258
8.2	Materials and methods	260

8.2.1	Aequorin and coelenterazine.....	260
8.2.2	Protein concentration	260
8.2.3	Concentration range of metal ions	261
8.2.4	Experimental setup and platereader settings	261
8.2.5	Method for double ion experiment.....	263
8.2.6	Data processing	263
8.2.6.1	Luminescence curves and concentration-response curves.....	263
8.2.6.2	Exponential model fitting	264
8.2.6.3	Estimation of glow luminescence kinetics.....	264
8.2.7	Artefacts from syringe injection.....	265
8.2.8	Assay objectives and limitations.....	266
8.3	Results and discussion	267
8.3.1	Luminescence curves: flash and glow-type	267
8.3.2	Concentration-response curves	271
8.3.3	Light yield comparison across all metals with wild-type and mutant aequorin	277
8.3.4	Analysis of bioluminescence kinetics	280
8.3.4.1	Flash-type luminescence: Model of double exponential decay	280
8.3.4.2	Glow-type luminescence kinetics	287
8.3.5	Double ion experiments: wild type plus Zn ²⁺ and Ca ²⁺	289
8.3.6	Correlation of ion size and shape of luminescence curve	294
8.4	Conclusion	297
9	Epilogue.....	301
9.1	Overall summary of this project.....	301
9.2	Overall conclusions.....	304
9.2.1	Cloning of apoaequorin gene	304
9.2.2	High-throughput screening process.....	304
9.2.3	Mutant library screening	305
9.2.4	Purification of aequorin.....	306
9.2.5	Study of purified wild-type aequorin and mutant Asn28Cys/Ser32His	306
9.3	Future work	308
9.3.1	Breaking down the problem.....	308

9.3.2	Mutational strategy.....	309
9.3.3	Repairing aequorin activity.....	310
9.3.4	Alternative scaffold.....	310
9.3.5	Alternative application.....	311
9.4	Analytics for future work.....	311
9.5	Value and risks of the project.....	313
Appendix.....		315
	Appendix Chapter 1.....	315
	Appendix Chapter 5.....	323
	Appendix Chapter 6.....	330
	Appendix Chapter 7.....	332
	Appendix Chapter 8.....	333
References.....		334

List of tables

Table 1-1 <i>Aequorea victoria</i> taxonomy	27
Table 1-2 Comparative sensitivity of detection methods.....	32
Table 1-3 Comparison of bioluminescent and fluorescent reporters	33
Table 1-4 Amino acid sequence preference in the EF-hand loop	45
Table 1-5 Aequorin tested against ions for triggering light-emitting reaction.....	49
Table 1-6 Identity between aequorin and other Ca ²⁺ -binding proteins.....	51
Table 1-7 Summary of preferred coordination geometries of metal ions derived from PDB and CSD	52
Table 1-8 Distances around metal ions from the PDB and CSD	53
Table 1-9 Metal-binding sites and experimental metal binding constants for ion-binding proteins.....	55
Table 2-1 PCR mix – general cloning.....	69
Table 2-2 PCR cycles – general cloning	70
Table 2-3 PCR mix – QuikChange [®]	72
Table 2-4 PCR cycles – QuikChange [®]	72
Table 2-5 Double digestion reaction mixture.....	74
Table 2-6 CloneSmart [®] ligation mixture	75
Table 2-7 TOPO [®] capture vector ligation mixture	75
Table 2-8 Agarose gel concentrations.....	77
Table 2-9 Metals in the form of hydrated salts	81
Table 3-1 Oligonucleotides for pET26b cloning route and sequencing	94
Table 3-2 PCR amplification of apoaequorin from pHAQ.....	94
Table 3-3 Oligonucleotides for TOPO [®] cloning route and sequencing.....	105
Table 3-4 PCR conditions for TOPO [®] cloning.....	105
Table 3-5 PCR mix for TOPO [®] cloning reactions	106
Table 4-1 Variation across the microplate with and without thermal insulation	131
Table 4-2 List of parameters chosen for the high-throughput assay.....	138
Table 5-1 Previous mutational work on aequorin. Selected literature.....	156
Table 5-2 Properties of aequorin with coelenterazine and some of its analogues	157
Table 5-3 Shifts of residue positions between Ca ²⁺ -free and Ca ²⁺ -bound EF-hands...	168
Table 5-4 Mutants Nr 1-23.....	178

Table 5-5 Mutants Nr 24–48.....	179
Table 6-1 PCR conditions for creating mutants Nr 1–48.....	185
Table 6-2 Metal ion stock concentrations and platereader gain settings for library screening	190
Table 7-1 Yield and purity of purified aequorin	245
Table 8-1 Aequorin kinetic constants in published literature	257
Table 8-2 Luminescence curve shapes.....	268
Table 8-3 Summary of characteristics of luminescence versus time and concentration-response curves	276
Table 8-4 Metal ion concentrations used for overall yield comparisons	278
Table 8-5 Fast time constant (τ_F) ranges for flash-type curves.....	286
Table 8-6 Slow time constant (τ_S) ranges for flash-type curves	286
Table 8-7 Ranges of $P_{15/Max}$ for glow-type curves	289
Table 10-1 Intracellular metal ion sensors	315
Table 10-2 Multiple alignment of calcium-binding proteins	317
Table 10-3 Preferred coordination geometries of metal ions in CSD structures	320
Table 10-4 Metal-ligand distances in the CSD and PDB.....	321
Table 10-5 Average metal-ligand distances for metal ions as a function of metal charge, coordination number (CN), donor atom's charge and coordination number of ligand.....	322
Table 10-6 Primers for mutants Nr 1–8	323
Table 10-7 Primers for mutants Nr 9–17	324
Table 10-8 Primers for mutants Nr 18–23	325
Table 10-9 Primers for mutants Nr 24–32	326
Table 10-10 Primers for mutants Nr 33–40	327
Table 10-11 Primers for mutants Nr 41–48	328
Table 10-12 Ionic radii of metals according to their coordination and spin state.....	329
Table 10-13 Library OD(600 nm) for aequorin wild-type and mutants	330
Table 10-14 Gain settings used for luminescence curves	333

List of figures

Figure 1-1 Classification of light-emitting phenomena	25
Figure 1-2 View of <i>Aequorea victoria</i> (a).....	28
Figure 1-3 View of <i>Aequorea victoria</i> (b)	28
Figure 1-4 Typical schematics of flash- and glow-type luminescence	29
Figure 1-5 Schematic of the bioluminescent reaction of photoprotein aequorin	39
Figure 1-6 Schematic of sigmoidal curve for light emission versus Ca^{2+} concentration.....	39
Figure 1-7 3-D model of aequorin with coelenterazine in the hydrophobic core	42
Figure 1-8 Apoaequorin with calcium ions in the Ca^{2+} -binding loops (EF-hands)	42
Figure 1-9 Conformation states of a photoprotein	43
Figure 1-10 The EF-hand calcium-binding loop.....	44
Figure 1-11 Proposed reaction for the bioluminescence of aequorin	48
Figure 1-12 Regeneration of aequorin	48
Figure 2-1 Schematic of QuikChange [®] II Site-Directed mutagenesis method	71
Figure 2-2 Protein ladders used in SDS protein electrophoresis.....	80
Figure 2-3 Platerreader layout – luminescence mode	83
Figure 3-1 Overview of worklow employed for cloning in this work	87
Figure 3-2 pHAQ vector	88
Figure 3-3 pBluescript vector.....	88
Figure 3-4 Plasmid vector pHAQ: agarose gel visualisation	90
Figure 3-5 pET26b vector and cloning expression region	91
Figure 3-6 pET26b cloning route schematic	92
Figure 3-7 Agarose gel visualisation of pET26b cloning route	96
Figure 3-8 pETAQ DNA sequencing results (a).....	97
Figure 3-9 pETAQ DNA sequencing results (b).....	98
Figure 3-10 pETAQ open reading frame	99
Figure 3-11 TOPO [®] capture vector.....	101
Figure 3-12 TOPO [®] cloning route schematic	102
Figure 3-13 Agarose gel visualisation of TOPO [®] cloning route	108
Figure 3-14 TOPAQ DNA sequencing results (a)	109
Figure 3-15 TOPAQ DNA sequencing results (b).....	110
Figure 3-16 TOPAQ DNA sequencing results (c)	111

Figure 3-17	TOPAQ open reading frame.....	112
Figure 4-1	Box diagram of the steps required for a complete HTS protocol.....	116
Figure 4-2	Experimental setup A for defining process parameters.....	120
Figure 4-3	Experimental setup B for defining process parameters.....	122
Figure 4-4	Experimental setup C for defining process parameters.....	123
Figure 4-5	Correlation of cell culture optical densities in microplates and cuvettes ...	125
Figure 4-6	Linearity of OD (600 nm) versus dilution.....	125
Figure 4-7	Choice of expression vector	126
Figure 4-8	Part of culture to be used for the screening process	127
Figure 4-9	Time of induction	1288
Figure 4-10	Choice of <i>E. coli</i> strain.....	128
Figure 4-11	Check for background signal.....	129
Figure 4-12	Activity profile across the microplate with and without thermal insulation	130
Figure 4-13	Luminescence activity and cell density of 96-DSW cultures versus time	132
Figure 4-14	Optimum coelenterazine concentration.....	133
Figure 4-15	Check for cross contamination during incubation and injection.....	134
Figure 4-16	Yield of aequorin in luminescence plate microwells	135
Figure 4-17	Detailed schematic of the finalised high-throughput process	136
Figure 5-1	Peroxidised coelenterazine in the hydrophobic core of aequorin.....	144
Figure 5-2	Coelenterazine binding site in the hydrophobic core of aequorin.....	144
Figure 5-3	Calcium binding in EF-I, EF-III and EF-IV hands of aequorin	145
Figure 5-4	Coupling of EF-I and EF-II of aequorin by means of hydrogen bonds	147
Figure 5-5	Coupling of EF-III and EF-IV of aequorin by means of hydrogen bonds .	147
Figure 5-6	EF-hand pair coupling residues, van der Waals distance range and cavity of aequorin.....	148
Figure 5-7	Aequorin and positions of interesting mutations.....	155
Figure 5-8	Coelenterazine and nine analogues	157
Figure 5-9	Atomic overlap	162
Figure 5-10	Structurally aligned calcium-free aequorin (1EJ3, chain B) and calcium-bound apoaequorin (1SL8).....	164
Figure 5-11	Shifts in the EF-hands upon calcium binding	167
Figure 5-12	Metal-ligand atomic overlap in EF-I of aequorin (a)	170
Figure 5-13	Metal-ligand atomic overlap in EF-I of aequorin (b).....	171

Figure 5-14 Metal-ligand atomic overlap in EF-I of aequorin (c)	172
Figure 6-1 Typical layout of each microplate of the mutant library	186
Figure 6-2 Layout of wild type and mutants in the four plates of the library	188
Figure 6-3 Illustration of measures of activity (C) and selectivity (D) in data processing.....	193
Figure 6-4 DNA agarose gel electrophoresis of mutant Asn28Cys.....	194
Figure 6-5 DNA agarose gel electrophoresis of mutant Asn28Cys/Ser32His.....	195
Figure 6-6 Visualisation of raw screening data, plates 1 and 2 of library	198
Figure 6-7 Visualisation of raw screening data, plates 3 and 4 of library	199
Figure 6-8 % Activity (C) compared to wild type against all ions tested, for the entire library	200
Figure 6-9 D as metric for ion selectivity shift of mutants	201
Figure 6-10 D on logarithmic scale.....	202
Figure 6-11 C (activity compared to wt) of single mutations	203
Figure 6-12 Selectivity (D) of single mutations.....	204
Figure 6-13 Matrix plot of D, arranged by number of EF-hand mutated	205
Figure 6-14 Matrix plot of D, arranged by position mutated in the EF-hand loop.....	206
Figure 6-15 Matrix plot of D, arranged by type of substitution in the EF-hand loop..	207
Figure 6-16 D versus C for all the mutants	208
Figure 6-17 Activity of central Gly mutants (impaired EF-hands).....	209
Figure 7-1 Luminescence activity and cell density of 500 mL cultures versus time...	230
Figure 7-2 Localisation of apoaequorin in the cell pellet – (a).....	231
Figure 7-3 Localisation of apoaequorin in the cell pellet – (b).....	232
Figure 7-4 Localisation of apoaequorin after charging with coelenterazine.....	233
Figure 7-5 Periplasmic extraction SDS gel – (a) of cultures carrying pETAQ and pET26b.....	235
Figure 7-6 Periplasmic extraction SDS gel – (b)	235
Figure 7-7 Clarification of aequorin suspension after coelenterazine incubation.....	236
Figure 7-8 Aequorin suspension concentrated using Minisart [®] filters.....	237
Figure 7-9 Mass spectrometry of aequorin sample prior to chromatographic purification	238
Figure 7-10 Elution profiles matched with activity of the elution fractions from size exclusion chromatography	240
Figure 7-11 SDS-PAGE visualisation of purified aequorin samples.....	241

Figure 7-12 Mass spectrometry analysis of purified wild-type aequorin	243
Figure 7-13 Mass spectrometry analysis of mutant Asn28Cys/Ser32His.....	243
Figure 7-14 Mass spectrometry analysis of commercial aequorin AquaLite [®]	243
Figure 7-15 Final protocol for the purification of aequorin.....	245
Figure 8-1 Bimolecular binding curves plotted on a linear x-axis and a logarithmic x-axis.....	251
Figure 8-2 Bimolecular binding curves illustrating two experimental extremes.....	252
Figure 8-3 Sigmoidal concentration-response curve – stimulation.....	255
Figure 8-4 Concentration-response curve – inhibition	255
Figure 8-5 Bell-shaped concentration-response curve.....	255
Figure 8-6 Light versus time – Wild-type aequorin.....	269
Figure 8-7 Light versus time – Mutant Asn28Cys/Ser32His.....	270
Figure 8-8 Concentration-response curves of wild-type aequorin.....	274
Figure 8-9 Concentration-response curves of mutant Asn28Cys/Ser32His	275
Figure 8-10 Wild type against each metal ion	279
Figure 8-11 Mutant Asn28Cys/Ser32His against each metal ion.....	279
Figure 8-12 Decay kinetics of wild-type aequorin.....	284
Figure 8-13 Decay kinetics of mutant Asn28Cys/Ser32His aequorin variant.....	285
Figure 8-14 Decay kinetics of commercial aequorin AquaLite [®]	286
Figure 8-15 Metric of reaction progress $P_{15s/Max}$ for glow-type curves.....	288
Figure 8-16 Wild type triggered by zinc and calcium	292
Figure 8-17 Mutant Asn28Cys/Ser32His aequorin triggered by zinc and calcium	293
Figure 8-18 Is radii overlap associated with sharpness of luminescence curves?	296
Figure 8-19 Type of luminescence, metal ion sensitivity and interatomic distance in the EF-I loop of wild-type aequorin.....	296
Figure 10-1 Sequencing results for mutant Asn28Cys	331
Figure 10-2 Optimisation of reducing agent in charging of aequorin.....	332

Abbreviations

AEX	Anion exchange (chromatography)
BFP	Blue fluorescent protein
BSA	Bovine serum albumin
CaB	Calcium-binding
CaBP	Calcium-binding Proteins
CaM	Calmodulin
CD	Circular dichroism
CN	Coordination number
CSD	Cambridge Structural Database
Cys, C	Cysteine
DMSO	Dimethyl sulfoxide
DSW	Deep square well (microplate)
dNTP	Deoxynucleoside triphosphates
DTT	Dithiotreitol
EC ₅₀	Half maxialeffective concentration
EDTA	Ethylenediaminetetraacetic acid disodium salt
EGTA	Ethyleneglycoltetraacetic acid
ER	Endoplasmic reticulum
FRET	Fluorescence resonance energy transfer
GFP	Green fluorescent protein
GPCRs	G-protein coupled receptors
GP-HPLC	Gel permeation high-performance liquid chromatography
HIC	Hydrophobic interaction chromatography
K _d	dissociation constant, μM
MCS	Multiple cloning site
MOPS	3-[N-Morpholino] propanesulfonic acid
MW	Molecular weight

MWCO	Molecular weight cut-off
Ni-NTA	Nickel-nitrilotriacetic acid
ompA	outer membrane protein A
ORF	Open reading frame
OD	Optical density
PCR	Polymerase chain reaction
PDB	Protein data bank
pET26b	Expression vector from Novagen
pETAQ	pET26b with apoaequorin gene insert
pHAQ	pBluescript vector carrying apoaequorin gene insert
PMSF	Phenylmethanesulfonyl fluoride
PP	Pyrophosphate
PrP	Prion protein
pTOPAQ	TOPO [®] with T7, rbs, <i>pelB</i> (pET26b), and apoaequorin gene.
RBP	Retinol binding protein (mammalian serum protein)
RBS	Ribosome binding site
RLU	Relative luminescence units
RO	Reverse osmosis
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
STDEV	Standard deviation
TAE	Tris·acetate EDTA
TBE	Tris·borate EDTA
TCA	Trichloroacetic acid
TnC	Troponin C
TOPAQ	TOPO [®] with T7, rbs, <i>pelB</i> (pET26b), and apoaequorin gene.
TOPO [®]	Open capture vector (Invitrogen Inc.) utilising the activity of DNA topoisomerase I

Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
WELAC	Western European Laboratory Accreditation Conference

1 Introduction

Protein aequorin has been used as the preferred intracellular calcium sensor for more than three decades (Shimomura, 1995b). This introduction aims to illustrate why it may be possible to expand the repertoire of aequorin to serve as a reporter for other ions such as zinc, cadmium, copper and lead, and thus help map brain disease states associated with such ions, like various forms of dementia (i.e. Parkinson's and Alzheimer's disease), anxiety and psychosis.

This chapter initially offers background information on what bioluminescence is and how it differs from other forms of light, e.g. fluorescence, and then reviews intracellular biosensing methods based on light emission. The ion-binding properties, structure and bioluminescence reaction of aequorin are described, along with its use as a calcium probe in biomedical research. An overview of geometries and distances of ion binding in metalloproteins is presented, in order to gain some insight of how metals are coordinated in various protein environments. Protein engineering methods and their potential applications for altering aequorin's ion selectivity are discussed along with examples of previous mutagenesis studies of calcium-binding sites. Finally, it is suggested that methods currently used for ion measurement in living cells are not sufficient to fully understand the role of divalent metals in cellular and tissue biochemistry and how mutants of aequorin could fill that gap.

1.1 Bioluminescence

1.1.1 Definitions

Luminescence is the emission of light by a substance that has not been heated and in bioluminescence it is typically the result of an oxidation reaction in which chemical energy is converted to light energy within a living organism. If the compounds that generate light are synthetic chemicals, the phenomenon is called chemiluminescence (Van Dyke *et al.*, 2002). To distinguish bioluminescence from other light emitting phenomena used in chemistry and biology, it would be appropriate to highlight the differences between them (Figure 1-1).

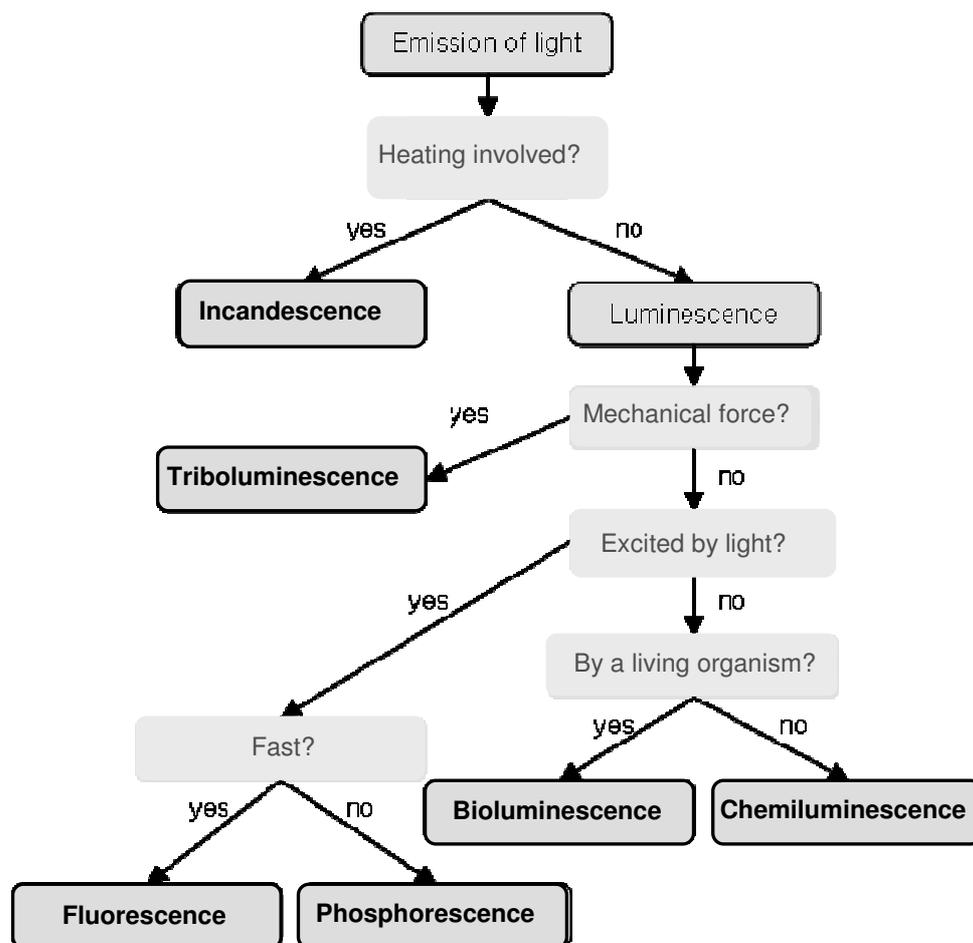


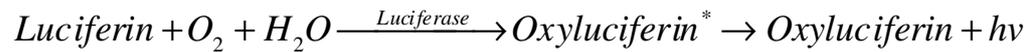
Figure 1-1 Classification of light-emitting phenomena (Wong, 2006).

Fluorescence is the emission of light that occurs where electromagnetic radiation, often ultraviolet light, is absorbed and “excites” an electron from a lower energy atomic orbital into a higher one. The electron then releases energy in the form of light when it falls back to a lower energy state. Light emission lasts only as long as the excitation energy is provided. Unlike fluorescence, phosphorescence continues for a short while after the source of excitation energy is removed and produces light at longer wavelengths (lower energy) (Atkins, 2006).

The two reactants in bioluminescence are often referred to as luciferins and luciferases. Luciferins are the substrate molecules reacted upon to produce light and are a family of heterocyclic compounds with structures that vary between organisms (Tran, 1993). Luciferases are the enzymes that catalyse the oxidation of the substrate luciferins. None of the major luciferases share sequence homology with each other, but

are grouped together as a wide range of enzymes that catalyse the oxidation of substrate luciferins (Greer and Szalay, 2002).

The bioluminescence reaction produces a chemical intermediate that can form at least one excited-state compound. The excited compound relaxes to ground state producing non-reactive oxyluciferin and releasing photons (Greer and Szalay, 2002). A simplified schematic of the bioluminescence reaction is provided below (Tran, 1993):



Only a handful of basic luciferin–luciferase systems account for the luciferase bioluminescence found in many phyla and hundreds of taxa in nature. This project concentrates on coelenterazine bioluminescence. This phenomenon is found in sea animals known as coelenterates. High-energy complexes that contain luciferin as a coelenterazine derived compound, luciferase as a protein chain, require molecular oxygen as a co-factor and calcium ions as a trigger of the enzymatic oxidation are known as photoproteins (Greer and Szalay, 2002).

1.1.2 Bioluminescence in nature

As a predominantly marine phenomenon, it was observed by man from times immemorial (Harvey, 1952). The ability to produce light is a feature of many different groups of living organisms, but not shared by any of the higher vertebrates (Herring, 1978). Perhaps the most striking biological fact regarding the emission of light by animals and plants is the great numbers of totally unrelated and diverse organisms that have developed this ability. Furthermore, one species in a genus may be luminous and another closely allied species may contain no trace of luminosity. Behaviour based on natural bioluminescence is attributed to baiting, startle or camouflage and courtship (Greer and Szalay, 2002).

Bioluminescent marine species include: cnidaria (hydrozoa-hydroids, scyphozoa (medusae) and anthozoa (corals, sea pens, etc)), ctenophora (comb-jellies), mollusca (includes snails, squid, clams etc), arthropoda (includes sea spiders, millipedes, centipedes, insects), echinodermata (starfish). On land it is most commonly seen as glowing fungus on wood (known as foxfire), in the railroad worm and in the few families of luminous insects such as fireflies (Haddock, 2004).

1.1.3 *Aequorea victoria*, the bioluminescent jellyfish

The jellyfish *Aequorea victoria* (Figure 1-2) is shaped like a hemispherical umbrella or bell, (Figure 1-2 and Figure 1-3) with up to 5-10 cm in diameter. The light emitting organs emit blue-green light and are located along the rim of the umbrella. This light captured the curiosity of scientists to extract and investigate the molecules responsible for light emission. In the early 1960s (Shimomura, 2005, Shimomura *et al.*, 1962). The work contributed two of the most widely used bioreporters to biological research: aequorin and green fluorescent protein (GFP), used respectively as an intracellular calcium probe and as an imaging tool for gene expression and intracellular processes monitoring (Prendergast, 2000, Shimomura, 2005, Miyawaki, 2008).

In the jellyfish, the function of aequorin and GFP are coupled: blue light produced by aequorin is absorbed by GFP and emitted as light shifted towards the green (~505nm). Why or how jellyfish use their bioluminescent capabilities is not well understood. They do not glow continuously in the field; in fact, light emission is rarely seen in undisturbed animals (Mills, 2008).

The evolutionary origins of bioluminescence remain obscure. Rees and co-workers (1998) have proposed that the anti-oxidative properties of the bioluminescence substrates (e.g. coelenterazine) were utilised in emergent bioluminescent systems for the detoxification of deleterious oxygen species. Coelenterazine, which is received by jellyfish through their diet (Haddock *et al.*, 2001), is highly reactive towards the superoxide anion or peroxides.

Table 1-1 *Aequorea victoria* taxonomy (Myers, 2009)

Kingdom:	Animalia (animals)
Phylum:	Cnidaria (obsolete term: Coelenterata)
Class:	Hydrozoa (hydralike animals, hydroids and hydrozoans)
Suborder:	Leptomedusae
Family:	Aequoreidae
Genus:	<i>Aequorea</i>
Species:	<i>Aequorea victoria</i>



Figure 1-2 View of *Aequorea victoria* (a). Picture reproduced from Shimomura (2005).

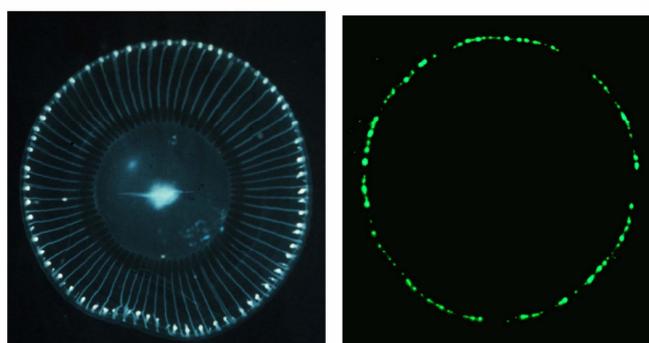


Figure 1-3 View of *Aequorea victoria* (b). Bottom view of the jellyfish taken in daylight and in the dark. Left: daylight view; the overall glow reflects the flash of the camera. Right: taken in the dark. The light organs of the jellyfish are located at the rim of the “umbrella”. The bright formation in the centre of the jellyfish is its highly expandable mouth (Mills, 2008). Picture reproduced from Shimomura (2005) and (2009).

1.1.4 Flash- and glow-type luminescence

The difference between flash- and glow-type luminescence is illustrated in Figure 1-4. Glow-type luminescence involves sustained equilibrium and kinetic reactions. Each type has distinct advantages and disadvantages. In terms of their use in bioanalytical techniques, flash reactions are more useful when studying quick reactions

real-time, however they are more difficult to reproduce consistently and require equipment that ensures rapid and thorough mixing (Van Dyke *et al.*, 2002).

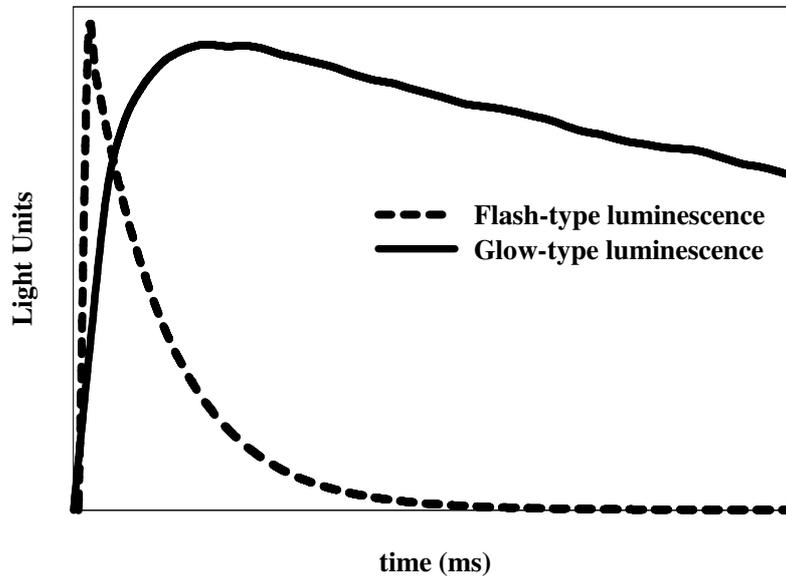


Figure 1-4 Typical schematics of flash- and glow-type luminescence.

1.2 Bioreporters

1.2.1 Definitions

A bioreporter is typically an enzyme or bacterium that produces a measurable signal in response to a specific chemical or physical agent in its environment. Biosensors usually refer to a device that incorporates a bioreporter molecule or organism with a transducer that is able to convert the recognition event into a measurable signal (Merriam-Webster, 2009, Van Dyke *et al.*, 2002). The terms biosensor and bioreporter are often used interchangeably in literature to refer to the molecules or organisms that produce the signal.

1.2.2 Need for ion bioreporters

Metal ions perform a wide variety of specific functions associated with life processes. In many cases, metal ions (e.g. Zn^{2+} , Mg^{2+} , Ca^{2+}) stabilise the structure of folded proteins, while in other cases they help to fix a particular physiologically active conformation of the protein. Metal ions are integral part of many enzymes and are indispensable in several catalytic reactions, e.g. hydrolytic, redox and isomerisation reactions. Alkali and alkaline earth ions, especially Na^+ , K^+ and Ca^{2+} also play a vital role in triggering cellular responses.

The determination of metal concentration for extracellular components is performed by well established methods such as gas chromatography (GC), high pressure liquid chromatography (HPLC), and other analytical techniques. Such methods require large equipment and/or highly trained personnel, which limits them to central facilities such as scientific institutes, hospitals and universities. The need for *in situ* environmental monitoring led to the development of portable handheld devices which process samples quickly and efficiently. Such devices are often “whole-cell biosensors” based on bacteria or fungi, that are genetically engineered to express bioluminescent or fluorescent proteins as a response to specific analytes in their environment, or to slow down their metabolic activity and stop emitting light in the presence of a deleterious analyte (Bachmann, 2003, Daunert *et al.*, 2000, Keane *et al.*, 2002, Kohlmeier *et al.*, 2007, Yagi, 2007, Belkin, 2003).

To study the transients of metal ions from one part of the cell to the other is a problem that cannot be solved with devices of the aforementioned scale. There have been extensive studies of intracellular Ca^{2+} (Allen and Blinks, 1978, Alvarez and Montero, 2002, Brini *et al.*, 1997, Brini *et al.*, 1995, Diarra *et al.*, 1999, Shimomura, 2005, Truong *et al.*, 2001). However, the study of biologically relevant trace metals such as zinc, and copper and contaminants such as cadmium and lead is far less advanced. The increasing interest in brain degenerative diseases and other mental disorders and the recent association of metal ions with brain disease states (Faller and Hureau, 2009, Garzon-Rodriguez *et al.*, 1999, Han *et al.*, 2008, Karr *et al.*, 2004, Miura *et al.*, 2000, Streltsov *et al.*, 2008, Talmard *et al.*, 2007, Tougu *et al.*, 2008), calls for new, *in vivo* and *in situ*, real-time intracellular metal ion sensors.

1.2.3 Desirable characteristics for bioreporters

Scientists can choose the cellular interrogation system to suit the requirements of their experiments, but the available toolbox for intracellular monitoring is still limited and a long list of analytes is awaiting their bioreporters (Czarnik, 1995, Martin, 2008a). Reporters aimed at monitoring a target analyte have the following preferable characteristics.

- **Selectivity:** Reporters ideally respond to the presence of one analyte only. In practice, complete selectivity is an unachievable ideal. However, many potentially interfering species are either not present in biological samples or in the cellular compartment of interest in sufficient concentrations to cause a detectable interference. The supply of competitive metals in cells is maintained in limited supply through the function of metal importers, exporters and metal stores, making proteins compete for metals rather than the other way around (Waldron *et al.*, 2009).
- **High sensitivity:** A large change in signal in response to small changes in analyte concentrations would enable better quantitation of cellular signals.
- **Dynamic range:** The reporter must respond to physiologically relevant concentrations of the target ion in various compartments of the cell. The concentration range may be very wide. For example, in the cytoplasm of resting cells Ca^{2+} is maintained at ~100 nM, while in the endoplasmic reticulum (ER) it is stored at hundreds of micromolar (Palmer, 2009).
- **High signal-to-noise-ratio:** As a rule of thumb, it is desirable that light readings are higher than three standard deviations of the background noise.
- **Cellular localisation:** Localisation of probes to individual compartments of the cell enables the dissection of the origin and destination of ion signals.
- **Non-invasive:** Probes are preferred that do not interfere with cellular functions.
- **Safety:** Non-toxic and non-radioactive probes are preferred.
- **Fast-response kinetics:** Rapid measurement allows real-time monitoring of cellular functions.
- **Quantitative:** While qualitative (presence or absence of analyte) data is useful, more advanced studies require time-dependent quantitation of analyte concentrations.

- **Signal wavelength:** Optical signals are often attenuated when monitored through certain thicknesses of tissue. Red and infrared wavelengths are preferable as they penetrate through tissue more effectively than the other colours.

1.2.4 Signal transduction methods

Fluorescent, bioluminescent and radioactive labels are used in wide range of bioanalytical techniques (Van Dyke *et al.*, 2002). Radioactive labels are less preferred than fluorescent or bioluminescent ones due to safety handling considerations. A comparison of sensitivity of methods is presented in Table 1-2.

Table 1-2 Comparative sensitivity of detection methods (Van Dyke *et al.*, 2002).

Methods	Limit of detection
Luminescence	10^{-19} M
Radioisotope	10^{-18} M
Fluorescence	10^{-12} M
Absorbance	10^{-9} M

The main constraints are linked to the fact that fluorescence requires light excitation, which generates autofluorescence, photobleaching, and phototoxicity. Background fluorescence (or autofluorescence) is an inherent problem in biological samples, and as the scale of the samples decreases and lower level of detection is required, background fluorescence becomes more evident. Bioluminescence is a relatively rare phenomenon in non-marine organisms and usually absent in the types of systems studied in biomedical sciences. This allows for the detection of the proteins at extremely low levels, making these photoproteins attractive labels for analytical applications (Lewis and Daunert, 2000).

Photobleaching and phototoxicity limit the duration of recording, while the autofluorescence prevents the recording of deep structures. Fast phenomena can be monitored, with short time imaging, while the imaging of deep structures still remains difficult. The real contribution of the phototoxicity to the signal itself is difficult to evaluate, since it is inherent to the technique (i.e. light excitation). Additionally, the

excitation required is in some cases within the ultraviolet range. UV exposure can be lethal to the cells in long running experiments. Bioluminescence does not require light excitation and consequently, does not generate auto-fluorescence. In terms of tissue depth, bioluminescence was used to monitor calcium ion activity in the ellipsoid body, which is located relatively deep within the brain (Martin, 2008b).

Table 1-3 Comparison of bioluminescent and fluorescent reporters (Martin, 2008b).

Bioluminescent reporter	Fluorescent reporter
Non-invasive (genetically encoded)	Invasive (dye) or non-invasive (genetically encoded)
Very good signal-to-noise ratio	Generally brighter, but low signal-to-noise ratio
No excitation required	Requires light excitation
Bifunctional: visualisation of localisation by fluorescence	Autofluorescence
Non-toxic	Phototoxicity
Long-term imaging (hours or few days of constant monitoring)	Photobleaching
Moderate spatial resolution	Very good spatial resolution
<i>In vivo</i> whole-animal imaging	High temporal resolution when imaging for short durations

A limitation of bioluminescence is that such reporters currently do not have the multiplexing capability that fluorescent labels do, where a wide variety of fluorescent markers can be utilised in a single assay and easily distinguished from one another. Attempts to change this are underway, and it is now possible to resolve signals from a dual bioluminescent assay based on the different kinetics of the bioluminescent reporter molecules (Rowe *et al.*, 2008a).

1.2.5 Current methods for intracellular ion monitoring

With non-destructive monitoring and accurate detection of weak optical signals, bioluminescent molecules have been used in a plethora of intracellular assays: monitoring important biological molecules (e.g. calcium, ATP); whole cell biosensor assays and tracking the survival of implanted cells in stem-cell based therapies e.g. regeneration of injured cardiac tissue, among others (Roura *et al.*, 2013, Scott *et al.*, 2011). A compilation of intracellular metal ion sensors – with calcium being the most exhaustively studied metal ion – is provided in Table 10-1 of Appendix Chapter 1.

1. Fluorescent dyes

Small fluorophores which bind to metal ions and cause either a change in fluorescence intensity or a spectral shift are available. This enables tracking of calcium ion dynamics in live cells using fluorescence microscopy. The dyes are added to the medium of the cells or tissue under study and allowed to diffuse into the sample, but typically lack a highly defined cellular localisation. To overcome this drawback reporter proteins were developed (Palmer, 2009).

2. Reporter proteins

Reporter proteins (e.g. aequorin) respond to the presence of an ion via a binding event that causes a structural change leading to bioluminescence or fluorescence of the protein. Transfection or transgenic technologies are used to express the protein in the cell, often fused to signal peptide sequences (e.g. ompA) to direct the protein to the desired intracellular location. Aequorin is the most widely used reporter protein for calcium ion monitoring; it is not fully functional without its synthetic chromophore substrate (coelenterazine or chemical analogues thereof) that must be added to the cell medium.

3. Reporter protein chimeras

Fluorescent protein chimeras are practical alternatives to bioluminescent proteins. GFP (Green fluorescent protein) or GFP variants are genetically fused to the functional domain of another protein (e.g. the calcium-binding region of calmodulin), or with a whole protein (e.g. aequorin). The resulting chimera can be localised

subcellularly with the help of an appropriate signal peptide. In this case, binding of an analyte (e.g. metal ion) alters the structure of the fused peptide which in turn affects the fluorescent properties of the GFP.

In recent years considerable progress is being made in this field. Genetically encoded FRET (Fluorescence resonance energy transfer) sensors are currently in development to monitor levels of magnesium, zinc and cadmium ions in transition metal homeostasis. Improvements are still needed; the success of these sensors depends on the sensitivity of the sensors, on achieving significant ratiometric changes of fluorescence signal upon ion binding and on successful intracellular calibration of the sensors, as *in vitro* calibration is not useful. Future scope is to achieve simultaneous imaging of multiple metals and in different locations of the cell (Vinkenburg, 2010, Vinkenburg *et al.*, 2009).

4. *In situ* fluorescent tagging of proteins

Another option employed recently has been to express and localise proteins that incorporate a peptide “linker”. The linker can bind covalently to an appropriate fluorophore that is added to the cell medium (Brun *et al.*, 2009). The fusion peptide is expressed in the cell and after a short incubation period the fluorophore labels the expressed localised peptide. This method can be used to tag and quantify the amount and location of the expressed protein, or act as a reporter that tracks changes of the fluorophore when an analyte, such as Ca^{2+} , binds to the protein or to the fluorophore. A wide range of fluorophores are available that emit light at different wavelengths (Covalys Biosciences, 2009).

5. Microinjection

Microinjection of ion dyes or reporter proteins into cells was used for the delivery of reporters. It is a laborious and limited method, now becoming obsolete due to the development of the systems described above (Creton *et al.*, 1999, Blinks, 1990).

1.3 Aequorin: a calcium-sensitive photoprotein

1.3.1 Introduction

Aequorin is calcium-sensitive photoprotein originating from the jellyfish *Aequorea victoria* (Shimomura, 2005, Shimomura *et al.*, 1962). The active protein is formed from apoaequorin, the luminophore coelenterazine and molecular oxygen (Tsuji *et al.*, 1986, Dikici *et al.*, 2009). The apoaequorin DNA encodes for 196 amino acids. Coelenterazine (MW 423) is an imidazolopyrazine that forms a hydroperoxy derivative with molecular oxygen, and is tightly but non-covalently bound within the photoprotein as an oxidation reaction intermediate (Ohmiya and Hirano, 1996, Vysotski and Lee, 2004).

Bioluminescent reaction in aequorin

Upon binding traces of Ca^{2+} aequorin undergoes a conformational change which triggers the oxidative decarboxylation of coelenterazine into coelenteramide (Figure 1-5). Products of this reaction are a flash of blue light ($\lambda_{\text{max}} \approx 469 \text{ nm}$), CO_2 and Blue Fluorescent Protein (BFP), which consists of the apoprotein loosely bound to coelenteramide (Shimomura and Inouye, 1999). The emission of light is caused by the decay of the bound coelenteramide from an excited state to the ground state (Ohmiya and Hirano, 1996). A schematic of the aequorin bioluminescent reaction is presented in Figure 1-5.

The reaction was found to follow first order kinetics (Shimomura *et al.*, 1962), with reaction rate constants of luminescence ranging from: $0.95\text{--}1.33 \text{ s}^{-1}$ for eight different purified isoforms of native aequorin (isoequorins). Luminescence activities for the same isoequorins were found to range from $4.35\text{--}5.16 \times 10^{15}$ photons/mg aequorin (Shimomura, 1986), which correspond to approximately 0.16–0.2 photons per molecule of aequorin.

Crystallographic structures

Aequorin has a compact globular shape with a hydrophobic cavity in the centre of the molecule. Two different conformations of aequorin have been crystallised, shown in Figure 1-7): PDB ID:1EJ3 is fully active aequorin, prior to any addition of

calcium, at 2.3 Å resolution (Head *et al.*, 2000) and PDB ID: 1SL8 is calcium-loaded apoaequorin, without coelenterazine at 1.7 Å resolution (Deng *et al.*, 2005). The latter is as of yet and to the best knowledge of the author, the closest structure to the “post-calcium-binding” conformation.

Calcium in the EF-hands and Coelenterazine in the hydrophobic core

The protein scaffold is formed by sets of four helices comprising helix-turn-helix (HTH) motifs called EF-hands. The EF-hands are arranged in pairs: EF-I and EF-II in the N-terminal region and EF-III and EF-IV in the C-terminal region (Figure 1-7, Figure 1-8). The anatomy of a typical EF-hand motif is presented in Section 1.3.4. In aequorin, three out of four EF-hand structures (EF-I, EF-III and EF-IV) are “canonical” EF-hands and serve as Ca²⁺-binding sites. Their presence was already known from analysis of the protein primary sequence (Inouye *et al.*, 1985, Charbonneau *et al.*, 1985) even before the crystallographic structure was resolved. Each loop contains twelve sequentially arranged amino acids, with residues in loop positions 1, 3, 5, 7 and 9 coordinating to each calcium ion (Head *et al.*, 2000). In the crystal structure of calcium-bound apoaequorin (Figure 1-8), each of the canonical EF-hand loops is occupied by one calcium ion, coordinated in the characteristic pentagonal bipyramidal configuration (Deng *et al.*, 2004). EF-II lacks the necessary amino acids to facilitate binding of calcium. Since it is apparently unable to bind calcium in aequorin, this domain might instead have a role in the enzymatic function, such as forming a stable scaffold against which the rest of the molecule moves upon the binding of calcium ions (Head *et al.*, 2000).

Hydroperoxy-coelenterazine (an “activated” coelenterazine) is accommodated in the hydrophobic core of the aequorin. The coelenterazine binding pocket is highly hydrophobic and is formed by residues originating from each of the helices (Vysotski and Lee, 2004).

Stoichiometry of calcium ions

Previous studies have suggested that occupancy of two of the three calcium-binding sites in aequorin is sufficient to trigger activation (Shimomura, 1995a, Shimomura and Inouye, 1996).

A log-log plot of light intensity as a function of calcium concentration over a range of 10⁻⁹ to 10⁻² M produced a sigmoidal curve with a lag phase, logarithmic phase

and saturated phase (Figure 1-6). The lag phase of the sigmoidal curve was attributed to binding of calcium by the EDTA in solution. Only when EDTA is saturated with calcium is any additional calcium available to aequorin which utilises it for the triggering of luminescence. The logarithmic phase occurred at physiologically relevant concentrations of calcium (10^{-7} – 10^{-5} M) and had a slope of 2.0–2.5 (Allen *et al.*, 1977). Based on the slope of the logarithmic curve the authors suggested that a molecule of aequorin must bind at least two calcium ions. The exponential part of the sigmoidal curve corresponds to the utilisation of the available calcium while the plateau of the sigmoidal curve shows that metal binding sites of aequorin are saturated and additional calcium does not contribute to increased activity.

Studies determining the stoichiometry of Ca^{2+} by luminometric titration proved that the light emission from aequorin is proportional to the amount of Ca^{2+} added and that two Ca^{2+} per protein molecule are needed to exhaust the luminescence capability of an aequorin sample (Shimomura, 1995a). The requirement for two calcium ions in the luminescence reaction justifies a slope of approximately 2.0–2.5 log-log plot of the relationship between calcium concentration and light intensity reported by Allen and co-workers (1977).

Further work by Shimomura and Inouye (1996) used titration of recombinant aequorin with calcium solution and measured concentration of free calcium by the use of a calcium-sensitive electrode. Three calcium ions were bound by aequorin, which is also consistent with the crystal structures of calcium-bound apoaequorin and other photoproteins (Deng *et al.*, 2005) and the two out of three binding sites were found to have 20-fold higher affinity for calcium ions compared to the third site, although it would not be possible to identify which site this was. Coupled with the fact that approximately two ions are needed to exhaust the luminescent reaction, the third binding site could be unrelated or simply not required for the luminescent reaction.

Slopes greater than 2 could be artifacts (Shimomura, 1995a); main cause for artifacts would be the presence of EDTA in the calcium buffer solutions when studying the relationship between light intensity and calcium concentration; the authors suggested that EDTA may inhibit light emission by binding to aequorin, the effect becoming more pronounced in buffers of low calcium concentration and higher EDTA concentration. Higher calcium concentrations may lead to recycling of calcium and slight increase of light emission occurring in the plateau.

Calcium-independent luminescence

Aequorin emits light even in the absence of calcium ions; at very low levels of calcium ($<10^{-8}$ M) the protein gives off a very low and calcium-independent concentration level of light termed as “calcium-independent luminescence”; however, the light intensity is increased up to one million-fold or more on the addition of calcium (Allen *et al.*, 1977). Detailed structural information on aequorin and theories on how binding of calcium ions may result in the de-stabilisation of the hydroperoxy-coelenterazine are presented in Chapter 5.

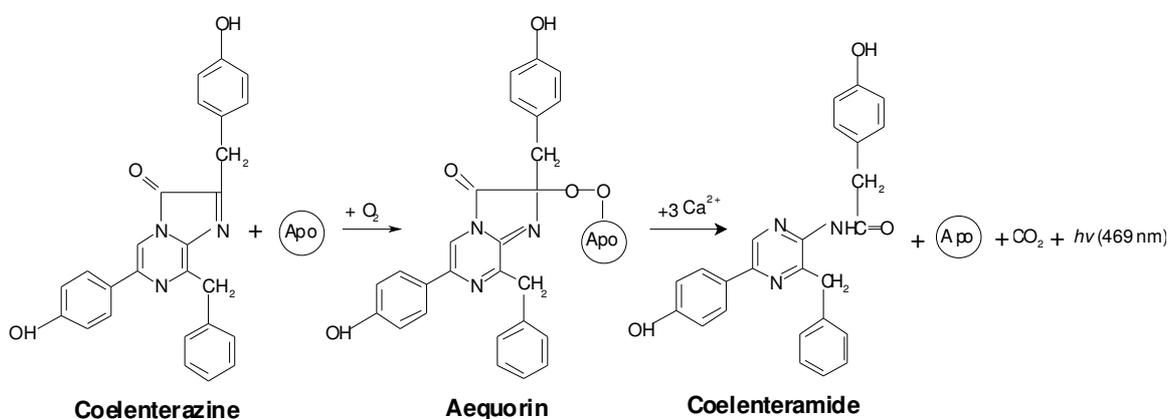


Figure 1-5 Schematic of the bioluminescent reaction of photoprotein aequorin (Ohmiya and Hirano, 1996). “Apo” refers to the apoequorin polypeptide chain.

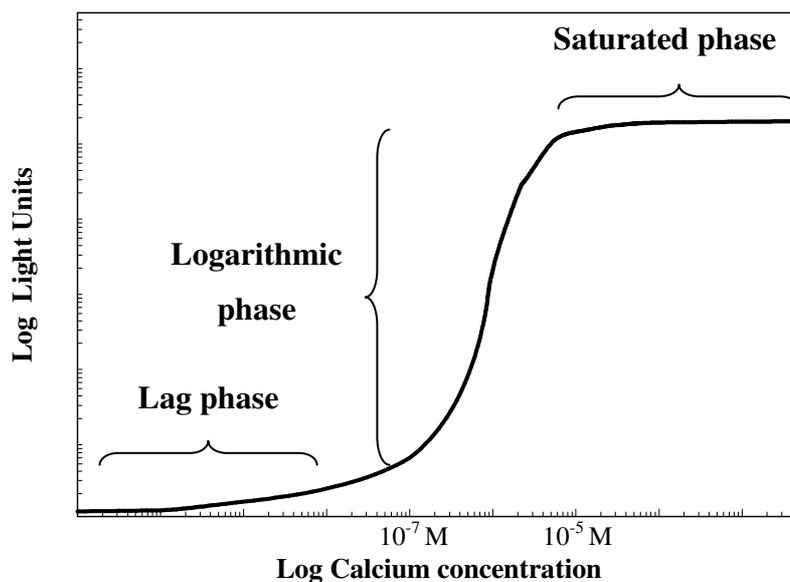


Figure 1-6 Schematic of sigmoidal curve for light emission versus Ca^{2+} concentration.

1.3.2 Uses of aequorin

Calcium concentration levels regulate a wide range of biological processes in eukaryotic cells, including gene expression, muscle contraction, glycolysis, cell division and growth (Dudev and Lim, 2003). Due to its high sensitivity to calcium, instant response to calcium concentration changes, its non-invasive nature and the lack of background interference signals, aequorin has been well suited in its use as a calcium indicator for more than three decades. The bioluminescent reaction proceeds within the physiological pH range and its rate varies steeply with calcium concentration between the biologically relevant range between 10^{-7} and 10^{-5} M (Hastings *et al.*, 1969).

Initial applications of aequorin required microinjection of protein purified from jellyfish, into cells (Allen and Blinks, 1978, Blinks *et al.*, 2000). Subsequent cloning of its gene (Inouye *et al.*, 1985, Prasher *et al.*, 1985, Charbonneau *et al.*, 1985) allowed for the recombinant expression of apoaequorin and the genetic transformation of bacteria, yeasts, plants, and animal cells. Fusion of apoaequorin with signal peptide sequences allowed its targeting to specific cellular compartments (Sala-Newby *et al.*, 2000). A few select examples are: the monitoring of gene expression in hamster ovary eggs (Inouye *et al.*, 1992), analysis of calcium ion homeostasis at the subcellular level using aequorin targeted to specific organelles (Chiesa *et al.*, 2001), use as a reporter enzyme in studying gene expression in mammalian cells (Tanahashi *et al.*, 1990) and analysis of Ca^{2+} homeostasis in primary cultures of skeletal muscle myotubes (Brini, 2008, Brini *et al.*, 1997).

Small peptides were fused at both the C- and N-terminal of aequorin, resulting in functional new proteins. It was shown that aequorin can tolerate fusions at both termini so long as the C-terminal proline is intact (Deo and Daunert, 2001, Lewis and Daunert, 2000). Thus it is possible to incorporate an affinity or fluorescent tag on one terminus and a peptide of interest at the other (Deo *et al.*, 2001). Using this capability, aequorin and its variants have been used as bioluminescent labels in immunoassays (Deo and Daunert, 2001, Deo *et al.*, 2001, Dikici *et al.*, 2009, Shrestha *et al.*, 2002) and as a tool for the search of GPCR (G-protein coupled receptors) ligands in the drug discovery (Dupriez VJ, 2002). The addition of an N-terminus hexahistidine tag allowed metal affinity purification resulting in highly pure and functional aequorin (Glynou *et al.*, 2003).

1.3.3 Properties of aequorin

Rapid loss of aequorin luminescence activity occurs at elevated temperatures (Shimomura *et al.*, 1962). The total light emitted is independent of pH over a wide range of 5.1–8.3. The reaction rate constants, however, are pH dependent. It increases with alkalinity from a near plateau between 6.5 and 7.5 and decreases in the same manner as enzyme activity when pH is lower than 6.5 (Shimomura *et al.*, 1962). The protein is unstable at pH lower than 4.0. When calcium and EDTA are equimolar, the velocity is about half of maximum, whereas only a slight excess of Ca^{2+} results in nearly maximum velocity.

Due to high sensitivity of the protein to Ca^{2+} , certain precautions need to be taken when handling aequorin or apoaequorin: use of only high purity chemicals is recommended, avoidance of any contact of solutions with soft (soda-lime) glass or metal parts where that is possible. EDTA and EGTA have been used to preserve the luminescent activity of aequorin until its use and to prepare calcium-buffer solutions for calibrating the light emission of aequorin (Shimomura, 1991). When the concentration of EDTA exceeds that of Ca^{2+} no luminescence is emitted.

As oxygen is incorporated in the holoprotein structure, Ca^{2+} -triggered light emission can occur in a vessel completely evacuated by air (Shimomura *et al.*, 1962), but an active aequorin complex cannot be regenerated in the absence of molecular oxygen (Shimomura and Johnson, 1975b). Coelenterazine is also retained by aequorin during purification by gel filtration, whereas coelenteramide is easily diffusible from the spent Blue Fluorescent Protein complex.

$(\text{NH}_4)_2\text{SO}_4$ is useful during long term storage due to the stabilising effect it has on the protein (Shimomura and Inouye, 1999). Aequorin solution aliquots can be stored at 4° C and –20° C for a few months without any loss of luminescent activity, but –80° C is required for long-term storage (Shimomura *et al.*, 1962).

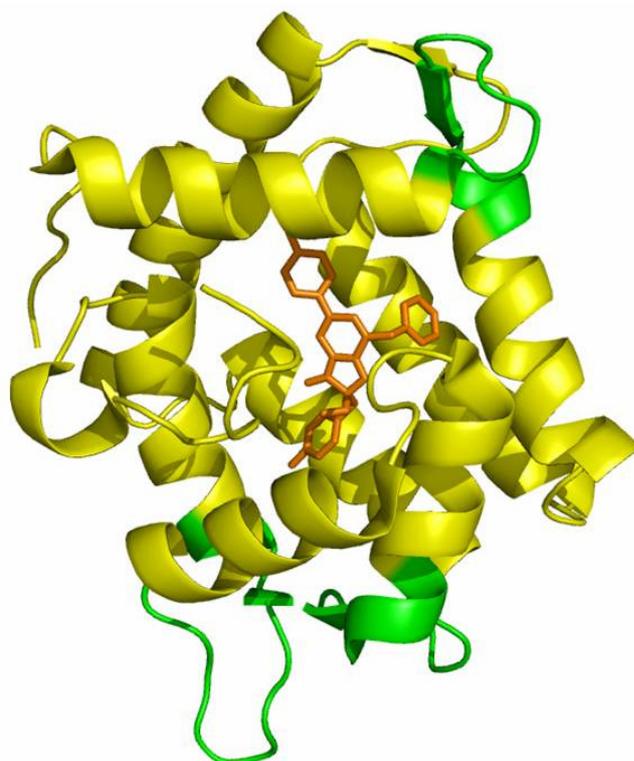


Figure 1-7 3-D model of aequorin with coelenterazine in the hydrophobic core. The EF-hand Ca^{2+} -binding loops are coloured green and bound coelenterazine hyperoxide is coloured orange. PDB ID: 1EJ3 (Head *et al.*, 2000), 3-D model rendered with Pymol (DeLano, 2002).

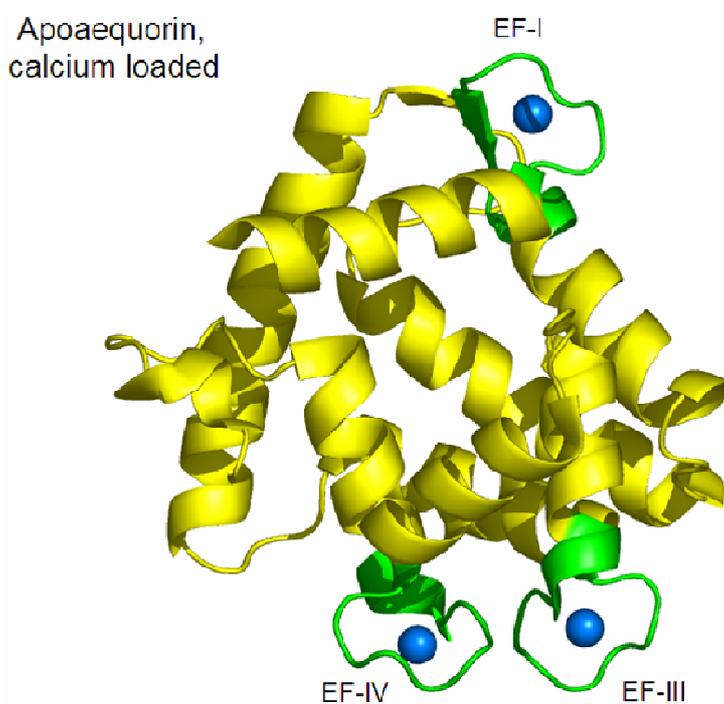


Figure 1-8 Apoaequorin with calcium ions in the Ca^{2+} -binding loops (EF-hands). PDB ID: 1SL8 (Deng *et al.*, 2005), 3-D model rendered with Pymol (DeLano, 2002).

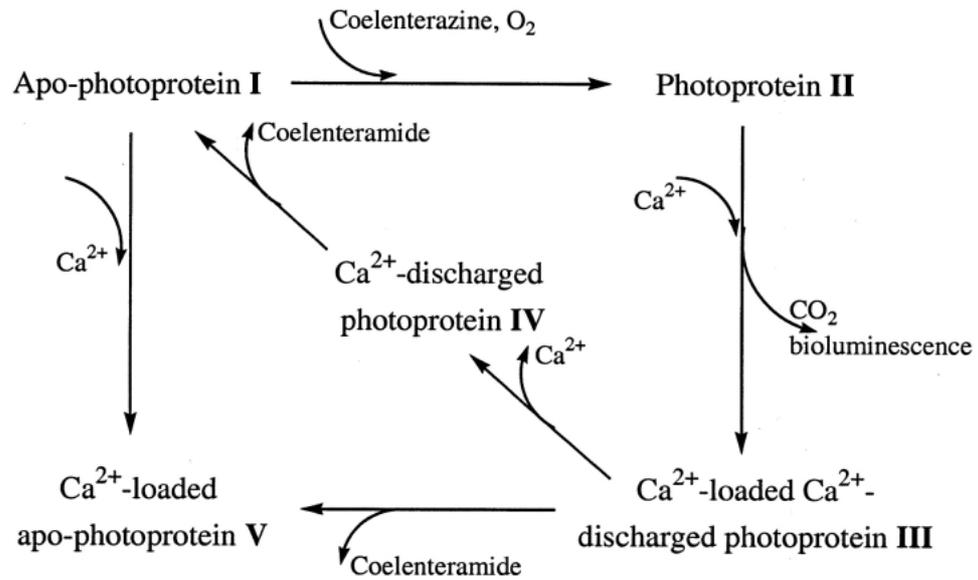


Figure 1-9 Conformation states of a photoprotein. Apoprotein (state I), photoprotein (with 2-hydroperoxycoelenterazine, without Ca²⁺) (state II), Ca²⁺-discharged photoprotein (protein with the reaction product, coelenteramide, and bound Ca²⁺) (state III), Ca²⁺-discharged photoprotein without Ca²⁺ (protein with coelenteramide without Ca²⁺) (state IV), Ca²⁺-loaded apoprotein (state V). The photoprotein conformational states were revealed by HSQC-NMR spectroscopy study of obelin (Deng *et al.*, 2005, Lee, 2001). Figure was reproduced from Deng and co-workers (2005).

1.3.4 EF-hand motif: general features

The EF-hand motif consists of two perpendicular alpha helices (each 10 to 12 residues long) with a 10 to 15 residue loop region between, forming a single calcium-binding site (helix-loop-helix). One calcium ion interacts with residues contained within the loop region. It is found in a broad range of functionally diverse calcium-binding proteins, known as the EF-hand protein superfamily. The EF-hand has been fine-tuned to selectively bind Ca^{2+} against the background of up to 105-fold higher concentrations of Na^+ , K^+ , and Mg^{2+} . Depending on length and the amino acid content of their loops, EF-hands are categorised as “canonical” and “non-canonical”.

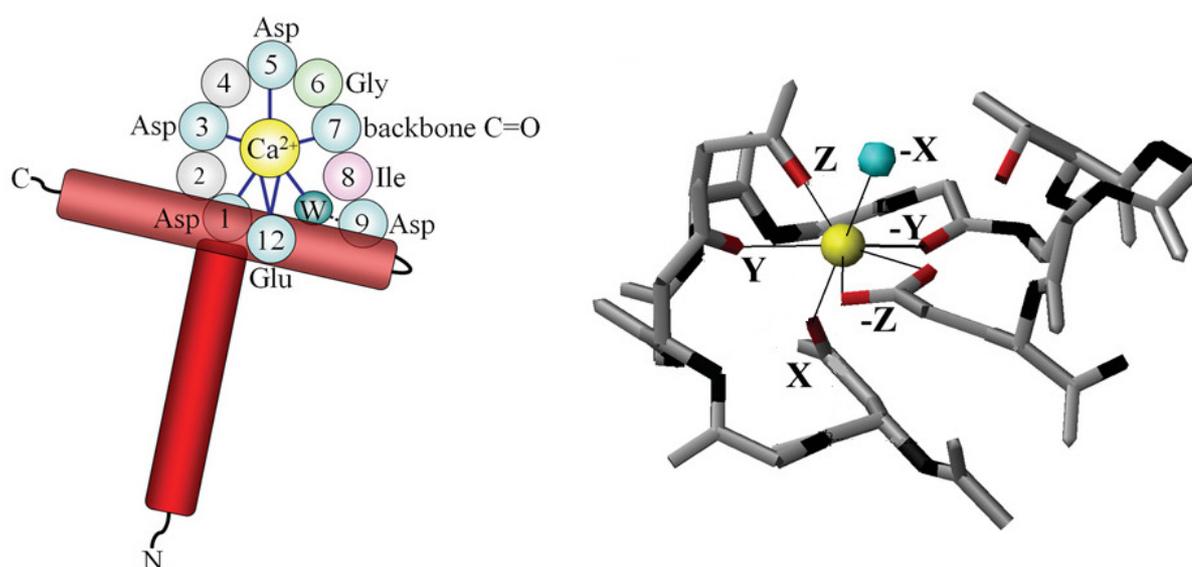


Figure 1-10 The EF-hand calcium-binding loop. On the left: A schematic of the calcium coordination sphere with the entering and exiting helices in red, the coordinating protein ligands in blue and coordinating water molecule (W) in blue. Purple highlights the conserved hydrophobic residue that forms the short β -sheet in the paired EF-hand. Also indicated are the most common amino acids found at the critical positions. On the right: calcium coordination by the canonical EF-hand (EF-I of CaM (PDB ID: 1EXR) illustrating the pentagonal bipyramidal coordination of the calcium ion. The calcium ion is in yellow, the side-chain oxygen atoms in red, and the coordinating water in blue. Backbone NH groups are indicated in black. Figure reproduced from Gifford *et al* (2007).

“Canonical” EF-hands represent the vast majority of the EF-hand motifs. Their loop contains twelve amino acids, each one being important for calcium coordination. Residues in positions 1, 3, 5, 9, and 12 of the loop region are highly conserved and provide oxygen ligands necessary for binding of calcium, while position 7 of the loop ligates the calcium ion with a main-chain carbonyl oxygen. The most common residues coordinating calcium and their respective positions in the canonical loop are listed in Table 1-4. In most EF-hand proteins the residue at position 12 is a glutamate, which contributes both its side-chain oxygens for calcium coordination. The calcium ion is coordinated in a pentagonal bipyramidal array with an average 2.4 Å separation to oxygen atoms (Figure 1-10). Frequently one or two water molecules are also involved in ligating calcium (Strynadka and James, 1989).

Table 1-4 Amino acid sequence preference in the EF-hand loop. Table recreated from Gifford *et al* (2007).

EF-loop position	1	2	3	4	5	6	7	8	9	10	11	12
Coordinating ligand	X sc		Y sc		Z sc		-Y bb		-X sc*			-Z sc2
Most common	Asp 100%	Lys 29%	Asp 76%	Gly 56%	Asp 52%	Gly 92%	Thr 23%	Ile 68%	Asp 32%	Phe 23%	Glu 29%	Glu 92%
Frequently observed		Ala Gln Thr Val Ile Ser Glu Arg	Asn	Lys Arg Asn	Ser Asn		Phe Lys Gln Tyr Glu Arg	Val Leu	Ser Thr Glu Asn Gln	Tyr Ala Thr Leu Glu Lys	Asp Lys Ala Pro Asn	Asp

The Ca²⁺ ligands are indicated by both their position in the EF-loop and in the coordinating array. (sc) indicates side chain coordination and (bb) indicates coordination via the backbone. The asterisk (*) indicates that the oxygen ligand is typically provided by a water molecule that is hydrogen-bonded to the side chain of the amino acid at position 9. The most common amino acids at each position are noted with their corresponding percentages of occurrence, followed by those that occur with a frequency greater than 5% in known EF-loops.

“Non-canonical” EF-loops can also bind Ca²⁺ in a pentagonal bipyramidal or octahedral coordination. The length of the loops ranges from eleven amino acids (rarest cases) to fourteen and sometimes fifteen amino acids. They may carry substitutions that disable Ca²⁺-binding or insertions that bring the coordinating residues too far apart to bind calcium in the canonical manner. However, they compensate for these adversities by offering oxygen from their main-chain carbonyl groups for ligating calcium, in some

cases even turning inside-out to accommodate this coordination and in some cases exhibit high affinity for calcium (Strynadka and James, 1989).

The role of EF-hands in proteins can be regulatory or structural. In regulatory EF-hand proteins binding of calcium induces a conformational change that is transmitted to their target proteins, and (often) ultimately results in catalysing enzymatic reactions. In structural EF-hand proteins, EF-hand domains do not undergo significant conformational changes but seem to play a role in buffering levels of intracellular calcium ions (Gifford *et al.*, 2007).

EF-hands tend to occur in pairs, which form a discrete domain so that most family members have two, four or six EF-hands. This pairing also enables communication, and many EF-hands display positive cooperativity. The ability of an EF-hand to bind Ca^{2+} depends on its selectivity over Mg^{2+} (a cation with similar chemical properties to Ca^{2+} and with a cytoplasmic concentration several orders of magnitude higher). Variation in calcium binding affinity for different EF-hand proteins is due to the amino-acid composition of the residues at the coordinating positions (dissociation constants for Ca^{2+} range from 0.5×10^{-3} to 0.3×10^{-9} M) (Dudev and Lim, 2003).

1.3.4.1 EF-hands and cooperativity

EF-hands tend to occur in pairs, which form a discrete domain so that most family members have two, four or six EF-hands. This pairing also enables communication, and many EF-hands display positive cooperativity. The conformational effects of Ca^{2+} binding are varied, function-dependent and in some cases, minimal. EF-hand proteins exhibit various sensitivities to Ca^{2+} , reflecting the intrinsic binding ability of the EF-hand as well as the degree of cooperativity in Ca^{2+} binding to paired EF-hands. Relatively little is known about the mechanisms of cooperativity or dynamic conformational effects of Ca^{2+} binding of the other EF-hand proteins as most research in this field to date has focused on the model members CaM, TnC and calbindin D9K (Gifford *et al.*, 2007).

1.3.5 Proposed mechanism for the bioluminescence reaction

Despite the availability of the tertiary structure, the precise catalytic mechanism of the photoprotein remains unsolved (Prendergast, 2000). The bioluminescence reaction mechanism has been partially deciphered on the basis of detailed structure-activity studies of both the apoprotein and coelenterazine (Ohmiya and Hirano, 1996) as well as on the resolved crystal structure of aequorin (Head *et al.*, 2000), and is assumed to follow the same mechanistic pathway as the proposed mechanism of a chemiluminescence reaction pathway proposed by Ohmiya and Hirano (1996). Figure 1-11 shows the mechanism proposed by Prendergast (2000): **a**) a reactive carbanion is formed as result of distortion of the ground state of coelenterazine **b**) the carbanion attacks molecular oxygen bound by the protein, forming coelenterazine hydroperoxide **c**) coelenterazine hydroperoxide is stabilised by Tyr184 of the protein chain **d**) a conformational change in the protein (caused by binding of calcium ions) allows the attack of hydroperoxy anion on the reactive carbonyl **e**) leading to the formation of an unstable, intermediate dioxetanone **f**) the dioxetanone undergoes scission to produce CO₂ and enolate ion in an excited state, which emits light when relaxing to ground state.

1.3.6 Regeneration of aequorin

The enzymatic function of aequorin was proven when it was shown that it could be recharged repeatedly and reversibly (Shimomura and Johnson, 1975b). BFP (Blue Fluorescent Protein) is a complex consisting of apoaequorin non-covalently bound to coelenteramide and calcium ions, which can be dissociated into its components. The regeneration procedure requires removal of Ca²⁺ with a chelating agent such as EGTA or EDTA, separation of the protein from coelenteramide and incubation with molecular oxygen and coelenterazine (Figure 1-12). BFP can be dissociated into apoprotein and coelenteramide by gel filtration or treatment by ether (Ohmiya and Hirano, 1996). A reducing (thiol) agent such as DTT or β-mercaptoethanol is added to facilitate the harbouring of coelenterazine into the hydrophobic cavity of the apoaequorin.

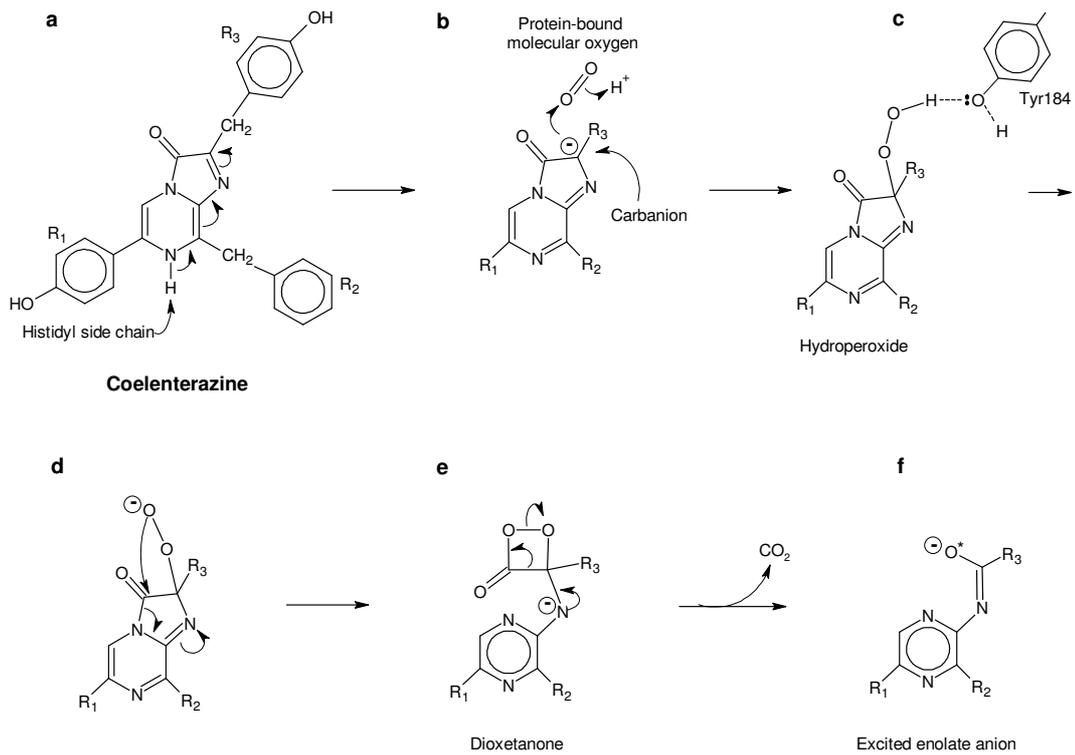


Figure 1-11 Proposed reaction for the bioluminescence of aequorin. Picture reproduced from Prendergast, 2000. Description in Section 1.3.5.

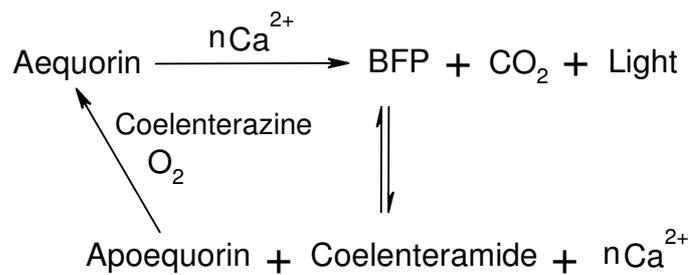


Figure 1-12 Regeneration of aequorin.

1.3.7 Ion selectivity of aequorin

The selectivity of aequorin bioluminescence against various cations has been the object of some debate (Shimomura *et al.*, 1962, Izutsu *et al.*, 1972, Shimomura and Johnson, 1973, Izutsu *et al.*, 1974). Originally believed to be triggered by calcium and rare earth metal ions only, it was shown that aequorin may also respond to a number of other cations, yielding lesser light outputs and depending on conditions such as pH and presence of various metal chelators. Table 1–5 is a synopsis of aequorin metal selectivity studies: fifteen cations in the form of salts of chloride, sulfate or acetate were tested for possible activation of the bioluminescence reaction.

Table 1-5 Aequorin tested against ions for triggering light-emitting reaction

Ion	Charge	Ionic radius (Å)	Atomic Number	Activity?
Sodium	+1	0.095	11	Inhibition
Potassium	+1	0.133	19	Inhibition
Ammonium	+1	0.143	7	Inhibition
Magnesium	+2	0.065	12	No
Copper	+2	0.069	29	Low
Cobalt	+2	0.075	27	Low
Iron	+2	0.076	26	No
Manganese	+2	0.08	25	No
Cadmium	+2	0.095	48	Yes
Calcium	+2	0.106	20	Yes
Europium	+2	0.112	63	Yes
Strontium	+2	0.113	38	Low
Barium	+2	0.135	56	Low
Lead	+2	1.19	82	Yes
Iron	+3	0.064	26	No

All ions were tested by Shimomura and co-workers (1962) and verified by Shimomura and Johnson (1973), unless stated otherwise. Inhibition caused by monovalent cations was reported by Moisescu and co-workers (1975). Strontium-triggered luminescence was used in monitoring changes of free Ca²⁺ concentration in the endoplasmic reticulum (ER) (Montero *et al.*, 1995).

Most divalent cations failed to initiate a light-emitting reaction, while the rare earth metals (here represented by Eu²⁺) and Sr²⁺ were found to successfully substitute for

Ca^{2+} (relative potencies of the different cations have been estimated as $\text{Eu}^{2+} > \text{Ca}^{2+} > \text{Sr}^{2+}$). The fact that Mg^{2+} , K^+ and Na^+ ions do not initiate luminescence, but inhibit the Ca^{2+} -triggered reaction suggests that they compete with Ca^{2+} for the binding sites without triggering luminescence (Ohmiya and Hirano, 1996). Previously contradicting results (Shimomura *et al.*, 1962, Izutsu *et al.*, 1972) were put down to the presence of contaminating calcium during protein handling and metal ion preparations and to pH buffering or lack thereof (Shimomura and Johnson, 1973).

As a conclusion, the luminescence reaction of aequorin is highly specific for Ca^{2+} at pH 7.5-8.0 and in the absence of rare earth metals or strontium (the latter condition is typically met when performing intracellular studies) (Shimomura and Johnson, 1973). Since then the usefulness of aequorin as an intracellular calcium indicator has been widely established (see Section 1.3.2). It may be possible that given the appropriate coordination environment, ions other than Ca^{2+} could potentially induce the required structural changes for evoking the luminescence yielding reaction.

1.3.8 Other photoproteins

Since the discovery of aequorin, more jellyfish photoproteins have been isolated, which use the same, or a very similar, chromophore (Tsuji *et al.*, 1995). Photoproteins mitrocomin, clytin and obelin showed very strong amino acid sequence identities with aequorin. The amino acid sequences of apo-photoproteins of aequorin, clytin (or phialidin), mitrocomin and obelin have been aligned with the amino acid sequences of non-luminescent Ca^{2+} -binding proteins: human and bovine calmodulin (a typical Ca^{2+} -binding protein), parvalbumin, troponin C, intestinal Ca^{2+} -binding protein and sarcoplasmic Ca^{2+} -binding protein (Table 1-6 and Table 10-2). The Ca^{2+} -binding sites of all these proteins share conserved positions (Table 10-2). The distance between the second and third EF-hand in aequorin and the third and fourth EF-hand structures in bovine calmodulin is conserved, suggesting that both have a common evolutionary origin (Tsuji *et al.*, 1986). The photoprotein group differed from other Ca^{2+} -binding proteins in that they contained a relatively large number of cysteine, tryptophan, histidine, proline and tyrosine residues, suggesting that these residues may have evolved as part of the light-emitting mechanism (Tsuji *et al.*, 1995).

Table 1-6 Identity between aequorin and other Ca²⁺-binding proteins. Mitrocomin, obelin and clytin are also photoproteins. The remaining proteins bind calcium but have no light emitting capabilities. Pairwise alignment and identity calculation was performed with BioEdit (Hall, 2007).

Pairwise alignment of aequorin with:	Reference	Identity
Mitrocomin	GenBank: AAA29298.1	69%
Obelin	PDB: 1SL7	63%
Clytin	GenBank: BAG49091.1	58%
Human calmodulin	GenBank: CAA36839.1	19 %
Bovine calmodulin	PDB: 1PRWA	19%
Human parvalbumin	GenBank: CAA44792.1	12%
Intestinal Ca ²⁺ -binding protein	PDB ID 1B1G	5%
Sarcoplasmic Ca ²⁺ -binding protein	PBD ID 2SAS	10%
Troponin C	GenBank: AAA30011.1	3%

1.4 Coordination of metals in proteins

Proteins that chelate metal ions specifically are called metalloproteins. One third of all proteins belong in this category (Palmer, 2002). Properties of the protein-metal interactions, such as coordination geometry, side chain interactions, overall charge and size of cavity, all play a role in ion binding and selectivity (Dudev and Lim, 2003). There are currently no generic rules on the binding and selectivity of metal ions in proteins. However, the factors required for metal binding are relatively well understood and metal-ion binding sites have been attractive targets in protein engineering (Lu and Valentine, 1997).

Metal-ligand (M–L) distances play a central role in metal discrimination in proteins. They reflect the properties of the metal complex, namely the electronic distributions of the metal and the ligands, their polarising abilities and interactions, as well as the metal coordination number (CN) and coordination geometry, which are key factors governing the structure and reactivity of a metal complex (Cotton and Wilkinson, 1988). It has been proposed that it is the protein matrix that dictates the specific role of a metal ion and M–L distances have been used to distinguish between catalytic versus structural role for Zn²⁺ in proteins (Kuppuraj *et al.*, 2009).

Native metal ions fulfil the best “fit” conditions in metalloproteins, and when substituted with an alien ion, their structures adopt different arrangements. For example, Cd^{2+} can replace the native Zn^{2+} in Zn-finger motifs in the same tetrahedral coordination as the native ion, but it cannot maintain the requisite conformation for the DNA-binding function of Zn-finger proteins due to increased metal-protein ligand distances (Kuppuraj *et al.*, 2009).

The preferred coordination geometry for some selected metal ions derived from statistical analysis (Harding, 2006) and data collection/observation (Dudev and Lim, 2003, Rulisek and Vondrasek, 1998), are summarised in Table 1-7 and more analytical information is given in the Appendix Chapter 1 (Table 10-3).

Analysis of the preferred coordination geometries and distances in the Cambridge Structural Database (CSD) for organic molecules and metal-organic compounds (Allen, 2002) shows that small metal ions exhibit a strong preference for a particular coordination but the larger metal ions are found in two or more geometries with comparable frequencies. In Table 1-9 some of the existing metal binding sites are listed and experimental dissociation constants are provided. Table 1-8 summarises distances of O, N and S around metal ions, taken from PDB and CSD structures. For Asn and Gln, M—O distances are expected to be similar or very slightly longer to monodentate carboxylates. For Ser and Thr, expected M—O distances are between those for water and for monodentate carboxylate. For Tyr, expected M—O distances are shorter by $\sim 0.1 \text{ \AA}$ than for monodentate carboxylate (Kuppuraj *et al.*, 2009).

Table 1-7 Summary of preferred coordination geometries of metal ions derived from PDB and CSD. The results of the original analysis are presented in Table 10-3.

Metal ion	Preferred coordination geometries
Co^{2+}	Octahedral
Cu^{2+}	Square pyramidal or square planar
Zn^{2+}	Tetrahedral
Cd^{2+}	Tetrahedral, often octahedral
Ca^{2+}	Octahedral
Pb^{2+}	Tetrahedral

Table 1-8 Distances around metal ions from the PDB and CSD. Resolution of PDB and CSD structures was less than 1.25 Å and 0.065 Å respectively. Table reproduced from Harding (2006).

	O, water	O, Asp or Glu monodentate	O, main-chain carbonyl	N, histidine	S, cysteine
Na	2.41**	2.41**	2.38**	—	—
Mg	2.07***	2.07**	2.26	—	—
K	2.81*	2.82*	2.74*	—	—
Ca	2.39**	2.36**	2.36**	—	—
Mn	2.19***	2.15***	2.19	2.21**	2.35
Fe	2.09**	2.04**	2.04	2.16*	2.30***
Co	2.09**	2.05***	2.08	2.14**	2.25*
Cu	2.13	1.99*	2.04	2.02**	2.15
Zn	2.09***	1.99***	2.07	2.03***	2.31**

*** most reliable values (standard deviation is ≤ 0.05 Å); ** standard deviation is ~ 0.10 Å; * standard deviation is 0.15–0.20 Å; No asterisk: least reliable

For some types of complexes, mainly those of Co^{2+} , Cu^{2+} and Zn^{2+} , several different coordination numbers are found. The complexes involving water and carboxylate donors have metal-ion coordination number six; some Zn^{2+} and Cu^{2+} complexes are four or five-coordinate and Ca^{2+} may also be seven- or eight-coordinate. Where imidazole is present, most Zn^{2+} complexes are four-coordinate. Cu^{2+} is found in four-, five- and six-coordinate examples. In thiolate complexes the common coordination numbers are 4 and 6 for Co^{2+} and mostly 4 for Zn^{2+} , while all the Cu^{2+} complexes are three-coordinate.

Generally, M–L distances increase with cationic radius but also with metal coordination number due to increased electronic and steric repulsion among the coordinating ligands. Donor atoms that can transfer more negative charge to the metal centre also result in shorter M–L distances. It was noted however that the size and charge donating ability of the coordinating atom is affected by the size, charge and volume of the other constituent atoms of the same ligand. Exceptional M–L distances can also occur. For example, zinc can make one or more abnormally long (2.3–2.5 Å) bonds in addition to the existing four or more typical bond lengths (Harding, 2006). Copper ions with coordination numbers 5 or 6 have also resulted in distances in the range 2.10–2.91 Å, depending on the ligand (Harding, 2006).

In order of decreasing prevalence, Ca^{2+} is typically coordinated to carboxylates, carbonyls, water and hydroxyl oxygen atoms, with a coordination number in proteins ranging from six to eight. The ionic radius of a calcium ion is always higher than that of Mg^{2+} for a given coordination number. In the EF-hands of aequorin calcium binds to

six ligands from amino-acid residues and one water molecule, with a coordination number of 7 in a pentagonal bipyramidal geometry.

Zn^{2+} is most commonly found in the zinc-finger family of proteins which are DNA-binding domains that contain zinc-binding motifs including Cys2His2, Cys3His and Cys4 metal-binding sites. Zinc prefers soft ligands such as the sulfide from Cys and imidazole nitrogen atom from His, but is also found coordinated to Asp and Glu side chains. In Zn-finger proteins the zinc usually prefers a tetrahedral coordination, but it can also adopt five- or six-coordinate geometries.

Mg^{2+} binds to ligands of low polarisability, with oxygen being the most preferred coordinating atom, followed by nitrogen. All Mg^{2+} -binding sites in proteins contain at least one carboxyl ligand. Non-charged protein ligands include the side chains of Asn/Gln and the backbone carbonyl groups, followed by the Ser/Thr, His, and Tyr side chains. Mg^{2+} nearly always prefers an octahedral ligand coordination geometry, which is complemented by water ligands. The few Mg^{2+} -binding sequence motifs that have been identified include -NADFDGD-, -YXDD- or -LXDD- Table 1-9 (Lu and Valentine, 1997).

Table 1-9 Metal-binding sites and experimental metal binding constants for ion-binding proteins.

molecule	Ca site	Zn site	Mg site	Cu site	Kd (M)	Reference
calmodulin	pseudo EF-hand ²				2.8×10^{-7}	Dudev and Lim, 2003
			pseudo EF-hand		3.7×10^{-4}	
	canonical EF-hand ³				5×10^{-8}	
			canonical EF-hand		1.7×10^{-3}	
calmodulin	pseudo EF-hand				1.8×10^{-5}	Porumb et al, 1977
	canonical EF-hand				2.4×10^{-6}	
parvalbumin	NA				3.7×10^{10}	Dudev and Lim, 2003
			NA		1×10^{-5}	
S100A7 (psoriasis) ¹		His-X4-His within pseudo EF-hand			1×10^{-4}	Brodersen, 1999
S100B		His-X4-His			NA	Brodersen, 1999
S100B				NA	4.6×10^{-7}	Nishikawa et al, 1997
PrP (priorin protein)				-HGGGW-	$<5 \times 10^{-6}$	Burns et al, 2002
Different RNA polymerases			-NADFDGD-		NA	Dudev, 2003
DNA polymerase I			-NADFDGD-		NA	Dudev, 2003
HIV reverse transcriptase			-NADFDGD-		NA	Dudev, 2003
Reverse transcriptase and telomerase			-YXDD- -LXDD-		NA	Dudev, 2003
Zinc finger family protein		Cys2His2			NA	Baudier et al, 1986
Zinc finger family proteins		Cys3His, Cys4			NA	Baudier et al, 1986
De novo RBP A		His-X-His and His 3			3.6×10^{-8}	Muller and Skerra, 1994
De novo RBP B		His-X-His and His 3			4.4×10^{-7}	Regan, 1995
Metallotheionins		NA			1.4×10^{-13}	Heizmann and Cox, 1998

1: Psoriasis is the only member of the S100 EF-hand CaB family that doesn't bind calcium. Instead, it exhibits affinity for zinc 2: Pseudo or variant-EF-hand is the N-terminal, 14-amino acid EF-hand loop found in members of the S100 protein family. The N-terminal EF-hand amino acid sequence in protein S100b is QYSGREGDKHKLKK (Brodersen *et al.*, 1999). 3: Canonical EF-hand is the C-terminal, 12-amino acid EF-hand loop normally found in all CaB proteins. The same formation serves as calcium binding site in aequorin (Charbonneau *et al.*, 1985). NA: Not available in the particular reference.

1.5 Protein engineering

Protein engineering is the process of creating new proteins with desirable traits and has evolved since the mid 1980s when the first protein mutants were created (Ulmer, 1983), into two main branches: rational design and directed evolution.

1.5.1 Rational design

Rational design is specific, deliberate design of a protein, based on knowledge of the protein's structure and function and/or use of structural homology in order to achieve new or improved properties. Rational design relies on well-developed mutagenesis techniques and typically involves a limited number of mutations.

Site-directed mutagenesis (SDM) is a PCR-based method that allows single point mutations in the amino acid sequence (replacement, deletion or insertion of a residue) (Old, 1994). Variations of the methods allow alterations in multiple adjacent amino acids (Stratagene, 2005) or mutations at up to five different selected sites of the amino acid sequence (Hogrefe *et al.*, 2002).

Seebeck and Hilvert (2003) reported the conversion of a pyridoxal phosphate (PLP)-dependent alanine racemase into an aldolase via a single active-site mutation Tyr265Ala. Lawson and co-workers (2009) increased solubility and decreased aggregation of 11 beta-hydroxysteroid dehydrogenase type 1 by replacing key hydrophobic surface residues with charged glutamic acid.

The prediction of specific mutations that elicit the desired effect is still very difficult in many cases (Dalby, 2003). Computational protein design algorithms, potential functions such as CharmM (Brooks *et al.*, 1983, MacKerel Jr *et al.*, 1998), and molecular dynamics simulations (Karplus and Kuriyan, 2005) are available to assist rational design, but their low accuracy often results in a need for repeated cycles of engineering and experimental testing.

A branch of protein engineering is the *de novo* protein design of new polypeptides from scratch (DeGrado, 1997). The number of possible amino acid sequence variations is enormous, but only a subset of these variations will fold reliably and quickly to a single native state. *De novo* protein design involves identifying such

sequences, and preferably those with a physiologically active native state. The holy grail of protein engineering is the “true” design of a protein to any specification; that is writing down a single sequence that will perform a desired function.

Successful results in rational protein design depend upon the availability of reliable structural information, an understanding of the mechanism of the protein function, and either a good predictive method for identifying the structure adopted by the newly mutated residues, or otherwise blind faith and intuition. In computation, the limitation is often in the experimental testing; the rate at which hypotheses, improvements or new approaches can be verified or rejected is low (Plückthun and Mayo, 2007). Much research currently is taking place into the understanding of protein folding and protein recognition for protein design principles, but the details that dictate function are still elusive. An alternative protein engineering approach requiring little or no understanding of the tertiary structure and function of the protein of interest is directed evolution.

1.5.2 Directed evolution

The detailed structural knowledge of a protein is often unavailable, and even when it is available, it can be extremely difficult to predict the effects of various mutations. Directed evolution borrows from nature's toolbox of mutational mechanisms to alter proteins without requiring any prior knowledge of their structure (Stemmer, 1994, Arnold, 2001, Arnold *et al.*, 2001, Arnold, 1993, Arnold, 1998). One directed evolution technique applies random point mutagenesis to a protein by for example, error-prone PCR (Zhou *et al.*, 1991, Zhao *et al.*, 1998, You and Arnold, 1996) followed by a screening regime to identify variants with the desired qualities. Further rounds of mutation and selection are then applied. This process mimics natural evolution and can produce results superior to rational protein design (Dalby, 2003, Dalby, 2011). DNA shuffling is a mix and match of DNA fragments from successful variants to produce improved variants (Stemmer, 1994). This method mimics recombination occurring naturally in sexual reproduction. More recent DNA shuffling techniques such as SHIPREC (Sieber *et al.*, 2001), ITCHY (Chopra and Ranganathan, 2003) and SCRATCHY (Lutz *et al.*, 2001) swap fragments between completely different non-homologous proteins. Using the latter approaches it is possible to explore and combine

distant regions of sequence space in combinations that would not have been attempted through rational design; new chimeric proteins with functionality and native folds can be created using this approach (Carbone and Arnold, 2007).

Directed evolution can produce variants with desired properties often through mutations that would not have been predicted from knowledge of the original protein's structure. To success probability of directed evolution depends on having large mutant libraries, high-throughput protocols for screening of desired qualities and expensive automation to support these functions. A challenge pertaining to any mutational approach – but more so to directed evolution – is that not all desired activities can be easily screened for (Goddard and Reymond, 2004).

1.5.3 Integration of random and rational approaches

Rational design and directed evolution techniques are not mutually exclusive; most researchers use various combinations of rational desing and directed evolution strategies to improve or alter proteins. The best approach is a matter of personal preference and decided judged on a case by case basis. Typically, the protein engineering strategy comprises of the following stages: (1) choice of locations for changes based on structural knowledge, homology modelling, sequence comparisons and/or computational modelling; (2) applying random or rational mutagenesis methods (using single or multiple substitutions, recombination, permutation, insertions and deletions) in the selected regions; (3) screening or selection for desired properties. (Bornscheuer and Kazlauskas, 2009).

Beneficial mutations for altering specificity and enhancing enantioselectivity are most often occurring in enzyme binding sites (Dalby, 2011, Morley and Kazlauskas, 2005, Kazlauskas, 2005) random mutagenesis within specific regions of binding sites can offer an efficient mutational strategy. Substitutions that increase the catalytic activity have been identified within binding/catalytic sites but also are scattered widely throughout the protein (Kazlauskas and Bornscheuer, 2009), suggesting that wider areas should also be targeted.

The knowledge of a protein's structure and computational methods are still not solely sufficient to predict and design the most effective mutations. DeSantis and co-workers used site saturation mutagenesis (SSM) to find a nitrilase variant with increased

enantioselectivity toward an intermediate for the synthesis of atorvastatin. Dumon and co-workers (Dumon *et al.*, 2008) achieved a 25 °C increase in thermostability of xylanase by using a combination of random and rational approaches (Gene site saturation mutagenesis, GSSM) followed by rounds of GeneReassembly Library construction, whereby combinatorial segments were blended; subtle but beneficial changes were created. The molecular basis for the increased thermostability was extraordinarily subtle and these changes would not have been predicted through knowledge of protein structure or confirmed computationally.

The capabilities of protein engineering are constantly expanding through advances in computation, mutational/evolutionary technologies, structural biology and screening/selection capabilities. Combination of all the relevant available options will be the winning strategy for protein design in the foreseeable future (Plückthun and Mayo, 2007).

1.5.4 Metalloprotein engineering

The area of design and redesign of metal binding sites is one with an increasing publication rate over the past two decades. The approach consists of starting with a naturally occurring, folded, and stable protein scaffold. By modifying local portions of the protein without causing significant perturbation to its folded structure and stability, a functional protein with altered or enhanced activity can be created. Nature has used the same approach successfully. Thousands of protein 3D structures can be classified into a limited number of basic scaffolds, while the active-site diversity (such as metal binding sites in metalloproteins) has been achieved by evolutionary fine-tuning (Lu and Valentine, 1997).

Metal-binding sites in proteins are attractive targets in protein engineering due to their importance in biological functions and relative simplicity compared to other protein ligands (Regan, 1995).

Metal-binding sites have been engineered into proteins for use in immobilised metal affinity chromatography. One example of this approach is a high-affinity zinc site engineered into recombinant serum retinol-binding protein (Schmidt *et al.*, 1996). Redesign of metal sites in order to change their specificity has also been achieved. Restriction endonuclease *EcoRV* metal-binding specificity was switched from Mg²⁺ to

Mn²⁺ by a single Ile→Leu mutation (Vipond *et al.*, 1996, Lu and Valentine, 1997). Another example was the redesign of the classical blue copper proteins azurin into purple CuA protein containing a mixed-valence binuclear copper site, through the use of loop-directed mutagenesis (Hay *et al.*, 1996) (Lu and Valentine, 1997). Cavity complementation is a method also used to redesign metal-binding sites. It involves changing one of the ligands to a smaller, non-coordinating ligand such as glycine or alanine, thus creating a 'cavity' within the metal-binding site. Adding different exogenous ligands can complement the cavity sterically. This can either restore the activity or result in a new structure and activity (Lu and Valentine, 1997).

Computational design approaches have also been used to design metal binding sites in protein scaffolds (DeGrado, 1997, Kaplan and DeGrado, 2004, Nanda *et al.*, 2005). However, this approach is still somewhat embryonic as the ability of the current generation of molecular dynamics force fields to properly describe metal pockets is severely lacking due to the intrinsic difficulty of handling polarisation and charge transfer contributions (Dal Peraro *et al.*, 2007).

1.6 Overview

1.6.1 Aequorin mutants as potential bioreporters

Aequorin has been the preferred intracellular calcium sensor, used for interrogating a wide range of biological processes in eukaryotic cells, including gene expression, muscle contraction, glycolysis, cell division and growth (Dudev and Lim, 2003). Many metal ions play a key role in the metabolism of the cell, and particularly zinc and copper have been found to participate in brain function and have been associated with degenerative brain diseases. Currently, the scientific toolbox lacks established intracellular reporters with a set of advantages similar to that of aequorin.

The wealth of knowledge available in the literature that explores the preferred coordination and parameters affecting ion binding in proteins could be very useful in altering the specificity of the protein for other ions. Previous mutational studies on aequorin and the EF-hand domains of other proteins suggest that it is possible to modulate the affinity of the calcium-binding loop, shift the spectral emission, and alter

the kinetics of the light-emitting reaction mechanism through mutagenesis of aequorin (Section 5.2).

1.6.2 Potential activator ions for aequorin mutants

The metal ions tested in the scope of this work as potential activators for aequorin and mutant variants thereof are: calcium, zinc, copper, cadmium, lanthanum, cobalt and lead. Since considerable time, effort and resources would be invested in creating both the library and the screening assay, it made sense to include as many ions as possible in order to allow space for serendipity. As it is not possible to predict the effects of the mutations, it may be that by serendipity, the preference of aequorin could be tweaked towards an ion that is not at the top of the list in terms of biological relevance, but still of great value for environmental studies.

Monovalent ions were excluded on the basis of their low charge and reported inhibition (Section 1.3.7, Table 1–5). Divalent ions would be more likely to bind tightly to the EF-hand loop and at least one member of the lanthanides (trivalent) would be used as these metals are known aequorin activators (Le Clainche *et al.*, 2003). Calcium, as the natural activator and lanthanum (as representative of the lanthanide group) were obvious choices. Zinc and copper became primary choices for study as they are associated with brain diseases which have become of ever increasing interest in the recent years (Barkalifa *et al.*, 2009, Faller, 2009, Gu *et al.*, 2009, Marino *et al.*, 2010). Cadmium prefers coordination geometry similar to that of calcium and has been used as a substitute for calcium in crystallographic studies, thus it would be interesting to observe its effect on aequorin mutants. Cobalt and lead were chosen on the basis that aequorin potentially showed some positive response to these ions. Iron (valence II and III) was excluded based on Table 1–5 and magnesium was excluded as it is known to stabilise aequorin and prevent calcium-independent luminescence (Ohashi *et al.*, 2005) which means it is bound in the EF-hands without triggering the light emitting reaction.

1.7 Aims and organisation of thesis

The main aim of this project was to create a novel metal ion sensor through protein engineering, using wild-type aequorin as template. The envisaged molecule would have new specificity for a divalent ion other than calcium but otherwise retain its ion-triggered light-emitting reaction. Hence, it would have the potential to offer as much insight into the role of different metal ions as aequorin has for intracellular calcium. In order to pursue the main aim, the work in this project was broken down into several stages, addressed in Chapters 3 to 8.

Subcloning of the apoaequorin gene into a suitable expression vector is described in Chapter 3. Chapter 4 discusses the development of a high-throughput screen capable of identifying the desired activity in mutant microwell cultures. Protein engineering can lead to dramatic changes (positive or deleterious in the majority of cases) or small, gradual steps towards a desired protein function. The sensitivity of such a screen must be maximised in order to identify slight changes within the mutant library. Steps taken for this purpose included: (1) selecting between expression vectors for the best candidate; (2) selecting the *E. coli* strain for the highest expression of the mutant library; (3) developing an effective procedure for reconstitution of the holoprotein mutants; (4) producing the highest possible activity levels from the functional protein variants; (5) extracting the maximum value from the screen.

The structure of aequorin in relation to calcium binding and light emission is discussed in Chapter 1 and additional insight gained from analytical structural analysis of aequorin (published and author's own) and published mutational studies of aequorin and other calcium-binding formations are presented in Chapter 5.

The mutational strategy is discussed in Chapter 5. Even though both rational and directed evolution methods, and various combinations of both are promising, an EF-loop targeted mutational strategy was chosen as it was considered more likely to drastically change the ion selectivity within the time constraints of the project. The focus is on the individual residues in the calcium coordinating positions of each EF-hand; they were substituted with cysteine and histidine, which are the prevalent residues in copper and zinc-binding sites. More mutations (alanine substitutions and impairment of EF-hand flexibility) were introduced to study the importance of specific amino acids or individual EF-hands. A total of forty eight mutants were designed.

Implementation of the high-throughput screening assay on the aequorin library is described in Chapter 6, where the complete library of mutants was tested against seven different metal ions. The screening results were analysed in detail and a few selected mutants were chosen as candidates for further analysis. Chapter 7 describes the preparation of high purity wild-type aequorin and one of the selected mutants (Asn28Cys/Ser32His) for testing against the set of the seven ions and Chapter 8 studies the response of the purified molecules to the same set of ions without the potential interferences from bacterial culture components of the crude screen. Conclusions drawn and future work arising from this research are presented in Chapter 9.

2 Materials and Methods

This chapter is organised as follows: Standard buffers and media, standard materials and methods in cell culture, DNA, protein and metal ion solution preparations. Where modifications of any of the procedures were necessary for method development or sample preparation, those will be described in the respective sections of Chapters 3-8. Automation-based methods (liquid handling and colony picking) are described in Chapter 4.

2.1 Preparation of buffers and media

All chemicals used were of Molecular Biology grade and purchased from Sigma-Aldrich Company Ltd, unless stated otherwise. Buffers and media for molecular biology were prepared according to (Sambrook *et al.*, 1989). Standard molecular biology protocols for glycerol stocks, overnight cultures, and streaked plates were from the same reference, unless otherwise stated. All solutions and buffers were prepared with R.O. water from an Elix Millipore purification system.

2.1.1 LB medium

LB (Luria Bertani) medium was prepared by dissolving 10 g tryptone, 5 g yeast extract and 10 g NaCl in 1 L of deionized water. The pH of LB medium was adjusted to 7.0 with concentrated NaOH solution and the medium was autoclaved at 121 °C and 1.2 atm for 20 min.

2.1.2 LB agar

LB agar was prepared by adding 2% w/v bacteriological agar in LB medium. It was used to fill Petri dishes (agar plates) for growth of solid cultures. The mix was sterilised by autoclaving as above. Petri dishes were partially filled with warm liquid LB agar containing added antibiotics.

2.1.3 YT agar

YT (yeast extract and tryptone) agar is a nutrient-rich medium designed for growth of recombinant strains of *E. coli* on agar plates. YT was prepared by dissolving 8 g tryptone, 5 g yeast extract and 5 g NaCl and 15 g of bacteriological agar in 1 L of deionised water, with adjustment to pH 7.0 and autoclaving as above.

2.1.4 SOC medium

SOC (Super Optimal broth with catabolite suppression) medium is a nutrient-rich bacterial growth medium which results in higher transformation efficiencies of plasmids. SOC medium was prepared by dissolving 20 g tryptone, 5 g yeast extract, 0.58 g NaCl and 0.19 g KCl in 1 L of deionized water and pH was adjusted to 7.0 with concentrated NaOH solution. After autoclaving, the following filter-sterilised solutions were added: 20 mM glucose and 10mM MgCl₂.

2.1.5 Kanamycin

Kanamycin stock solutions were prepared at 10 mg/mL in RO water, sterilised by filtration through a 0.2 µM Minisart[®] filter (Sartorius Stedim Biotech), aliquoted into 1.5 mL eppendorfs and stored at -20 °C. Working concentration of kanamycin in liquid cultures and agar plates was 50 µg/mL. Preparations containing 50 µg/mL of kanamycin are denoted Kan+ in the text.

2.1.6 Ampicillin

Ampicillin stock solutions were prepared at 10 mg/mL in RO water, sterilised by filtration through a 0.2 micron Minisart[®] filter and aliquoted into 1.5 mL eppendorfs and stored at -20°C. Working concentration of ampicillin in liquid cultures and agar plates was 150 µg/mL. Preparations containing 150 µg/mL of ampicillin are denoted Amp+ in the text.

2.1.7 X-gal

X-gal (5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside) is used to indicate whether a cell expresses the β -galactosidase enzyme, which is encoded by the lacZ gene, in a technique called blue/white screening. If X-gal and an inducer of β -galactosidase (usually IPTG) are added in a bacterial culture, colonies carrying a plasmid which encodes for the lacZ gene are able to produce the enzyme β -galactosidase which can then cleave the X-gal present within the nutrient agar, resulting in a blue colony. Bacteria which are not transformed with a plasmid encoding for the lacZ gene or where the lacZ gene has been interrupted by a cloned DNA sequence remain white. Blue/white screening was not applicable in the case of *E. coli* strain BL21(DE3)1SHOT which encodes for lacZ in its genomic DNA.

X-gal stock solution was prepared at 20 mg/ml by dissolving into DMSO (dimethyl sulfoxide). Tubes containing the stock solution were wrapped in foil for protection against light and stored at $-20\text{ }^{\circ}\text{C}$. X-gal solutions do not require sterilisation (Sambrook *et al.*, 1989). X-gal was either spread on top of pre-made agar plates (40 μL of stock solution) or poured into melted agar at $\sim 55\text{ }^{\circ}\text{C}$ (final concentration 40 $\mu\text{L}/\text{mL}$). IPTG was added to 0.1 mM final concentration in the media.

2.1.8 Tris-HCl buffer

1 M stock solutions of Tris-HCl (tris(hydroxymethyl)aminomethane hydrochloride) of pH 7.2, 7.5 and 7.8, molecular biology grade, were purchased from Sigma-Aldrich. This ensured the lowest possible (and measured) contamination from various metal ions which may activate aequorin or aequorin variants.

2.1.9 EDTA

0.5 M EDTA (ethylenediamine tetraacetic acid) stock solution at pH 8.0 was purchased from Sigma-Aldrich at molecular biology grade. EDTA solution is prepared by dissolving the appropriate mass of EDTA disodium salt in deionized water and adjusting the pH to 8.0 with NaOH concentrated solution.

2.2 Cell culture

2.2.1 *E. coli* strains

2.2.1.1 TOP10 electrocompetent

E. coli TOP10 electrocompetent cells (Invitrogen) were used for routine cloning: transformation and storage of plasmids, production of plasmid DNA. Their genotype is: F⁻ mcrA Δ (mrr-hsdRMS-mcrBC) Φ 80lacZ Δ M15 Δ lacX74 recA1 araD139 Δ (ara leu) 7697 galU galK rpsL (StrR) endA1 nupG.

2.2.1.2 BL21Star™(DE3)

BL21Star™(DE3) electrocompetent cells from Invitrogen and Lucigen were used in this work. *E. coli* BL21Star™ (DE3) electrocompetent cells (Invitrogen) were used in protein expression applications. This strain is suitable for production from target genes cloned into T7 driven expression vectors and deficient in RNaseE (me131). The Invitrogen BL21Star™ (DE3) genotype is: F⁻ ompT hsdSB(rB⁻, mB⁻) gal dcm rne131 (DE3).

BL21(DE3) *E. cloni*™ electrocompetent cells from Lucigen were also used in protein expression applications. This strain exhibits high transformation efficiency and is suitable for production from target genes cloned into T7 driven expression vectors and deficient in the *lon* and *ompT* proteases. The Lucigen BL21(DE3) *E. cloni*™ genotype is: F⁻ mcrA Δ (mrr-hsdRMS-mcrBC) Φ 80dlacZ Δ M15 Δ lacX74 endA1 recA1araD139 Δ (ara, leu)7697 galU galK rpsL nupG λ - tonA.

2.2.2 Streaked plates

Streaking a plate allows the bacteria to be spread out so that a colony from a single bacterium can be isolated from a culture sample. Cultures were streaked out on Petri dishes containing Kan⁺ (or Amp⁺ where appropriate) LB agar. Each plate was incubated at 37 °C overnight and stored at 4-6 °C for approximately one week.

2.2.3 Overnight cultures

A single colony was picked from a Petri dish using a sterile loop or sterile toothpick into 5 mL of LB medium in a 50-mL Falcon tube. The tube was incubated overnight at 37 °C with 225 rpm agitation.

2.2.4 Shake flask cultures

Appropriate volume of the overnight culture was added to sterile LB medium in 250-, 500-mL or 2-L Erlenmeyer flask and incubated for 8 h at 37°C with 225 rpm agitation. For a 5% inoculation 5 mL of overnight culture was added to 95 mL fresh medium in a 500-mL flask.

2.2.5 Glycerol stocks

Glycerol stocks at 20% (v/v) were used for long term storage of DNA constructs. To create a glycerol stock, a single colony of each clone was picked from an agar plate and grown overnight as described in Section 2.2.3. Samples from the overnight culture and 40% v/v sterile glycerol solution were mixed in a one to one volume ratio and aliquots were stored at –80°C, in sterile eppendorf tubes or microplates.

2.2.6 Sonication

Sonication was performed in an MSE Soniprep 150 (Sanyo Europe Ltd.). One milliliter of culture was transferred into a 1.5-mL eppendorf tube which in turn was packed into a small beaker containing ice. The sonicator probe was placed in the culture and the cells were lysed on ice using a protocol of 10 s on/off for 10 cycles at 8 µm amplitude. Cell debris was removed by centrifugation at 1,600 rpm for 20 min.

2.3 DNA

2.3.1 Preparation of plasmid DNA

Isolation of plasmid DNA from *E. coli* cultures was performed using Qiaprep mini- and midi-preparation kits, purchased from Qiagen Ltd. and the procedure followed was according to the manual.

Purification of DNA products from polymerase chain reaction mixtures was performed using Qiaquick PCR purification kit and according to Qiagen's manual. DNA gel extraction from low melt agarose electrophoresis gels was performed using Qiaex II Gel Extraction kit from Qiagen Ltd.

2.3.2 PCR amplification

PCR reactions were performed on a Techgene Thermo Cycler or a Techne thermal cycler, when gradient PCR was required. Samples were loaded and prepared with DNase and RNase-free, low-retention tips and eppendorfs from Molecular Bioproducts Inc. *PfuTurbo*[®] DNA polymerase was from Stratagene Inc. and *ProofStart* DNA polymerase was purchased from Qiagen. The template method used is presented in Table 2-2; where any deviations were applicable, those will be presented in the relevant sections.

Table 2-1 PCR mix – general cloning

Component	Volume
10× Reaction Buffer	5 µl
5–50 ng of dsDNA template	X µl
125 ng of oligonucleotide primer #1	1.25 µl (stock 100ng/µl)
125 ng of oligonucleotide primer #2	1.25 µl (stock 100 g/µl)
dNTP mix	1 µl
ddH ₂ O	up to 50 µl
<i>PfuTurbo</i> DNA polymerase (2.5 U/µl)	1 µl

Table 2-2 PCR cycles – general cloning

PCR Step	Duration (min)	Temperature (°C)
Initial activation (required with HotStart enzymes, e.g. <i>ProofStart</i>)	5	95
3-step cycling, 35 cycles:		
Denaturation	1	94
Annealing	1	69*
Extension	1	72
Final extension	7	72
End of PCR cycling:	Indefinite	4

* a range of annealing temperatures was tested (50-72 °C). 69 °C were found to be the optimum for all amplifications of the apoaequorin gene (Section 3.3.3 and 3.4.2.2), while 60 °C were sufficient for the amplification of pET components (Section 3.4.2.2).

2.3.3 QuikChange[®] site-directed mutagenesis

QuikChange[®] II SDM (Stratagene Ltd.) is a rapid four-step procedure (Stratagene, 2004b). The basic procedure is shown in Figure 2-1. It utilises supercoiled dsDNA vector as template and two oligonucleotide primers containing the desired mutation. *PfuTurbo*[®] DNA polymerase is used to extend the primers and generate a mutated plasmid containing staggered nicks. The product is then treated with endonuclease *Dpn* I to digest the methylated parental DNA template and to select for mutation-containing synthesised DNA (non-methylated). The nicked vector DNA containing the desired mutations is transformed into electrocompetent cells. After transformation the nicks in the mutated plasmid are repaired by the host cell.

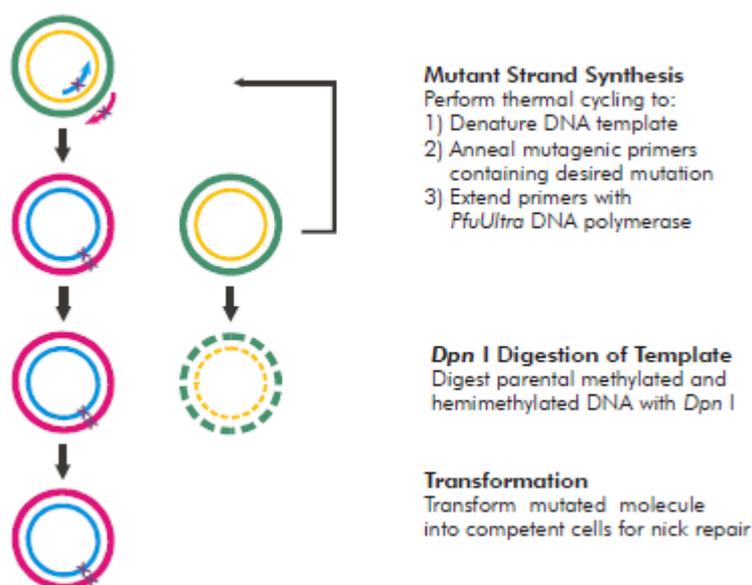


Figure 2-1 Schematic of QuikChange[®] II Site-Directed mutagenesis method. Image reproduced from Stratagene (2004b).

In this work QuikChange[®] Site-Directed Mutagenesis (SDM) protocol was used with only minor variations. *PfuUltra* High Fidelity DNA polymerase (used instead of *PfuTurbo* for its higher fidelity) was used to extend the primers and generate a mutated plasmid containing staggered nicks. Primers were approximately 45 bases in length, with a melting temperature of $\geq 78^{\circ}\text{C}$. A 25 μL QuikChange[®] reaction mix (Table 2–3) was submitted to a program of temperature cycling (Table 2–4) in a PCR thermal cycler (Techne Ltd.). Typically, a master mix would be created for the number of reactions carried out on the day and aliquoted out; DMSO, MgSO_4 , Qsolution or water would be added, where applicable, for a final volume of 25 μL . The QuikChange[®] reaction was digested with 0.6 μL (or 1.53 U) of *Dpn* I (Stratagene Ltd.) for one hour at 37°C , followed by a second addition of 0.6 μL *Dpn* I in the mix and incubation for an additional hour at 37°C . One microliter of the digestion was transformed into competent *E. coli* TOP10 electrocompetent cells (Invitrogen) by electroporation (Section 2.2.1.1 and 2.3.8). The transformed cells were spread on LB Kan+ agar. Plasmid DNA was prepared (Section 2.3.1) from overnight cultures (Section 2.2.3) of one colony per PCR reaction, checked for correct size on DNA agarose gels (Section 2.3.9) and mutations were verified through sequencing (Section 2.3.4).

Table 2-3 PCR mix – QuikChange®

Component	Volume (μL)
10x reaction buffer *	2.5
dNTP mix (10 mM of each)	1
Sense Primer (62.5 ng per 25 μl reaction)	Variable
Antisense Primer (62.5 ng per 25 μl reaction)	Variable
<i>Pfu</i> UltraHigh Fidelity (2.5 U/μL)	0.25
Template DNA (5 ng in 25 μl)	0.25
25 mM MgSO ₄	Optional
DMSO	Optional
Qsolution	Optional
RNAse-free H ₂ O	Top up to final volume
Total Volume (μL)	25

* 10× Reaction Buffer contains: 100 mM KCl, 100 mM(NH₄)₂SO₄, 200 mM Tris-HCl (pH 8.8), 20 mM MgSO₄, 1% Triton® X-100, 1 mg/ml nuclease-free bovine serum albumin (BSA). Qsolution: PCR additive (proprietary content) by Qiagen, for amplification of templates that are GC-rich or have extensive secondary structure. Additional concentrations in the final mix of the optional chemicals for each reaction are presented in Table 6 1.

Table 2-4 PCR cycles – QuikChange®

PCR Step	Duration (min)	Temperature (°C)
Initial activation (applies only to <i>ProofStart</i>)	30 s	95
3-step cycling, 35 cycles:	12–18	
Denaturation	30 s	94
Annealing	1 min	*
Extension	1 min/kb of plasmid length	68 or 72
Final extension	5 min	72
End of PCR cycling:	Indefinite	4

* a range of annealing temperatures was tested (50–69 °C). Final temperatures per individual reaction for all QuikChange® mutants in Table 6-1.

2.3.4 Sequencing

All DNA samples were sequenced at the Scientific Support Services of the Wolfson Institute for Biomedical Research, UCL. For the sequencing reactions the following amount of DNA was prepared: 100 ng/ μ L for plasmids, 1 ng/ μ L per 100 bp for PCR product, at 10 μ L of template per reaction. Primers were prepared at 2-5 pmoles/ μ L at 6 μ L per reaction.

2.3.5 Oligonucleotides for amplification and sequencing

Primer sequences were designed using Bioedit software. Calculation of annealing temperatures and testing of oligos for dimers and hairpin loops was performed on AnnHyb (Friard, 2004). All primers were from Operon Biotechnologies GmbH, with the exception of primers M13-20, M13-40 (supplied by Wolfson Institute) and primers T7 promoter, T7 terminator purchased from Novagen. Sequences of the primers used for PCR and sequencing reactions are shown in Table 3-1 and Table 3-3. Primers M13-For and M13-Rev were supplied in the TOPO[®] ligation kit (Invitrogen). Where restriction sites were to be introduced (refer to TOPO[®] cloning Section 3.4.2) the design of primers met the following criteria: (1) restriction sequences should be unique both in the template and in the destination plasmid; (2) corresponding restriction enzymes would leave cohesive ends, (3) restriction sites subject to Dcm, Dam and EcoR methylation were excluded and (4) all selected restriction enzymes are active in Universal Buffer by NEB, so that double digests can be performed.

2.3.6 Restriction digests

Restriction digests were performed to unravel circular, supercoiled and plasmid DNA dimer forms of pHAQ, pET26b and pETAQ and to shave off bases from PCR constructs described in Chapter 3. A typical mixture of restriction digest is shown in Table 2-5. All other digestion reactions performed were adapted to this protocol. Diagnostic digests were usually performed in 25 μ L of final volume, while preparative

digests were performed in a total volume of 50 μ L. *Xho* I, *Msc* I, *BamH* I, *Sma* I and *Nhe* I restriction enzymes were purchased from New England Biolabs Inc.

Table 2-5 Double digestion reaction mixture

Component	Diagnostic	Preparative
	Volume (μ L)	
H ₂ O	x	2 \times x
NEB Buffer 4	2.5	5
BSA solution	2.5	5
DNA template	y	2 \times y
R.E.1 (1 U/ μ L)	1	2
R.E.2 (2 U/ μ L)	0.5	1
Final volume	25	50

x: added water to a final volume of 25 μ L

y: appropriate volume containing 1000 ng of PCR2 or 500 ng of pET26b

R.E.1: Restriction enzyme 1 (e.g. *Msc* I)

R.E.2: Restriction enzyme 2 (e.g. *Xho* I)

2.3.7 Ligation

2.3.7.1 Standard ligation

UltraClone DNA ligation and transformation kit, containing *E. cloni*TM electrocompetent cells was purchased from Lucigen, USA, were used for the pET cloning approach (see Chapter 3).

The cut vector and insert were quantified by measurement of UV Absorbance. The appropriate amount of insert and vector was determined in order to satisfy the 3:1 molar ratio of insert to vector. The ligation reaction was prepared as seen in Table 2-6 and the components mixed by gentle pipetting. The reaction mixture was incubated at room temperature for 10 min and immediately after the ligation the mixture was incubated at 70 $^{\circ}$ C for 15 min to terminate the reaction.

Table 2-6 CloneSmart® ligation mixture

Component	Volume (µL)
Vector	x
Insert	y
CloneDirect 10xLigation Buffer	1
CloneSmart DNA ligase	1
H ₂ O	z
Final volume	10

x: appropriate volume containing 0.03 pmol vector

y: appropriate volume containing 0.1 pmol insert

z: added water for final volume of 10 µL

2.3.7.2 Ligation of TOPO® open vector

Zero Blunt® TOPO® vector kit with One Shot electrocompetent cells was purchased from Invitrogen and used for the TOPO® cloning approach (Chapter 3, Section 3.4.2).

Table 2-7 TOPO® capture vector ligation mixture

Component	Volume (µL)
Fresh PCR product (PCR4)	3
Dilute salt solution (provided in kit)	1*
H ₂ O	2
pCR® II-Blunt-TOPO®	1
Open TOPO® capture vector	1
Final volume	6

* recommended: 0.5–4 µL

The ligation mix consisted of 3 µL of blunt ended PCR product, dilute salt (provided in Kit), water and TOPO® open vector as shown in Table 2–7. The mix was incubated for 5 min at RT. The transformation procedure was: 2 µL of the ligation mixture were transferred into 50 µL of Top10 One Shot® electrocompetent *E. coli* that had been completely thawed on ice.

Falcon tubes with 5-10 mL Luria-Bertani media containing 100 mg/L of ampicillin were inoculated with transformants and placed in a shaking incubator at 37°C, 200 rpm agitation for 12 h. Cells that contain non-recombinant vector (no insert)

are killed upon plating as successful ligation would cause disruption of the lethal *E. coli* gene *ccdB* (Bernard *et al.*, 1994) (Invitrogen, 2004).

2.3.8 Electrocompetent cell transformation

Electroporation was performed on a Micro-Pulser Electroporator from Bio-Rad Laboratories, according to the manufacturer's instructions for the electroporation of bacterial cells. 0.1 cm electroporation cuvettes were from Bio-Rad Laboratories.

*E. cloni*TM electrocompetent cells (Lucigen) were left to thaw on wet ice and aliquoted in pre-chilled eppendorf tubes (25 µL of cells per tube). 1 µL of the heat denatured reaction mixture was added to the cells and stirred briefly using a sterile pipette tip. The cell/DNA mixture was pipetted into pre-chilled electroporation cuvettes. The cuvette was flicked downward so as to deposit the cells across the bottom of the well, followed by electroporation. Within 10 s of the pulse, 980 µL of SOC medium were added to the cuvette and cells were resuspended by gentle pipetting. The mixture was transferred to a 15 mL centrifuge tube and incubated at 250 rpm and 37 °C for 1 hour to allow expression of the antibiotic resistance genes. The transformed cells were spread on agar plates containing 50 µg/mL kanamycin and plates were incubated at 37°C overnight. This culture was used for further growth of the transformed clones for preparation of glycerol stocks and plasmid extraction with Qiagen miniprep kit.

In a similar manner, One Shot TOP10[®] ElectrocompTM electrocompetent cells were transformed according to the Invitrogen Instruction Manual (Invitrogen, 2004).

2.3.9 DNA agarose gel electrophoresis

Agarose gel preparation and casting was according to (Sambrook *et al.*, 1989). The most commonly used buffers for electrophoresis of DNA are TAE (Tris-acetate-EDTA) and TBE (Tris-borate-EDTA). A concentrated (10x) stock solution of TBE was supplied by Bio-Rad Laboratories. Higher concentrations of agarose allow separation of lower molecular weight DNA molecules, while low agarose concentrations allow resolution of higher molecular weight DNA molecules. Table 2–8 presents the agarose gel concentration used for various DNA fragment sizes.

Table 2-8 Agarose gel concentrations

DNA fragment size (Kb)	% agarose gel
1–30	0.5
0.8–12	0.75
0.5–10	1.0
0.4–7	1.25
0.2–3	1.5
0.01–0.5	2–5

Agarose and low-melt agarose were purchased from Sigma-Aldrich. Agarose was used for visualisation and low-melt agarose for preparatory electroporation (followed by gel extraction).

Typically 0.8% or 1% were used for visualisation of plasmids and 1.2% gels were used for visualisation of PCR constructs. The wells of the gel lanes were loaded with 3 μ L or 6 μ L of DNA samples mixed with 6x DNA loading buffer (Novagen). The wells of the lanes containing DNA ladders were loaded with 5 μ L Perfect 1kb DNA ladder (0.5–12 Kb) or PCR markers (50 bp–2 Kb), as required, both purchased from Novagen. Visualisation of DNA electrophoresis gels was performed on a Gel Doc 2000 from Bio-Rad Laboratories.

2.3.10 Absorbance measurement of DNA

Measurement of DNA absorbance was performed on a UV2 spectrophotometer by Unicam Ltd, in quartz cuvettes by Sigma-Aldrich Company Ltd. Absorbance of DNA was measured at 260 nm and 280 nm after appropriate dilutions to absorbance values <1.0 AU (Absorbance Unit).

2.4 Protein

2.4.1 Determination of protein

2.4.1.1 Determination by UV absorbance

Protein concentration of purified protein samples (used in Chapter 8) was determined by measuring the absorbance (A) at 280 nm and converting to concentration through the Beer-Lambert law:

$$A = E_{1cm}^{0.1\%} \cdot l \cdot C \quad \text{Equation 2.1}$$

where $E_{1cm}^{0.1\%}$ is the mass extinction coefficient for a 0.1% or 1 mg/mL protein solution, l is the pathlength (cm) and C is the protein concentration (mg/mL). This is an excellent method for measuring protein concentrations provided that an accurate value of $E_{1cm}^{0.1\%}$ is available and that the sample is of high purity. A280 absorbance measurements of purified protein were performed on a Nanodrop (Thermo Scientific). The value for molar extinction coefficient of aequorin was derived from previously published literature $E_{1cm}^{0.1\%}$ at 280 nm = 3.0 (Shimomura *et al.*, 1990) and verified from 280 nm absorbance measurements of pure, commercially available aequorin (Aqualite[®] Molecular probes).

An established and well-trusted method for estimating molar extinction coefficient ϵ for proteins is the Edelhoch method (Pace *et al.*, 1995), based on the protein's content of tryptophan, tyrosine and cystine (disulfide bonds). An extinction coefficient of aequorin could not be calculated using the Edelhoch method as it that designed for polypeptide chains; aequorin is comprised of a polypeptide chain and a prosthetic chromophore and the extinction coefficient of aequorin is not the product of a simple addition of the extinction coefficients of the participating molecules.

2.4.1.2 Determination by Bradford assay

For the determination of total protein concentration the Bradford assay was used (Bradford, 1976). Bradford is a colorimetric protein assay based on an absorbance shift of the dye Coomassie Brilliant Blue G-250 under acid conditions when a redder form of the dye is converted into a bluer form on binding to protein. Binding of the protein

stabilises the blue form of Coomassie dye, thus the amount of complex present in solution is a measure for the protein concentration by use of an absorbance reading at 595 nm. Protein standard solutions must be prepared in the same buffer as the samples to be assayed.

Coomassie Brilliant Blue G-250 stock solution was purchased from Bio-Rad and used in an appropriate dilution. A standard curve was made using bovine serum albumin (BSA) from Sigma-Aldrich. When the standard procedure was used, 1 mL of each sample or BSA standard solution was mixed with 1 mL of Bradford reagent diluted 1:2, and absorbance at 595 nm was recorded. BSA standard solutions ranged from 0.015 to 0.125 mg/ml. When performing the assay in microtiter plates, 10 μ L of sample or BSA standard solution were mixed with 200 μ L of Bradford reagent diluted 1:5 and protein standards were ranged from 0.05 mg/ml to 0.5 mg/ml. In the Nanodrop determination 10 μ L of sample or standard solution were mixed with 10 μ L of 1:2 diluted Bradford reagent and protein standards were 0.015–0.125 mg/ml.

2.4.2 Protein size – SDS-PAGE

Electrophoresis of protein samples was carried out using 12% and 15% w/v polyacrylamide gels. Gels were purchased from Bio-Rad while SDS running buffer was prepared as in (Sambrook *et al.*, 1989). Each sample was diluted 1:1 (or at other appropriate ratio) with SDS loading buffer, purchased from Bio-Rad. Samples were heated at 95 °C for 5 min prior to electrophoresis. To determine the size of a sample protein, a series of standards (proteins of known molecular weight) were loaded in adjacent lanes of the gel.

Precision Plus Protein™ Standards, containing ten recombinant protein bands (10 kDa–250 kDa) pre-stained with Coomassie Blue were purchased from Bio-Rad Laboratories, Inc. These standards contain three bands (25 kDa, 50 kDa and 75 kDa) which are three times as intense as the other bands and pre-quantified protein bands which can be used for a rough approximation of a protein of interest.

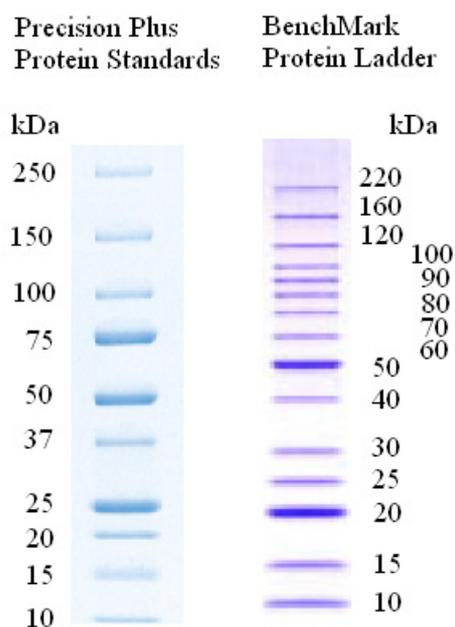


Figure 2-2 Protein ladders used in SDS protein electrophoresis.

2.4.2.1 Protein quantity by SDS-PAGE

The Precision Plus Protein™ Standards contain three bands (25 kDa, 50 kDa and 75 kDa) which are three times as intense as the other bands and pre-quantified protein bands which can be used for a rough approximation of a protein of interest by densitometry. This method was used only as a rough guide and not relied upon for accurate quantification of protein.

2.4.2.2 Protein sample purity by SDS-PAGE

Pre-cast Ready Gel® Tris-HCl, 12% and 15% polyacrylamide gels for protein electrophoresis were purchased from Bio-Rad Laboratories. The SDS running buffer was 50 mM Tris-HCl, 0.38 M glycine, 0.1% (w/v) SDS in RO water, adjusted to pH 8.8. The staining solution was 0.05% (w/v) Coomassie Brilliant Blue, 50% (v/v) methanol, and 10% (v/v) acetic acid in RO water. Protein ladders used were Precision Plus Standards (Bio-Rad Laboratories) and Benchmark Protein Ladder (Invitrogen) (Section 2.4.2.1, Figure 2-2).

2.5 Metal ion concentration

2.5.1 Ion solution preparation

All metal ion salts and Tris-HCl stocks were purchased from Sigma Aldrich. Stock solutions were created by weighing chloride ion salts (nitrate salt in the case of lead) into polypropylene volumetric flasks, suitable for atomic absorption assays. All metal salts were diluted in 50 mM Tris-HCl pH 7.5 buffer with the exception of lead nitrate, which precipitates in Tris-HCl and R.O. water was used instead. Serial dilutions were performed to obtain solutions in a wide range of nominal concentrations down to a few micromolar (μM). To correct for possible dilution errors and salt precipitation of the stock solutions, the solutions were tested for ion concentration with atomic absorption (Section 2.5.2) and the nominal metal ion concentrations were corrected where necessary. The corrected ion concentration values were used in Chapters 7 and 8. Blank buffers containing 50 mM Tris-HCl 10 μM EDTA pH 7.5 were also tested for total calcium contamination which was found to be 2.1×10^{-7} M.

Table 2-9 Metals in the form of hydrated salts.

Salt		Purity
Calcium chloride hexahydrate	$\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$	99+%
Cobalt chloride hexahydrate	$\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$	99-102%
Cadmium chloride	CdCl_2	99.99+%
Copper sulfate pentahydrate	$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	99.99+%
Lanthanum chloride heptahydrate	$\text{LaCl}_3 \cdot 7\text{H}_2\text{O}$	99.99+%
Zinc chloride	ZnCl_2	99.999+%
Lead nitrate	$\text{Pb}(\text{NO}_3)_2$	99.99+%

2.5.2 Atomic absorption

Atomic absorption measurements were conducted by ICP-AES (Inductively Coupled Plasma Atomic Emission Spectrometry) on a Varian VISTA PRO by the Natural History Museum Analytical imaging facilities department. This method enables the detection of major, minor and trace elements in sample solutions. Samples are dispersed into a stream of argon gas through a nebuliser and carried to an ICP. As the sample atoms are excited in the ICP they lose electrons. As they regain their lost electrons the sample atoms emit photons of light, with wavelengths characteristic of the elements present. A spectrometer is used to separate the light emitted into the various wavelengths and these are detected and recorded simultaneously using a solid-state detector.

2.6 Automation

2.6.1 Automated colony picking

A Qpix2 robot (Genetix Ltd.) was used to pick colonies of *E. coli*. The robot was programmed to inoculate every well of a 96-DSW (deep square well) plate with colonies of *E. coli* from petri dishes. Each plate was sealed by adjusting firmly an inverted shallow 96-well plate (Sarstedt Inc) and taping it over the top of the DSW plate. The sealed plates were incubated for 16 hours at various agitation speeds on a Variomag Teleshake unit (Camlab Ltd.) at 37 °C.

2.6.2 Automated liquid handling

Automated liquid dispensing into microplates was carried out using a TECAN Genesis (TECAN Ltd, Reading, UK) robotic platform fitted with an 8-channel multipipette, sterile disposable filtered tips within a clean cabinet pre-sterilised using one in-built and one floor-mounted UV lamp.

2.7 Luminescence measurements

2.7.1 Platereader

The light measurements were performed using the FluoStar Optima microplate reader (BMG Lab technologies Ltd, Bucks, UK), a luminescence, fluorescence and absorbance reader equipped with two syringe pumps for reagent injection. The luminescence mode was used (no excitation as it is not required in bioluminescence). Emission of light was obtained through the emission lens (no filters were used) in order to collect light from the all the visible spectrum. This option offers the advantage of not missing light emitted in a shifted wavelength, in case this may occur with certain protein variants. The BMG Fluostar Optima has a limit of detection lower than 50 amol/well ATP, spectral range between 240 and 740 nm and a dynamic range of nine decades. The instrument was checked annually for recalibration using radioactive luminescence standards by BMG Lab Technologies.

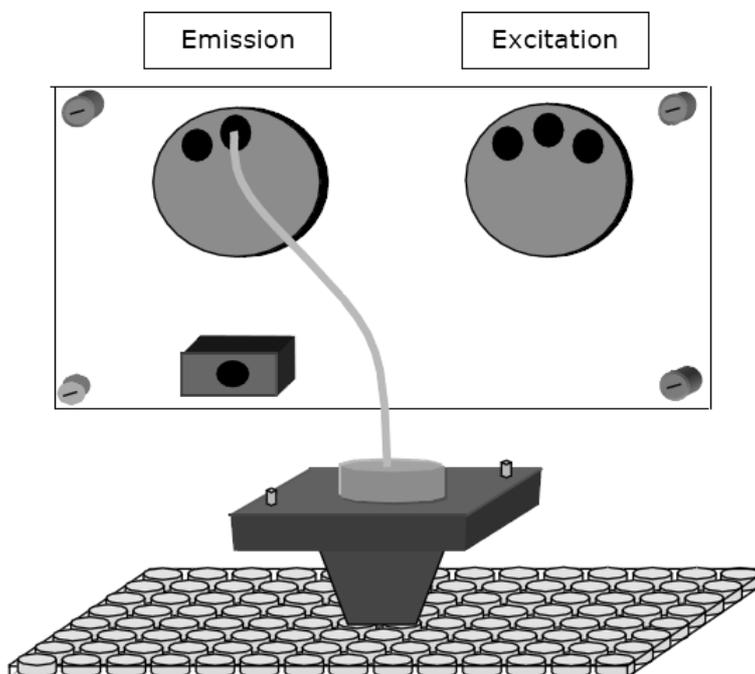


Figure 2-3 Platereader layout – luminescence mode.

2.7.1.1 Visualisation of results

The typical flash luminescence activity of aequorin vs time is shown in Figure 1.4. The activity is usually assessed by measuring the peak (and less often by measuring the total amount) of light derived from the flash luminescence curve. The platereader software enables simultaneous visualisation of the luminescence curves of all of the wells of the microplate, thus allowing direct comparison between them. In most cases this visual output option was adequate to inform on decisions such as which operating parameter was optimal or whether an operating condition met the purposes of the assay.

3 Cloning of apoaequorin gene

This chapter describes the efforts to create an expression vector which facilitates production of apoaequorin in a way that allows for easy reconstitution with the heterocyclic compound coelenterazine in high-throughput assays and also for the purification of the protein (apo- or holo-form of aequorin). Two separate cloning routes are described. For clarity, this chapter in particular is divided in four main sections: In Introduction (Section 3.1) the envisaged expression vector and possible cloning routes are identified. Following this, Section 3.2 describes the verification and sequence of the apoaequorin gene in its original plasmid vector. Each of the two subcloning routes is presented separately (Section 3.3 and Section 3.4) with methods, results and discussion. The outcome and expectations for the competing approaches will be discussed in Conclusions (Section 3.5).

3.1 Introduction

3.1.1 Envisaged expression vector

A suitable expression vector would contain the following features:

- (1) gene encoding for apoaequorin
- (2) promoter site and ribosome-binding site (RBS) for transcription and translation
- (3) signal peptide sequence, e.g. *pelB* leader sequence: for localisation of the expressed apoaequorin in the periplasmic space of the cell, where it can be easily reconstituted into fully charged holoprotein complex by incubation with coelenterazine
- (4) 6xHis-tag encoding sequence for potential metal affinity purification in Ni-NTA columns and/or Ni-NTA coated microwell plates
- 5) unique restriction sites flanking the apoaequorin gene, for the future option of “lifting” the construct from the vector with restriction digest, in order to ligate into another vector.

3.1.2 Cloning strategies

In order to ensure successful cloning of the apoaequorin gene in the construct described, two subcloning routes were pursued:

The first route (pET cloning route) aimed to transfer the apoaequorin gene into the pET26b commercial expression vector by double digestion and ligation.

The second route (TOPO[®] cloning route) aimed to create the desired construct by using a consecutive series of PCR amplifications, resulting in a blunt ended construct that could be ligated into an open TOPO[®] vector.

The first route (pET cloning route) not immediately successful; the second route (TOPO[®] cloning) was pursued in parallel, as an alternative. In the end both routes yielded successful results at the same time.

3.2 Apoaequorin in original capture vector

This section describes studies conducted to verify the integrity of the apoaequorin gene which would be used as a template for further cloning. All general materials and methods used in this Section and the two Sections thereafter are described in Chapter 2, Section 2.1. Cell culture methods and *E. coli* strains are described in Section 2.2 and DNA handling methods are described in Section 2.3. Modifications and optimisations to the methods described in Chapter 2 are presented in the relevant sections. The overall implementation of methods and verification of outcomes of both cloning routes was based on a general framework which is described in Figure 3–1.

The gene encoding for apoaequorin was a kind gift from Professor Trewavas of the Institute of Cell and Molecular Biology, University of Edinburgh. The gene came inserted in plasmid pHAQ (Figure 3–2). pHAQ is a pBluescript vector (Figure 3–3), with the apoaequorin gene ligated between *Sal* I and *Pst* I restriction sites. In the maps of all vectors shown, the orientation of each gene is represented by the direction of its arrow. pBluescript is a high copy number plasmid vector which was designed for DNA cloning, DNA sequencing, in vitro mutagenesis and in vitro transcription (Stratagene, 2004a) as shown on the relevant plasmid maps it lacks features encoding for periplasmic localisation and metal affinity.

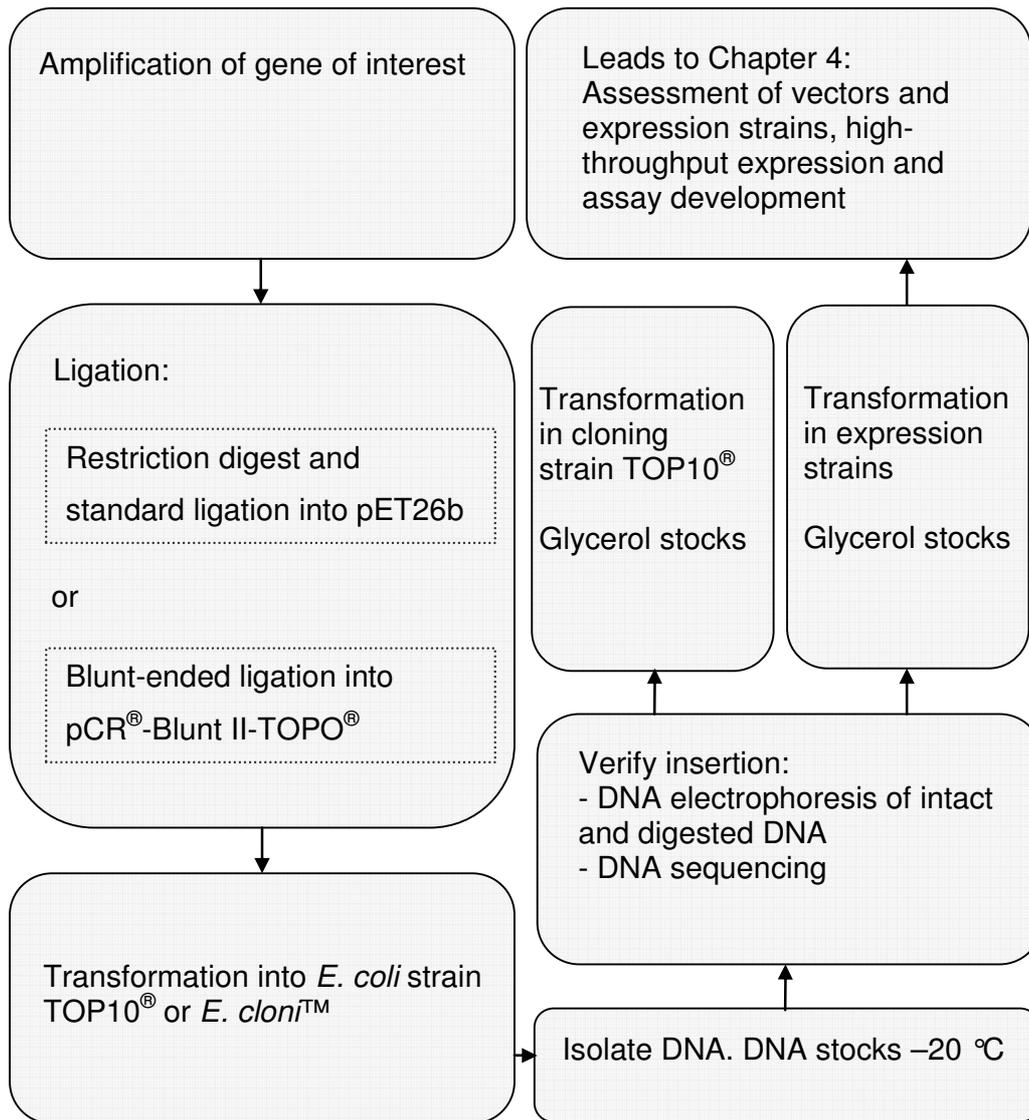


Figure 3-1 Overview of workflow employed for cloning in this work.

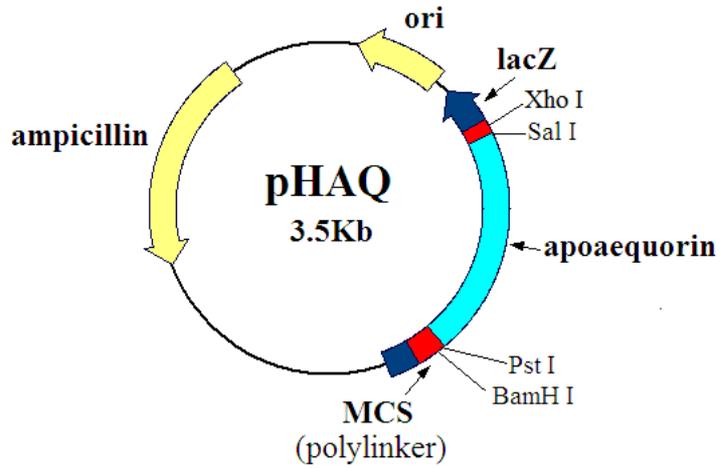
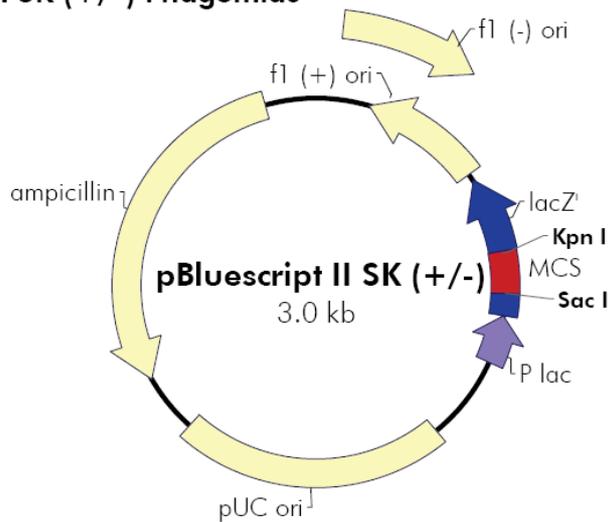


Figure 3-2 pHAQ vector: pBluescript backbone with apoaequorin gene inserted in the polylinker region between *Sal* I and *Pst* I recognition sites.

pBluescript II SK (+/-) Phagemids



**pBluescript II SK (+/-) Multiple Cloning Site Region
(sequence shown 598–826)**

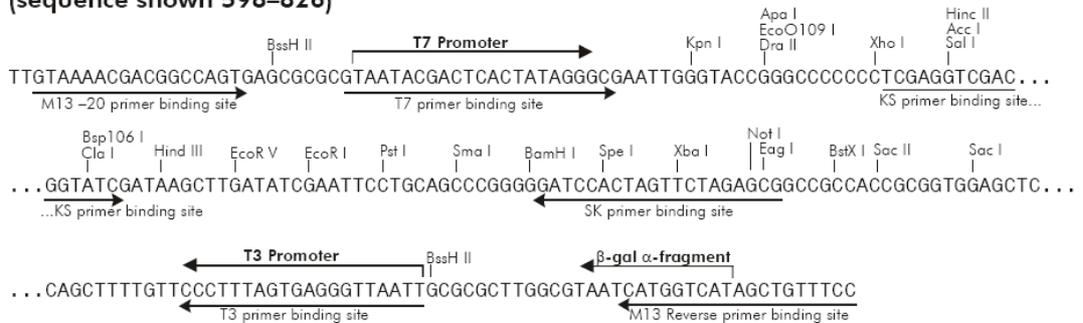


Figure 3-3 pBluescript vector. Image reproduced from Stratagene.com (Stratagene, 2004a). Apoaequorin gene was cloned into the MCS, between *Sal* I and *Pst* I restriction sites.

3.2.1 Verification of apoaequorin gene in pHAQ vector

Electrocompetent *E. coli* TOP10 were transformed with pHAQ vector and the transformed cells were plated on LB agar-kanamycin plates. Colony selection, cultivation in LB media, centrifugation of the culture and plasmid DNA purification were performed as described in Sections 2.2 and 2.3. A double digest with restriction enzymes *BamH* I and *Xho* I was performed on the purified plasmid, followed by gel electrophoresis separation (1.2 % agarose gel). This was done to verify that the insert of the expected size was indeed within the specific restriction sites. The *Xho* I site is 6 bp upstream of the *Sal* I site and the *BamH* I site is 15 bp downstream of the *Pst* I site (Figure 3-2, Figure 3-3). Visualisation of the gel showed two fragments of the expected sizes in Figure 3-4.

The unrestricted plasmid was sequenced with standard T3 and T7 promoter primers which flanked the apoaequorin gene. Using Bioedit (Hall, 2004), the resulting sequence was aligned with the sequence of apoaequorin retrieved from Genbank, Accession Number M16104 (Benson *et al.*, 2005). The alignment is presented in Figure 3-8 amongst the sequencing results from pET system cloning.

Figure 3-8 shows the sequence of the apoaequorin gene as provided in pHAQ (sequence name: denoted: apoaequorin_gene) aligned with the M16103 apoaequorin gene sequence in Genbank, Accession Number M16104 (Benson *et al.*, 2005). One base difference is observed between the two sequences: position 424 TCTGAT in M16103 versus TCTGCT in the pHAQ sequence translates to a D (Asp) in M16103 instead of an Ala in the sequence used in this work. In Chapter 5 the translated protein sequence of the apoaequorin gene (used this work) is aligned for comparison with the translated M16103 and the published crystallographic PDB structures of the apoaequorin chains 1EJ3 (Head J.F., 2000) and 1SL8 (Deng *et al.*, 2005).

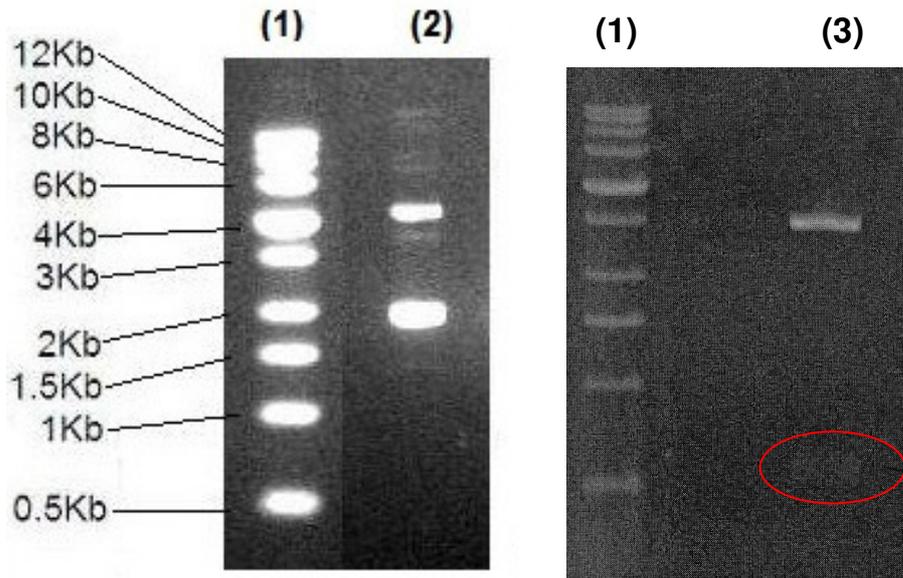


Figure 3-4 Plasmid vector pHAQ: agarose gel visualisation. 1.2% agarose (left) and 2.5% low melt agarose (right) gel visualisation of: (1) 0.5-12 Kb DNA ladder, (2) open circular, linear and supercoiled forms of pHAQ uncut plasmid 3,506 bp, (3) pHAQ double digest with *BamH* I and *Xho* I restriction enzymes. The faint lower band at 609 bp corresponds to apoaequorin gene plus a few flanking base pairs and upper band at ~2,900 bp corresponds to the remaining vector.

3.3 pET cloning route

The first cloning approach towards obtaining the apoaequorin gene within a desired expression vector (Section 3.1.1) was a “traditional” approach: the gene of interest would be cut from its carrier plasmid using the appropriate restriction enzymes (if the carrier did not incorporate the expression features required and in the correct orientation). The destination vector, one which includes the desired features and in the correct orientation would also be cut using the same restriction enzymes so that its multiple cloning site (MCS) has flanking restricted ends complementary to the ones of the restricted DNA fragment of interest. The cut vector and the cut gene of interest (insert) are ligated in order to form a complete expression vector and the ligated product is transformed into competent cells. The schematic of the pET cloning concept and its details is described in Figure 3-6.

3.3.2 Schematic of pET cloning route

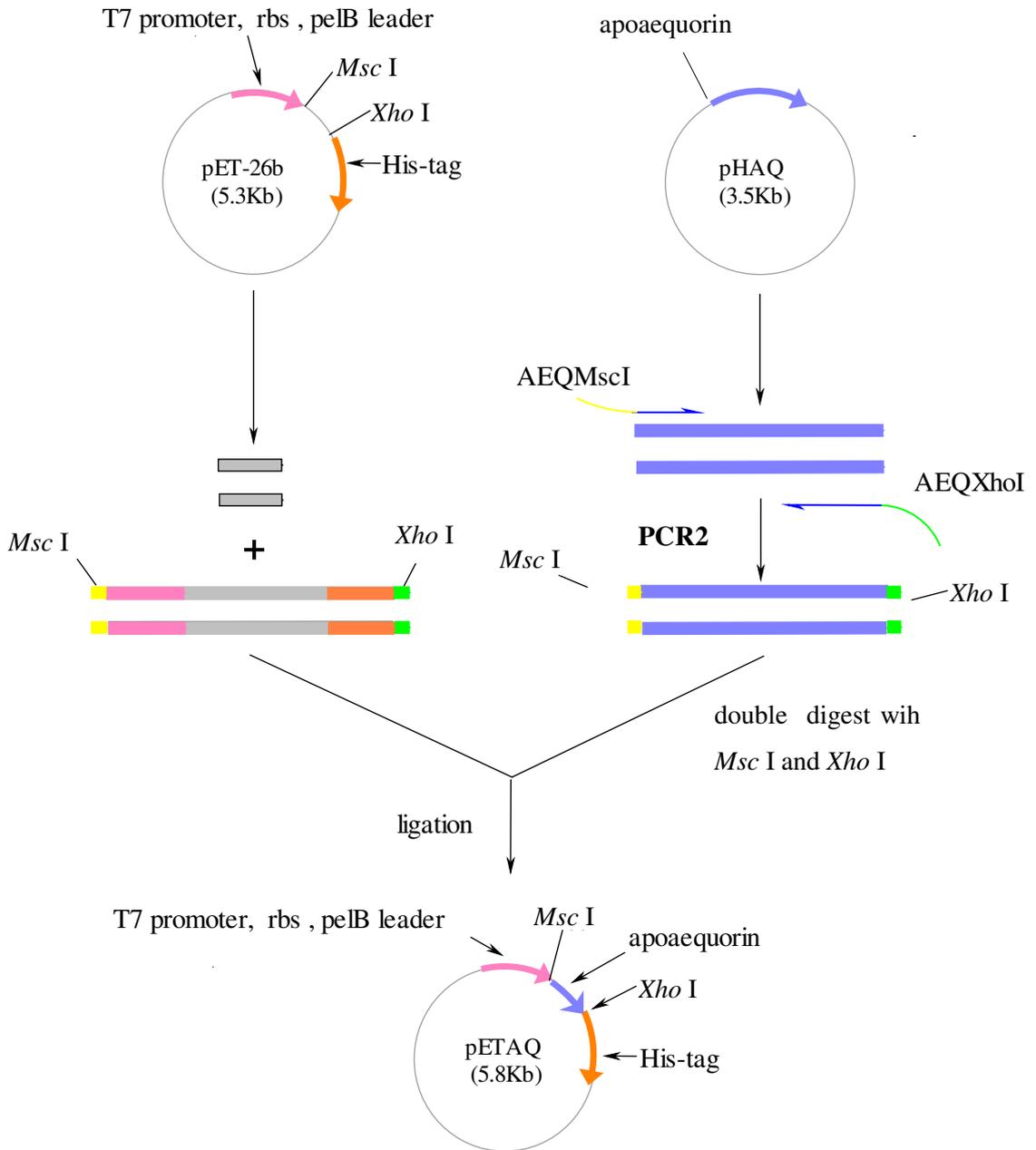


Figure 3-6 pET26b cloning route schematic

3.3.3 Methods for pET cloning

Primer sequences used in this Section and their melting temperatures are shown in Table 3–1. Primers AEQMSCI and AEQXHOI were designed to introduce the restriction endonuclease site for *Msc* I and *Xho* I upstream and downstream of the apoaequorin sequence, respectively. The two primers were used for the amplification of apoaequorin gene from pHAQ while obtaining overhangs for *Msc* I upstream and *Xho* I downstream of the gene. Primers For10+26, For252+20, Rev450-20 were designed for sequencing purposes. Their names indicate their position and direction in the apoaequorin gene, for example For10+26 is a forward (sense) primer which starts at position 10 of the apoaequorin gene and is 26 bases long.

Amplification PCR2 was carried out in 50 μ L final volume and according to the general methods described in Table 2-2 and 3-1, Section 2.3.2.

Vector pET26b and the product of reaction PCR2 were digested with restriction enzymes *Msc* I and *Xho* I. The digested DNA was electrophorised on low-melt agarose gels and the bands of desired size were extracted and purified using Qiaex II DNA extraction kit by Qiagen. Vector (Figure 3-7, lane 3) and insert (Figure 3-7, lane 5) were ligated at 3:1 molar ratio using CloneSmart Ligase (Lucigen, USA) as described in Section 2.3.7.1. Successful ligation using the Lucigen kit followed a series of unsuccessful ligations using T4 DNA ligase at 3:1 and 5:1 insert to vector ratio. The product of ligation was used to transform electrocompetent *E.coloni*TM cells (Table 2–5). Plasmid DNA was isolated from overnight cultures and sequenced using the following primers: T7 Promoter and T7 Terminator (Novagen), For10+26, For252+20 and Rev450-20 (Operon).

Table 3-1 Oligonucleotides for pET26b cloning route and sequencing

Primer		Sequence	T _m (°C)
AEQMSCI	sense	5'CGGCGATGGCCATGACCAGCGAACAATACTCAGTCA ³	82.0
AEQXHOI	antisense	5'GTGGTGTCTCGAGGGGGACAGCTCCACCGTAGAGC ³	84.8
T7 promoter	sense	5'TAATACGACTCACTATAGGG ³	46.8
T7 terminator	antisense	5'GCTAGTTATTGCTCAGCGG ³	52.9
For10+26	sense	5'GAACAATACTCAGTCAAGCTTACACC ³	63.0
For252+20	sense	5'TGAATGGCCTGAATACATCG ³	58.6

Table 3-2 PCR amplification of apoaequorin from pHAQ

PCR Nr	Enzyme	Template	(mg)	Primers	DMSO (v/v)	Mg ²⁺ (mM)	Qsol	Cycles	T _a (°C)
PCR2	Pfu	pHAQ	20	AEQMSCI AEQXHOI	-	-	-	30	60

3.3.4 Results and discussion for pET cloning

The incorporation of the apoaequorin gene in pET26b vector resulted in the plasmid named pETAQ. Figure 3-7 shows the DNA agarose gel electrophoresis of the pET cloning route. Figure 3-8 and Figure 3-9 show the sequencing results obtained for pETAQ, aligned with the apoaequorin sequence as was expected from the lab of Professor Trewawas and with the apoaequorin gene sequence obtained from Genbank database. The expected one base pair difference with the Genbank entry is highlighted in red. The sequencing results indicate that apoaequorin gene was successfully ligated between *Msc* I and *Xho* I recognition sites in pET26b, between *pelB* leader sequence and 6xHis-tag sequence. The end codon TGA follows immediately after the 6xHis-tag sequence. Six base pairs between the apoaequorin gene and the 6xHis-tag came as part of the pET26b vector; they add two amino acids (Leu, Glu) to the C-terminal of the translated protein sequence before the 6xHis-tag.

The ORF of pETAQ was explored using the ExPASy Translate tool (Gasteiger *et al.*, 2003) and the translated protein sequence was aligned with the M16103 translated apoaequorin in Figure 3-10. The translated protein sequence of the pETAQ ORF was compared to the translated M16103 gene. The *pelB* peptide leader is indeed expressed upstream of the apoaequorin gene (black font, amino acid numbering 1-23). The alignment shows that the one base pair difference leads to one amino acid difference at position 143, counting from the beginning of the apoaequorin gene (or position 165 counting from the beginning of the *pelB* peptide) and that 2 extra amino acids are added at the C-terminus upstream of the six histidines.

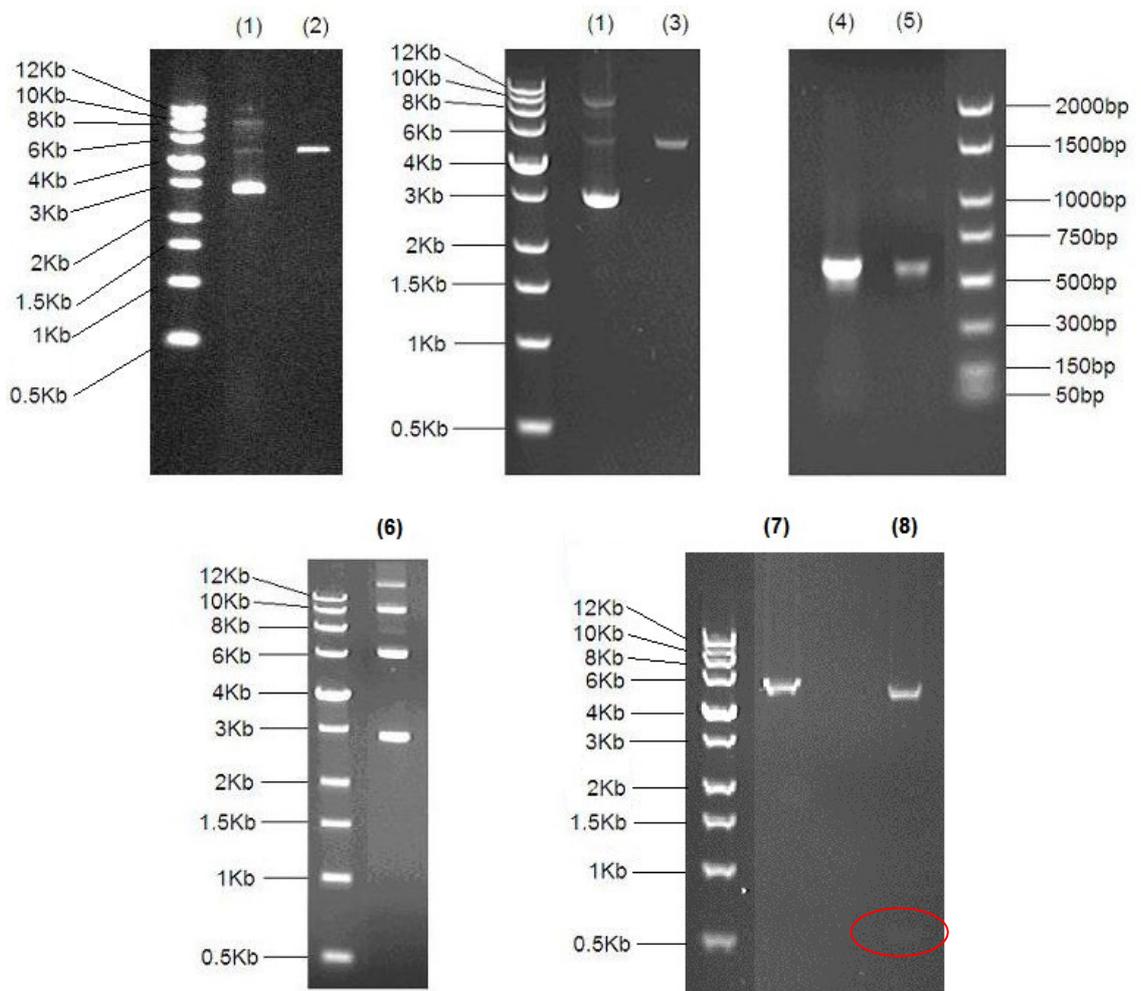


Figure 3-7 Agarose gel visualisation of pET26b cloning route. (1) open circular, linear and supercoiled forms of pET26b uncut plasmid (5,360bp), (2) pET26b single digest with *Msc* I restriction enzyme (5,360bp) (3) vector for ligation (5,293 bp): pET26b plasmid following double digest reaction with *Msc* I and *Xho* I restriction enzymes, (4) gel: blunt ended PCR2 construct (611bp), (5) insert for ligation (605 bp) PCR2 following double digest reaction with *Msc* I and *Xho* I restriction enzymes, (6) pETAQ plasmid: open circular, linear and supercoiled forms. Lanes (1) to (5) were of 1.2% agarose gels and (6) was of 1.4% agarose gel. (7) pETAQ single digest with *Xho* I, band at ~5.9 Kb(8) pETAQ double digest with *Xho* I and *Msc* I restriction enzymes. The faint lower band at ~588 bp corresponds to apoaerugin gene plus a few flanking base pairs and upper band at ~5.3 Kb corresponds to the remaining vector.

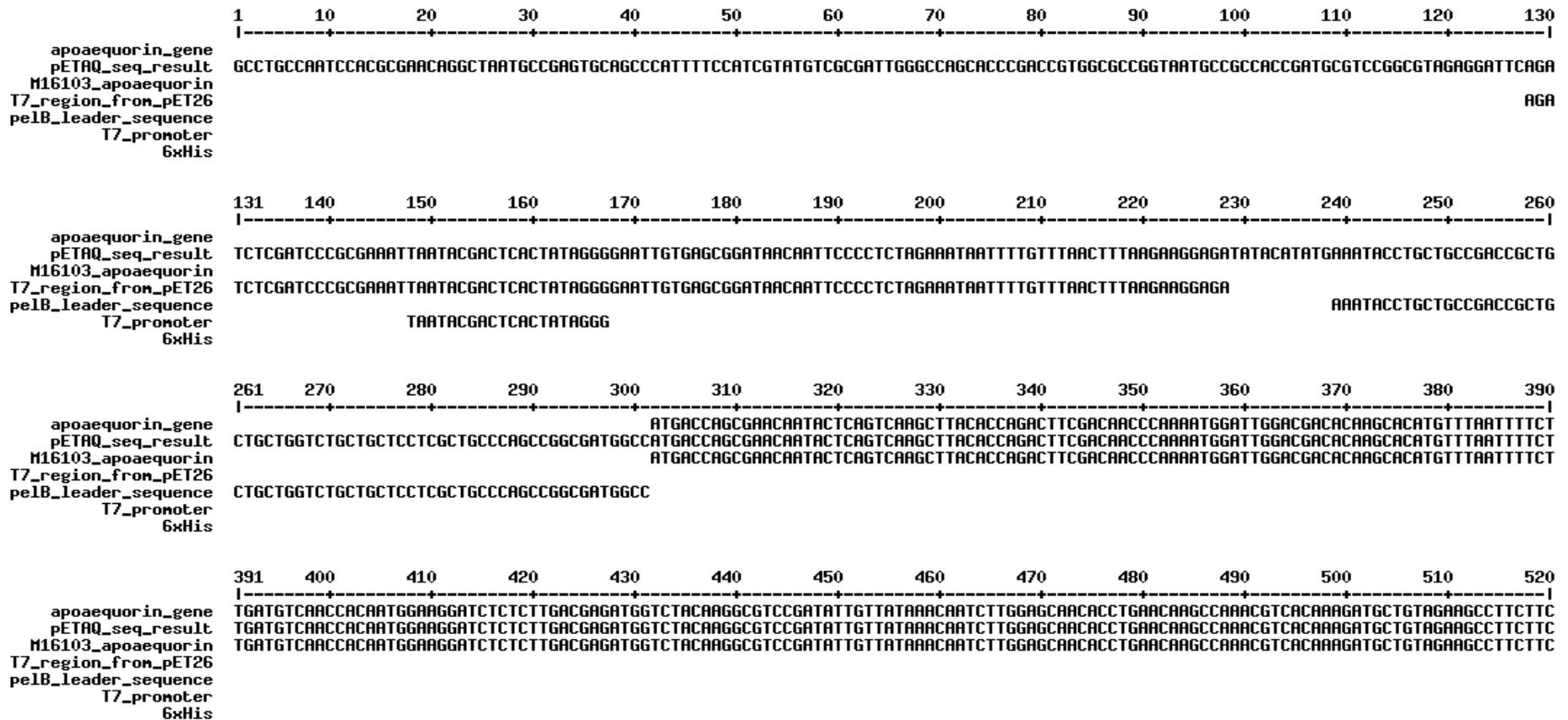


Figure 3-8 pETAQ DNA sequencing results (a). Multiple alignment of pETAQ plasmid sequence with the apoaequorin gene sequence received from Professor Trewawas lab, apoaequorin gene sequence described in Genbank M16103 and the desired features of T7 promoter region, *pelB* leader sequence and 6xHis-tag. Multi-alignment was performed using software Bioedit (Hall, 2004) and MultAlin (Corpet, 1988).

```

521   530   540   550   560   570   580   590   600   610   620   630   640   650
|-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----|
apoaequorin_gene  GGAGGAGCTGGAAATGAATATGGTGTAGAAACTGAATGGCCTGAATACATCGAAGGATGGAAAAGACTGGCTTCCGAGGAAATGAAAAGGTATTCAAAAACCAATCACACTTATTCGTTTATGGGGTG
pETAQ_seq_result  GGAGGAGCTGGAAATGAATATGGTGTAGAAACTGAATGGCCTGAATACATCGAAGGATGGAAAAGACTGGCTTCCGAGGAAATGAAAAGGTATTCAAAAACCAATCACACTTATTCGTTTATGGGGTG
M16103_apoaequorin  GGAGGAGCTGGAAATGAATATGGTGTAGAAACTGAATGGCCTGAATACATCGAAGGATGGAAAAGACTGGCTTCCGAGGAAATGAAAAGGTATTCAAAAACCAATCACACTTATTCGTTTATGGGGTG
T7_region_from_pET26
pe1B_leader_sequence
T7_promoter
6xHis

651   660   670   680   690   700   710   720   730   740   750   760   770   780
|-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----|
apoaequorin_gene  ATGCATTGTTTCGATATCATTGACAAGACCAAATGGAGCTATTTCACTGGATGAATGGAAAGCATACACCAATCTGCTGGCATCATCCAATCGTCAGAGATTGCGAGGAACATTCAGAGTGTGCGA
pETAQ_seq_result  ATGCATTGTTTCGATATCATTGACAAGACCAAATGGAGCTATTTCACTGGATGAATGGAAAGCATACACCAATCTGCTGGCATCATCCAATCGTCAGAGATTGCGAGGAACATTCAGAGTGTGCGA
M16103_apoaequorin  ATGCATTGTTTCGATATCATTGACAAGACCAAATGGAGCTATTTCACTGGATGAATGGAAAGCATACACCAATCTGCTGGCATCATCCAATCGTCAGAGATTGCGAGGAACATTCAGAGTGTGCGA
T7_region_from_pET26
pe1B_leader_sequence
T7_promoter
6xHis

781   790   800   810   820   830   840   850   860   870   880   890   900   910
|-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----|
apoaequorin_gene  TATTGATGAAAGTGGACAGCTCGATGTTGATGAGATGACAAGACAAACATTTAGGATTTTGGTACACCATGGATCCTGCTTGCGAAAAGCTCTACGGTGGAGCTGTCCC
pETAQ_seq_result  TATTGATGAAAGTGGACAGCTCGATGTTGATGAGATGACAAGACAAACATTTAGGATTTTGGTACACCATGGATCCTGCTTGCGAAAAGCTCTACGGTGGAGCTGTCCCC
M16103_apoaequorin  TATTGATGAAAGTGGACAGCTCGATGTTGATGAGATGACAAGACAAACATTTAGGATTTTGGTACACCATGGATCCTGCTTGCGAAAAGCTCTACGGTGGAGCTGTCCCC
T7_region_from_pET26
pe1B_leader_sequence
T7_promoter
6xHis
CACCACCACCACCAC

911   920   930   940   950   958
|-----+-----+-----+-----+-----|
apoaequorin_gene  CACTGAGATACCGGCTGCTAACAAAGTCCCGAAGGCAAGTGAGGCACG
pETAQ_seq_result  CACTGAGATACCGGCTGCTAACAAAGTCCCGAAGGCAAGTGAGGCACG
M16103_apoaequorin  CACTGAGATACCGGCTGCTAACAAAGTCCCGAAGGCAAGTGAGGCACG
T7_region_from_pET26
pe1B_leader_sequence
T7_promoter
6xHis  CAC

```

Figure 3-9 pETAQ DNA sequencing results (b). Figure 3-8 continued. Multiple alignment of pETAQ plasmid sequence with the apoaequorin gene sequence received from Professor Trewawas lab, the apoaequorin gene sequence described in Genbank M16103 and the desired features of T7 promoter region, *pe1B* leader sequence and 6xHis-tag. Multi-alignment was performed using software Bioedit (Hall, 2004) and MultAlin (Corpet, 1988).

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	
M16103_translate	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----														
pETAQ_translate	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----														
Consensus	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----														
	131	140	150	160	170	180	190	200	210	220	226				
M16103_translate	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----														
pETAQ_translate	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----														
Consensus	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----														

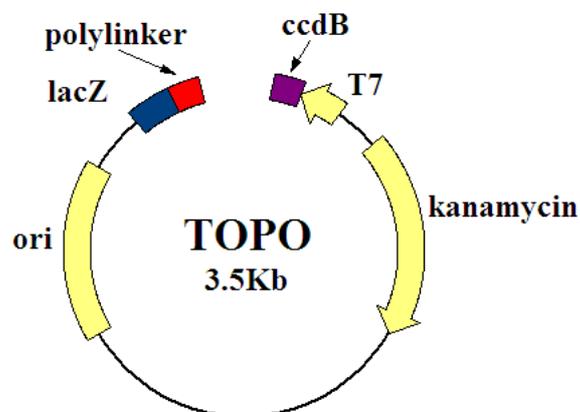
Figure 3-10 pETAQ open reading frame. M16103_translate is the translated protein sequence of M16103, used as reference sequence.

3.4 TOPO[®] cloning route

The second cloning approach towards obtaining the apoaequorin gene within a desired expression vector (Section 3.1.1) was decided as backup to the pET cloning route and as a result of initially unsuccessful pET vector and apoaequorin insert ligation attempts (Section 3.3.3).

This approach would avoid restriction digests and sticky (cohesive) end ligation methods. The envisaged DNA construct (Section 3.1.1) incorporating the apoaequorin gene behind a strong promoter, *pelB* leader sequence for periplasmic localisation and 6xHis-tag for metal affinity purification would be constructed through a series of PCR amplifications. The DNA construct produced this way would be inserted into a suitable capture vector using blunt ended ligation. The pCR[®]-Blunt TOPO[®] vector had successfully been used for blunt ended cloning in the lab and hence was the capture vector of choice. The TOPO[®] cloning route is illustrated in Figure 3-12.

The pCR[®]-Blunt II-TOPO[®] open capture vector incorporating genes for replication, transcription and antibiotic resistance was from Invitrogen Ltd. shown in Figure 3-11. TOPO[®] already contains a T7 promoter for eukaryotic RNA; the T7 promoter region from pET system would be utilised for *E. coli* expression.



```

M13 Reverse priming site                                     SP6 promoter/priming site ↓
201 CACA CAGGAA ACAGCTATGA CATGATTAC GCCAAGCTAT TTAGGTGACA CTATAGAATA
    GTGT GTCCTT TGTCGATACT G TACTAATG CGGTTCGATA AATCCACTGT GATATC TTAT

      Nsi I  Hind III  Asp718 I  Kpn I  Ecl136 II Sac I  BamH I   Spe I
CTCAAGCTAT GCATCAAGCT TGGTACCGAG CTCGGATCCA CTAGTAACGG CCGCCAGTGT
GAGTTCGATA CGTAGTTCGA ACCATGGCTC GAGCCTAGGT GATCATTGCC GCGCGTCACA

      EcoR I
GCTGGAATTC GCCCTT Blunt PCR Product AAGGGCGAATTCT GCAGATA
CGACCTTAAG CGGGAA TTCCCGCTTAAGA CGTCTAT

      Not I   Xho I           Nsi I Xba I   Dra II   Apa I   T7 promoter/priming site
TCCATCACAC TGGCGGCCGC TCGAGCATGC ATCTAGAGGG CCCAATTCC CCTATAGTGA
AGGTAGTGTG ACCGCCGGCG AGCTCGTACG TAGATCTCCC GGGTTAAGCG GGATATCACT

M13 Forward (-20) priming site
GTCGTATTAC AATTCACTGG CCGTCGTTTT AC AACGTCGT GACTGGGAAA ACCCTGGCGT 476
CAGCATAATG TTAAGTGACC GGCAGCAAAA TGTTCAGCA CTGACCCTTT TGGGACCGCA .

```

Figure 3-11 TOPO[®] capture vector. Image adapted from www.invitrogen.com.

3.4.1 Schematic of TOPO[®] cloning route

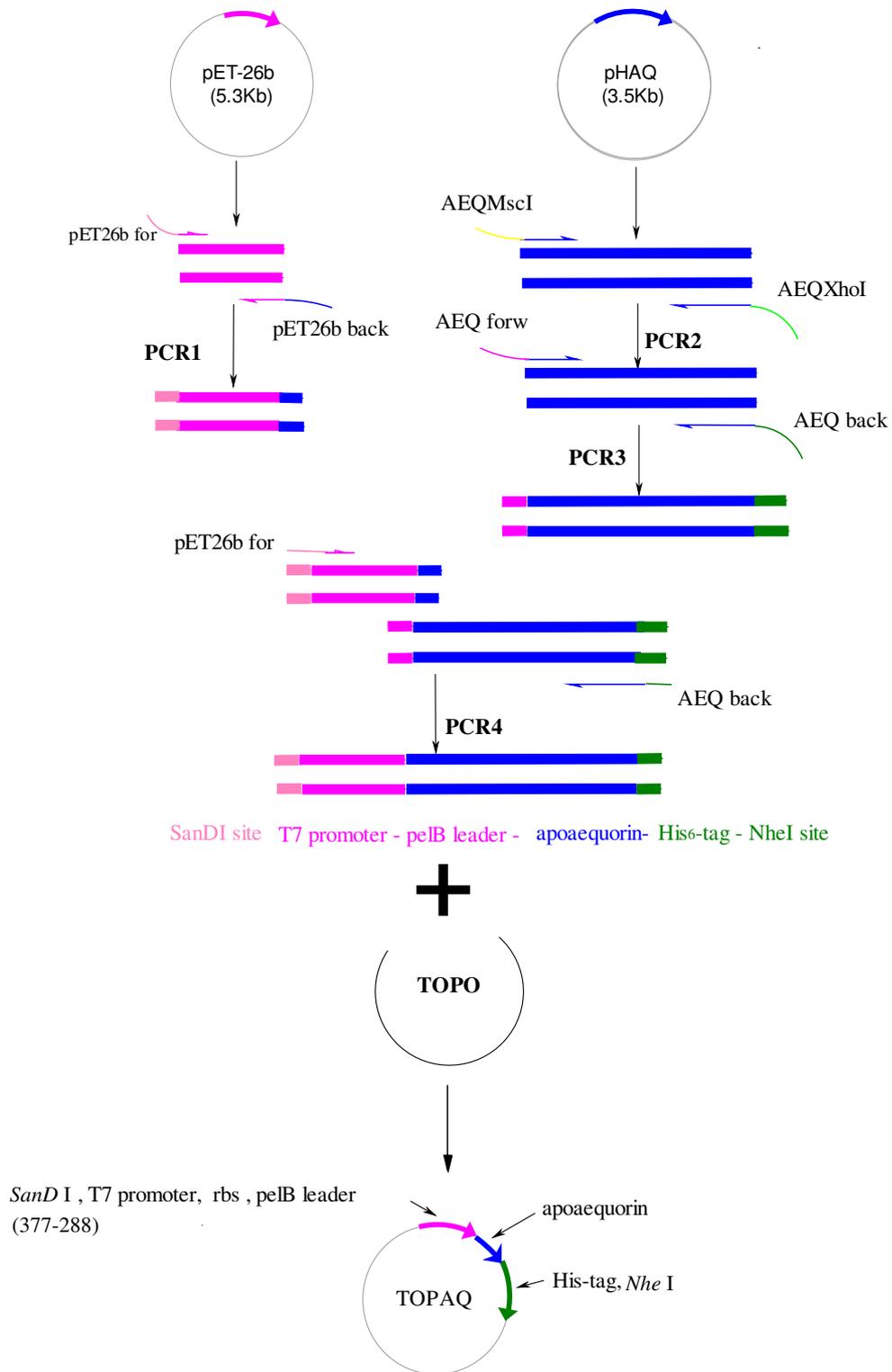


Figure 3-12 TOPO[®] cloning route schematic

3.4.2 Methods for TOPO[®] cloning

3.4.2.1 Design of primers

Primer sequences and melting temperatures are shown in Table 3-3. Primers PET26bFOR and PET26bREV are specific to the T7 promoter region and *pelB* sequence of plasmid pET26b respectively. Primer PET26bFOR contains an overhang with *Nhe* I restriction site and T7 promoter at the end of the gene encoding for the N-terminus of the protein. Primer PET26bREV (antisense) contains an overhang with the first 29 bases of the apoaequorin gene.

Primers AEQMSCI, AEQXHOI, AEQfor and AEQrev are specific to the apoaequorin gene. AEQMSCI and AEQXHOI had been previously designed to introduce the restriction endonuclease site for *Msc* I and *Xho* I upstream and downstream of the apoaequorin sequence, respectively and had been used for pET cloning route. AEQfor was designed to add the last 25bp of *pelB* leader of pET26b to the upstream of the apoaequorin gene. AEQrev was designed in order to add 6xHis-tag, end codon TAA and restriction site *SanD* I to the C-terminus of the protein. Restriction sites *SanD* I and *Nhe* I were chosen as unique restriction sites at the N- and C-terminus end of the construct respectively, according to the criteria set in Section 2.3.5.

3.4.2.2 PCR amplifications

PCR amplifications were carried out in mixtures of 25 or 50 μ L final volumes, using the general methods described in Table 2-2 and 3-1, Section 2.3.2. The reactions required several optimisation attempts and the final optimum conditions for each reaction are shown in Table 3-4. The success of each step of the series of PCR reactions was initially verified by DNA agarose gel electrophoresis. When more than one size of products was generated despite optimisation attempts, the band of the desired size was excised and extracted from low melting-point agarose gel and purified (Section 2.3.1) for use in subsequent PCR steps.

PCR1: The T7 promoter and *pelB* signal encoding region of template plasmid pET26b were amplified using primers PET26bFOR and PET26bREV. In the resulting blunt-ended construct, this region is flanked by restriction site *Nhe* I at the 5' end and by the first 25 bp encoding for the apoaequorin gene at the 3' end.

PCR0: Attempt to amplify the apoaequorin gene, flanked by the last 25 bp of the *pelB* signal sequence at the 5' end and a sequence encoding for 6 His-tags, end codon and recognition site for restriction enzyme *SanD* I using pHAQ as template and primers AEQFOR and AEQREV. Despite numerous optimisation attempts this reaction was unsuccessful, which led to PCR3. For the reaction PCR3 the template used was the product of PCR2 (previously used in pET cloning route), instead of the full plasmid pHAQ.

PCR2: The apoaequorin gene was amplified from template pHAQ, flanked by *Msc* I and *Xho* I restriction enzyme recognition sites at the 5' and 3' end respectively, using primers AEQMSCI and AEQXHOI. This construct served as an intermediate template for reaction PCR3.

PCR3: This reaction resulted in the amplification of the apoaequorin gene, flanked by the last 25bp of the *pelB* signal sequence at the 5' end and a sequence encoding for 6 His-tags, end codon and recognition site for restriction enzyme *SanD* I. It was not possible to acquire a product of the desired length by using pHAQ as a template with primers AEQFOR and AEQREV, despite numerous optimisation attempts. For reaction PCR3 the template used was the product of PCR2, instead of plasmid pHAQ.

PCR4: Creation of the final construct using primers PET26bFOR and AEQREV. This has been the most successful of a series of several attempts to optimise this reaction. Gel visualisation suggests possible formation of more constructs than the one of the desired size (721 bp). Band of the desired size was extracted from low melt agarose gel for further ligation step with TOPO[®] open vector.

PCR4 product and TOPO[®] were ligated and the resulting plasmid was used to transform electrocompetent *E. coli* Top10 (Table 2–7). Plasmid DNA was isolated from overnight cultures of the transformed cells and sequenced using standard primers M13 For and M13 Rev. The sequencing results indicated that the desired construct was successfully ligated into TOPO[®] vector. The resulting plasmid was named TOPAQ.

Table 3-3 Oligonucleotides for TOPO[®] cloning route and sequencing

Name		Sequence	T _m (°C)
PET26bFOR	sense	5'GCTAGCATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATA ^{3'}	78.2
PET26bREV	antisense	5'TGACTGAGTATTGTTCGCTGGTCATGGCCATCGCCGGCTGGGCAGCG ^{3'}	79.6
AEQMSCI	sense	5'CGGCGATGGCCATGACCAGCGAACAATACTCAGTCA ^{3'}	82.0
AEQXHOI	antisense	5'GTGGTGCTCGAGGGGGACAGCTCCACCGTAGAGC ^{3'}	84.8
AEQFOR	sense	5'CCTCGCTGCCAGCCGGCGATGGCCATGACCAGCGAAC AATACTCAGTCA ^{3'}	80.3
AEQREV	antisense	5'GGGTCCCTTAGTGGTGGTGGTGGTGGTGGGGACAGCTCCACCGTAGAGCTTTT ^{3'}	80.7
M13-for	sense	5'GTAAAACGACGGCCAG ^{3'}	49.6
M13-rev	antisense	5'CAGGAAACAGCTATGAC ^{3'}	46.1
For10+26	sense	5'GAACAATACTCAGTCAAGCTTACACC ^{3'}	63.0

Table 3-4 PCR conditions for TOPO[®] cloning

Reaction	Sense primer	Antisense primer	Templates	Enzyme	DMSO (v/v)	Qsol (v/v)	Mg ²⁺ (mM)	Nr cycles	T _a (°C)
PCR1	PET26bFOR	PET26bREV	pET26b	Pfu	2%		1.5	30	60
PCR2	AEQMSCI	AEQXHOI	pHAQ	Pfu			1.5	30	60
PCR3	AEQFOR	AEQREV	PCR2	<i>ProofStart</i>	4%	20%	1.5	30	68
PCR4	PET26bFOR	AEQREV	PCR1 and PCR3	<i>ProofStart</i>			2.5	30	68

3.4.2.3 TOPO[®] ligation

TOPO[®] ligation was performed according to the manufacturer's instructions and as described in Section 2.3.7.2.

Table 3-5 PCR mix for TOPO[®] cloning reactions

Component	Volume (μL)	PCR1	PCR2	PCR3	PCR4
10x optimal buffer	5	5	5	10	5
25mM MgSO ₄	Variable	-	-	-	3
dNTP mix (10 mM of each)	1.5	1	1	3	1.5
Sense Primer	Variable	1	1	2	1
Antisense Primer	Variable	1	1	2	1
<i>ProofStart/Pfu</i> polymerase (2.5U/ μL)	1	1 (<i>Pfu</i>)	1 (<i>Pfu</i>)	2 (<i>Proof</i>)	1 (<i>Proof</i>)
RNAse-free H ₂ O	Variable	39.4	40.6	55.4	25.7
Template DNA (20ng final)	Variable	0.6	0.4	1.6	1
Template DNA 2 (applies only to creation of construct PCR4)	Variable	-	-	-	0.8
DMSO	Variable	1	-	4	-
Qsolution	Variable	-	-	20	10
Total Volume	50	50	50	100	50

3.4.3 Results and discussion for TOPO[®] cloning

The incorporation of the apoaequorin gene in TOPO[®] open vector resulted in the plasmid named TOPAQ. Figure 3-13 shows the DNA agarose gel electrophoresis of the TOPO[®] cloning route. Figure 3-14 to Figure 3-16 show the sequencing results obtained for TOPAQ, aligned with the apoaequorin sequence as was expected from the lab of Professor Trewawas and with the apoaequorin gene sequence obtained from Genbank database. As with pETAQ results, the expected one base pair difference with the Genbank entry is highlighted in red. The sequencing results indicate that apoaequorin gene was PCR amplified downstream of T7 promoter region and *pelB* leader sequence and 6xHis-tag sequence. The end codon TGA follows immediately after the 6xHis-tag sequence. The construct is flanked by *SanD* I and *Nhe* I restriction sites. The ORF of the expressed sequence in TOPAQ is aligned with the theoretical expected apoaequorin sequence in Figure 3-17. TOPO[®] vector already contains a T7 promoter priming site downstream of the blunt ended cloning site. The apoaequorin DNA construct created in Section 3.4.2.2 contains its own T7 region and lac operator as those were amplified by PCR from pET26b template. The direction of the insert in TOPO[®] was reverse-oriented to the vector numbering shown in Figure 3-11.

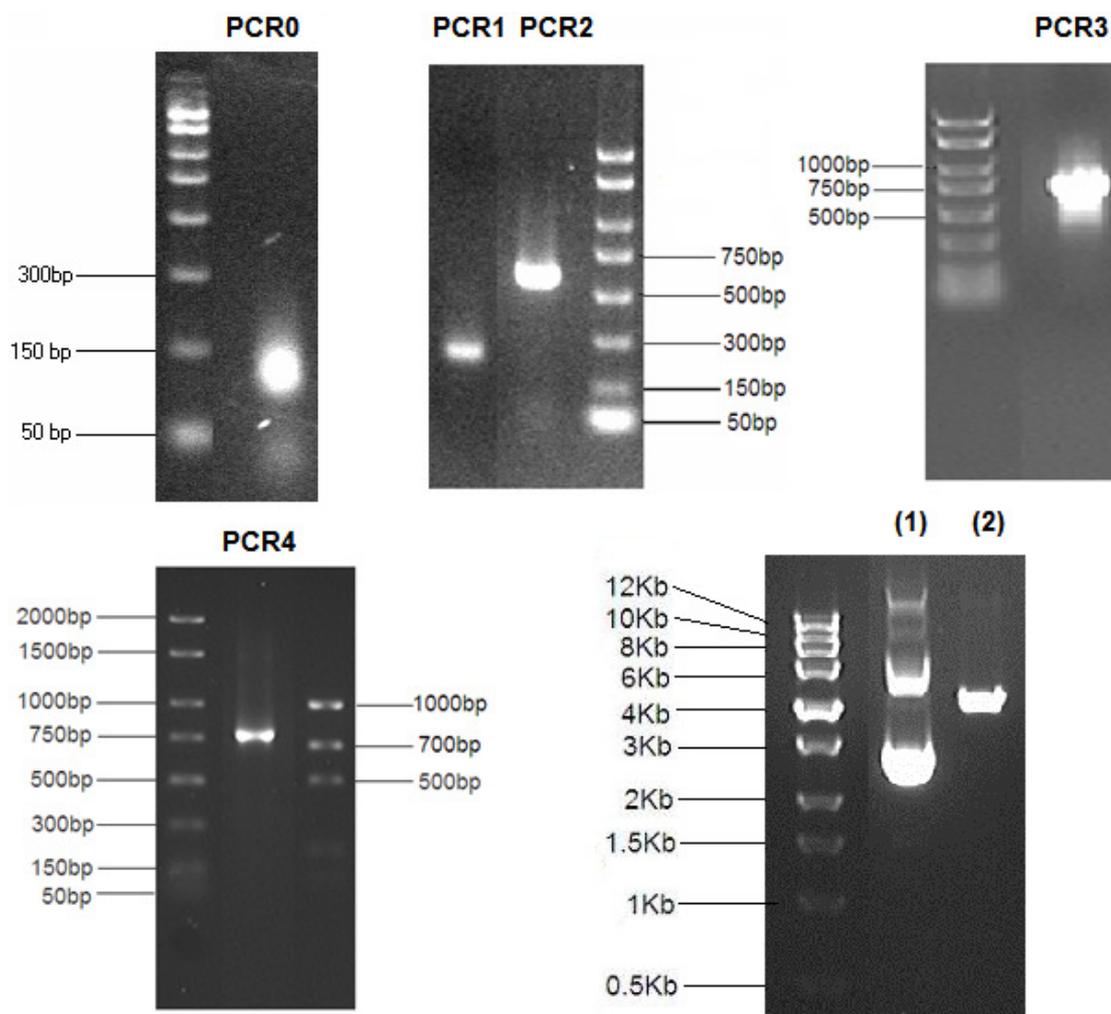


Figure 3-13 Agarose gel visualisation of TOPO[®] cloning route. PCR0: Attempts to amplify apoaequorin gene using primers AEQFOR and AEQREV, using pHAQ as template. There is no product of the desired size (642 bp). Contamination seen on gel picture is suspected to be primer-dimer formations, possibly due to a specificity problem of the primers and template. This problem was circumvented by consecutive reactions PCR2 and PCR3. PCR1: Amplification of T7 promoter and *pelB* leader using primers PET26bFOR and PET26bREV. PCR product was of expected size (209 bp). PCR2: Amplification of apoaequorin gene using primers AEQMSCI and AEQXHOI. Product of the desired size (611 bp) seems to have been produced. This fragment was used as template for PCR3. PCR3: Amplification of apoaequorin gene using primers AEQFOR and AEQREV. Product of the desired size (634 bp) seems to have been produced. This product was used in combination with PCR1 product, as template for PCR4. PCR 4: Production of the final construct using primers PET for and AEQREV. This has been the most successful of a series of several attempts to optimise this reaction. Faint bands near the product of the desired size (721 bp) indicate presence of some contamination. The band of desired size was extracted from the gel in order to be used for ligation into the open capture vector. Lane (1): TOPAQ uncut plasmid open circular, linear and supercoiled forms, Lane (2): TOPAQ digested with *Msc* I, fragment ~4.22 Kb.

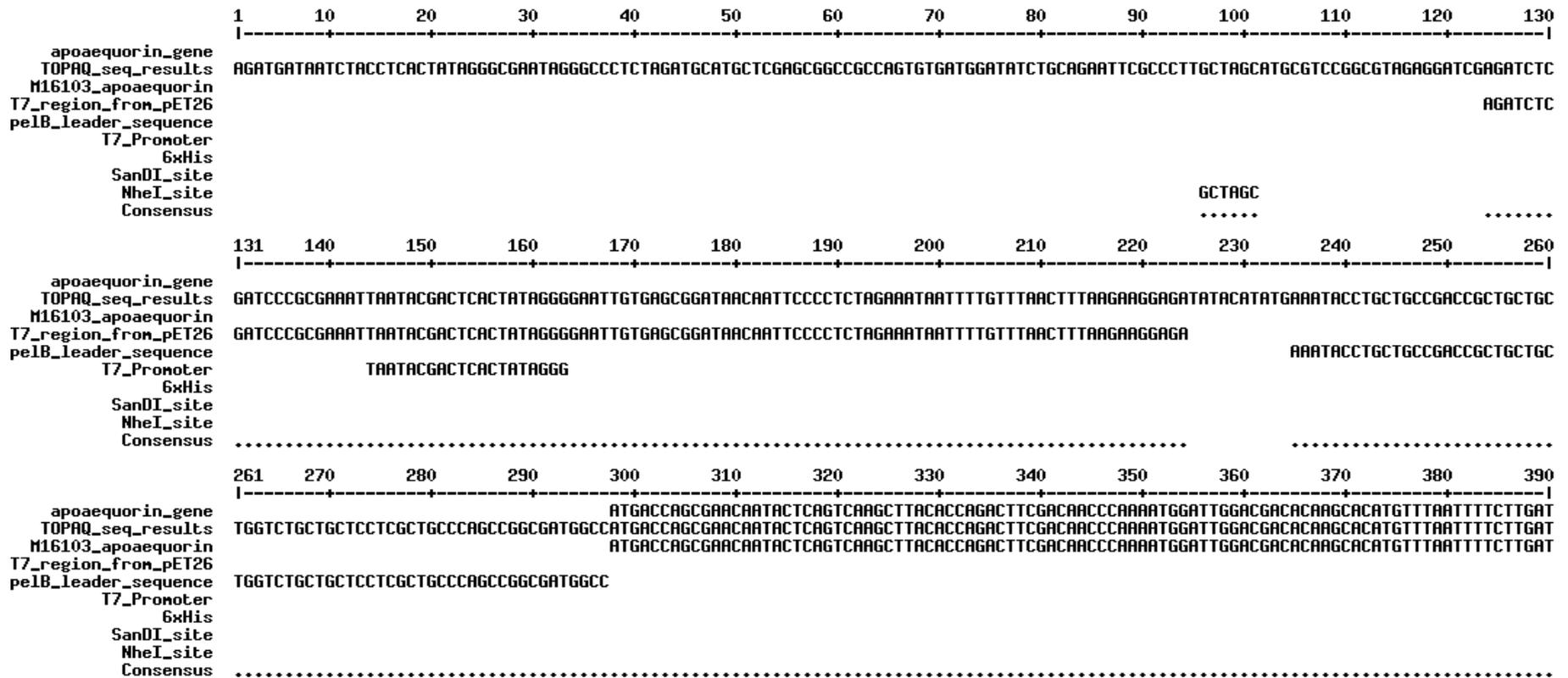


Figure 3-14 TOPAQ DNA sequencing results (a). Multiple alignment of TOPAQ plasmid sequence with the apoaequorin gene sequence received from Professor Trewawas lab, the apoaequorin gene sequence described in Genbank M16103 and the desired features of T7 promoter region, *pe1B* leader sequence and 6xHis-tag. Multi-alignment was performed using software Bioedit (Hall, 2004) and MultAlin (Corpet, 1988).


```

781   790   800   810   820   830   840   850   860   870   880   890   900   910
|-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----|
apoaequorin_gene  GATGAAAGTGGACAGCTCGATGTTGATGAGATGACAAGACAACATTTAGGATTTTGGTACACCATGGATCCTGCTTGCGAAAAGCTCTACGGTGGAGCTGTCCCC
TOPAQ_seq_results GATGAAAGTGGACAGCTCGATGTTGATGAGATGACAAGACAACATTTAGGATTTTGGTACACCATGGATCCTGCTTGCGAAAAGCTCTACGGTGGAGCTGTCCCC
M16103_apoaequorin GATGAAAGTGGACAGCTCGATGTTGATGAGATGACAAGACAACATTTAGGATTTTGGTACACCATGGATCCTGCTTGCGAAAAGCTCTACGGTGGAGCTGTCCCC
T7_region_from_pET26
pelB_leader_sequence
T7_Promoter
6xHis
SanDI_site
NheI_site
Consensus .....

911   920   928
|-----+-----+-----|
apoaequorin_gene  CCCAAGGGCGAATTCAG
TOPAQ_seq_results
M16103_apoaequorin
T7_region_from_pET26
pelB_leader_sequence
T7_Promoter
6xHis
SanDI_site CCC
NheI_site
Consensus ...

```

Figure 3-16 TOPAQ DNA sequencing results (c). Figure 3-15 continued. Multiple alignment of TOPAQ plasmid sequence with the apoaequorin gene sequence received from Professor Trewawas lab, the apoaequorin gene sequence described in Genbank M16103 and the desired features of T7 promoter region, *pelB* leader sequence and 6xHis-tag. Multi-alignment was performed using software Bioedit (Hall, 2004) and MultAlin (Corpet, 1988).

	1	10	20	30	40	50	60	70	80	90	100	110	120	130
M16103_translate	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----													
TOPAQ_translate	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----													
Consensus	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----													
	131	140	150	160	170	180	190	200	210	220	224			
M16103_translate	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----													
TOPAQ_translate	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----													
Consensus	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----													

Figure 3-17 TOPAQ open reading frame. M16103_translate is the translated protein sequence of M16103, used as reference sequence.

3.5 Conclusions

The gene encoding for apoaequorin was successfully subcloned into two different plasmid vectors: pETAQ is a pET26b derivative and TOPAQ is a pCR-Blunt II-TOPO[®] derivative, both containing the apoaequorin cDNA (one base pair difference from GenBank Accession No. M16103), the *pelB* signal peptide sequence and 6xHis-tag fusion downstream of the T7 promoter. Both are inducible expression systems.

T7 was chosen in both cloning approaches as it is a very strong promoter and T7 RNA polymerase transcribes the ORF very quickly. However there are problems associated with overproduction of recombinant proteins and not efficient expression/release at the periplasmic space. Translocation across the cell membrane of *E. coli* is still not well understood; a leader (e.g. *pelB*, *ompT*, CBD or DsbA/C) is necessary, but not sufficient for export into the periplasm. Translocation also can depend on the mature domain of the target protein, which is recognised by *SecB*, the major chaperone of export (Wickner *et al.*, 1991).

Based on the features of the two vectors, their expression performance could differ; for example, the pET vectors are optimised for high expression levels. TOPO[®], as well as pBluescript vectors containing T3 and T7 bacteriophage promoters allow efficient *in vitro* synthesis of strand-specific RNA.

Even with the T7 promoter present in both vectors, the pET system is specifically designed for protein expression having some specific features that allow and maximize the performances in the host cell strain. pET vectors ensure high level of activity of the polymerase and high translation efficiency mediated by the T7 gene 10 translation initiation signals, which is not present in TOPO[®] vector. In pET the gene encoding the protein of interest may be cloned directly after the gene 10 initiation codon using cloning sites engineered for optimal expression of the cloned protein-coding sequence; the gene 10 transcription terminator is also included downstream of the cloning sites to allow efficient termination of transcription, preventing transcriptional read-through of unwanted plasmid sequences and increasing the RNA polymerase density on the sequence of interest—allowing high level accumulation of the specific protein-coding RNA transcripts. TOPO[®] also lacks the ribosome binding site (RBS).

As both approaches were successful, the resulting vectors were both tested for expression of luminescent activity in microplate studies (Chapter 4) and a high throughput expression assay was developed with the vector providing the highest levels of expression (Chapter 4) and subsequently used in library creation (Chapter 5), screening (Chapter 6) and protein kinetics (Chapter 7 as preparation for Chapter 8).

An alternative cloning approach would be to engineer the *pelB* leader sequence and 6xHis-tag sequence into pHAQ (pBluescript derivative). As pBluescript lacks a strong STOP sequence and RBS, protein expression would likely be lower compared to the pET system.

4 High-throughput expression and screening

This chapter describes the development of a high-throughput method which facilitates the production of aequorin wild type and mutants by *E. coli* in 96-well plates, followed by an activity assay against several potential activators (metal ions). The envisaged assay was required to serve as a preliminary screen in order to identify promising mutants for purification and further in-depth analysis. It builds upon the work described in Chapter 3, which yielded two plasmid vectors (pETAQ and TOPAQ) encoding for the apoaequorin gene. Hierarchical optimisation was employed for the development of the final high-throughput protocol.

4.1 Introduction

4.1.1 Steps to developing a high-throughput process

The high-throughput process should include cell growth, expression of apoprotein, incubation with coelenterazine for the reconstitution of functional aequorin in the cell medium and testing of aequorin mutant libraries for activation by seven different ions (as chosen in Chapter 5): calcium, zinc, copper, lead, cobalt, cadmium and lanthanum. The protocol should describe all the steps between handling of bacterial colonies in agar plates and measuring activity of the produced aequorin variants. Figure 4-1 presents the required stages of such a method and contains some main considerations pertaining to each stage. As the assay is meant to be a preliminary/crude screen for the assessment of aequorin mutant libraries, the design of the individual steps must satisfy a set of general high-throughput screen requirements adapted to the properties and behaviour of the aequorin system (see Section 4.1.2).

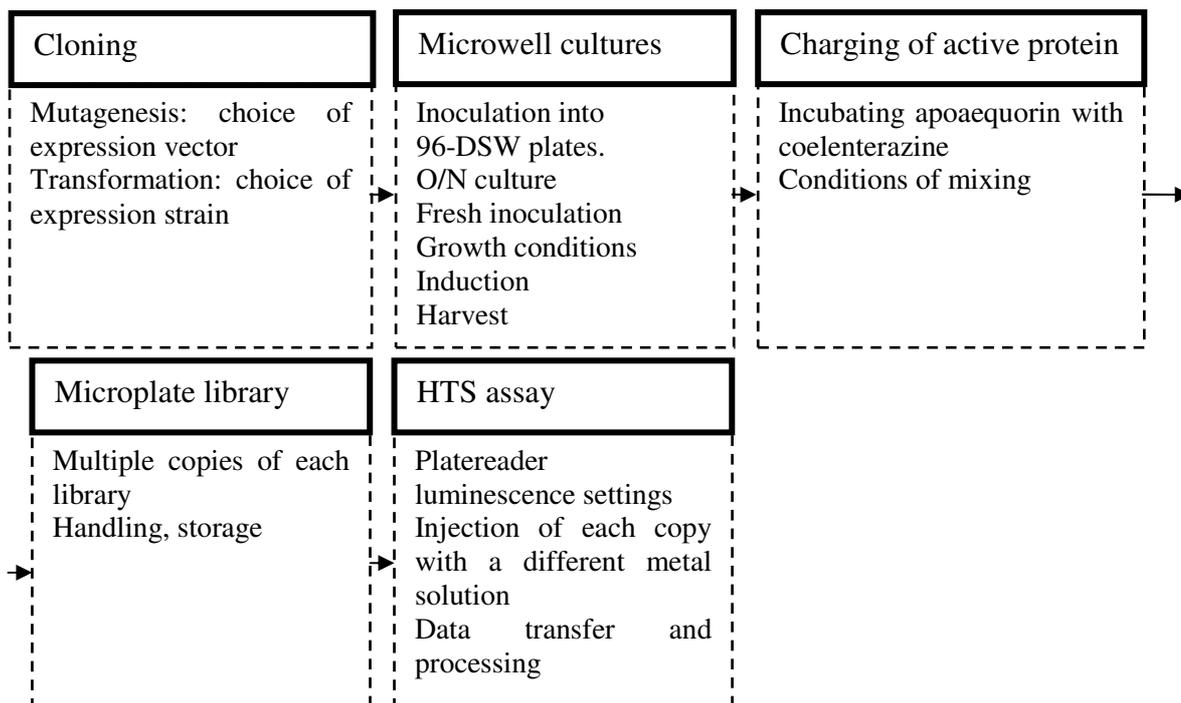


Figure 4-1 Box diagram of the steps required for a complete HTS protocol. Some main considerations pertaining to each step are included.

4.1.2 Requirements for a high-throughput assay

Several factors contribute to a suitable high-throughput assay for aequorin activity as expressed from *E. coli* in a microplate format:

- (1) High expression of native apoaequorin
- (2) Ease of activation to holo-aequorin with simple addition of coelenterazine to the medium. This requires either cell lysis or secretion of the apo-aequorin into the periplasm.
- (3) Practicality/ease of use. Assays with the fewest processing steps are generally more robust, accurate, cheaper to use and easier to automate.
- (4) Low limits of detection (high signal-to-noise ratio). Some emergent activities from new ions may be very low at first. Sensitivity will allow these to be detected.

(5) Reproducibility and low standard deviation. Few repeats are carried out for high-throughput assays initially as “hits” can be re-evaluated later. High reproducibility and low error makes it less likely that “hits” will be missed.

(6) Size of library and automation. The microplate format and use of automation enables the processing of large library sizes. Flexibility of this protocol to adapt to different mutational strategies in this and future work will allow the processing of large libraries (directed evolution) as well as selected individual mutants (rational approach).

4.2 Materials

4.2.1 Chemicals

Standard buffers and media and reagents used in this chapter are described in Chapter 2. Coelenterazine (1 mg) in lyophilised form was purchased from Sigma Aldrich and stored in -20 °C. When ready to use, 1 mL of ice cold ethanol was used to resuspend and mix the coelenterazine by pipetting and transferred into a pre-chilled eppendorf tube and kept on ice. The tubes were covered with foil and dim light conditions were used in order to protect the chromophore from light. Polypropylene volumetric flasks suitable for atomic absorption were used in order to minimise metal ion contamination of the protein samples. For the preparation of coelenterazine solutions the flasks and eppendorf tubes used were pre-chilled and covered in foil for protection against light and heat and kept on ice.

4.2.2 Consumables

All microplates used were of the 96-well format. Deep square well (DSW) microplates used for bacterial cultures and mixing of apoprotein with coelenterazine were from Sarsted. Clear Shallow clear round-well flat bottom (F type) microplates used for optical density (OD) measurements of bacterial cultures and as lids for the DSW plates were also from Sarsted. Luminescence measurements were performed in Lumitrac™ white polystyrene, F-type (flat-bottom) 96 well microtiter plates (Greiner Bio One Ltd., Gloucestershire, UK).

4.3 Method development strategy

4.3.1 Building the assay protocol

The assay design was based on published work by Shimomura and Inouye (1999) who described a simple method for the production of aequorin at bench scale. The method exploits the fact that apoaequorin is genetically fused with a signal peptide, is transferred to the periplasm and can be extracted into the medium and reconstituted with coelenterazine into fully functional aequorin in one step. In summary, *E. coli* cells with a plasmid expression vector that included the apoaequorin cDNA downstream of the *ompA* signal peptide sequence were cultivated in LB medium using standard cell growth conditions. Luminescent activity was measured using one of two ways, depending on the requirements:

(1) 0.1 mL of bacterial suspension was incubated with 1 mL of 20 mM Tris buffer (pH 7.5) containing 10 mM EDTA, 5 mM 2-mercaptoethanol, and 1 mM coelenterazine at 0 °C for 3 h to regenerate aequorin, while allowing oxygen diffusion to occur.

(2) The bacterial suspension was centrifuged and the cell pellets are cooled and dispersed in a buffer containing EDTA, dithiothreitol and coelenterazine, then left to incubate for 1 h, while allowing diffusion of oxygen to occur.

In both (1) and (2) options, after overnight storage at 0 °C and centrifugation, a clear supernatant was produced that contains aequorin (Shimomura and Inouye, 1999).

The method was adapted to the 96-well microplate scale, based on the expression levels of the bacterial system, the operating capabilities of equipment and other considerations (i.e. sensitivity, robustness, cost, time restraints and availability). A range of other methods have been described for obtaining purified aequorin and these are considered in Chapter 7 (Protein Purification, Section 7.1).

Some of the principal questions concerning the application of microwell format for aequorin production were: (1) could adequate luminescent activities be achieved in the microwells? (2) could microwell cultures produce consistent activities across the microplate with reproducibility across microplates? (3) what were the optimal conditions in order to achieve maximum activity? (4) was there external or well-to-well cross-contamination?

The following sections explore what is possible and practical in achieving a HTS protocol for aequorin activity which satisfies the criteria mentioned in Section 4.1.2. A number of experimental setups were used to test different conditions and to gradually optimise the process for sensitivity, signal-to-noise and assay variability. Each experimental setup is described in the relevant section.

4.3.2 Experimental setup A

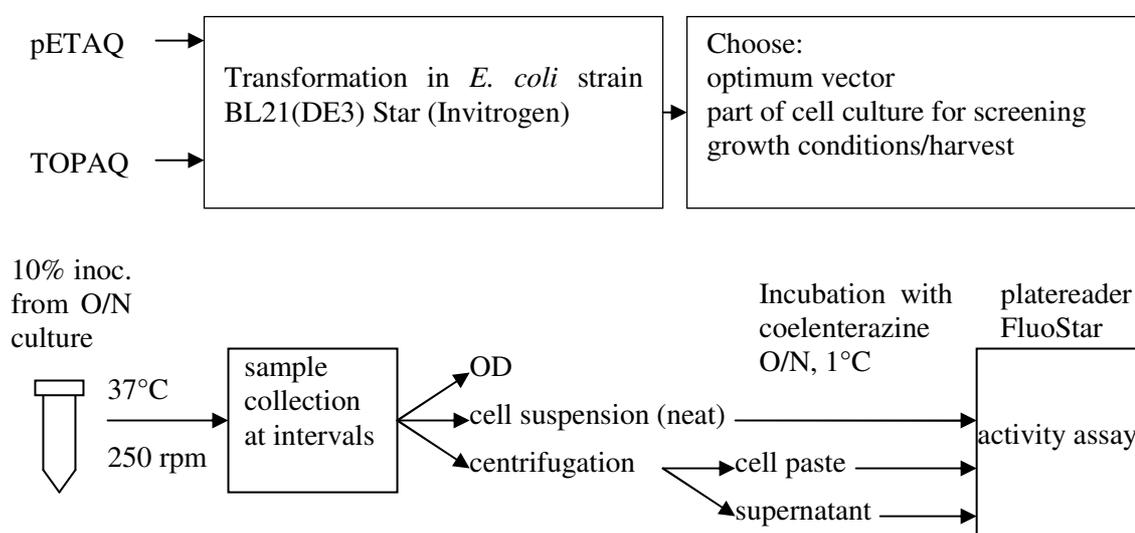


Figure 4-2 Experimental setup A for defining process parameters.

E. coli BL21(DE3) Star (F⁻ ompT hsdSB(rB⁻, mB⁻) gal dcm rne131 (DE3)) (Invitrogen) electrocompetent cells were transformed with plasmid expression vectors pETAQ or TOPAQ, brought forward from Chapter 3. Single colonies were used to inoculate 10 mL of LB medium in Falcon tubes containing 50 µg/ml of kanamycin and left to incubate overnight at 37 °C in an orbital shaker at 250 rpm. One milliliter of each overnight seed culture was used for a 10% inoculation into fresh medium (10 mL final volume, 50 µg/ml of kanamycin) in falcon tubes and grown at 37 °C and 250 rpm for up to 7 h.

One culture was induced by addition of IPTG at the start of incubation (time = 0 h), one culture was induced at 2.5 h and one culture was not induced. During the incubation 1 mL samples were collected at times 2.5, 5 and 7 h. The samples were

measured for cell optical density (OD), 100 μ L of neat culture was saved for further processing, and 500 μ L of culture were centrifuged at 12,000 rpm for 10 min and the resulting supernatant and cell pellets were stored separately for further processing. The neat culture, cell pellets and supernatant were stored on ice in order to hinder further growth and protein production.

When samples from all time points were collected, a cold (0–4°C) solution was prepared that contained 2 μ M coelenterazine, 20 mM Tris-HCl, 10 mM EDTA and 5 mM 2-mercaptoethanol, pH 7.7. The wells of a pre-chilled 96-DSW plate were filled with 950 μ L of the coelenterazine solution. One hundred microliters of each of the cell culture samples and of the culture supernatant were transferred into wells of the DSW plate according to their category (type of vector, time of harvest, time of induction). The cell pellets were dissolved by pipetting in 1 mL of coelenterazine solution and transferred into wells of the DSW plate as described previously. The position of each sample in the 96-well plate is highlighted in Figures Figure 4-7 to

Figure 4-9). The DSW microplate was covered with an inverted shallow 96-well plate (Sarstedt Inc) and placed on a table top thermomixer (Eppendorf Inc) with intermittent shaking at 1000 rpm for 3 h, temperature control at 1 °C and the whole set up placed in a cold cabinet (4–8 °C).

After the end of the cold incubation with coelenterazine solution, 100 μ L of each well were transferred into the respective wells of a Lumitrac™ shallow well plate using an 8-channel multipipette. The activity of the regenerated aequorin in the 100 μ L mixtures was assayed in the FluoStar platereader, by measuring the light emission of the sample when an equal volume of 10 mM CaCl₂ solution was injected.

4.3.3 Experimental setup B

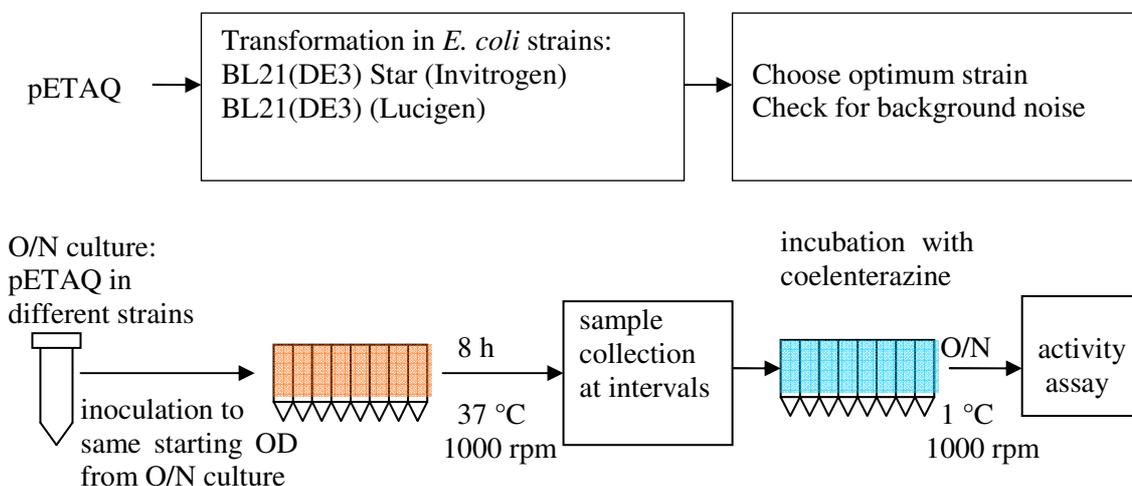


Figure 4-3 Experimental setup B for defining process parameters

Two *E. coli* BL21(DE3) electrocompetent high expression strains available in the lab were transformed with plasmid expression vector pETAQ. The transformed cells were spread on agar plates (Sections 2.2.2 and 2.3.8). Single colonies were picked from the agar plates manually and were used to inoculate 10 mL of Kan+ LB medium in falcon tubes. These seed cultures were left to incubate overnight at 37 °C and 250 rpm.

The wells of a 96-DSW plate were filled with LB Kan+ medium (450 μ L). this plate would serve as the “day culture plate”. The overnight seed cultures in the falcon tubes were used for a ~10% inoculation into a new DSW plate containing 450 μ L of Kan+ LB medium (500 μ L final volume), making sure that the starting OD of the two different strain microcultures was equal. The LB medium of the microwells was topped up to the final volume of 500 μ L. This plate was incubated at 37 °C and 1000 rpm for up to 7 h. Induction with IPTG was performed at 1 h after inoculation (corresponding to OD 0.6–0.8). Samples of the culture were collected at various time points and kept on ice for further processing.

At the end of incubation, 100 μ L of each of the collected samples were transferred into the corresponding wells of a new DSW plate, each containing 950 μ L of cold coelenterazine solution 4 μ M, 20 mM Tris-HCl, 10 mM EDTA and 5 mM 2-mercaptoethanol, pH 7.7. The DSW was covered with an inverted shallow 96-well

plate (Sarstedt Inc), placed on a table top thermomixer (Eppendorf Inc) with intermittent stirring at 1000 rpm and 1 °C overnight and placed in a cold cabinet (4–8 °C). After the end of the cold incubation with coelenterazine solution, 100 µL of each well was transferred into the corresponding wells of a Lumitrac™ shallow well plate using a the liquid handling robotics of Tecan. The activity of the regenerated aequorin in the 100 µL mixtures was assayed in the FluoStar platereader, by measuring the light emission of the sample when an equal volume of 100 mM CaCl₂ solution in 20 mM Tris-HCl, pH 7.5 was injected.

4.3.4 Experimental setup C

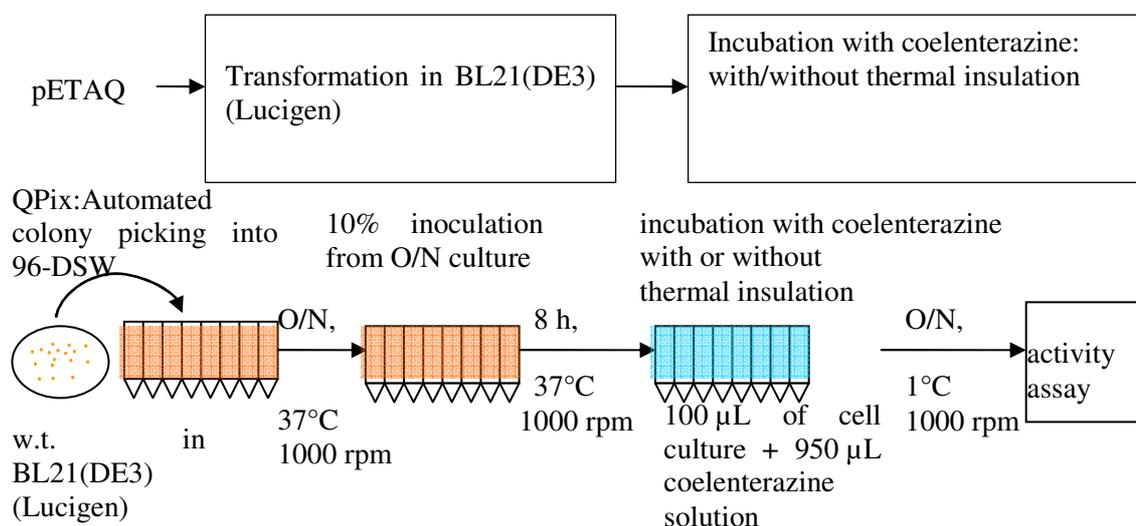


Figure 4-4 Experimental setup C for defining process parameters

E. coli BL21(DE3) electrocompetent cells (Lucigen) were transformed with plasmid expression vector pETAQ (expressing wild type apoaequorin). The transformed cells were spread on agar plates (Sections 2.2.2 and 2.3.8). The wells of a 96-DSW plate were filled with LB Kan+ medium (950 µL). Qpix2 robot was used to inoculate each well of this “seed culture plate” with colonies from the agar plate. The plate was left to incubate overnight at 37 °C and 1000 rpm. The overnight seed culture was used for a 10% inoculation into a new DSW plate containing 450 µL of Kan+ LB medium (500 µL final volume). This “day culture plate” was incubated at 37 °C and 1,000 rpm for up to 10 h. Induction with IPTG was performed at OD 0.6–0.8 or 1 h

after inoculation. Samples of the culture could be collected at various time points or at the end of the day culture (depending on the type of question being answered) and saved on ice for further processing.

At the end of incubation, 100 μ L of each of the collected samples were transferred in the wells of the same corresponding position in a new DSW plate, containing 950 μ L of cold coelenterazine solution 4 μ M, 20 mM Tris-HCl, 10 mM EDTA and 5 mM 2-mercaptoethanol, pH 7.7. The DSW was covered with an inverted shallow 96-well plate (Sarstedt Inc) and placed on a table top thermomixer (Eppendorf Inc) with intermittent stirring at 1000 rpm and 1 $^{\circ}$ C overnight and placed in a cold cabinet (4–8 $^{\circ}$ C). After the end of the cold incubation with coelenterazine solution, 100 μ L of each well was transferred into the respective wells of a Lumitrac™ shallow well plate using the Tecan liquid handling robot. The activity of the regenerated aequorin in the 100 μ L mixtures was assayed in the FluoStar platereader, by measuring the light emission of the sample when an equal volume of 100 mM CaCl₂ solution was injected.

4.4 Results and discussion

4.4.1 Measurement of cell density in microplates

Measurement of optical density in microplate format is necessary in order to monitor and validate a high-throughput expression method. Optical density measurements for a serial dilution of bacterial culture were made in microplates (using the Tecan Magellan absorbance reader) and in standard cuvettes in benchtop spectrophotometer. The association between the measurements is presented in Figure 4-5 and was found to correlate linearly with an R² value of 0.998. The slope of approximately 1.5 reflects the different pathlengths used in the two formats. These data show that the platereader selected for plate-based measurements was sufficiently accurate to reproduce values obtained in a cuvette based spectrophotometer.

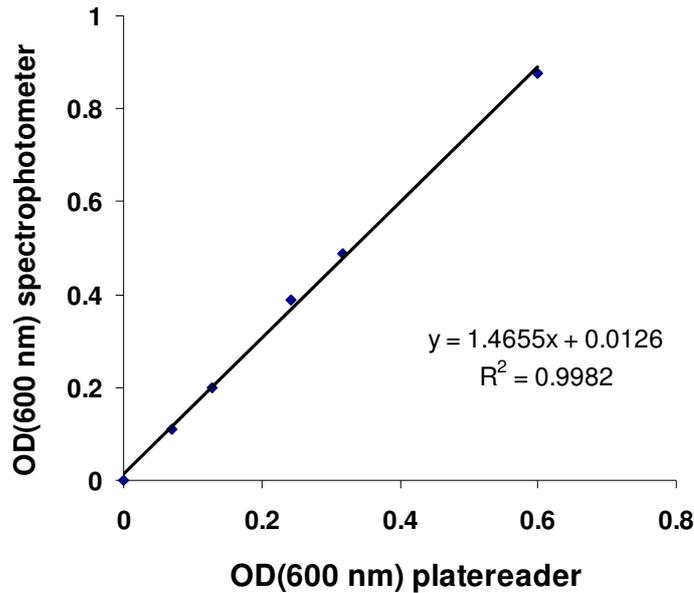


Figure 4-5 Correlation of cell culture optical densities in microplates and cuvettes. OD(600 nm) values of cell culture were measured in 96-well microplates using Magellan platereader and the corresponding OD (600 nm) were measured in a benchtop spectrophotometer.

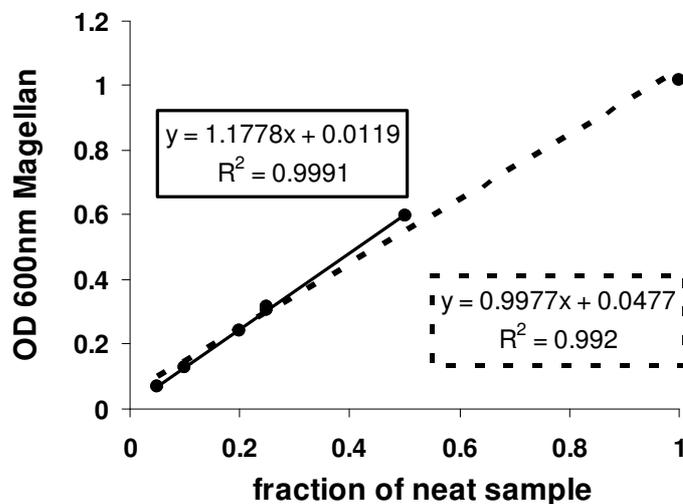


Figure 4-6 Linearity of OD (600 nm) versus dilution. The original culture was of OD(600 nm)=2.0 in Magellan platereader, pathlength corresponding to 200 μ L total volume in a 96-microplate well. The dashed line includes the neat culture, whereas the continuous line excludes the neat sample.

Figure 4-6 further explores the linearity of readings obtained in the Tecan Magellan platereader as the bacterial samples were diluted. These data show that the dynamic range of the platereader was sufficiently accurate for cell density

measurements in the range of OD(600nm) of 0–1.0. Having determined this, the plate reader was adopted for all further optimisations of strain and vector development, cell culture and bioluminescent assays.

4.4.2 Process parameters

4.4.2.1 Choice of vector

This initial experiment to compare different plasmid vectors for the aequorin bioluminescence assay was based on experimental setup A (Section 4.3.2). The two expression vectors were assessed in conjunction with two different *E. coli* high expression strains in order to determine the expression system with the highest luminescence activity for testing mutant variants of aequorin in microwell assays. The variation in luminescence profiles obtained as a function of time are shown in Figure 4-7 using the different strains. It can be seen that BL21(DE3)Star cells produce higher luminescence when carrying pETAQ vector than TOPAQ vector. The results of the comparison are consistent using cell suspension, cell culture supernatant or resuspended cell paste obtained from cell culture centrifugation.

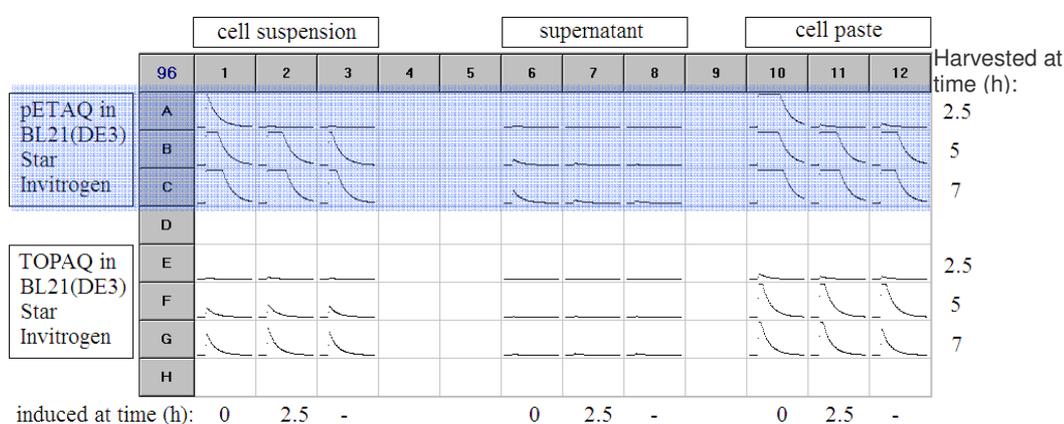


Figure 4-7 Choice of expression vector. Graphical output of luminescence traces versus time. Each well represents a different experimental condition. In each trace depicted, y-axis represents Relative Luminescence Units and the x-axis represents time (8 s total measurement time, readings taken at 0.1 s intervals). The highest producers (highlighted in blue) are cultures expressing the pETAQ vector. Platereader gain set at 2,000. Experiment carried out as described in Section 4.3.2 (Experimental setup A) and generic Section 2.7 (Luminescence measurements).

4.4.2.2 Part of culture as screening material

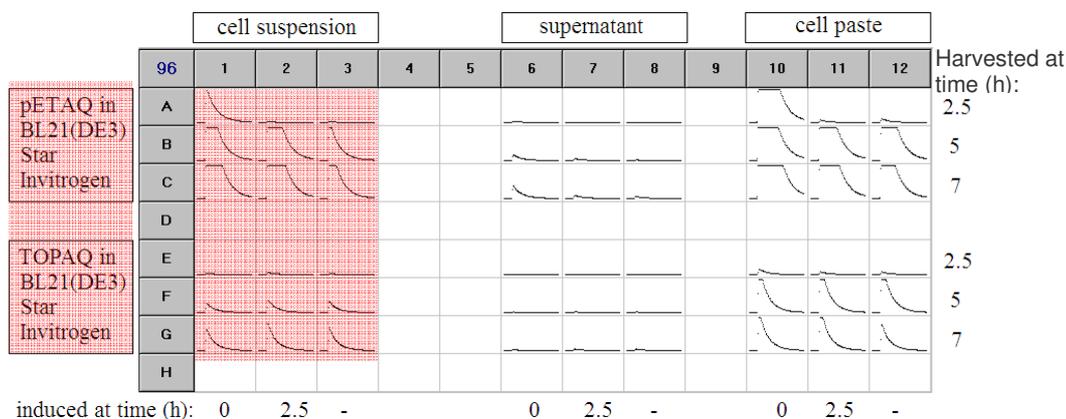


Figure 4-8 Part of culture to be used for the screening process. Experimental settings as in Figure 4-7. Comparing luminescence curves from cell suspension, supernatant and (resuspended) cell paste, sufficient amount of luminescence is produced from cell suspension (highlighted in pink) to use as high-throughput screening platform.

This experiment was based on experimental setup A. It answers the question: which part of the culture produces adequate luminescence in the most practical, easily automatable method? As seen in Figure 4-8, the cell suspension of the microwells produces enough signal for screening. Cell paste produces the maximum signal, but requires the extra steps of microplate centrifugation, supernatant removal and resuspension of pellets, which would prolong and complicate the assay. Therefore the cell suspension is an adequate and practical choice as HTS material.

4.4.2.3 IPTG induction

This experiment was based on experimental setup A. It answers the question: “when to induce with IPTG?” Nine hundred microliters of 100 mM IPTG stock were diluted with 2.1 mL LB and 10 μ L of the mix were transferred into each well using a multipipette. As seen in Figure 4-9, microwell cultures were induced at time 0, 2.5 h and one set was not induced at all. The relatively high activity at early induction is most likely attributable to the high inoculation (10%) from overnight culture, whereas “leaky” T7 expression is likely to be responsible for activity in non-induced cultures. Highest activity was observed at IPTG induction at 2.5 h, at OD~1.0, although earlier induction times would be preferable and were assessed in the later steps of assay development. The maximum harvest time out of those tested in this set (7 h) was

optimum, although longer times would be possible to achieve higher protein yields and luminescence activity.

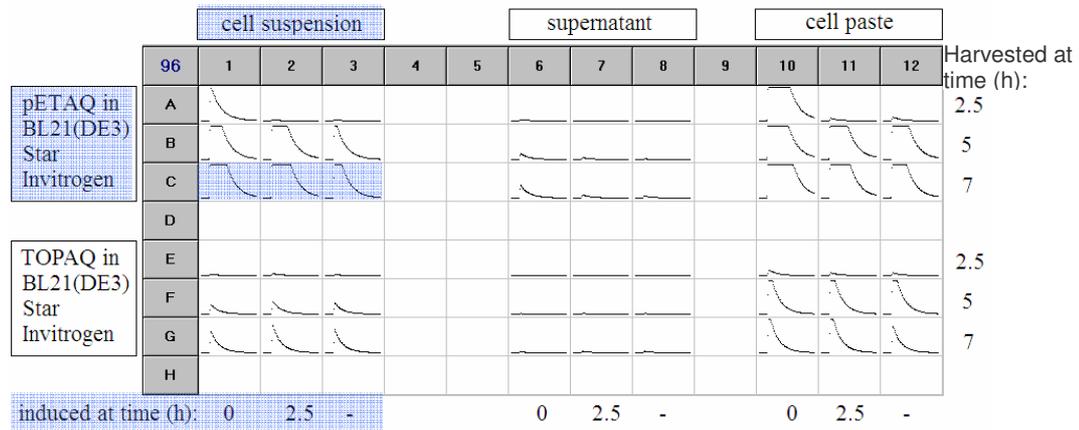


Figure 4-9 Time of induction. Experimental settings as in Figure 4-7 and Section 4.3.2. Comparison of luminescence curves from cell cultures induced with IPTG at time 0, 2.5 h after inoculation and not induced. Maximal luminescence activity from the cell suspension was observed when inducing at 2.5 h.

4.4.2.4 Choice of bacterial strain

This experiment was based on experimental setup B. *E. coli* BL21(DE3) Star (F- ompT hsdSB(rB-, mB-) gal dcm rne131 (DE3)) (Invitrogen) and *E. coli* (*E. coli* BL21(DE3) (F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15ΔlacX74 endA1 recA1araD139 Δ(ara, leu)7697 galU galK rpsL nupG λ- tonA) (*E. coli*™, Lucigen) electrocompetent cells were compared for expression levels of apoaequorin. *E. coli*™ cells exhibited slightly higher activity than the Invitrogen strain, possibly due to more rapid growth (ODs not shown).

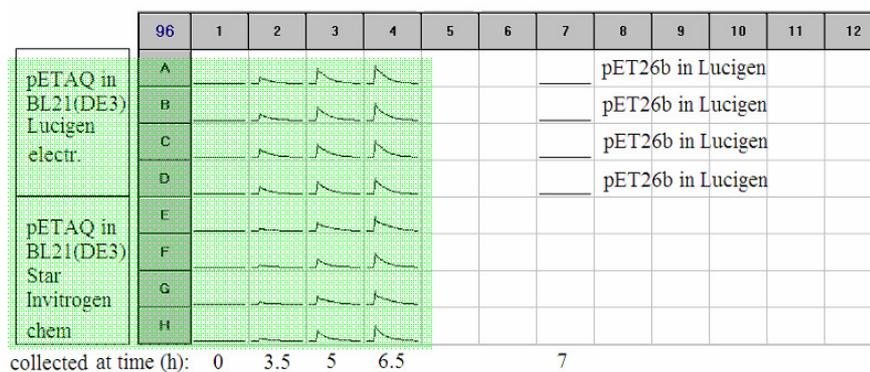


Figure 4-10 Choice of *E. coli* strain. Graphical output of luminescence traces versus time. In each trace depicted, the y-axis represents Relative Luminescence Units and the x-axis represents time (8 s total measurement time, readings taken at 0.1 s intervals). This figure compares luminescence activity produced in the cell suspension of *E. coli* BL21(DE3) cells from Lucigen and Invitrogen, both carrying expression vector

pETAQ. The highest producers were cultures of BL21(DE3) from Lucigen. Experiment carried out as described in Section 4.3.3 (Experimental setup B) and generic Section 2.7 (Luminescence measurements). Plater reader gain set at 2,000.

4.4.2.5 Check for background signal

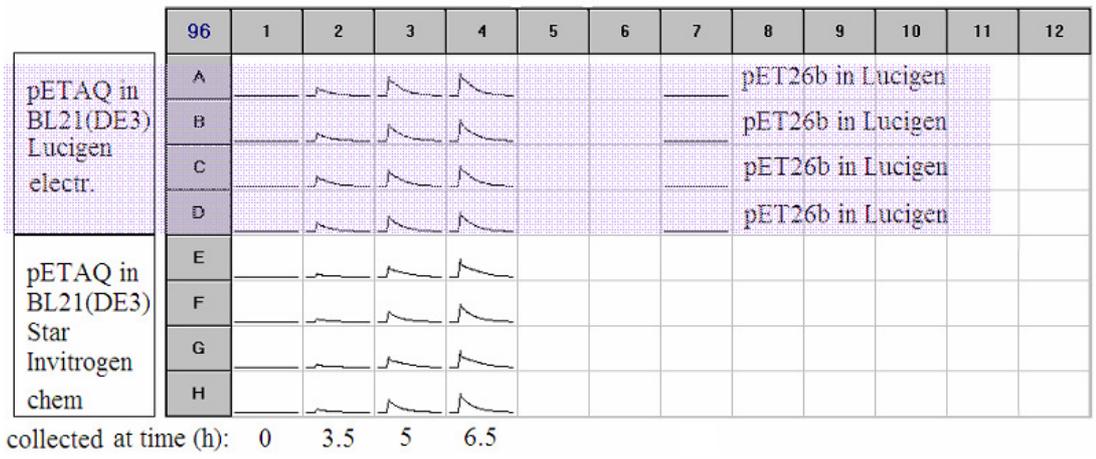


Figure 4-11 Check for background signal. Same experimental setup as Figure 4-10; Comparison of luminescence curves from Lucigen BL21(DE3) cell cultures carrying pETAQ versus those of the same cells that incorporate pET26b vector without the apoaequorin gene insert. Luminescence traces in purple show that luminescence activity is only observed in cells that encode for apoaequorin. Residual coelenterazine that is present in all experimental preparations does not cause detectable background luminescence.

This experiment was based on experimental setup B. The test was performed in order to verify that no background signal is produced by any component encoded in the host strain or expression vector pET26b (pET26b is pETAQ without the apoaequorin gene). As expected, there is no background signal; bioluminescence is not naturally observed in *E. coli*.

4.4.2.6 Activity profile across the microplate

This experiment was based on experimental setup C. Significant variation of activity across the geometry of the microwells was observed. As seen in Figure 4-12, the outer wells of the microplate exhibit significantly lower activity than the central wells of the microplate. A preliminary test (not shown here) proved that the step responsible for the variation of activity across the microplate was the overnight cold incubation step with coelenterazine and not the microwell culture. In light of the temperature-sensitive nature of coelenterazine, there was need for better temperature control during the cold incubation step.

Originally, during this step the microplate was adjusted on a table top thermomixer where temperature control is delivered through the base of the microplate holder, with cooling capabilities of up to 15 °C lower than ambient temperature, the lower attainable setting being 1 °C. In order to achieve the lowest possible ambient temperature, the thermomixer carrying the microplate is placed in a cold cabinet (4–8 °C). These settings were clearly not sufficient for achieving uniform temperature across the plate.

In order to achieve (more) uniformity of temperature (and therefore activity) across the microplate, thermal insulation was applied to the plate, so that the wells proximal to the microplate edges may retain the lowest possible temperature. Slices of polystyrene foam (thickness of approximately 0.5 cm) were carved out of polystyrene ice boxes and taped around the 96-DSW (deep square well) plate and on top of the shallow well plate that served as its lid.

As seen in Figure 4-12 and Table 4–1, uniformity of activity improved significantly with insulation of the microplate during O/N reconstitution. Excluding the outer lines and columns further reduces variation across the remaining microplate area. Improved uniformity across the plate allows maximal usage of the plate area, hence testing more mutant variants per plate and easier direct comparison between wells of the same microplate.

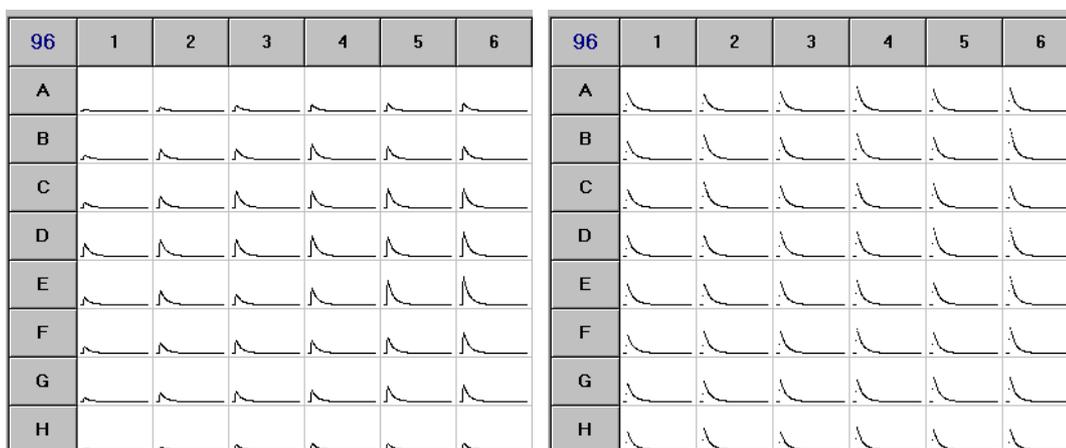


Figure 4-12 Activity profile across the microplate with and without thermal insulation. Graphical output of luminescence traces versus time from *E. coli* BL21(DE3) (Lucigen) carrying expression vector pETAQ. In each trace depicted, y-axis represents Relative Luminescence Units and the x-axis represents time. Profile across plate (left) and with insulation (right). Shown here are columns 1–6 of each microplate, symmetrical to columns 7–12 (not shown). Experiment carried out as described in Section 4.3.4 (Experimental setup C) and generic Section 2.7 (Luminescence measurements). Plater reader gain set at 2,000.

Table 4-1 Variation across the microplate with and without thermal insulation

Microplate area	Variation of activity without insulation	Variation of activity with insulation
	Relative STDEV	Relative STDEV
all wells		
Peak	54%	15%
Total light	52%	12%
all wells excluding 1 outer line and 2 outer columns		
Peak	29%	12%
Total light	27%	10%
all wells excluding 2 outer lines, 2 outer columns		
Peak	24%	10%
Total light	22%	9%

4.4.2.7 Cell growth and activity versus time

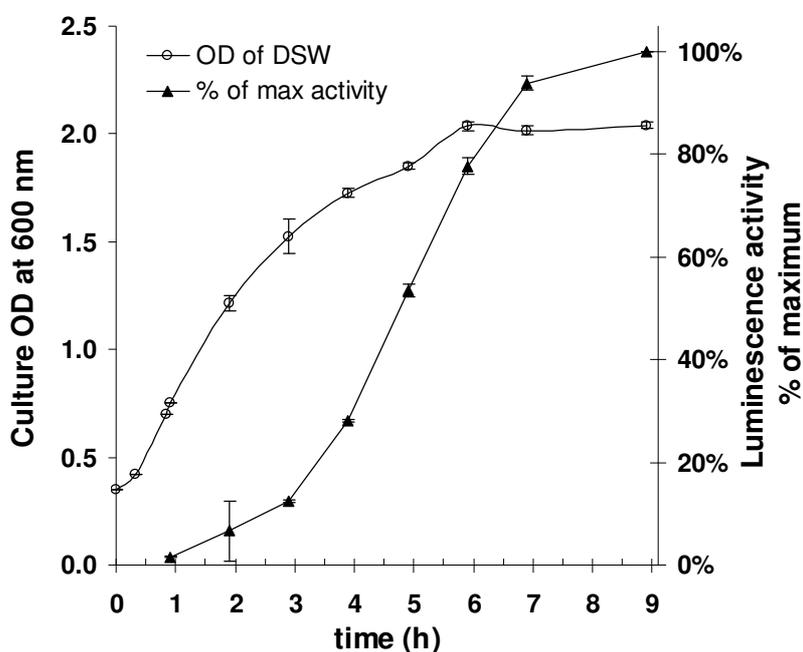


Figure 4-13 Luminescence activity and cell density of 96-DSW cultures versus time. At time zero a 10% inoculation from O/N culture was performed and IPTG induction was at time 0.9 h at OD=0.7.

This experiment was based on experimental setup C. It correlates cell growth with activity levels in the microwell cultures. Samples from various positions of the “day culture plate” were collected at time intervals and stored on ice in order to delay cell growth. At the end of a nine-hour incubation the samples were mixed with coelenterazine solution in the central wells of a “cold incubation” plate and left overnight to reconstitute active aequorin. Maximum activity was reached at 8–9 h of incubation and after OD had reached a plateau.

4.4.2.8 Optimum coelenterazine concentration

To ensure that the highest possible activities were achieved in the screening reactions, the concentration of coelenterazine needed to be optimised. Apoaequorin and coelenterazine bind at 1:1 molar ratio, but published literature has suggested an excess of coelenterazine during the cold incubation step (coelenterazine being 4–6 times or 1.2–1.3 times the calculated amount) (Shimomura and Inouye, 1999). Bacterial suspension was collected from the “day culture plate” after nine hour incubation.

Coelenterazine solutions were prepared at a range of concentrations (1–100 μM) in 20 mM Tris-HCl, 10 mM EDTA and 5 mM 2-mercaptoethanol buffer, pH 7.7. One hundred microliters of the bacterial suspension were added to 950 μL of the prepared coelenterazine solutions in central wells of a pre-chilled 96-DSW plate and left for a cold overnight incubation, as previously described. The optimum concentration of coelenterazine for the apoaequorin produced in the “day culture plate” is $\sim 10 \mu\text{M}$.

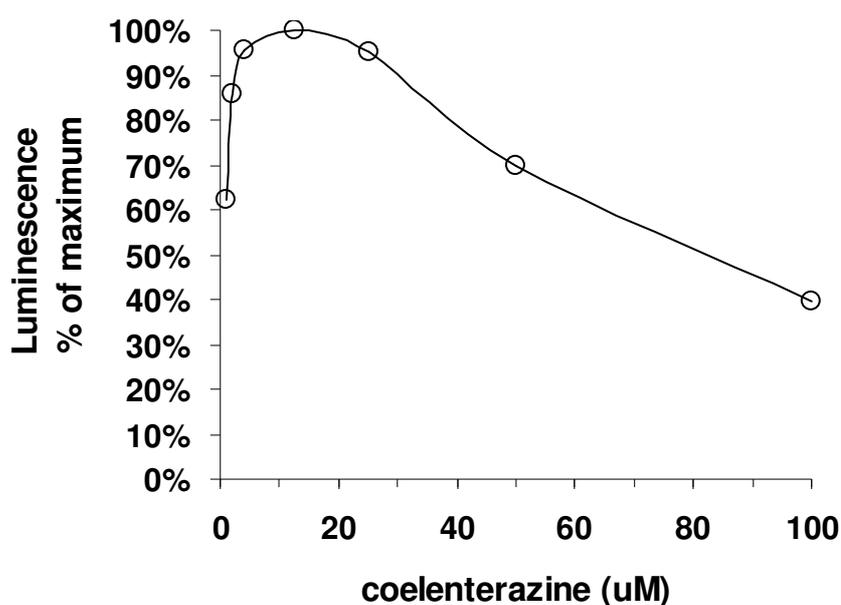


Figure 4-14 Optimum coelenterazine concentration. Shown here are average activity values of duplicates.

4.4.2.9 Injection, mixing

Flash reactions require rapid and thorough mixing, which depends upon variables such as the force, angle, and volume of injection; the geometrical relationship of the injector to the container; and the speed and completeness of mixing relative to the kinetics of reaction (Van Dyke *et al.*, 2002). The highest injection speed of the FluoStar Platerreader was used (420 $\mu\text{L/s}$) for 1:1 volume ratios of mixing in order to achieve most efficient and rapid mixing.

4.4.3 Check for cross-contamination between wells

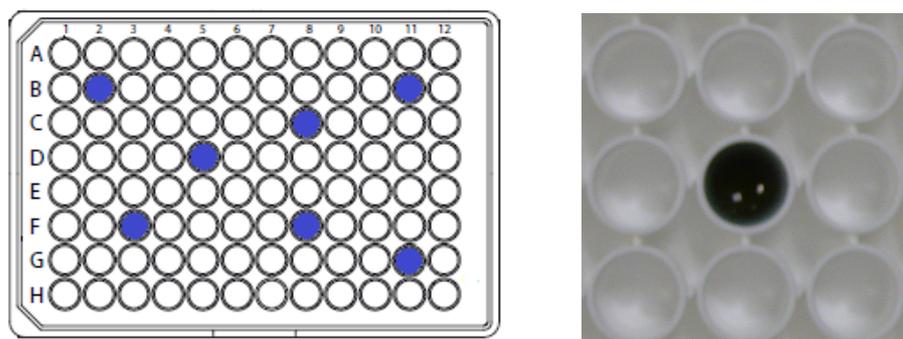


Figure 4-15 Check for cross contamination during incubation and injection. Concept of layout (left) and picture of microwell injected with loading buffer containing dye. The neighbouring wells were not affected. The same concept of testing applied to bacterial cultures in microwells.

This section checks for cross-contamination between wells in all of the stages where it might occur: automated colony picking, microwell culture, liquid handling and injection during the luminescence assay in BMG.

Figure 4-15 features a Lumitrac™ well plate for luminescence assays. The wells of the microplate were filled with 100 µL of RO water. Marked wells were injected with blue loading buffer (1 volume of SDS loading buffer to 8 volumes of water) at maximum injection speed 420 µL/s. The absorbance of the wells adjacent to the wells injected with blue dye was compared to the absorbance of RO water at 595 nm. There was no observed difference in absorbance between the measured samples, which indicates there was no spillage/transfer to neighbouring wells due to plater reader injection.

In a similar manner, tests using blue dye were performed in 96 DSW plates in order to check that 1,000 rpm shaking, liquid transfer/pipetting would not cause cross-contamination. Microwells of O/N culture plates were tested for bacterial contamination by testing OD at 600 nm: Wells containing LB nutrient and adjacent to wells inoculated with *E. coli* did not exhibit any bacterial growth.

4.4.4 Yield of aequorin in microwells

Concentration of soluble protein in the microwells was calculated using the Bradford method (Section 2.4.1.2) by adding 50 μL of 1:2 diluted Bradford reagent to 50 μL of protein sample. The total soluble protein present in the microwells of the “cold incubation step” was $\sim 26.3 \mu\text{g/mL}$, while the supernatant of a microwell at the end of the “cold incubation step” was $12.3 \mu\text{g/mL}$. Densitometry analysis of SDS-PAGE protein gels showed that the protein sample present in the cold supernatant contains approximately 78% apoaequorin, hence $\sim 9.6 \mu\text{g/mL}$ are present in the supernatant. Each microwell of the cold incubation step contains $\sim 0.4 \mu\text{M}$ aequorin. Taking into account the dilution of protein suspension in the cold incubation step by 10.5 times, $4.3 \mu\text{M}$ of functional wild-type aequorin is present in each “day culture” well. In terms of total soluble protein and taking into account the dilution of protein suspension in the cold incubation step by 10.5 times, the total soluble protein in each microwell of the “day culture” microplates was $\sim 276.2 \mu\text{g/mL}$, $\sim 36\%$ of which is apoaequorin with cleaved *pelB* signal peptide (therefore potentially available for release into the medium) and $\sim 24\%$ has retained the peptide leader sequence and will remain in the intracellular environment. Analysis of the localisation of the protein in cell suspension, cell pellets and cold incubation is available in Chapter 7, Section 7.4.2.

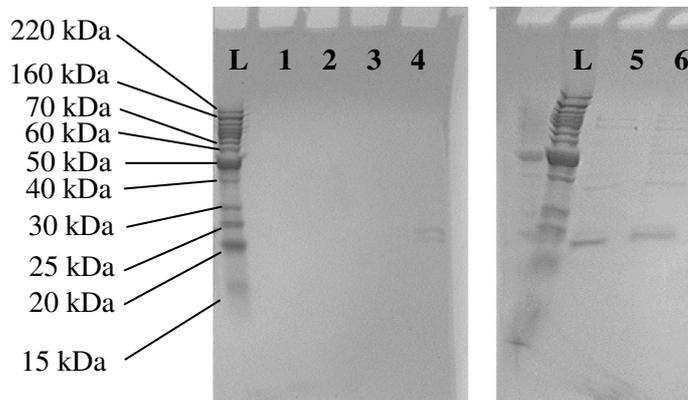


Figure 4-16 Yield of aequorin in luminescence plate microwells. L: Benchmark™ Protein Ladder (Invitrogen, Figure 2-2); Lanes 1 and 4: neat protein sample in the luminescence screening plate microwells, represents total protein in the well; Lane 2: supernatant of centrifuged sample of Lane 1, represents extracellular protein; Lane 3: supernatant after sonication and centrifugation of sample of Lane 1. This represents the total soluble protein in the well. Lane 5: same as Lane 2, concentrated 35 \times (aequorin purity: 78%); Lane 6: same as Lane 3, concentrated 26 \times (aequorin purity: 66%). 15% polyacrylamide gel. All lanes were loaded with 10 μL . Samples were prepared as in Section 2.4.2. Sample concentration was performed with a 10MWCO spin filter.

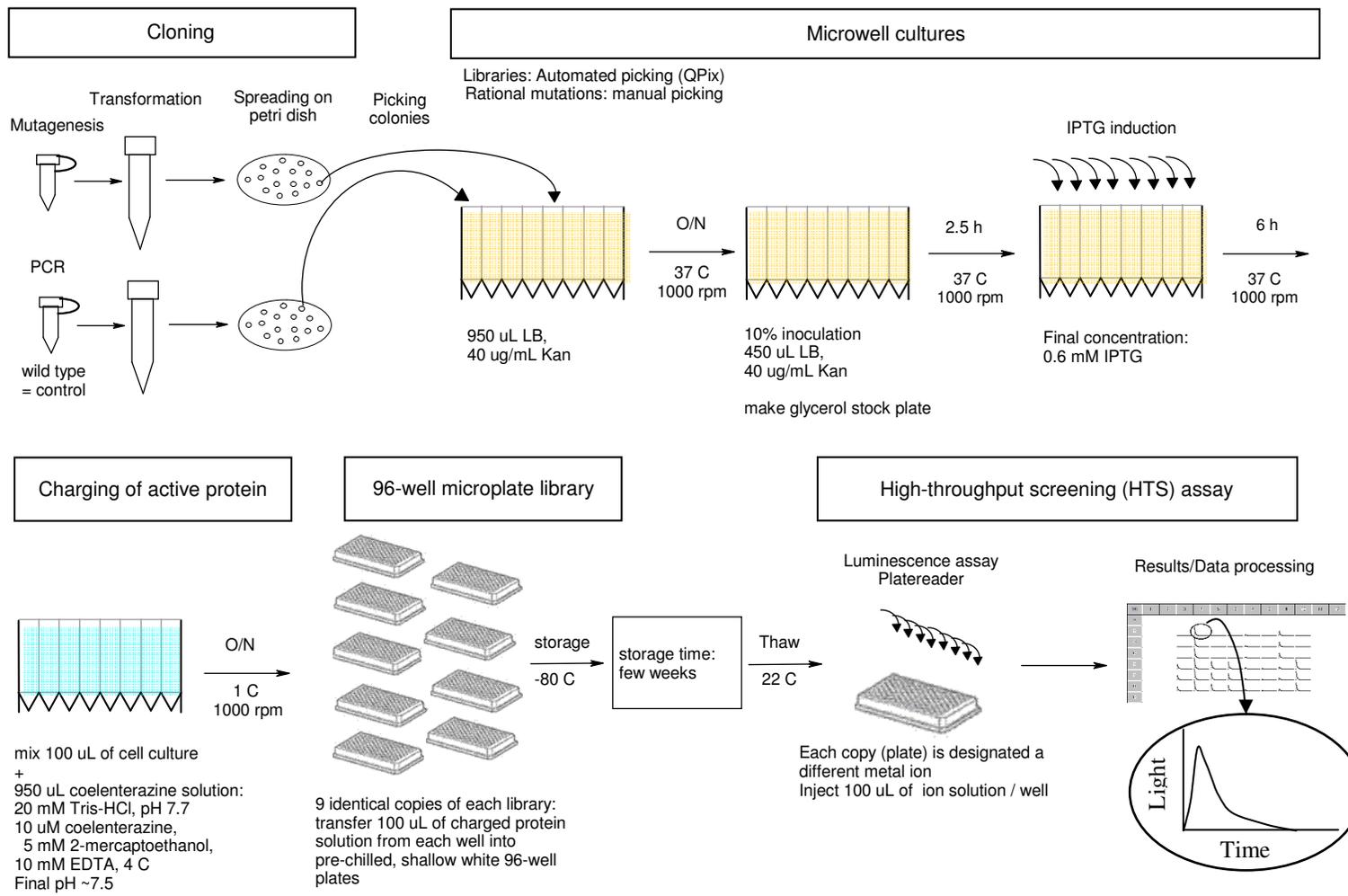


Figure 4-17 Detailed schematic of the finalised high-throughput process

4.5 Conclusions

This chapter describes the building of a high-throughput system for expression of apoaequorin in 96-microwell plates, the incubation of the apoprotein with the chromophore coelenterazine for charging of the active protein aequorin and activity screening against up to nine different potential activators (metal ion solutions). The assay described here is a practical, automated and scalable assay which allows for the screening of small to large number of mutants and displays high activity signals with zero background noise.

4.5.1 Final process

Based on the results of this experiment, the high-throughput production and activity screening of aequorin libraries is presented in the schematic of Figure 4–17 and the detailed parameters and conditions are presented in Table 4–2. The process is divided in five major steps:

- (1) cloning
- (2) microwell cultures
- (3) charging of active protein
- (4) 96-well microplate library
- (5) high-throughput screening (HTS) assay

Table 4-2 List of parameters chosen for the high-throughput assay.

Parameters	Results	Section
Vector	pETAQ	4.4.2.1
<i>E. coli</i> strain	BL21(DE3) Lucigen	4.4.2.4
Part of culture used	Bacterial suspension	4.4.2.2
Incubation time	8 h	4.4.2.7
Induction time	~1 h	4.4.2.3
Coelenterazine	10 μ M	4.4.2.8
Background signal	None at Gain 2,000	4.4.2.5
Variation across plate	Apply insulation	4.4.2.6
Injection	Max speed 420 μ L/min	4.4.2.9
Cross-contamination	none	4.4.3

4.5.1.1 Cloning and microwell cultures

Colonies of *E. coli* BL21(DE3) cells (Lucigen) carrying plasmid pETAQ (for wild-type apoaequorin) or variants thereof (for mutant apoaequorin) are picked from (50 μ g/mL) Kan+ agar plates into conical bottom 96-DSW (deep square well) plates (“seed culture plates”). Each well contains LB Kan+ medium (950 μ L, 50 μ g/mL kanamycin). Colony picking can be manual, using sterile inoculation loops, or automatic, using the QPix2 robotic colony picker, depending on whether the library consists of selected mutants or a large number of random mutants. Each individual microplate is inoculated with at least three wild-type colonies in assigned wells and these positions are kept constant throughout the entire library. These wells containing wild type will serve as internal controls and the proposed library layout is explained analytically in Section 6.2.1.

All liquid handling of microplates is automated using the Tecan Genesis robot. The top of each 96-DSW plate is covered by an inverted shallow 96-well plate (Sarstedt Inc), sealed onto the DSW plate using autoclave tape. The 96-DSW plate is then secured on a table top thermomixer (Eppendorf Inc) with shaking at 1,000 rpm, which in turn is placed in a 37 °C incubator/cabinet for the duration of the fermentation. This

method of sealing prevents evaporation, but sufficient air is trapped above each culture to permit aerobic growth.

The “seed culture plate” is incubated overnight at 37 °C and 1,000 rpm. Following the overnight incubation, the seed culture is used for a 10% inoculation into a new DSW plate containing 450 µL of Kan+ LB medium (500 µL final volume). This “day culture plate” is incubated at 37°C and 1000 rpm for 8 h. Induction with IPTG is performed at 1 h after inoculation, which corresponds to OD ~0.8.

If a library of random mutants were to be assessed, the seed culture plate would be copied into a glycerol stock plate for the later identification of interesting mutant variants by sequencing.

4.5.1.2 Charging of active protein

At the end of the 8 h fermentation, the plate is left to cool to room temperature and 100 µL of cell suspension are transferred into a pre-chilled 96-DSW plate, containing 950 µL of cold coelenterazine solution 10 µM, 20 mM Tris-HCl, 10 mM EDTA and 5 mM 2-mercaptoethanol, pH 7.7. This “cold incubation plate” is covered with an inverted shallow 96-well plate (Sarstedt Inc), insulated thermally using polystyrene foam and secured on a table top thermomixer (Eppendorf Inc) with overnight intermittent stirring at 1,000 rpm and 1 °C, which in turn is placed in a cold cabinet (4-8 °C).

4.5.1.3 96-well microplate library

After the end of the cold incubation with coelenterazine solution, Tecan robotics are used to make nine identical library copies in white shallow Lumitrac™ 96-well plates. More analytically, 100 µL from the wells of the “cold incubation plate” are transferred into the respective wells of each of the Lumitrac™ microplates. This allows for the same library to be tested against up to a maximum of nine potential activators. In practice seven potential activators were tested while the two out of the nine replicate plates were kept as backup in case a screening run required repetition (e.g. if luminescence readings were too high or too low and gain required adjustment).

4.5.1.4 High-throughput screening assay

The activity of the libraries against the potential activators is assayed using the FluoStar Plater reader (BMG Lab technologies Ltd, Bucks, UK). Light emission from

each well is measured when an equal volume (100 μL) of an activator ion solution is injected at maximum speed (420 $\mu\text{L/s}$). Light measurements started before injection (in order for any background signal to be detected) and the duration of reading was 15 s. At the Gain setting equal to 2,000, wild-type aequorin in microplates produced light counts in the scale of 10^6 against zero background signal.

Data deriving from luminescent assays performed in microplate format were exported in .xls format and copied onto template Microsoft Excel worksheets for data processing.

4.5.2 Further improvements to the method

A seamlessly automated process could be achieved with further adjustments to the setup, such as temperature control within the enclosure of the Tecan robot (for both heating to 37 °C and cooling to 1 °C). The ionic strength and pH conditions of the luminescence assay were not explored in this chapter, but they were revisited in Chapter 7 which deals with purification of the proteins. A potentially useful further development would be to explore or establish an association between the kinetics of aequorin luminescence in the crude library conditions with the kinetics of purified protein.

5 Aequorin structure and mutant library design

Aim of this chapter is to enhance the understanding of the structure of aequorin and its response to calcium ions. This knowledge will be useful in:

a) selecting an appropriate strategy for mutant library design. The envisaged variants would exhibit altered ion-binding properties. This library will be screened against new potential activator ions in Chapter 6.

b) discussion and interpretation of the library screening results presented in Chapter 6 and of further studies on purified fractions in Chapter 8.

Section 5.1 is a literature review of the detailed structural analysis published on aequorin. It builds on the general information given in Chapter 1, Section 1.3 and highlights important positions and amino acids in the protein structure and their interdependent relationship/nature. The published crystallographic structures of aequorin without calcium and apoaequorin with calcium are examined in detail in terms of the repositioning of the residues as a result of population of the EF-loops by calcium ions.

Sections 5.2 and 5.3 are a compilation of previous mutagenesis studies performed on aequorin or other calcium-binding proteins and peptides in order to study or enhance properties such as metal affinity, intensity, sensitivity and thermostability. As function of the protein is attributed to changes in its structure this compilation is a library of what has already been tried by researchers and a possible basis on how to proceed with further mutagenic studies.

In Section 5.4 the aequorin crystallographic structures are further analysed by the author, particularly relating to the metal-ligand distances within each of the EF-hands and the extent of structural rearrangement undergone by the protein in order to accommodate calcium ions.

Assuming the “native” octahedral coordination is required for optimal triggering of the bioluminescent reaction; seven metal ions were modelled by simple replacement into the coordinate position of calcium in the ion binding loops of EF-I to examine their potential for coordination in this specific protein conformation. Finally, the library of mutants is chosen in Section 5.6.3.

5.1 Aequorin: detailed structure

This section zooms into the crystallographic structures of the fully functional protein containing coelenterazine (PDB ID: 1EJ3, Figure 1-7) and the apoaequorin molecule with three calcium ions bound (PDB ID: 1SL8, Figure 1-8) shown in Section 1.3.1. The coelenterazine binding cavity is illustrated in Figure 5-1 and Figure 5-2 and the calcium ions bound within the EF-hands are shown in Figure 5-3. Coupling of the EF-hands is illustrated in Figure 5-6 and Figure 5-7. Some images were reproduced from published work and some were created in Pymol (DeLano, 2002) in order to aid in visualisation and analysis.

Numbering of amino acids in aequorin structure is according to the 1EJ3 structure numbering throughout the text, unless stated otherwise. From alignment of the amino acid sequences of 1EJ3, 1SL8 and the translated sequence of wild-type aequorin used in this thesis, the numbering of amino acids across the three sequences is:

$$\text{Numbering}_{1SL8} = 2 + \text{Numbering}_{1EJ3}$$

$$\text{Numbering}_{\text{wild-type_aequorin}} = 7 + \text{Numbering}_{1EJ3}$$

For example, Glu35 in 1EJ3 is Glu37 in 1SL8 and Glu42 in wild-type apoaequorin expressed in this work.

5.1.1 Coelenterazine-binding cavity and calcium-binding EF-hands

The hydrophobic cavity has a volume of approximately 600 \AA^3 and binds coelenterazine in the form of coelenterazine-2-hydroperoxide. The hydroperoxide group is attached at the C2 position of the ligand, as shown in Figure 5-1 and Figure 5-2 (Head *et al.*, 2000). It appears to be stabilised by hydrogen bonding to the phenolic oxygen of Tyr184 which is itself hydrogen bonded to the N ϵ 2 of His169. The imidazole of His169 is situated close to the carbonyl oxygen of C3 on the ligand and to the indole of Trp173 (Head *et al.*, 2000).

The following side-chain interactions in the binding cavity may position coelenterazine in the binding site and contribute to the mechanism of aequorin action: (1) the phenolic OH of Tyr132 hydrogen bonds to N1. It is also linked through hydrogen bonds by a water molecule to His58. This histidine imidazole is adjacent to the indole ring of Trp108 which partly overlays the imidazopyrazinone ring system,

coming as close as 3.6 Å; (2) The *p*-OH of the phenol attached at C6 of coelenterazine is positioned at the centre of a triangle made by the third set of Tyr–His–Trp, consisting of the phenolic oxygen of Tyr82, the Nε1 of His16 and the Nδ1 of Trp86; (3) The *p*-OH group on the benzyl substituent at C2 of the ligand is hydrogen bonded to a water molecule which itself interacts with the carbonyl oxygen of Ile105 and the side-chain oxygen of Thr166 (Figure 5-1 and Figure 5-2).

Tyr184 is part of an extended loop structure at the C-terminal of aequorin (residues 177-189, numbering according to 1EJ3) which lies in a space between the first helix of EF hand I and the first helix of EF hand IV and closes the coelenterazine-binding cavity. Numerous hydrogen bonds are formed in this region anchoring the C-terminal chain in its conformation and positioning the side chain of Tyr184 adjacent to both His169 and the ligand. Some of these hydrogen bonds also couple the C-terminal chain with both of the flanking helices (Head *et al.*, 2000).

The calcium-loaded apoproteins retain the same compact scaffold and overall fold as the unreacted photoproteins containing the bound substrate, 2-hydroperoxy-coelenterazine (see Section 5.5). Binding of Ca²⁺ into the loops of the EF-hands leads to subtle structural shifts in the photoprotein that destabilise hydroperoxy-coelenterazine and lead to decarboxylation of the hydroperoxide and rapid emission of bioluminescence (Liu *et al.*, 2006).

Figure 5-3 shows a close up of each of the calcium binding EF-hands coordinating a calcium ion. Metal-ligand distances were measured using Pymol (DeLano, 2002). NMR and X-ray data by Ohashi and co-workers showed that the loop conformation of EF-IV exhibits more structural flexibility than EF-I and EF-III (Ohashi *et al.*, 2005).

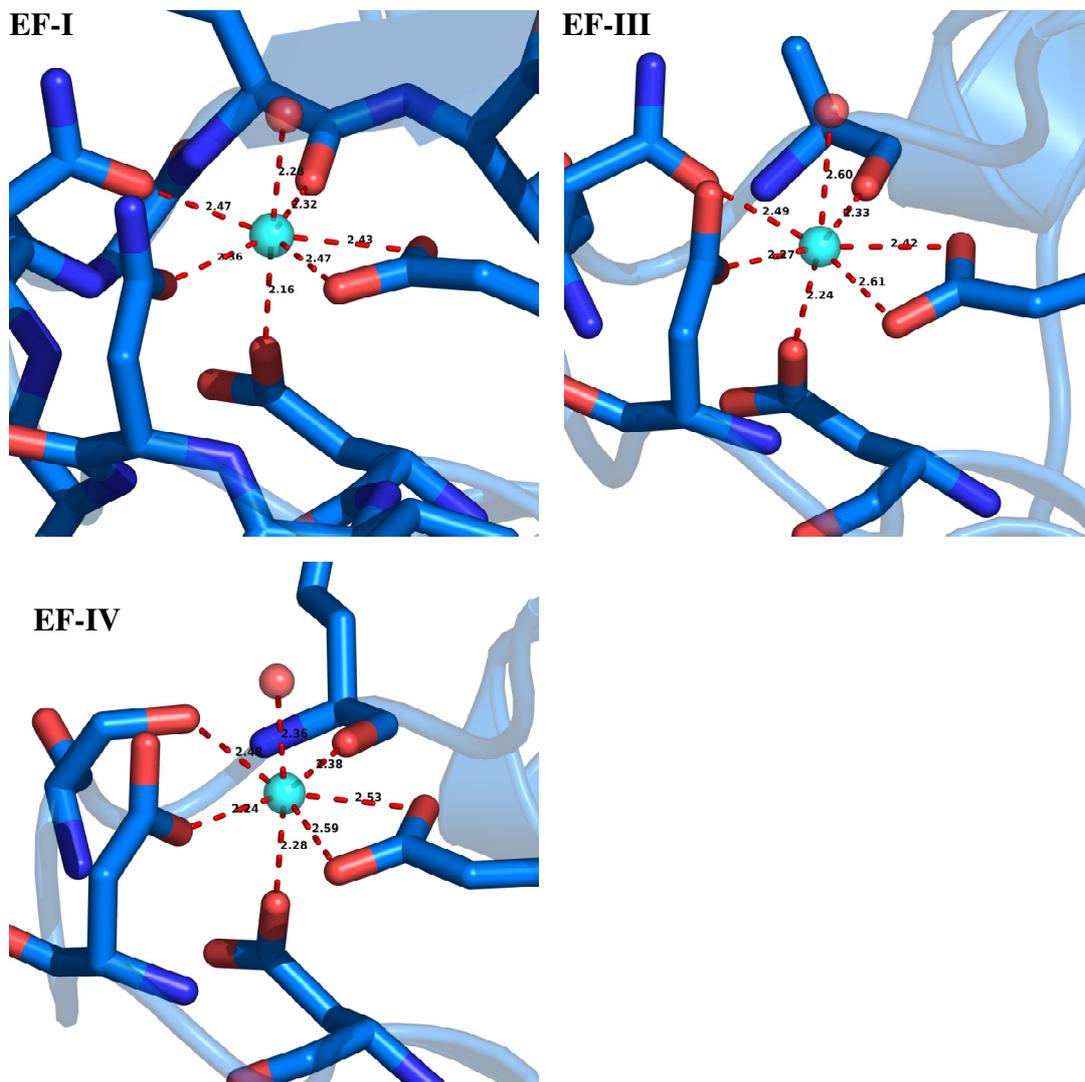


Figure 5-3 Calcium binding in EF-I, EF-III and EF-IV hands of aequorin. One calcium ion (cyan coloured sphere) is bound in pentagonal bipyramidal configuration. Coordinating oxygens from residues at positions 3,5,7 and 12 form a planar pentagon and position 1 and one oxygen from a water molecule (hydrogen bonded to position 9) are axial coordinating residues. Picture rendered and interatomic distances between calcium ions and each coordinating oxygen atom were calculated using Pymol (DeLano, 2002).

5.1.2 H-bond coupling of EF-hands

Deng and co-workers (2005) described the extensive hydrophobic interactions between the two sets of EF-hand pairs EF-I with EF-II and EF-III with EF-IV and the effect of calcium binding to the coupling (Deng *et al.*, 2005).

The loops of EF-I and EF-II are bound by means of hydrogen bonds between main-chain nitrogen and carbonyl oxygen atoms of Ile31 and Thr77, the side-chain O δ 1 of Asp34, and the main-chain nitrogen atom of Gly74, and between N ζ of Lys30 and the side-chain oxygen atom of Glu76 (Figure 5-4) (Deng *et al.*, 2005).

The binding of calcium ion in loop I of aequorin abolishes the hydrogen bond between Lys30 and Glu76. The accommodation of calcium in the loop makes the hydrogen bond distances between the main-chain nitrogen atom of Ile31 and the main-chain carbonyl oxygen atom of Thr77, and the main-chain carbonyl oxygen atom of Ile31 and the main-chain nitrogen atom of Thr77 shorter. The hydrogen bond distances between O δ 1 of Asp34 and the main-chain nitrogen atom of Gly74 that bind helices B and C of aequorin increase in distance upon calcium binding.

The interaction between the loops of EF-III and EF-IV occurs mainly through hydrogen bonds between main-chain atoms of Ile124 and Leu160 and between the main-chain carbonyl oxygen atom of Gly122 and the main-chain nitrogen atom of Val162 (Figure 5-5). In the calcium-loaded state of aequorin, the hydrogen bond distances between the main-chain nitrogen atom of Ile124 and the main-chain carbonyl oxygen atom of Leu160, and the main-chain carbonyl oxygen atom of Ile124 and the main-chain nitrogen atom of Leu160 become shorter than in the calcium-free state. The hydrogen bonding between Gly122 and Val162 is abolished after calcium binding (Figure 5-5). Gly122, Ile124 occupy positions 6 and 8 of loop EF-III and Leu160 and Val162 occupy positions 8 and 10 of the loop, respectively.

Within van der Waals interaction range (3.6–4.5 Å) from the residues participating in H-bond coupling between the EF-hand loops lie residues which directly coordinate coelenterazine or are within close proximity (Figure 5-6). Tyr82 (directly coordinates coelenterazine) and Met19 (close proximity) are within the range in the N-terminus EF-hand pair and Thr166 (directly coordinates coelenterazine and is in close proximity to His169 and Trp129, Phe113, Met165 and Trp108).

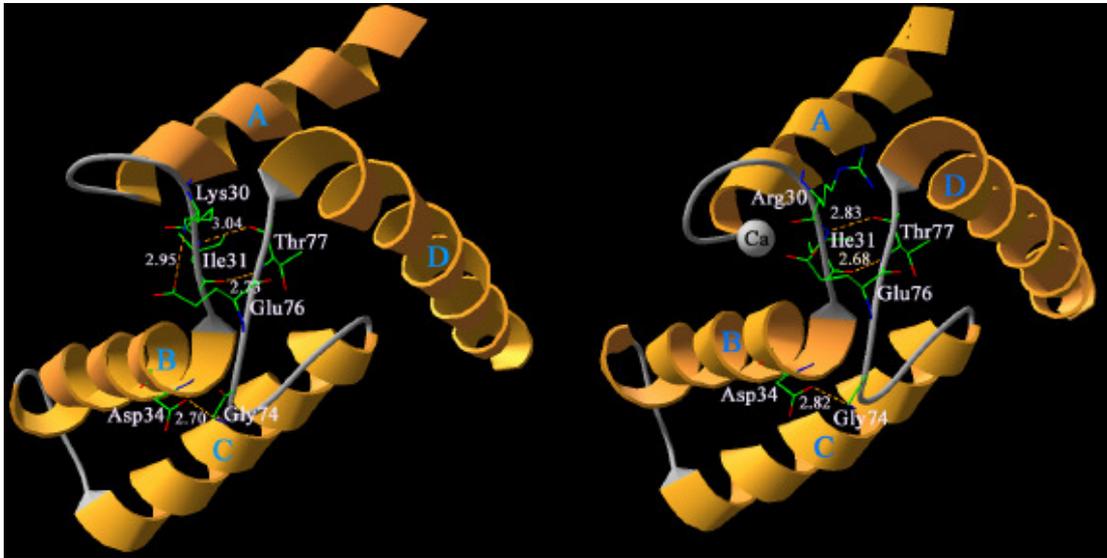


Figure 5-4 Coupling of EF-I and EF-II of aequorin by means of hydrogen bonds. The left hand side shows the calcium-free conformation and the right hand side the Ca^{2+} -bound conformation. Image reproduced from Deng *et al* (2005).

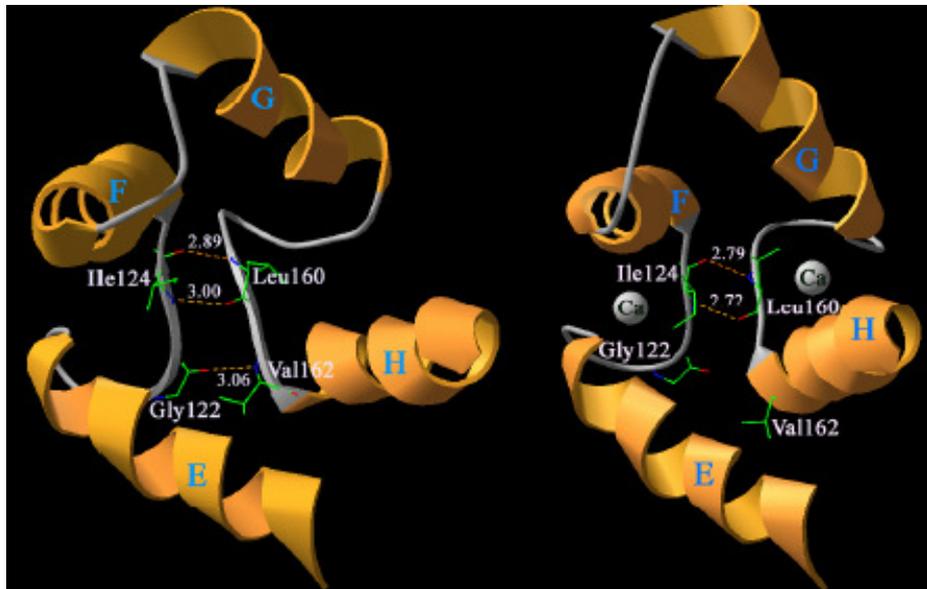


Figure 5-5 Coupling of EF-III and EF-IV of aequorin by means of hydrogen bonds. The left hand side shows the calcium-free conformation and the right hand side the Ca^{2+} -bound conformation. Image reproduced from Deng *et al* (2005).

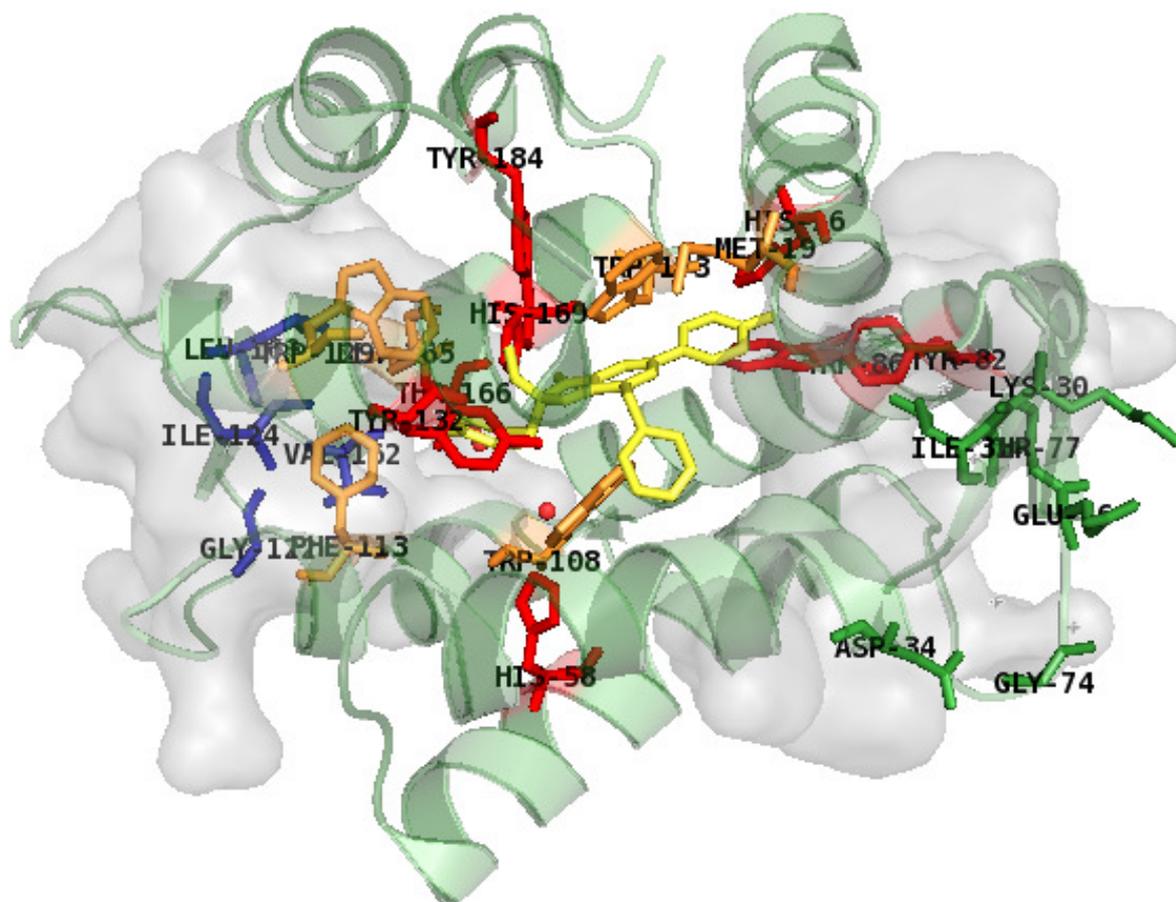


Figure 5-6 EF-hand pair coupling residues, van der Waals distance range and cavity of aequorin. The EF-I-EF-II hydrogen bonding residues are shown as green sticks, the EF-III-EF-IV hydrogen bonding residues are shown as dark blue sticks. In the protein core coelenterazine is shown as yellow sticks. As in Figure 5-2, red sticks are amino acids that directly bind to coelenterazine and orange sticks are in close proximity of the chromophore. The grey transparent surfaces indicate regions of van der Waals interactions with the amino acids involved in EF-hand coupling hydrogen bonds. Some of the amino acids which stabilise coelenterazine (Tyr82, Met19, Thr166, Trp129, Phe113, Met165, Trp108) lie within these regions. Image created using Pymol (DeLano, 2002).

5.1.3 Theories on signal transduction following calcium binding

The mechanism of signal transduction upon calcium binding is still under discussion and the exact number of Ca^{2+} ions required for bioluminescence is discussed in Section 1.3.1. The holo-aequorin and apoaequorin structures provide some insights. The sequence of molecular events occurring in a photoprotein molecule has been hypothesised based on structural data and the assumed mechanism for triggering the bioluminescence.

When calcium binds at either or both EF-I and EF-IV the helices of these ‘hands’ change their relative orientations. Displacement of the helices flanking the C-terminal tail disrupts the local hydrogen-bonding network, resulting in a relocation of the side chain of Tyr184 (see Figure 5-10). This in turn disrupts the hydrogen bonds to His169 and the peroxide. No longer stabilised, the peroxide would be free to attack the adjacent carbon C3 (Figure 5-2) to initiate the light-emitting reaction (Figure 1-11). Depending on the extent of the shift in the flanking helices during calcium activation, the C-terminal ‘tail’ could become partly or completely uncoupled from the helices (see Figure 5-10), thus affecting the rate of bioluminescence (Head *et al.*, 2000).

A concurring hypothesis by Deng and co-workers (2005) suggests that displacement of His169 that lies in helix H is a crucial step for triggering the bioluminescence (Figure 5-11). Binding of one Ca^{2+} to the loop of EF-IV preceding helix H, and needing a significant repositioning of the coordinating residues to properly accommodate the Ca^{2+} , would propagate to a repositioning of His169. Due to the existence of hydrogen-bonding networks and other types of interaction between the EF-hand motifs (Section 5.1.2), particularly between loops III and IV, binding of Ca^{2+} to each of the loops will not be independent (Deng *et al.*, 2005). A possible sequence of molecular events has been proposed to connect binding of calcium to the triggering of bioluminescence in photoprotein obelin and a similar cascade of events is expected for aequorin.

The events proposed are as follows, with residues numbered according to the aequorin sequence PDB ID 1EJ3:

- (1) One calcium ion is preferentially bound to loop I as observed in the soaked obelin structure (Liu *et al.*, 2003) and expected from the “preformed” nature of this loop itself (Deng *et al.*, 2005).

(2) The binding of calcium ion to the EF-I loop and optimisation of the pentagonal, bipyramidal geometry (Figure 5-3) produces a “twist” of the EF-I around a pivot point by means of changes in hydrogen bond distances between the main-chain atoms of Ile31 and Ile77. The accommodation of the calcium ion also changes slightly the inter-helical angle between helices A and B.

(3) All these changes produce a pulling of helix A in the direction of the N-terminus of the protein.

(4) Since helix A is tightly bound with helix H and the C-terminus through numerous hydrogen bonds, the changes in helix A will result in a displacement of helix H and the C-terminus. For this reason, the binding of only one calcium ion into the Ca^{2+} -binding loop of the EF-I could be sufficient to trigger bioluminescence.

(5) The binding of a calcium ion to loop of EF-IV and the optimisation of the pentagonal, bipyramidal geometry produces a “twist” of the EF-IV around a pivot point by means of changes in hydrogen bond distances between the main-chain atoms of Ile124 and Leu160. A small change of inter-helical angle is induced between helices H and G.

(6) The displacement of helix G produces a rearrangement of helix F, which is hydrogen-bonded with helix G. The displacement of helix F can therefore adjust the Ca^{2+} -binding loop of EF-III, increasing its affinity for calcium and facilitating its binding. The accommodation of calcium into this Ca^{2+} -binding loop completes the rearrangements of helices F and E of EF-III and again leads to an additional stimulation of the bioluminescence. The accommodation of the third calcium ion and the rearrangements in EF-III complete all the structural rearrangements in the photoprotein molecule, producing a final conformation that is optimal for effective bioluminescence (Deng *et al.*, 2005).

Based on this sequence of events, Deng and co-workers (2004) and Vysotski and Lee (2004) have suggested that binding of even one calcium ion in EF-I will be enough to set off bioluminescence and that binding of the other two calcium ions is a co-operative event leading to greater stimulation of bioluminescence. In terms of binding preference, Liu and co-workers (2003) found that after exposing obelin crystals to a trace of calcium, one calcium ion was bound in the loop of EF-I and according to Deng and co-workers (2004) and Vysotski and Lee (2004) the second calcium ion binds to the loop of EF-IV.

The likelihood that each binding event will trigger a bioluminescence response depends on three factors: (1) the on-rate for Ca^{2+} binding, which is probably proportional to the binding affinity; (2) the degree of residue shift for Ca^{2+} ligation and the rate of propagation of this change to helix H and 3) the amount of any position shift that His169 undergoes as a result (Deng *et al.*, 2005).

5.2 Previous mutagenesis on aequorin

After an in-depth study of the protein structure, this section is dedicated in compilation of some interesting mutational studies conducted on aequorin. Selected mutations are also presented in Table 5-1.

5.2.1 Calcium sensitivity

It is possible to alter the sensitivity of aequorin's response to trace amounts of calcium by changing the chromophore attached to the protein chain, rather than by altering the protein itself. Analogues of coelenterazine were previously synthesised to produce semi-synthetic aequorins, namely cp-, i-, br- and n-aequorin (Table 5-2 and Figure 5-8) taken after the names of the chromophore moieties. Crystal structures of their complexes with aequorin at resolutions of 1.6–1.8 Å show that their overall structures were almost identical to native aequorin, though some significant differences were found in the interactions between the substituents of the coelenterazine moiety and the amino acid residues in the binding pocket. The differences of various semi-synthetic aequorins in Ca^{2+} -sensitivity and reaction rate are explained by the capability of the involved groups and structures to undergo conformational changes in response to the Ca^{2+} -binding (Toma *et al.*, 2005). Depending on the bulk of their C2-substitutions, the coelenterazine analogues can interfere with or increase the conformational freedom of the protein to promote the light-emitting reaction, with implications on the light yield and rate of light production (Toma *et al.*, 2005). These results add to previous work by Shimomura and co-workers, (1993) who created semi-synthetic aequorins using 15 different chromophore analogues, which also displayed a wide range of Ca^{2+} -sensitivity.

Shifts of aequorin sensitivity were also achieved through mutations in the protein chain. This work is presented in the next section (Section 5.2.2).

5.2.2 Impaired EF-hands / central Gly→Arg mutations

From mutational studies in each of the three binding sites of aequorin, it has been elucidated that not all of the EF-hands share the same importance in calcium-triggered luminescence. To determine the importance of each of the Ca²⁺-binding sites, amino acid substitutions were made at the three EF-hand loops.

Tsuji and co-workers (1986) targeted the highly conserved central glycine at position 6 of the EF-hand loop (see Section 1.3.4 and Figure 1-10). This site was independently mutated to arginine at each of the three functional EF-hands in aequorin. Mutation of the third calcium binding site (EF-IV) gave no significant loss of activity relative to the wild type, suggesting that the binding of Ca²⁺ to this site may be unnecessary for light emission. By contrast, the same mutations directed to EF-I and EF-III, significantly reduced luminescence activity by 100% and 50% respectively (Tsuji *et al.*, 1986). Therefore, the authors concluded that two calcium ions may be sufficient for initiating the bioluminescence reaction.

In a similar study, Tricoire and co-workers (2006) created aequorin mutants with the first or last residue of the Ca²⁺-binding loops (an Asp or Glu) replaced by Gly; these substitutions removed oxygen ligands essential for calcium coordination (Tricoire *et al.*, 2006). Interestingly, the sensitivity of aequorin towards calcium increased for mutants with impaired EF-I and decreased for mutants with impaired EF-III and EF-IV. Among the mutants with only one EF-hand intact, sensitivity for calcium decreased for EF-I⁺ EF-III⁻ EF-IV⁻ and increased for EF-I⁻ EF-III⁺ EF-IV⁻ and with EF-I⁻ EF-III⁻ EF-IV⁺ suggesting that EF-I has lower affinity for calcium (Tricoire *et al.*, 2006).

The two studies appear at odds with each other, which may be a result of indirect effects due to the different mutational strategies as well as different types of reported outputs; Tsuji and co-workers measured intensity but not sensitivity whereas Tricoire and co-workers (2006) presented sensitivity but no intensity results. However, the latter did publish half-times of light decay curves for each mutant. Coupled with the premise that the total light remained constant, the half-life of the decay curves could provide an indication of relative intensity amongst the mutants: higher values of half-life would suggest lower initial light intensity. Based on this premise, mutations which removed essential oxygen ligands in EF-I resulted in greater loss of initial intensity than similar mutations in EF-III and EF-IV. It is also worth noting that mutations at different sites

within the loop yielded varied half-life times, e.g. Asp117Gly at position 1 of the EF-III loop yielded a longer half-life decay (28.6 ± 1.3 s) than Glu128Gly at position 12.

Even though both mutational strategies aim to impair EF-hands it is worth considering the nature of the substitutions and their potential effect on protein function.

The central glycine of the loop at position 6 is highly conserved in most calcium-binding proteins and in all known photoproteins (see Section 1.3.4). A survey on the architecture of metal coordination groups showed that glycine has a particular significance in calcium coordination groups, especially in the position adjacent to a donor residue (Harding, 2004). It is found adjacent to donor residues more frequently (20% probability) than random statistics would predict (6.9% probability); sometimes it provides a 'turn' in the protein-chain direction and at high coordination numbers (e.g. calcium binding) or small chelate loops its small size may be helpful in allowing the protein chain to make the required coordinate links to a metal atom (Harding, 2004).

It is reasonable to assume that mutation of central glycine to a longer and more bulky arginine introduces rigidity in a position which is flexible throughout the family, thus affecting ion binding, signal transduction or both. Replacing calcium-coordinating Glu or Asp with Gly replaces conserved binding residues, with a small and flexible glycine which is unlikely to be sterically perturbing, so the protein can possibly "work around it" using its remaining oxygen ligands. However it does remove one (Asp→Gly) or two coordinating oxygen atoms (Glu→Gly) from the loop, a loss expected to decrease the loop's affinity to calcium ion. Lack of binding at the entrance and exit of the loop (positions 1 and 12) would limit the response of the loop upon calcium binding and its effect on the entering and exiting helices, which in turn serve to transduce the signal (or binding effect) to the rest of the molecule (Section 5.1.3).

As the EF-hands are coupled in pairs and "communicate" with the coelenterazine binding site through a hydrogen bond network (Sections 5.1.2 and 5.1.3), both strategies can cause structural disruption beyond the loop and hence disrupt the signal transduction to the protein core.

5.2.3 Spectral shift

Rowe and co-workers (2008a) developed a method for discriminating two semi-synthetic aequorin variants (each incorporating a different coelenterazine analogue) from one another using time resolution. Two aequorin variants were paired with different coelenterazine analogues and their signals were resolved from one another using the difference in decay kinetics and half-life times. The resulting aequorins were used to develop simultaneous, dual-analyte, single-well immunoassays.

Taking this approach one step further, Dikici and co-workers (2009) changed the emission characteristics of aequorin by mutating the aequorin protein chain as well as by pairing the mutants with ten different coelenterazine analogues to yield semi-synthetic aequorins. The mutations targeted residues of aequorin known to play a role in the light emitting reaction (residues His16, Met19, Tyr82, Trp86, Trp108, Phe113 and Tyr132). The result was a set of semi-synthetic photoprotein mutants with significantly altered bioluminescent properties: emission wavelengths, decay kinetics, and stability (Dikici *et al.*, 2009).

5.2.4 Intensity

In vitro evolution experiments by Tsuzuki and co-workers (2005) obtained high intensity yielding aequorin mutants. Mutations neighbored the His16 or His169 coelenterazine-binding residues or were located in the first EF-hand (amino acid numbering according to PDB ID: 1EJ3). In particular, high intensity mutants were Lys17Arg Val25Ala, Asn26Asp, Gln168Arg and Leu170Ile. Other mutations located in the EF-hand loops or their vicinity resulted in lower peak intensities and slow decay times. A prevailing characteristic of these mutants was removal of at least one oxygen donor, e.g. Glu35Gly, Asp117Gly, Glu128Gly, and Asp135Gly.

5.2.5 Thermostability

The screen described in Section 5.2.4 produced two aequorin mutants with increased thermostability (Gln168Arg and Leu170Ile) and one mutant (Phe149Ser) with decreased thermostability (Figure 5-7). Two random libraries were created based on Gln168 and Leu170 respectively and screened for the impact of these positions on thermostability. The experimental results in conjunction with crystal structures of aequorin in published literature suggested that both Phe149 and Gln168 fulfill a dual purpose: they stabilise the coelenterazine peroxide and contribute to triggering of light by linking EF-III to the coelenterazine-binding residues Trp129 and His169 (Tsuzuki *et al.*, 2005).

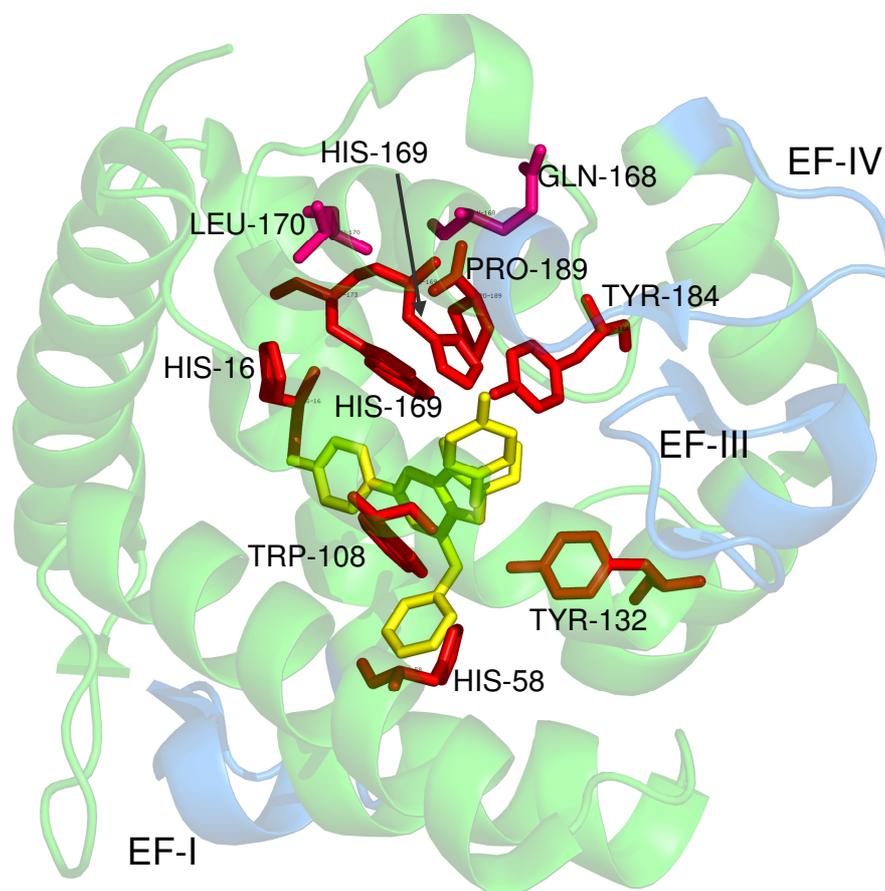


Figure 5-7 Aequorin and positions of interesting mutations. Three-dimensional representation of aequorin with bound coelenterazine (yellow sticks). The Ca^{2+} -binding loops are transparent blue and the rest of the protein is transparent green. Amino acids that were found to contribute to the bioluminescent activity and/or stabilisation of bound coelenterazine are highlighted as red sticks (His16, His58, His169, Trp108, Trp173). Amino acids that have been linked to thermostability are highlighted as magenta sticks (Gln168, Leu170). PDB 1EJ3, 3D model rendered with Pymol (DeLano, 2002).

Table 5-1 Previous mutational work on aequorin. Selected literature.

Type of alterations	Location in aequorin	Property altered	Effect on activity	References
Q168R, L170I	Hydrophobic core	Thermostability	↑	(Tsuzuki <i>et al.</i> , 2005)
F149S	Hydrophobic core	Thermostability	↓	(Tsuzuki <i>et al.</i> , 2005)
Coelenterazine analogues	Hydrophobic core	Calcium sensitivity	↑, ↓	(Shimomura <i>et al.</i> , 1993, Toma <i>et al.</i> , 2005)
Central Gly→Arg	Centre of EF-I loop	Flexibility	↓	(Tsuji <i>et al.</i> , 1986)
Central Gly→Arg	Centre of EF-III loop	Flexibility	↓	(Tsuji <i>et al.</i> , 1986)
Central Gly→Arg	Centre of EF-IV loop	Flexibility	↓	(Tsuji <i>et al.</i> , 1986)
Fusion with 6xHis	N-terminal	Metal affinity tag	–	(Glynou <i>et al.</i> , 2003)
Fusion	N- and/or C-terminal	Metal affinity tag	–	(Deo <i>et al.</i> , 2001, Lewis and Daunert, 2000)
Deletion of Pro	C-terminal	Light emission, stability	↓	(Nomura <i>et al.</i> , 1991, Watkins and Campbell, 1993)

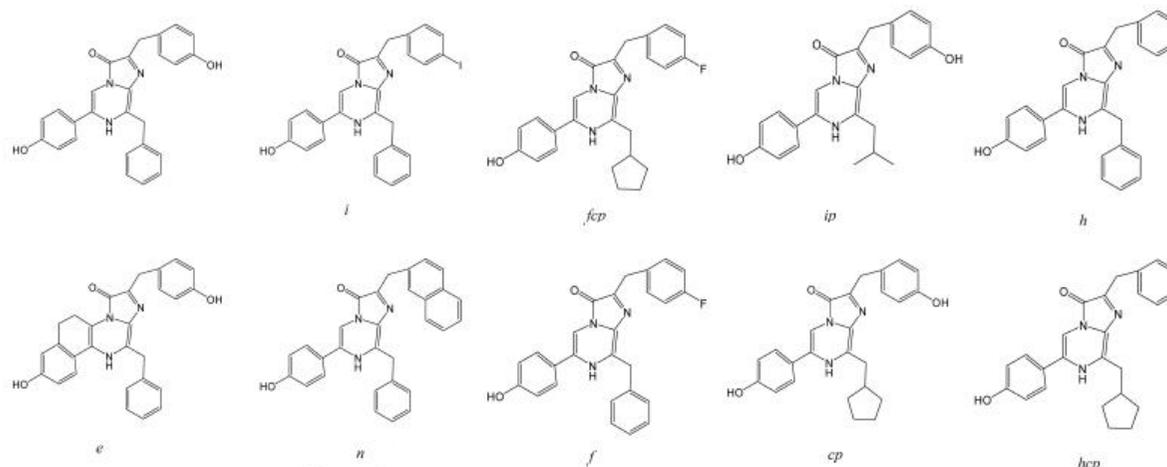


Figure 5-8 Coelenterazine and nine analogues. Different coelenterazine analogues were incorporated in mutant variants of photoproteins aequorin and obelin to create photoproteins with altered light emission characteristics. Figure reproduced from Rowe *et al* (2008b).

Table 5-2 Properties of aequorin with coelenterazine and some of its analogues. Data taken from Shimomura and co-workers (1993).

Coelenterazine Analogue	Emission maxima (nm)	Relative Luminescence Capacity ¹	Relative Intensity ²	Half-Rise Time ³ (ms)
native	466	1	1	6–30
cp	442	0.63	28	2–5
f	472	0.8	20	6–30
h	466	0.75	16	6–30
hcp	445	0.65	500	2–5
n	468	0.25	0.15	6–30

1: Relative luminescence capacity = total time-integrated emission of aequorin in saturating Ca^{2+} relative to native aequorin = 1.0 2: Relative intensity at 100 nM Ca^{2+} . 3: Half-rise time = time for the luminescence signal to reach 50% of the maximum after addition of 1 mM Ca^{2+} to a standard of aequorin reconstituted with the coelenterazine analogue of interest.

5.3 Mutation of other calcium-binding structures

Valuable information on the metal-binding requirements of EF-hands has come from studies with synthetic peptides, *de novo* proteins and even non-EF-hand proteins that have calcium coordination similar to that of the EF-hands. Such systems come with benefits such as lack of the complexities associated with cooperativity, engineered tyrosines within or near the loop for fluorescence monitoring and use of terbium instead of calcium for its phosphorescence properties.

5.3.1 Short peptides with calcium-binding ability

Researchers have attempted to study the ion binding affinities of the EF-hands isolated from the complex protein scaffold. Le Clainche and co-workers (2003) created 33-34 amino acid peptides corresponding to the helix-turn-helix EF-hand motif of the calcium binding site I from *Paramecium tetraurelia* calmodulin. Isolated from the rest of the protein these peptides were able to acquire native-like conformations due to the incorporation of a disulfide bond to bridge the two helical regions. The disulfide-stabilised peptides acquired a calcium-dependent helical conformation and native-like affinity for calcium and lanthanide ions (Le Clainche *et al.*, 2006, Le Clainche *et al.*, 2003). Their work showed that metal selectivity of the calcium binding loop can be modulated by specific mutations in metal coordinating and non-coordinating positions.

Replacement of Glu by Asp at position 12 in the calcium binding loop resulted in loss of calcium affinity but preserved lanthanide affinity. The mutation is presumed to leave more space in the binding loop for larger ions. The presence of carboxylate oxygens at positions 1, 3, 5 and 12 of the loop allowed for high affinities for terbium and particularly the presence of Asp at position 5 of the loop was decisive in maintaining high affinity for terbium.

Interestingly, while at position 3 Asp→Asn abolished calcium affinity of the peptide system, the equivalent sites in aequorin are normally Asn. However this may indicate the effects of the neighbouring sequence, including the different structural constraints on the loop for the small peptide such as the impact of the engineered

disulfide bond and also the introduction of a tyrosine residue at position 7 (used as a fluorescent probe of metal binding in the peptides).

Alanine substitutions at non-coordinating residues did not abolish calcium binding to the peptide, but significantly reduced the affinity, as did alanine substitutions in the helix preceding the loop.

Finally, Oishi and co-workers (1992) investigated the affinities of EF-hands I, III and IV for Ca^{2+} by using 20-22 amino acid synthetic peptide fragments. The dissociation constants for Ca^{2+} showed the binding affinity order of III, I and IV (Oishi *et al.*, 1992).

5.3.2 EF-hand loops in model proteins

A study by Drake and co-workers (1996) examines the contribution of the ninth position of the Ca^{2+} -coordinating EF-loop of the *E. coli* galactose binding protein to the tuning of calcium affinity and selectivity. It was proposed that position 9 (see Section 1.3.4 Figure 1-10) serves in: (1) charge selectivity, which may stem from the electrostatic repulsion between the coordinating oxygens and (2) size selectivity, which may involve complex interactions between multiple coordinating side chains.

Position 9 of the EF-hand was substituted with residues commonly found in other EF-hand loop sequences. The Ca^{2+} affinities and ionic selectivities of the new loop variants were compared with those for ions of groups Ia, IIa, and IIIa of the Periodic Table and for the lanthanides. Variants with neutral side chains of different sizes (Gly, Ala, Ser, Thr, Asn, Gln) exhibited similar affinities for calcium ions and excluded Na^+ , K^+ , and Mg^{2+} . Acidic residues (Glu, Asp) reduced affinity for calcium and significantly enhanced affinity toward trivalent cations. All mutants exhibited a partial loss of ion size selectivity and favoured lanthanide binding. In a relevant study, Drake and Falke (1996) used the same EF-hand model to propose that position 9 of the EF-hand loop serves as a “gateway” to modulate the kinetics of Tb^{3+} binding.

A similar study was carried out for the *E. coli* receptor for D-galactose and D-glucose (GGR), which contains one single Ca^{2+} -binding site. The effects of metal ion size and charge on the affinity (K_d) of metal binding was studied for spherical ions from groups IA, IIA, IIIA of the periodic table and some of the lanthanides (Snyder *et al.*, 1990). GGR was at the time of publication structurally the closest known relative of the

EF-hand class. It has the significant experimental advantage that only one Ca^{2+} -site is present in the molecule, hence was considered ideal for a systematic study of Ca^{2+} -site affinity.

Monovalent ions exhibited very low affinities. Divalent group IIA exhibited affinities related to their size with optimal binding at an effective ionic radius between those of Mg^{2+} (0.81 Å) and Ca^{2+} (1.06 Å) (Snyder *et al.*, 1990). Comparison of the dissociation constants for the binding of spherical ions from groups IA, IIA and IIA and the lanthanides indicated that both charge and size were important parameters in determining the specificity of the GGR Ca^{2+} -binding site, with monovalent ions being excluded on the basis of charge.

5.4 Structural analysis methods

5.4.1 Crystallographic structure examination

The structures of calcium-free aequorin (PDB ID: 1EJ3) and calcium-bound apoaequorin (PDB ID: 1SL8) were visualised in Pymol (DeLano, 2002) and used to calculate distances between the centres of the metal ion and coordinating oxygen atoms. The upper limit for hydrogen bond distances between donor and acceptor atom was set to 3.2 Å and the range for accepted van der Waals distances was set to 3.6–4.5 Å. Structural alignment between 1SL8 and 1EJ3 was also performed in Pymol.

5.4.2 Interatomic overlap calculations

Interatomic overlap (D-r-R) was defined as the interatomic distance between calcium and each of the oxygen ligands in the pentagonal bipyramidal coordination (PDB ID: 1SL8) subtracted by the sum of their atomic radii (Figure 5-9). When $D-r-R < 0$ there is overlap of the atoms which suggests tight binding. In Pymol this would be indicated as steric clash between the metal ion and the coordinating rotamer of the residue (image not shown). The ionic radii of metal ions tested were taken from Shannon (1976) (Appendix Chapter 5, Table 10-12). Some metal ions were assigned more than one ionic radii in line with the metal-ligand distances found in different coordination geometries in protein metal-binding sites. When more than one ionic radius was used for the atomic overlap calculations, the highest ionic radius of the metal was included in the range.

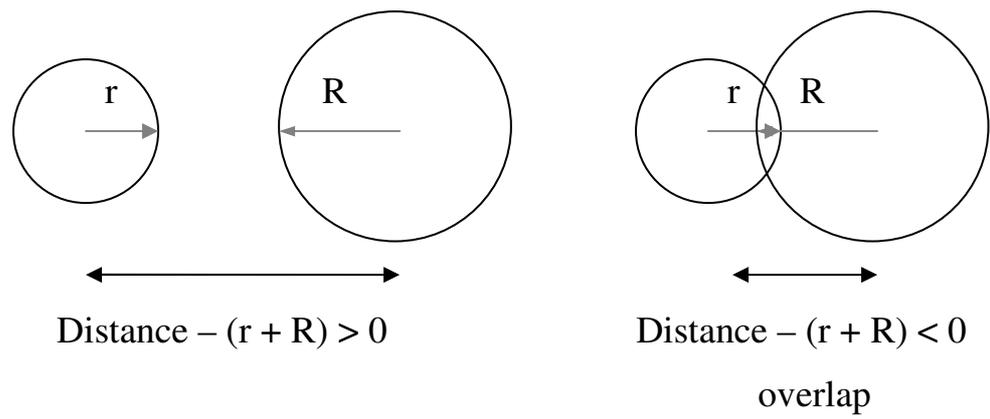


Figure 5-9 Atomic overlap. Distances and overlap between neighbouring atoms. The distance is calculated from the crystal structures and the ionic radii are from Shannon (1976).

5.5 Results and discussion

This section designs the mutations to be made and screened in Chapter 6. Observations arising from structural analysis of Sections 5.5.1-5.6 combined with the literature review conducted in Sections 5.1-5.3 could help elucidate properties of the EF-hand and response towards calcium and potentially new ions and subsequently help identify targets for mutational studies. Where information is retrieved from published bibliography, this is clearly stated in the text.

5.5.1 Calcium-free versus calcium-bound aequorin structure

Figure 5-10 shows the structural alignment of the two crystallographic structures 1EJ3 and 1SL8 and Figure 5-11 presents detailed views of the EF-hands with and without calcium bound.

There is very little gross structural difference among Ca²⁺-bound apoaequorin (1SL8) and Ca²⁺-free aequorin (1EJ3), but closer examination reveals that there are local changes (Deng *et al.*, 2005). The overall conformation of the molecules remains intact with one notable difference in the C-terminus where the C-terminus “tail” is released into the solvent leaving the hydrophobic core exposed and the bond between Tyr184 and hydroperoxy-coelenterazine disrupted.

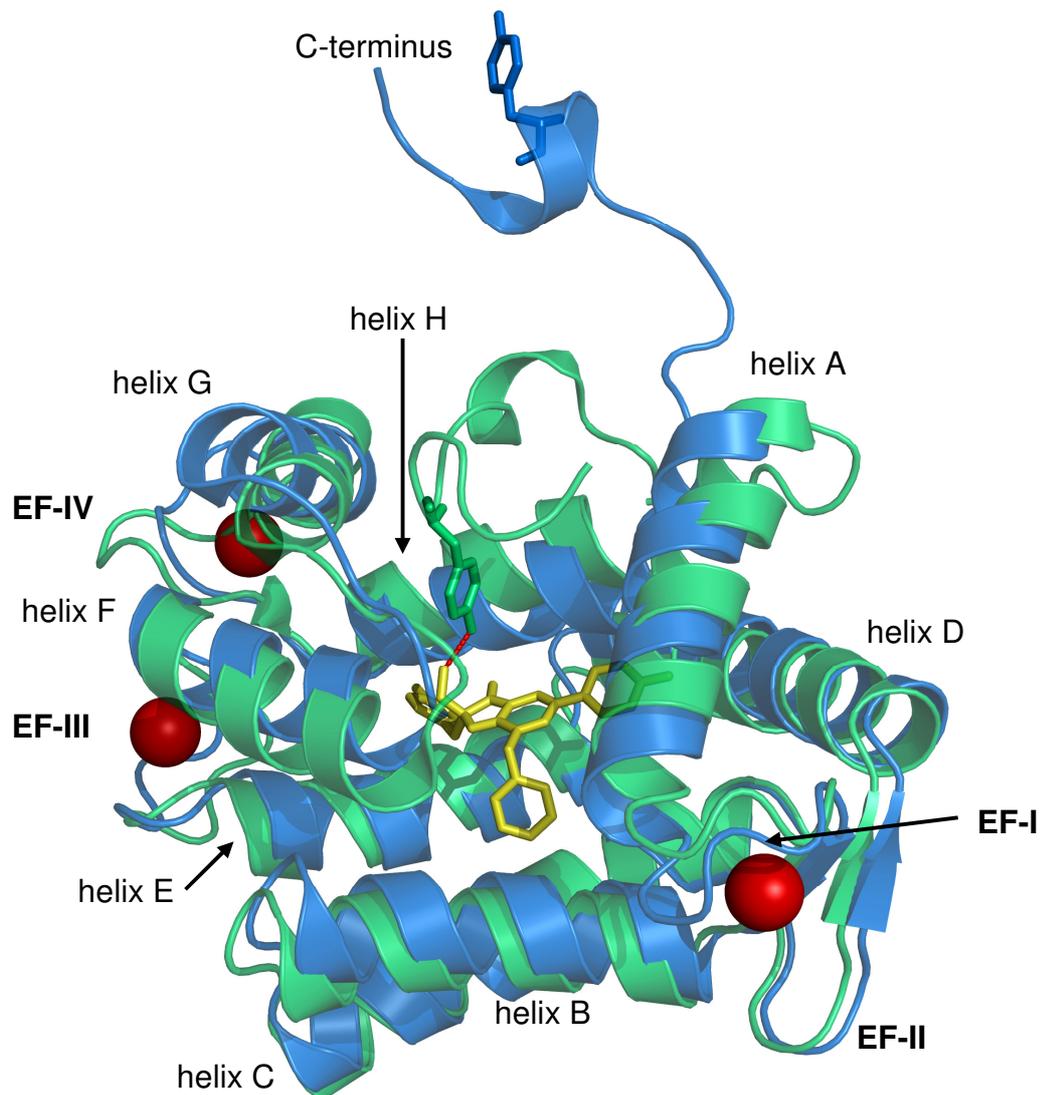


Figure 5-10 Structurally aligned calcium-free aequorin (1EJ3, chain B) and calcium-bound apoaequorin (1SL8). 1EJ3 is coloured green and 1SL8 is coloured blue. The calcium ions of 1SL8 are represented as red spheres. Stabilised coelenterazine in the centre of 1EJ3 is coloured yellow. Tyr184 of the C-terminus is shown as sticks; in 1EJ3 (green) it is hydrogen bonded to hydroperoxy-coelenterazine. The hydrogen bond is shown as red dashes.

5.5.2 Movement of residues to coordinate calcium

This section compares the three functional EF-hands of the two crystallographic structures 1SL8 and 1EJ3. Figure 5-11 shows that some of the coordinating residues move further than others in order to achieve coordination of calcium. The distances travelled before and after calcium binding by the coordinating oxygen (O) and the α -carbon of the respective residue are presented in Table 5-3.

Observations on the changes in conformation of the EF-hands due to binding of calcium ions

Upon binding of calcium ion by the loop of EF-I, both helices are “pulled” upwards towards the loop. Asp24 is pulled closer to calcium to coordinate it. The distance between the entry and exit of the loop, assigned as the exit of helix A and entrance of helix B in the Pymol representation of the molecule, was 11.81 Å and 9.75 Å in 1EJ3 and 1SL8 respectively, showing a contraction of the loop. The most structural rearrangement occurs in the first half (the N-terminus) of the loop (positions 1-5) with Asn26 (position 3 in the loop) changing orientation to face inwards and Asn28 (position 5 in the loop) rotating to coordinate the calcium ion.

In EF-III the entrance and exit of the loop are identified as Asp117 and Thr125. The distances between them are 7.68 Å and 8 Å in 1EJ3 and 1SL8 respectively, indicating little or no contraction at the entry and exit points of the loop. Residues at positions 3 and 5 rotate to bind calcium ion, but overall there is minimal backbone distortion. As in EF-I, the most obvious rearrangements occur in the N-terminal half of the loop.

EF-IV is overall more disordered prior to calcium binding than are EF-I or EF-III. As in EF-III, the maximum structural re-adjustment occurs in the N-terminal half of the loop. The entering helix (helix G) is pulled upwards towards calcium. Interestingly, the entrance and exit of the loop are identified at different positions in 1EJ3 and 1SL8. In 1EJ3, the entrance and exit of the loop are Cys152 and Val162 (distance 10.7 Å) whereas in 1SL8 they are Asp155 and Asp163 (distance 8 Å), exhibiting considerable contraction. This observation however is made with caution, as the limits between helix and loop in the Pymol representation of the molecule are indicative and depend on the

definitions of secondary structure used by the software. Aspartate at position 9 changes orientation with the side chain oxygens facing towards the loop.

As seen in Figure 5-11 and summarised in Table 5-3, in EF-I and EF-III: residues at positions 3 and 5 rotate or change orientation upon calcium binding. In EF-IV residue at position 1 rotates but residues at positions 3 and 5 are displaced in parallel orientation compared to their calcium-free state. Interestingly, position 9 in EF-IV is occupied by aspartate which changes orientation upon calcium binding, which is not observed in serine at position 9 of EF-I and EF-III. Potentially position 9 is of special significance in EF-IV than in the other EF-hands. In EF-IV residues at position 1 and 12 respond by rotation/change in orientation, while residues at positions 3 and 5 move in parallel compared to their calcium-free state. Overall, residues at positions 1 and 12 are likely to be critical in achieving loop contraction upon binding and thus “pulling” the entering and exiting helices of each loop.

The local differences in response to calcium binding by the EF-hand loops are likely to be associated with their intrinsic affinities and their ability to transduce the conformational shifts towards destabilisation of hydroperoxy-coelenterazine.

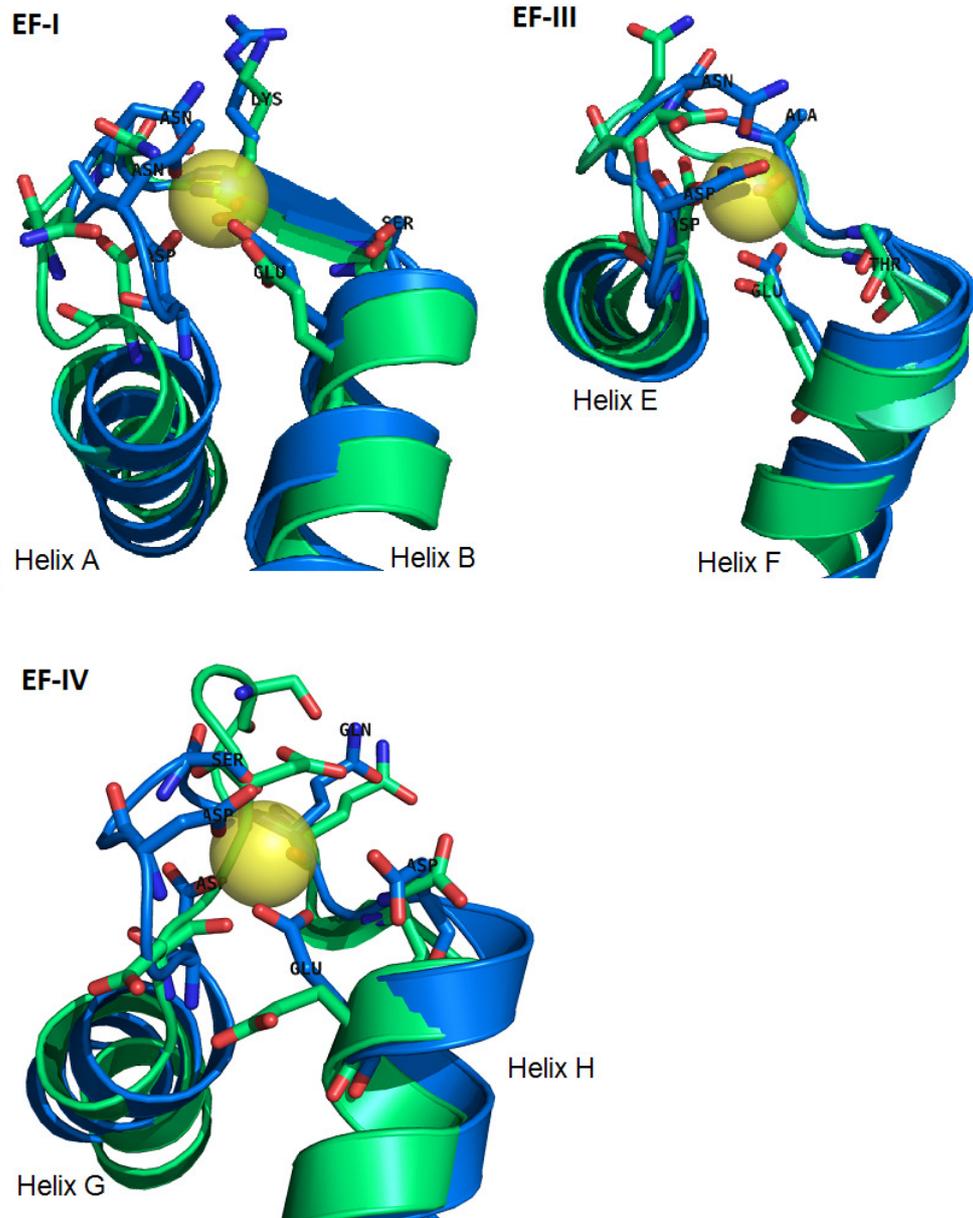


Figure 5-11 Shifts in the EF-hands upon calcium binding. Isolated fragments from Figure 5-10. 1EJ3 (calcium-free) is coloured green and 1SL8 (calcium-bound) is coloured blue. The entering (N-terminus) helix of each EF-hand is shown on the left hand side and the exiting helix (C-terminus) is shown on the right hand side. Calcium ions are shown as transparent yellow spheres. The coordinating residues are shown as sticks, both in the Ca^{2+} -bound and Ca^{2+} -free conformation. All oxygen atoms are coloured red. Coordinating water molecules were omitted from this picture. Picture was rendered in Pymol (DeLano, 2002).

Table 5-3 Shifts of residue positions between Ca²⁺-free and Ca²⁺-bound EF-hands. Distances were calculated between the α -carbon atoms of the superimposed 1EJ3 (calcium-free) and 1SL8 (calcium bound) structures for each residue in the loop. For the coordinating residues (marked in bold letters), distances between oxygen coordinating atoms were calculated. Amino acid numbering shown for both structures.

EF-hand	a.a. in 1EJ3	a.a. in 1SL8	Position	O-O, Å	C α -C α	comments
EF-I	N24	N26	1	1.8	2.5	
	V25	V27	2		4.41	Same orientation
	N26	N28	3	4.9	3.3	Orientation out→in
	H27	H29	4		2.3	Some twist
	N28	N30	5	3.8	0.5	Rotates
	G29	G31	6		0.59	Same orientation
	K30	R32	7	1.0	1.0	
	S32	S34	9	0.9	1.1	
	E35	E37	12	1.6 (O1) 1.3 (O2)	1.51	
EF-III	D117	D119	1	0.9	0.9	
	K118	K120	2		2.3	Change orientation
	D119	D121	3	2.9	3.2	Rotates
	Q120	Q122	4		2.6	Slight rotation
	N121	N123	5	6.6	3.9	Rotates
	G122	G124	6		3.1	
	A123	A125	7	0.7	1.3	
	I124	I126	8		1	
	T125	S127	9	1.1	1.1	
	L126	L128	10		1.4	parallel
	D127	D129	11		1.3	
	E128	E130	12	1.0 (O1) 1.1 (O2)	1.1	
EF-IV	D153	D155	1	6.4	2.4	Rotates
	I154	I156	2		5.2	Change orientation
	D155	D157	3	5.8	5.6	Move parallel
	E156	E158	4		6.0	Same orientation
	S157	S159	5	5.4	5.4	Move parallel
	G158	G160	6		1.8	
	E159	E161	7	1.0	1.1	Move parallel
	D161	D163	9	3.3	1.3	Orientation out→in
	V162	V164	10		1.9	
	D163	D165	11		0.65	Change orientation
	E164	E166	12	5.5 (O1) 4.1 (O2)	0.8	Orientation out→in

5.5.3 Overlap of atomic radii in the ion-binding loop

It is reasonable to assume that the octahedral bipyramidal coordination of calcium as seen in structure 1SL8 is the optimal for exerting the maximum effect upon ion binding. This section replaces the calcium ion in the loop of EF-I (PDB ID: 1SL8) with each of the additional six metal ions chosen in Section 1.6.2 and calculates the distance D and $D-r-R$ (see Section 5.4.2). For all ions other than calcium this is a hypothetical calculation and illustrates the effect of the cavity size in potential for tight binding, or lack thereof. Two basic assumptions are made: (1) the metal ions are spherical and (2) the local protein environment would “choose” the coordination and place the ions in the native calcium-binding geometry, despite of the fact that the preferred coordinating geometries of these metals in existing metalloproteins are different (Section 1.4). In Figure 5-12 to Figure 5-14 $D-r-R$ is calculated between the ion and every coordinating oxygen atom at positions 1, 3, 5, 7, 9 and 12 in the loop. At position 9 the distance is calculated from the oxygen of a water molecule (Section 1.3.4 and Figure 5-3) and position 12 offers two oxygen atoms from glutamate, assigned in the graphs as pos 12-1 and pos 12-2.

Negative values of $D-r-R$ show potential for overlap between the oxygen ligands and the metal ion, hence tight binding.

As shown in Figure 5-12, calcium, lanthanum and cadmium can overlap with the coordinating oxygens of the loop. Lead also shows potential for atomic overlap (Figure 5-13) which varies depending on the coordination number the ionic radius corresponds to. Smaller ions cobalt, zinc and copper cannot be reached by all the oxygen ligands (Figure 5-14). It is worth noting that oxygen is not the preferred ligand for these metals (Section 1.4). It is possible that in reality these ions would bind in different coordination geometries engaging fewer ligands from the ion EF-hand loop. This in turn would affect the signal transduction to the protein core where luminescence is produced.

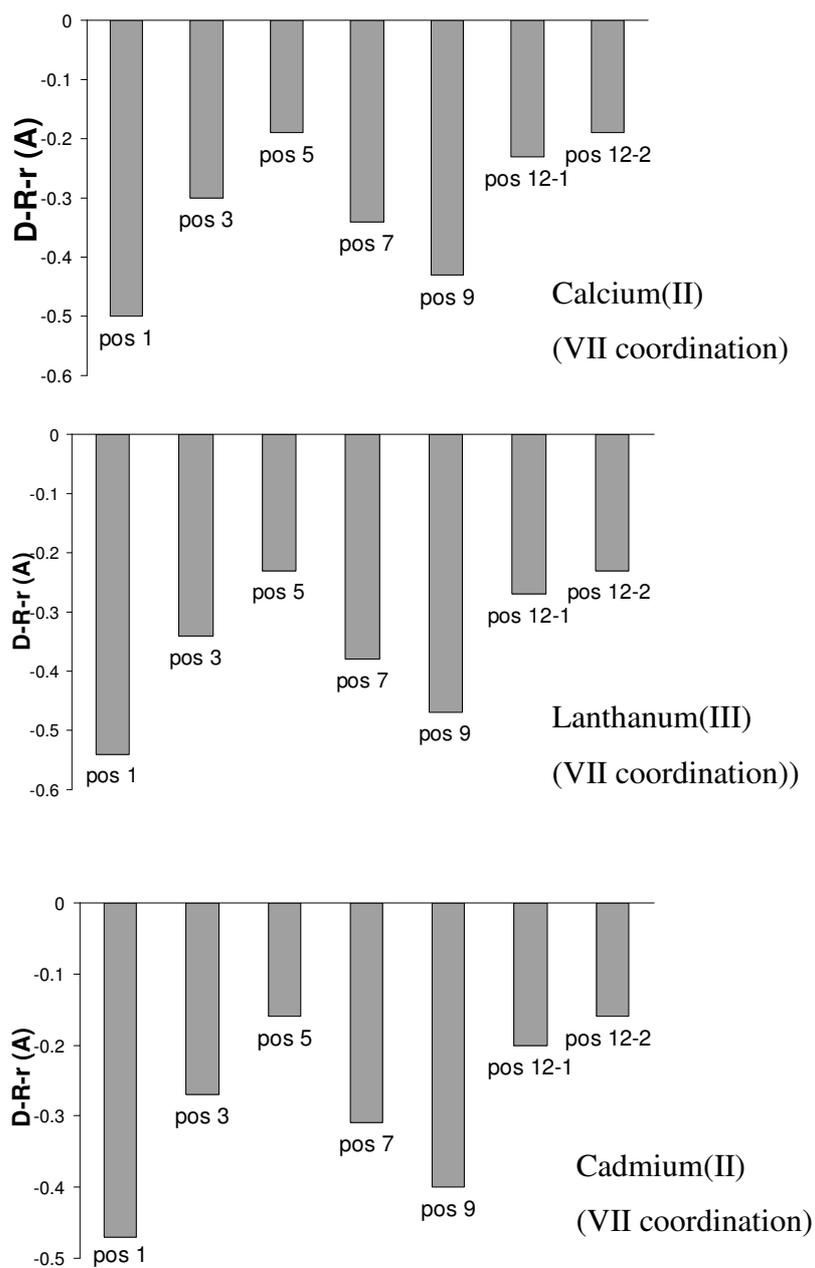


Figure 5-12 Metal-ligand atomic overlap in EF-I of aequorin (a). Interatomic distances taken from 1SL8 crystallographic structure of calcium-bound EF-I and applied to calcium, lanthanum and cadmium ions. Ionic radii for each ion correspond to the coordination numbers shown in each graph.

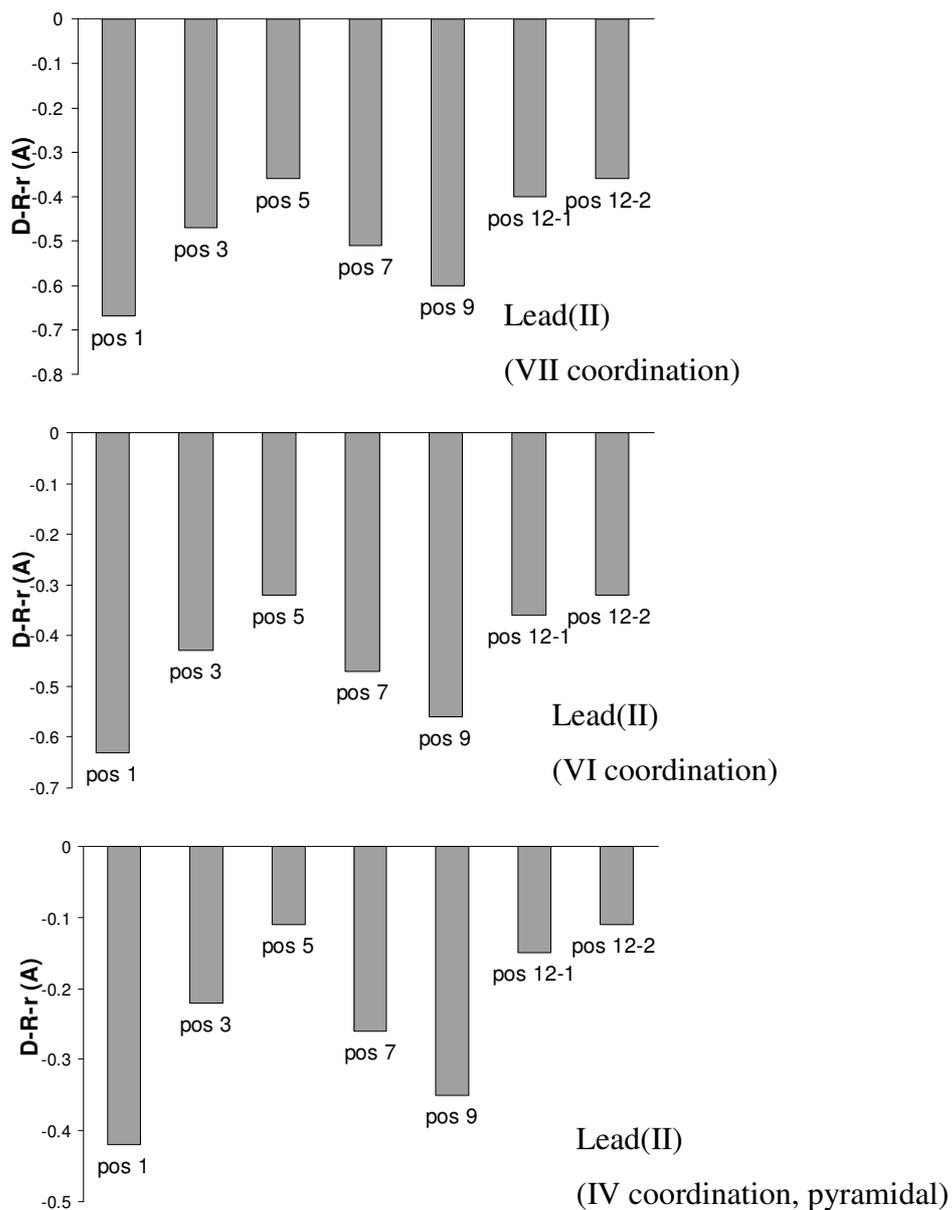


Figure 5-13 Metal-ligand atomic overlap in EF-I of aequorin (b). Interatomic distances taken from 1SL8 crystallographic structure of calcium-bound EF-I and applied to lead ions for ionic radii corresponding to three different coordination numbers/geometries.

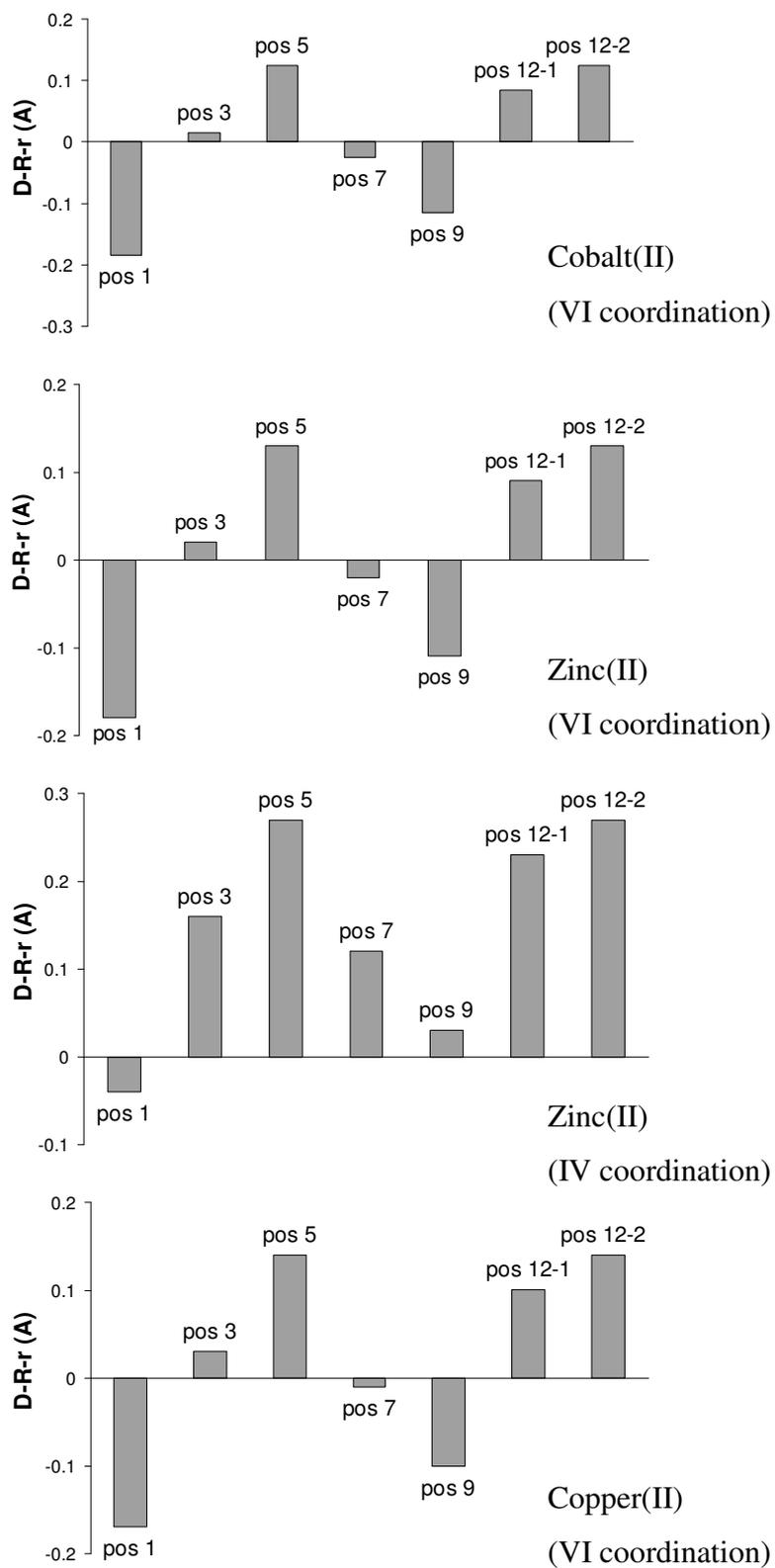


Figure 5-14 Metal-ligand atomic overlap in EF-I of aequorin (c). Interatomic distances taken from 1SL8 crystallographic structure of calcium-bound EF-I and applied to cobalt, copper and two different coordination geometries of zinc ions.

5.6 Conclusions

5.6.1 Conclusions from structural analysis

The decisive conformational shift for the light emitting reaction is associated with the release of the C-terminus “tail” of the protein, which carries Tyr184, a residue that stabilises the hydroperoxy of coelenterazine and interacts with His169. This C-terminus displacement of aequorin is the only major conformational change obvious at first glance. It is also clear that the more subtle local conformational shifts upon binding of calcium ultimately propagate into the major shift and bioluminescence reaction (Section 5.5).

The binding of calcium ion by the EF-hands requires different levels of re-adjustments of the loops. EF-IV seems to undergo the most extensive local rearrangements to accommodate calcium amongst the three calcium binding EF-hands. Residues of EF-III seems to be already pre-positioned for calcium binding judging from the extent of loop contraction and residue rearrangement while EF-I and EF-IV may be carrying the most weight in the transduction of the conformational shifts upon binding of calcium (Section 5.5.2).

In each of the loops the N-terminal side (positions 1-5) undergoes more extensive rearrangements upon calcium binding than does the C-terminal side (positions 6-12). In terms of EF-hand coupling, it occurs between the C-terminal part of the participating loops (EF-I with EF-II and EF-III with EF-IV). The residues participating in hydrogen bonding between the pairs of EF-hands are within van der Waals range from residues which directly stabilise coelenterazine or are within close proximity of it (Section 5.1.2 and Figure 5-6). The coordinating residues may rotate, change orientation or move in parallel compared to their calcium-free state to achieve calcium binding, and there is some variability in the type of response among the three EF-hands. This variability could be associated with their metal binding affinity and their role in transduction of the conformational shift. Overall, positions 1, 3 and 5 are important in binding of a new ion while positions 1 and 12 mark the entry and exit of the loop.

The size of the EF-loop cavity favours metals of similar ionic size as calcium. A theoretical substitution of calcium by six other metal ions in the loop of EF-I (PDB ID: 1SL8) shows that metal ions with ionic size similar to that of calcium (lead, lanthanum, cadmium) are more likely to assume tighter binding in the native coordination geometry than smaller ions which will not be reached by all required ligands for this geometry (Section 5.5.3). Instead the smaller ions could attract fewer ligands from the protein environment and hence trigger reduced signal transduction.

5.6.2 Conclusions from previous mutational studies

The central glycine of each EF-hand is highly conserved and its replacement with the more rigid and bulky arginine can result in loss of luminescence, possibly by impeding adequate structural response to calcium (Section 5.2.2). This strategy can be used to potentially “knockout” EF-hands one by one and hence elucidate their contribution to overall activity. Removal of coordinating side chain oxygens from positions 1, 3, 5 and 12 of the loops also results in reduction of aequorin’s activity by reducing affinity for calcium (Sections 5.2.2, 5.2.4, 5.3.1). Position 9, which coordinates indirectly to calcium through a water molecule, could serve for size and charge selectivity as well as a “gateway” to modulate binding kinetics (Section 5.3.2). Alanine substitutions at non-coordinating positions in the EF-hand loop or its preceding helix reduced affinity for calcium (Section 5.3.1). Spectral tuning is possible by mutating residues within the coelenterazine binding site or in close proximity to it as well as by using coelenterazine analogues, however this topic is outside the scope of this work.

Finally, the effect of mutations is highly dependent on the surrounding protein environment; one type of substitution in a certain position of the EF-hand would not necessarily have the same effect in two different EF-hand proteins.

5.6.3 Choice of mutant library

The goal was to create an aequorin library with a high likelihood of including desirable mutations in key positions of the molecule. Ideally these mutations should lead to altered metal ion selectivity while preserving luminescence activity.

While directed evolution (Section 1.5.2) and random-rational approaches (Section 1.5.3) present possible routes to altering the ion selectivity of aequorin, the availability of the crystallographic structures 1SL8 and 1EJ3 enables the direction of mutations to the ion binding sites. The examination of crystallographic structure and literature review presented in the previous sections were combined with common sense and curiosity. The priority was to introduce specific mutations in specific positions of the EF-hand loops. The mutations designed can be divided in the following categories: (1) histidine and cysteine substitutions, (2) alanine substitutions, (3) central Gly→Arg mutations, (4) double mutations. The forty eight mutants that were designed are presented in Table 5-4 and Table 5-5 and annotated in regards to their position in the protein and the nature of their substitution. The reasoning behind the choices is outlined below.

Preference amongst EF-hands

EF-I and EF-III were targeted first as they are more optimally pre-positioned to bind calcium than EF-IV, which suggests they may carry more weight in ion selectivity. Amongst the two, EF-I was studied more extensively because in published literature its impairment was reported to be the most detrimental to aequorin activity amongst the other EF-hands.

Mutations of coordinating and non-coordinating residues

Coordinating residues were prime targets for mutations as they are directly responsible for binding of calcium; however, non-coordinating residues were also targeted to a lesser extent. The logic behind this was that mutations at non-coordinating positions could potentially introduce a new rearrangement for metal binding. These positions could also be more “forgiving” to radical mutations as their amino acids are not highly conserved (Table 1-4). This statement excludes glycine at position 6 of the EF-hand loop which was targeted for different reasons, as explained in this section.

Histidine and cysteine substitutions

In aequorin there are three calcium-binding EF-hands (EF-I, EF-III and EF-IV), with six calcium-coordinating residues in each, four out of which are highly conserved and two not as highly conserved (Section 1.3.3 and Table 1-4). All published mutational work has highlighted substitutions of these residues by glycine (e.g.

Asp→Gly, Glu→Gly) or by other conserved coordinating residues (e.g. Asp→Asn, Glu→Asp).

To the best of the author's knowledge, the selectivity of the EF-hands has never been engineered towards binding of smaller (e.g. zinc, copper and cobalt) or softer ions, (cadmium, lead). Cysteine and histidine would normally be found in metal-binding sites of zinc and copper but are "exotic" to the EF-hand. For this reason these amino acids were deemed as highly likely to alter the ion selectivity of the EF-hand loops.

Each of the coordinating residues of EF-I and EF-III were individually substituted by histidine and cysteine. Some non-coordinating residues were also targeted in both EF-hands. Histidine and cysteine have never been purposely introduced in the EF-hand loop before, nor have such mutants been isolated from previous high-throughput mutagenesis screens.

Alanine substitutions (or alanine scan)

In this group of mutants each of the coordinating, conserved residues at positions 1, 3, 5, 9 and 12 (position 9 is less conserved) of EF-I were replaced by alanine. These substitutions can help appreciate the significance of each of the essential side chains by removing them.

Central Gly →Arg mutations

Impairing the EF-hands by introducing rigidity in the loops (Gly→Arg at position 6) has been used to assess the importance of each EF-hand in the luminescence reaction, however these mutants were only tested against calcium (Section 5.2.2). These mutants were re-created in order to be screened against each of the seven metal ions of interest. This set consists of mutants with one or both impaired EF-hands. In the initial (mutant design and screening) stages of this work, the set of Gly→Arg mutants were referred to as "knockout mutants" due to the expectation from published literature that this type of mutation would abolish activity. This choice of nomenclature however was revisited upon examination of the screening results and comparison with the effect of other mutations with similar or more dramatic effect.

Double mutations

Variants of this set carried two mutations each: two substitutions within EF-I or one substitution in EF-I and one in EF-III. These combinations always target the

coordinating residue at position 5 of EF-I (Asn30) which was found to significantly rearrange itself in space in order to bind calcium (Section 5.5.2). The use of combined mutations was deemed as a more drastic measure to alter the protein's ion selectivity.

Options not pursued – ideas for future work

Due to time constraints the library did not include more mutants in EF-IV, which was a prospect for future studies. Further rounds of either localised random mutagenesis or random mutagenesis – both achievable with alterations of the MEGAWHOP (megaprimer PCR of whole plasmid) method (Miyazaki and Takenouchi, 2002) – could possibly serve for restoration of activity and/or further fine-tuning of selectivity of the molecule. Such a strategy would potentially address aspects such as cooperativity between the EF-hands and intricate hydrogen bonding network interactions, which are notoriously hard to predict and engineer with rational mutagenesis at our current level of insight on aequorin function.

Table 5-4 Mutants Nr 1–23. Amino acid numbering according to PDB IDs 1SL8 and 1EJ3 and numbering of mutated positions within each EF-hand.

Nr#	Amino acid Nr according to 1EJ3	Amino acid Nr according to 1SL8	position in EF-I	position in EF-III	position in EF-IV
Alanine scan: replaces coordinating and conserved positions 1,3,5,9 and 12 with alanine					
1	Asp24Ala	Asp26Ala	1		
2	Asn26Ala	Asn28Ala	3		
3	Asn28Ala	Asn30Ala	5		
4	Ser32Ala	Ser34Ala	9		
5	Glu35Ala	Glu37Ala	12		
Central Gly→Arg mutations in EF-I, EF-III and EF-IV					
6	Gly29Arg	Gly31Arg	6		
7	Gly122Arg	Gly124Arg		6	
8	Gly158Arg	Gly160Arg			6
9	Gly29Arg/Gly122Arg	Gly31Arg/Gly124Arg	6	6	
10	Gly29Arg/Gly158Arg	Gly31Arg/Gly160Arg	6		6
11	Gly122Arg/Gly158Arg	Gly124Arg/Gly160Arg		6	6
Histidine substitutions in EF-I					
12	Asp24His	Asp26His	1		
13	Val25His	Val27His	2		
14	Asn26His	Asn28His	3		
15	Asn28His	Asn30His	5		
16	Ser32His	Ser34His	9		
17	Glu35His	Glu37His	12		
Cysteine substitutions in EF-I					
18	Asp24Cys	Asp26Cys	1		
19	Val25Cys	Val27Cys	2		
20	Asn26Cys	Asn28Cys	3		
21	Asn28Cys	Asn30Cys	5		
22	Ser32Cys	Ser34Cys	9		
23	Glu35Cys	Glu37Cys	12		

Table 5-5 Mutants Nr 24–48. Amino acid numbering as in Table 5-4.

Nr#	Amino acid Nr according to 1EJ3	Amino acid Nr according to 1SL8	position in EF-I	position in EF-III	position in EF-IV
Histidine substitutions in EF-III					
24	Asp117His	Asp119His		1	
25	Asp119His	Asp121His		3	
26	Asp121His	Asp123His		5	
27	Ala123His	Ala125His		7	
28	Ser125His	Ser127His		9	
29	Asp127His	Asp129His		11	
30	Glu128His	Glu130His		12	
Cysteine substitutions in EF-III					
31	Asp117Cys	Asp119Cys		1	
32	Asp119Cys	Asp121Cys		3	
33	Asp121Cys	Asp123Cys		5	
34	Ala123Cys	Ala125Cys		7	
35	Ser125Cys	Ser127Cys		9	
36	Asp127Cys	Asp129Cys		11	
37	Glu128Cys	Glu130Cys		12	
Double mutations in EF-I					
38	Asn28Cys/Glu35His	Asn30Cys/Glu37His	5, 12		
39	Asn28His/Glu35Cys	Asn30His/Glu37Cys	5, 12		
40	Asn28Cys/Ser32His	Asn30Cys/Ser34His	5, 9		
Double mutations in EF-I and EF-III					
41	Asn28Cys/Asp119His	Asn30Cys/Asp121His	5	3	
47	Asn28Cys/Asp121His	Asn30Cys/Asp121His	5	5	
42	Asn28Cys/Ser125His	Asn30Cys/Ser127His	5	9	
43	Asn28Cys/Asp127His	Asn30Cys/Asp129His	5	11	
44	Asn28Cys/Asp119Cys	Asn30Cys/Asp121Cys	5	3	
45	Asn28Cys/Asp121Cys	Asn30Cys/Asp123Cys	5	5	
46	Asn28Cys/Asp127Cys	Asn30Cys/Asp129Cys	5	11	
48	Asn28Cys/Ser125Cys	Asn30Cys/Ser127Cys	5	9	

6 Library screening and mutant selection

This chapter describes the creation of the mutant DNA library brought forward from Chapter 5 (see Table 5–4 and 5–5) using PCR mutagenesis methods, followed by high-throughput screening of the library against seven potential activator ions: Co^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} , Ca^{2+} , La^{3+} and Pb^{2+} using the methods developed in Chapter 4. The mutants will be assessed in terms of retained activity and relative selectivity compared to wild-type aequorin and a select number of mutants will be proposed as potential candidates for further study. One mutant plus wild-type aequorin will be carried forward for purification (Chapter 7) and study of their response towards the same set of metal ions (Chapter 8) without background interferences present in the crude lysate used in this chapter.

6.1 Introduction

6.1.1 Creation of mutant library

The available mutational approaches are described in Sections 1.5.1–1.5.3. Site-directed mutagenesis (SDM) was the sensible mutagenic method for the creation of the library specified in Chapter 5 (Table 5-4 and Table 5–5). The most practical SDM protocol available was QuikChange[®] (Stratagene, 2004) (Section 1.5.1, Section 2.3.3). The manufacturer's protocol for QuikChange[®] is supplied in general Materials and Methods Chapter 2, Section 2.3.3 and its implementation (including variations from the basic protocol) is described in Section 6.2.1. Wild-type apoaequorin gene in plasmid pETAQ (Chapter 4, Table 4.2) was used as template for mutagenesis and *E. coli* TOP10 cells were used for DNA preparations.

6.1.2 Previous work in methods of aequorin mutagenesis

Due to its advantages, QuikChange[®] has been employed in the most recent work involving rational mutagenesis; Row and co-workers (2008b) and Dikici and co-

workers (2009) used it to target aequorin residues involved in the stabilisation of hydroperoxy-coelenterazine.

Before this method was available, Tsuji and co-workers (1986) employed site-specific mutagenesis in order to introduce single mutations in aequorin (Section 5.2.2) based on a method by Morinaga and co-workers (1984); the desired mutations were incorporated in a synthesised oligonucleotide strand (mutagen). The apoaequorin gene was encoded in a plasmid vector also encoding for β -lactamase (allowing ampicillin resistance). One sample of this plasmid was digested in order to remove a portion of the gene for β -lactamase and the cleaved fragment treated with the Klenow fragment of DNA polymerase I and dNTPs for creating blunt ends. Another sample was digested in order to remove a portion of the aequorin gene to be mutagenised. Both fragments (plasmid with portion of β -lactamase gene cleaved off and plasmid with portion of apoaequorin cleaved off) were mixed with the synthetic oligonucleotide and the mixture was denatured at 100 °C for 3 min, followed by gradual cooling for the complementary fragments to re-anneal. The mixture of re-natured DNA was treated with the Klenow fragment and dNTPs (for creating blunt ends), T4 ligase for blunt ended ligation and transformed into *E. coli* and the ampicillin-resistant transformants were isolated. This method yielded 13% yield for a single base substitution and was simpler than previous methods requiring an M13 phage vector for gene expression (Morinaga *et al.*, 1984).

Tsuzuki and co-workers (2005) used random mutagenesis approach in order to find aequorin mutants with improved thermostability and random-rational approach in order to further explore specific sites highlighted from the random approach (Section 5.2.2 and 5.2.5). *In vitro* evolution by DNA shuffling was the method used for random mutagenesis. The cDNA sequence encoding for apoaequorin was cleaved with *DNAse* I and fragments 50-300bp isolated through extraction from a DNA electrophoresis gel. The isolated fragments were used in a primerless PCR for DNA shuffling to occur. The product of this PCR was used as template for a second round of PCR with primers designed to flank the start and finish of the entire apoaequorin gene while adding appropriate restriction sites (*Kpn* I and *EcoR* I) at each end. Digestion with *Kpn* I and *EcoR* I followed and the digested fragment was inserted into a plasmid vector under control of *Plac* promoter (for β -galactosidase production) and further transformed into *E. coli*. The transformants were plated on agar plates containing ampicillin and their colonies picked into wells of 96-well microplates. The colonies were grown into the

microplates and aequorin was reconstituted by addition and of coelenterazine into each well and screening for luminescence after injection with calcium solution.

The forty brightest mutants were selected and grown into new cultures for the isolation of plasmid DNA. These plasmids were mixed in equal proportions and co-amplified using primers that flanked the start and finish of the entire apoaequorin gene (used also in the second round of PCR). The product of this amplification was entered in a new round of DNA shuffling for *in vitro* evolution.

The same authors (Tsuzuki *et al.*, 2005, Tricoire *et al.*, 2006) employed rational-random approach by creating randomised mutants in two specific positions in the apoaequorin sequence (Gln168 and Leu170) appearing to be crucial in bioluminescence yield.

As a first step, mutagenesis of the Gln168 and Leu170 was performed by a PCR step with a mutagenic sense primer containing NNS codons at the two positions of interest. The same primer introduced a silent mutation one codon upstream of codon Gln168 encoding for a unique restriction site (*Mlu* I). The antisense primer did not introduce any mutations but included one unique restriction site (*EcoR* I) which was part of wild-type apoaequorin sequence.

The products of the first PCR step were used as template for a second round, in which the population of all mutagenised fragments was amplified using a new sense primer that included the unique restriction site (*Mlu* I) upstream of the mutated codons but did not include the mutated codons and the same antisense primer as the first round PCR. A unique *Mlu* I site was also created upstream of the Gln168 codon by introducing silent mutations in the wild-type apoaequorin cDNA within the plasmid vector. The product of the second round of PCR was digested using *Mlu* I and *EcoR* I as was the plasmid carrying wild type apoaequorin and the silent mutation allowing *Mlu* I digestion and the digested mutagenised fragments of aequorin were ligated into the vector. The resulting population of plasmids was transformed into *E. coli* and protein expression and screening was similar as in the DNA shuffling mutants.

6.1.3 Screening of mutant library

This is a library of 48 mutants and wild type in a total of four 96-well microplates. Each plate was injected with each of the seven test ions: calcium,

cadmium, cobalt, copper, lanthanum, lead and zinc. The screening process was developed in Chapter 4 (Section 4.5.1, Figure 4.17) and implemented as described in Section 6.2.2. Raw data from each microwell of the library will consist of relative luminescence units (RLUs) versus time (milliseconds).

Before processing the raw data arising from the high-throughput screen, it is important to set the metrics for meaningful comparison and clarifying the relevant terminology. Selectivity is key in assessing the response of the mutants towards calcium and new ions. The terms selectivity or specificity are in many cases used interchangeably in analytical chemistry, biochemistry and enzymology. Selectivity describes “whether the analyte can be measured without interferences from other components in the mixture” (Persson and Vessman, 2001). According to IUPAC (2001) “specificity is the ultimate of selectivity” and only a method which is perfectly selective for an analyte is said to be specific (den Boef and Hulanicki, 1983).

In previous work on aequorin mutagenesis the performance of aequorin variants has been presented as relative to that of wild type (Tsuji *et al.*, 1986, Tsuzuki *et al.*, 2005) or in terms of the half-life of the luminescence decay curves compared to that of wild type in the context of kinetics studies (Tricoire *et al.*, 2006). Relative activity of mutants compared to that of wild type has also been considered in this chapter but additional metrics stem from the expectations arising from the mutations. Aequorin has three calcium binding EF-hands containing six calcium-coordinating amino acids each. In this work, only one or two amino acids were mutated at a time. Realistically, single point or double mutations are unlikely to turn around the ion selectivity of a molecule that relies on up to three sites and eighteen amino acids for this function.

Each mutant of the library will be assessed in terms of (1) yield of mutant versus wild type for every metal ion tested and (2) shift of a mutant’s response from calcium towards a new ion compared to the equivalent shift of wild type. Details on the calculation of the metrics used are presented in Section 6.2.3. Ranking of the mutants according to these metrics is aided by the use of matrix plots and bar plots (Figure 6-8 to Figure 6-17). Bar plots allow visual assessment of the performance of the mutants and identification of salient traits. Matrix plots were constructed to visualise possible correlations between traits and to identify determinants of selectivity between different metal ions.

6.2 Methods

6.2.1 Mutant library construction

Mutagenic primers were designed to flank the residues to be replaced and as described in Section 2.3.5. Due to the large number of primers created, these are presented in Appendix Chapter 6, Tables Table 10-6 to Table 10-11. Their length was 50–60 bp, which exceeds the usual requirement of 18–25 bp; this was found to reduce the occurrence of by-products due to multiple insertions of primer sequence in the final construct.

The QuikChange[®] Site-Directed Mutagenesis (Stratagene Ltd.) protocol and PCR reaction mix were as described in in Section 2.3.2 and Tables 2-2 to 2-4. The was made up as described. PCR cycles are presented in Table 2-4. Any variations or deviations applicable to individual PCR reactions (e.g. additional Mg²⁺ and DMSO or changes in annealing temperature) are captured in Table 6-1. The QuikChange[®] reaction was initially performed in a TechGene thermal cycler (Techne Ltd.). Due to an initial low success rate and in order to speed up the process of attaining successful PCR reactions, multiple reactions were run per mutant, including variations in the PCR mix and annealing temperatures. A temperature gradient PCR thermal cycler was used (Techne Ltd.) for this purpose.

After the PCR reaction, the stages of digestion with *Dpn* I, transformation in *E. coli* and selection of transformants were performed as described in section 2.3.2. Growth of bacterial cultures and preparation of plasmid DNA was performed as described in Section 2.2.1 to 2.2.4 and Section 2.3.1. Existence of the desired mutations was verified by sequencing (Section 2.3.4). Where the QuikChange[®] reaction yielded more than one product (visualised through DNA gel electrophoresis) the entire miniprep of the transformed cells would then be run on a low-melt agarose gel and only the band of the desired size would be extracted and sent for sequencing. An example of such a case is presented in Figure 6-4. Sequencing results for the reactions are supplied in their original Bioedit file format as Supplementary Material.

Expression of apoaequorin mutant in *E. coli* cultures and cold incubation with coelenterazine for production of aequorin variants were conducted in microplate format as dictated in Chapter 4, Section 4.5.1 and Figure 4-17.

Table 6-1 PCR conditions for creating mutants Nr 1–48. Annealing temperatures are deviations from the template protocol of Table 2–4. Concentrations of magnesium sulfate, DMSO and Qsolution in the final reaction mix are additional to the concentration shown in Table 2-3. Numbering of mutants as in Table 5-4 and Table 5-5.

Nr	T _a (°C)	Mg ²⁺ (mM)	DMSO (v/v)	Qsol (v/v)	Nr	T _a (°C)	Mg ²⁺ (mM)	DMSO (v/v)	Qsol (v/v)
1	69	0.5	2%		25	68	0.5	2%	
2	69	0.5	2%		26	68	0.5	2%	
3	69	0.5	2%		27	67.5	0.5	2%	
4	69	0.5	2%		28	68	0.5	2%	
5	69	0.5	2%		29	68	0.5	2%	
6	67.5	0.5	2%		30	68	0.5	2%	
7	67.5	0.5	2%		31	68	0.5	2%	
8	51				32	68	0.5	2%	
9	67.5	0.5	2%		33	68	0.5	2%	
10	53		0.8%		34	67.5	0.5	2%	
11	53		0.8%		35	68	0.5	2%	
12	61.5	0.5			36	68	0.5	2%	
13	62	0.5	2%		37	68	0.5	2%	
14	64		2%	10%	38	62	0.4	2.2%	
15	56.3	0.75			39	62	0.4	2.2%	
16	69	0.5	2%		40	62	0.4	2.2%	
17	63			4%	41	67	0.5	2%	
18	63				42	67	0.5	2%	
19	62	0.5	2%		43	67	0.5	2%	
20	57	0.5	2%		44	67	0.5	2%	
21	65	1	2%		45	67	0.5	2%	
22	62	0.5	2%		46	67	0.5	2%	
23	64		2%		47	67	0.5	2%	
24	68	0.5	2%		48	67	0.5	2%	

Mutant Nr 21: DNA polymerase KOD (Novagen) was used for a successful reaction instead of PfuUltra High Fidelity polymerase.

6.2.2 Mutant library layout and screening

The forty eight mutants of the library were arranged in a total of four sub-libraries in 96-well microplate format. A typical layout of each microplate is presented in Figure 6-1. In each microplate, the wells marked with X1,X2,...,X48 were inoculated with a single colony of *E. coli* cells expressing either wild-type or a mutant of apoaequorin. Each mutant and wild type was arranged in triplicates as shown in Figure 6-1. In order to avoid position-related bias within the plate (e.g. effects of temperature and humidity gradient which can ultimately affect protein yield and activity), bacteria expressing wild-type apoaequorin were always grown in wells X4,X5,X6 (positions B4,C4,D4 in microplate coordinates) and in wells X28,X29,X30 (positions E4,F4,G4 in microplate coordinates), while the rest were used for mutant variants. The outer wells of each plate were not inoculated, as they are associated with high errors, as illustrated in Chapter 4. They were however filled with the same liquid media as the inoculated wells at all stages of microwell aequorin production in order to help maintain constant temperature throughout the microplate. Based on this layout, each microplate can hold a maximum of 14 triplicate sets of aequorin variants and two triplicate sets of the wild type. Wild type produced in wells B4,C4,D4 was used as internal reference for each plate.

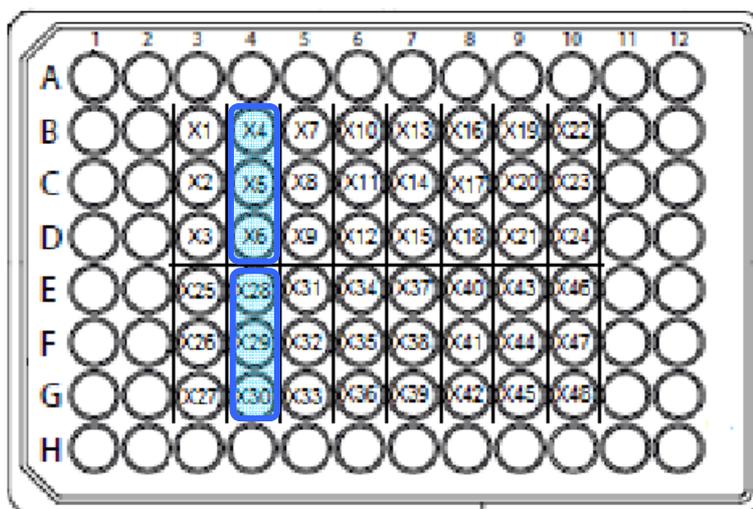


Figure 6-1 Typical layout of each microplate of the mutant library. Wells highlighted in blue contain wild-type protein throughout the library. The remaining wells contain mutant variants of aequorin. The black line grid between the wells marks the positions of triplicates.

Wild type produced in E4,F4,G4 of each plate was harvested and pooled. This sample was used for the determination of optimal luminescence reading parameters (e.g. gain settings, optimal concentration of each metal ion to be injected) on the BMG microplate reader. In some plates (Figure 6-2) additional wild type was produced in wells X22,X23,X24 (microplate coordinates B10,C10,D10) or X46,X47,X48 (microplate coordinates E10,F10,G10) in case additional wild type would be required. It was important to set the luminescence screening settings based on the properties of aequorin produced under the same conditions as the microplate library under study. These wild type samples were stored at -20°C separately from the microplates so that optimisation experiments for screening parameters could be carried out without need to thaw the library.

Figure 6-2 shows the exact position of the mutants in the each of the four plates of the library during the entire process from inoculation into microwells through to high-throughput screening against metal ions. The numbering of the forty eight mutants is according to Table 5-4 and Table 5-5 and they are arranged in triplicates.

plate-1	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			12	WT	13	14	15	16	17	WT		
C			12	WT	13	14	15	16	17	WT		
D			12	WT	13	14	15	16	17	WT		
E			18	WT	19	20	21	22	23	WT		
F			18	WT	19	20	21	22	23	WT		
G			18	WT	19	20	21	22	23	WT		
H												

plate-2	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			24	WT	25	26	27	28	29	30		
C			24	WT	25	26	27	28	29	30		
D			24	WT	25	26	27	28	29	30		
E			31	WT	32	33	34	35	36	37		
F			31	WT	32	33	34	35	36	37		
G			31	WT	32	33	34	35	36	37		
H												

plate-3	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			1	WT	2	3	4	5	6*	WT		
C			1	WT	2	3	4	5	6*	WT		
D			1	WT	2	3	4	5	6*	WT		
E			6	WT	7	8	9	10	11	WT		
F			6	WT	7	8	9	10	11	WT		
G			6	WT	7	8	9	10	11	WT		
H												

plate-4	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			38	WT	39	40	41	43	44	45		
C			38	WT	39	40	41	43	44	45		
D			38	WT	39	40	41	43	44	45		
E			46	WT	47	42	48	28	35	WT		
F			46	WT	47	42	48	28	35	WT		
G			46	WT	47	42	48	28	35	WT		
H												

Figure 6-2 Layout of wild type and mutants in the four plates of the library. Numbering of mutants as in Table 5-4 and 5-5. In Plate-3: In mutant denoted 6* the codon for arginine was CGA and in mutant 6 the codon was CGT. Both were used for construction of the mutant library but only readings for mutant Nr 6 were used in data processing. Mutants Nr 28 and 35 were repeated in Plate-4.

6.2.2.1 Metal ion concentrations for library injection

For practical reasons only one concentration per metal ion could be used to screen the entire library. It was suspected that high concentrations of certain metals could prove deleterious for aequorin activity but there was no published work on the subject. It was also understood that the optimal concentration of each ion for each of the forty eight mutant variants may differ. Determination of optimal conditions (metal ion concentration and photomultiplier gain) would be based on performance of wild type only.

The experiments were performed using wild type in the presence of a range of concentrations of each metal. The wild type was from wells E4,F4,G4 of the microplate library (Section 6.2.2).

Metal ion preparations were according to Section 2.5.1. Solutions of calcium chloride, zinc chloride, cadmium chloride, copper sulfate, lanthanum chloride and cobalt chloride were prepared at nominal concentrations of 1, 5, 10, 12.5, 20 and 30 mM in 50 mM Tris-HCl pH 7.5. Lead nitrate was prepared in RO water. Concentrations of metals were at the millimolar range as the protein suspension was in ~10 mM EDTA (Section 4.5, Figure 4.17).

100 μ L of protein solution was injected with metal ion solution at 1:1 volume ratio and the light output measured in the FluoStar Platerreader (Section 2.7). At constant gain settings, protein solutions were injected with the range of metal salt concentrations in order to find the optimum. The light readings from all the wells tested were plotted against the nominal concentrations of the metal ion solutions. The resulting curves are provided in Supplementary Material.

The titration curves showed that there was indeed an optimum concentration for each metal; low concentrations would trigger low response whereas for certain metals higher concentrations seemed to cause aequorin inhibition. The optimal concentration of each metal was used for the screening of the mutant library and is shown in Table 6-2.

6.2.2.2 Gain settings

Initial tests of Section 6.2.2.1 demonstrated that one uniform gain setting for all metals might result in missing low activity levels from poor performers (if the gain is set relatively low) or in over the linear range readings from high performers.

The point mutations introduced in aequorin were likely to bring about minor shifts to protein response towards various metals accompanied by major loss in light production. The desired output of the screening process was to capture even minor improvements occurring at low light levels. This required using high gain settings for ions that cause low level aequorin response. Conversely, the combination of wild type with calcium was expected to yield high levels of activity and would need to be read at low gains.

Gain settings for each ion were determined by trial and error at photomultiplier gains within the range of 1,650 to 2,200 in order to ensure that readings would not exceed the instrument's range. The settings chosen for the screening of the library are shown in Table 6-2.

Table 6-2 Metal ion stock concentrations and platereader gain settings for library screening.

Metal Ion salt	Concentration for library injection (mM)	Platereader Gain
CaCl ₂	25	1,650
CdCl ₂	20	1,900
CoCl ₂	12.5	2,200
CuSO ₄	12.5	2,150
ZnCl ₂	25	2,000
LaCl ₃	2.5	1,850
Pb(NO ₃) ₂	12.5	2,150

6.2.3 Data processing of high-throughput screening

A schematic of the concept behind data processing is shown in Figure 6-3. Every microplate contains an internal reference of wild-type aequorin in triplicate (wells B4,C4,D4). The average mean of the luminescence peak (or in lack of a clear peak, the highest luminescence reading) is calculated and divided by the respective mean of the internal control wild type. This ratio is called C and is a measure of the performance of the mutant compared to wild type, for each ion. Using an internal control takes into account factors that may influence the production of aequorin variants in each microplate (Figure 4-17). Hence, measure of activity C is comparable between different plates and across the library (Equation 2):

$$C_{mu}|X^{2+} = \frac{Peak_{mut}|X^{2+}}{Peak_{w.t.}|X^{2+}} \quad \text{Equation 6.1}$$

The next measure is selectivity. The envisaged mutants would detect a new ion and discriminate the other ions present in the test sample. The abolishment of calcium affinity would be an ideal result. Measure of selectivity (D) was introduced to assess the extent to which a mutant responds favourably to a new ion versus calcium and to simultaneously compare this response to the extent to which wild type responds to a new ion versus calcium. Measure of selectivity D is also comparable across the library as shown by Equation 3:

$$D_{mu}|X^{2+} = \frac{\frac{Peak_{mut}|X^{2+}}{Peak_{mut.}|Ca^{2+}}}{\frac{Peak_{w.t.}|X^{2+}}{Peak_{w.t.}|Ca^{2+}}} = \frac{Peak_{mut}|X^{2+} \times Peak_{w.t.}|Ca^{2+}}{Peak_{w.t.}|X^{2+} \times Peak_{mut.}|Ca^{2+}} = \frac{\frac{Peak_{mut}|X^{2+}}{Peak_{w.t.}|X^{2+}}}{\frac{Peak_{mut.}|Ca^{2+}}{Peak_{w.t.}|Ca^{2+}}} = \frac{C_{mu}|X^{2+}}{C_{mu}|Ca^{2+}} \quad \text{Equation 6.2}$$

D=1 means no change in selectivity compared to calcium.

D<1 means selectivity towards Ca²⁺ was enhanced as affinity for other ions was even further diminished comparatively to calcium.

D>1 is preferential selectivity of mutant for X²⁺ (or La³⁺) versus Ca²⁺.

True selectivity is defined when the competing substances are present in the test sample, which is not the case in this initial screen. Derivation of measures C and D are described schematically in Figure 6-3. The calculations were carried out using Excel (Microsoft Ltd) and the results were arranged in Excel and Minitab (Minitab Ltd). In Figure 6-2 it is shown that all mutants were arranged in triplicates. If the % deviation amongst the measured triplicate values was $\geq 20\%$ the raw data was examined and outliers amongst the triplicates were excluded. In such a case the final values of C and D would represent the average C and D for a duplicate set.

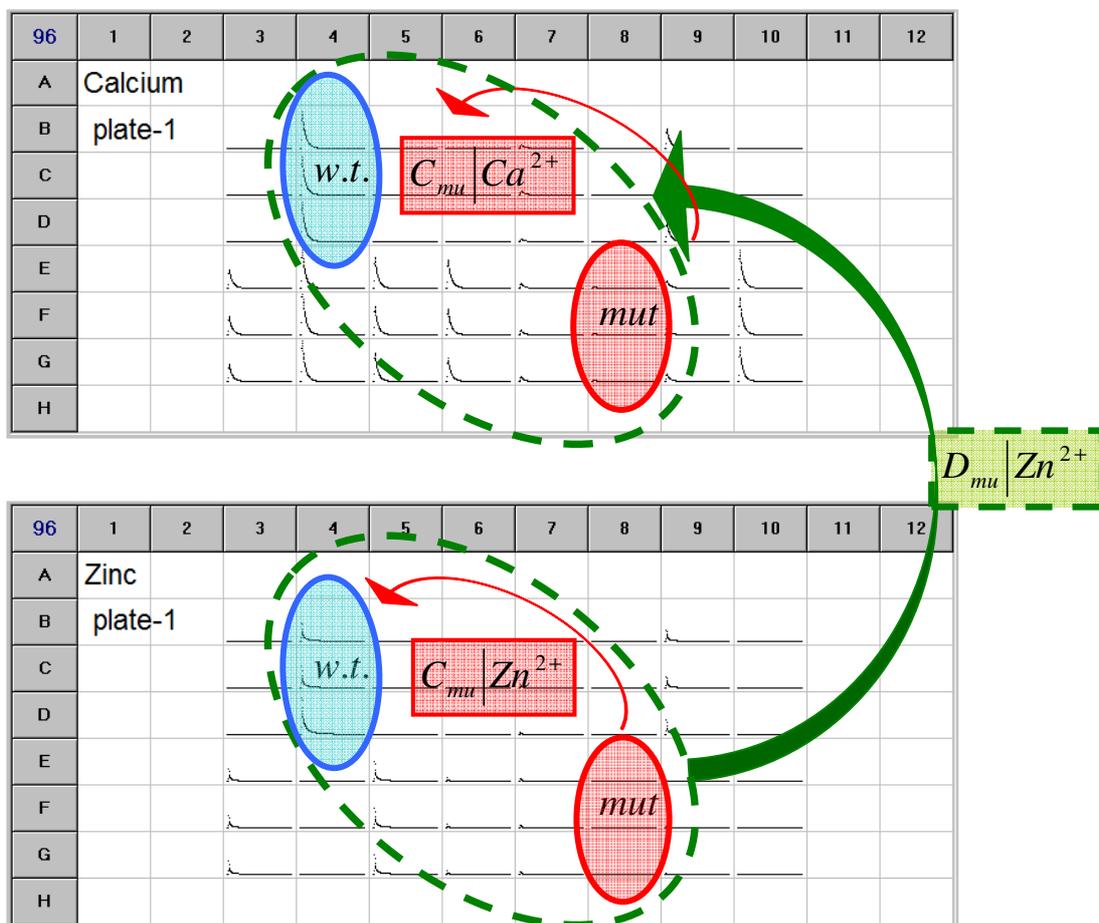


Figure 6-3 Illustration of measures of activity (C) and selectivity (D) in data processing. As an example, Plate-1 injected with calcium and Plate-1 injected with zinc are compared. In blue oval shape are the triplicate sets of wild type and in red ovals are the triplicate sets of one mutant variant. For each mutant, the ratio of the mean average peak luminescence of the mutant (red) over the wild type (blue) gives value C corresponding to the ion tested in each plate (symbol in red square boxes). The ratio of the C value for any given ion over C value for calcium is value D corresponding to that particular ion (symbol in square green box). Metrics C and D can be used for comparisons across the entire library due to the internal references in each plate.

6.3 Results

6.3.1 Library creation results

The size of the plasmids carrying mutated apoaequorin sequence was verified by agarose DNA gel electrophoresis and successful mutations were verified with sequencing. Figure 6-4 and Figure 6-5 present a typical set of DNA gel visualisations of mutants Asn28Cys and Asn28Cys/Ser32His respectively.

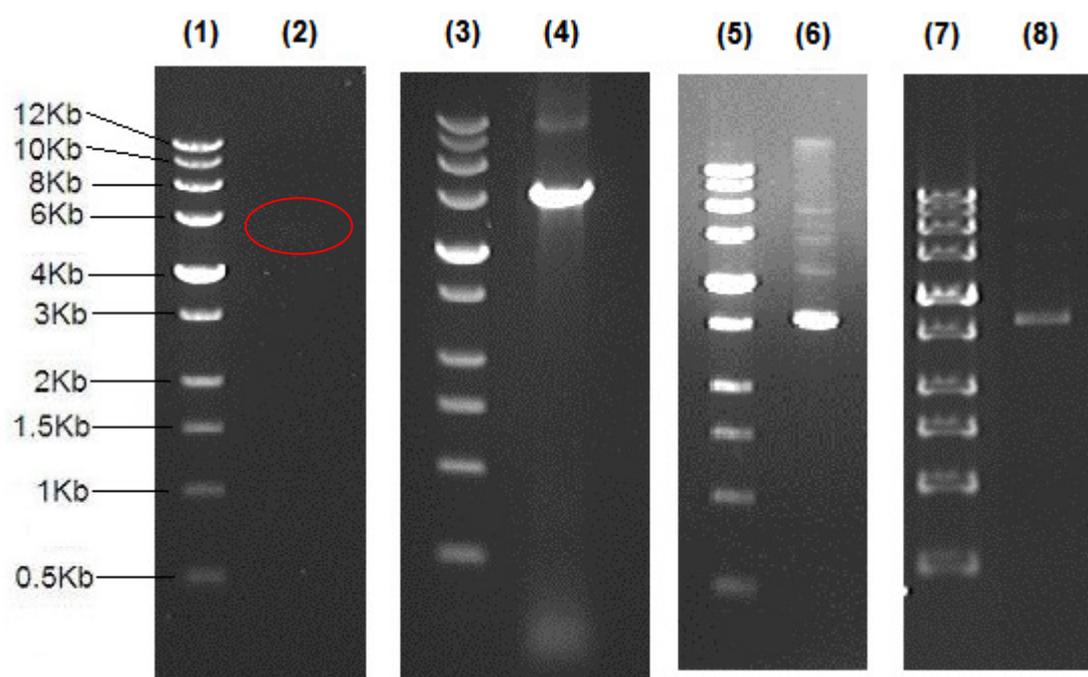


Figure 6-4 DNA agarose gel electrophoresis of mutant Asn28Cys. 0.8% agarose gel. Lanes 1, 3, 5 and 7: 0.5–12 Kb DNA marker; Lane 2: primerless PCR (control) containing template pETAQ (wild type) in very faint band; Lane 4: product of PCR with primers “Asn28Cys for” and “Asn28Cys rev” (primer sequences in Appendix Chapter 6, Table 10-7 and Table 10-11). The product of QuikChange[®] amplification is an open circular plasmid, hence the dominant band of ~6 Kb; Lane 6: 5 μ L of plasmid miniprep of transformed *E. coli* with PCR product of Lane 4. Remaining amount of plasmid was run on low-melt agarose gel and the lowest band (supercoiled form) extracted using QIAquick[®] gel extraction kit. Lane 8: product of gel extraction low-melt agarose gel. This sample was sequenced and upon verification of desired results stored as part of the mutant DNA library.

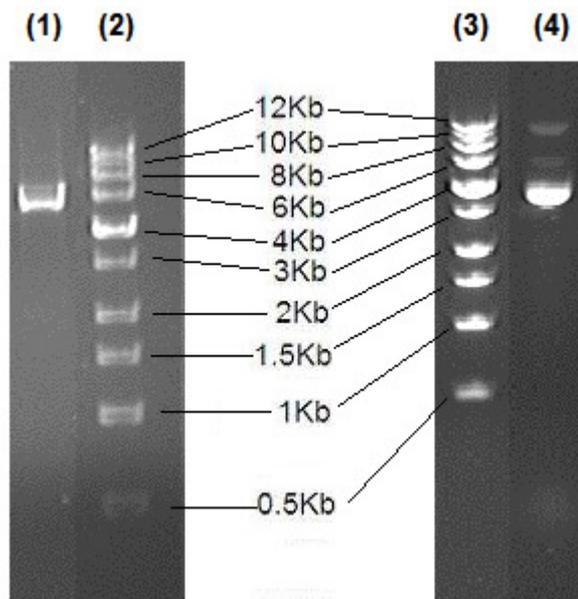


Figure 6-5 DNA agarose gel electrophoresis of mutant Asn28Cys/Ser32His. 0.8% agarose gel. Lanes 2 and 3: 0.5–12 Kb DNA ladder; Lane 1: product of PCR with template plasmid containing mutation Asn28Cys and primers “Ser32His for” and “Ser32His rev” (primer sequences in Appendix Chapter 6, Table 10-8 and 10-11). The product of the QuikChange[®] PCR amplification is an open circular plasmid, hence the dominant band of ~6 Kb; Lane 4: plasmid miniprep from *E. coli* cells transformed with material of Lane 1.

Figure 10-1 (Appendix Chapter 6) shows a typical set of sequencing results for Asn28Cys mutant aligned with the sequence of wild-type apoaequorin. The sequencing results for the library of mutants are presented in Supplementary Material in Bioedit file format.

6.3.2 Library screening results

The schematic output of screening results (four plates, seven ions per plate) as displayed by the platereader software after the completion of each test run is shown in Figure 6-6 and Figure 6-7. In this style of representation each well contains the trace of luminescence counts versus time. The raw screening data is presented in Supplementary Material.

Processing of the raw data show metrics of activity C and selectivity D (definitions in Section 6.2.3) for all mutants and the results are shown in Figure 6-8 to Figure 6-10, which will be discussed in Section 6.4. Figure 6-8 shows the activity of the mutants compared to that of wild type for each of the seven ions tested. The mutants are grouped according to type of mutation and location in the molecule (also refer to Table 5-4 and Table 5-5): cysteine and histidine mutants in EF-I or EF-III; single and double Gly→Arg mutants; alanine replacements in EF-I; double mutations in EF-I and double mutations in EF-I and EF-III combined. Their position in the molecule and substitution type is shown on the x-axis.

Figure 6-9 plots D in the y-axis. D takes into account not only the activity of mutants in response to each new ion, but also their response towards calcium, in that way serving as a measure of selectivity. As high D values do not necessarily guarantee high protein activity, the x-axis is also highlighted to mark mutants with C values >10% and >50% so it would be possible to select mutants with high D and reasonable C. Figure 6-10 is an alternative presentation of D for each mutant. Logarithmic scale is used for values of D in order to visually aid assessment of the entire library: D=1 corresponds to wild type and the higher the D value the greater the shift of selectivity towards new ions and further away from calcium.

Figure 6-11 and Figure 6-12 are alternative representations of values C and D respectively. The same data as in Figure 6-9 and Figure 6-10 is presented in individual value plots and arranged according to the EF-hand targeted (EF-I or EF-III), the final substitution (His or Cys) and the exact position in the EF-hand loop targeted (positions 1, 2, 3, 5, 7, 9, 11, 12). This type of representation can reveal potential trends associated with these groupings. Central Gly→Arg mutants were excluded from these graphs.

Figure 6-13 to Figure 6-15 are matrix plots of the metric of selectivity D for all the metal ions tested. In each of the three graphs the data points are coloured according to EF-hand mutated (Figure 6-13), position in the EF-hand loop (Figure 6-14) and type of mutation (Figure 6-15). Matrix plots can help identify pairwise associations between parameters, clustering of groups or presence of outliers.

Figure 6-16 plots the metrics of selectivity D against retained activity C for the entire library and against all of the metals tested. An ideal outcome would be a mutant that rates highly in both metrics.

Figure 6-17 plots the retained activity for the single and double Gly→Arg mutants in order to assess the importance of each of the EF-hands I, III and IV in the activity of aequorin with each of the seven metal ions tested.

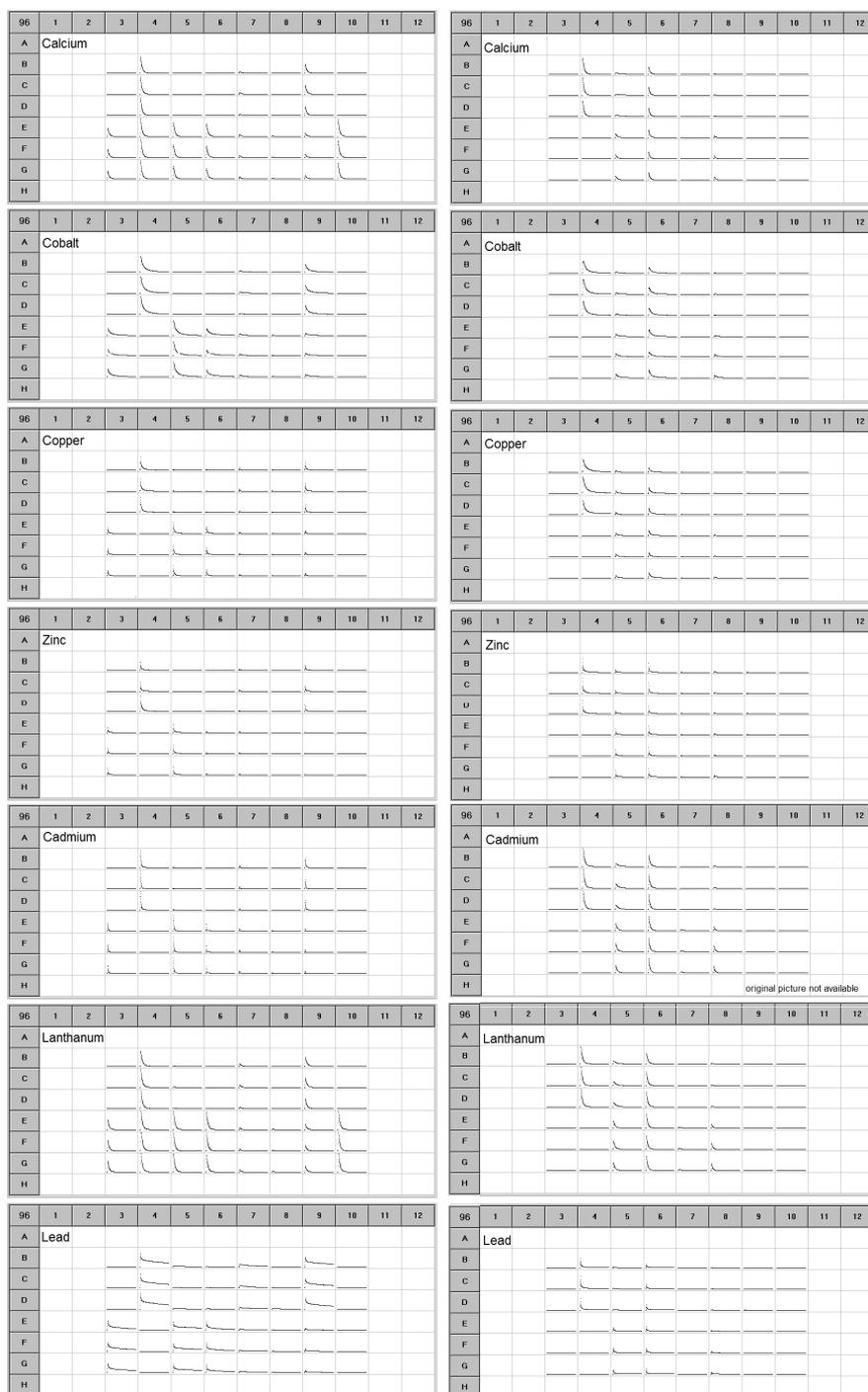


plate-1

plate-2

Figure 6-6 Visualisation of raw screening data, plates 1 and 2 of aequorin library. Each well represents a luminescence versus time curve. The wells of plates 1 to 4 contain protein injected with each of the seven metal ion solutions as candidates for triggering luminescence. The experimental design is described in Section 6.2.2 and discussed in 6.3.2. The position of wild type and mutants in each plate was according to the layout in Figure 6-1 and 6-2. Gain settings for luminescence measurements and concentrations for each metal ion injected are presented in Table 6-2. Data analysis was according to Figure 6-3.

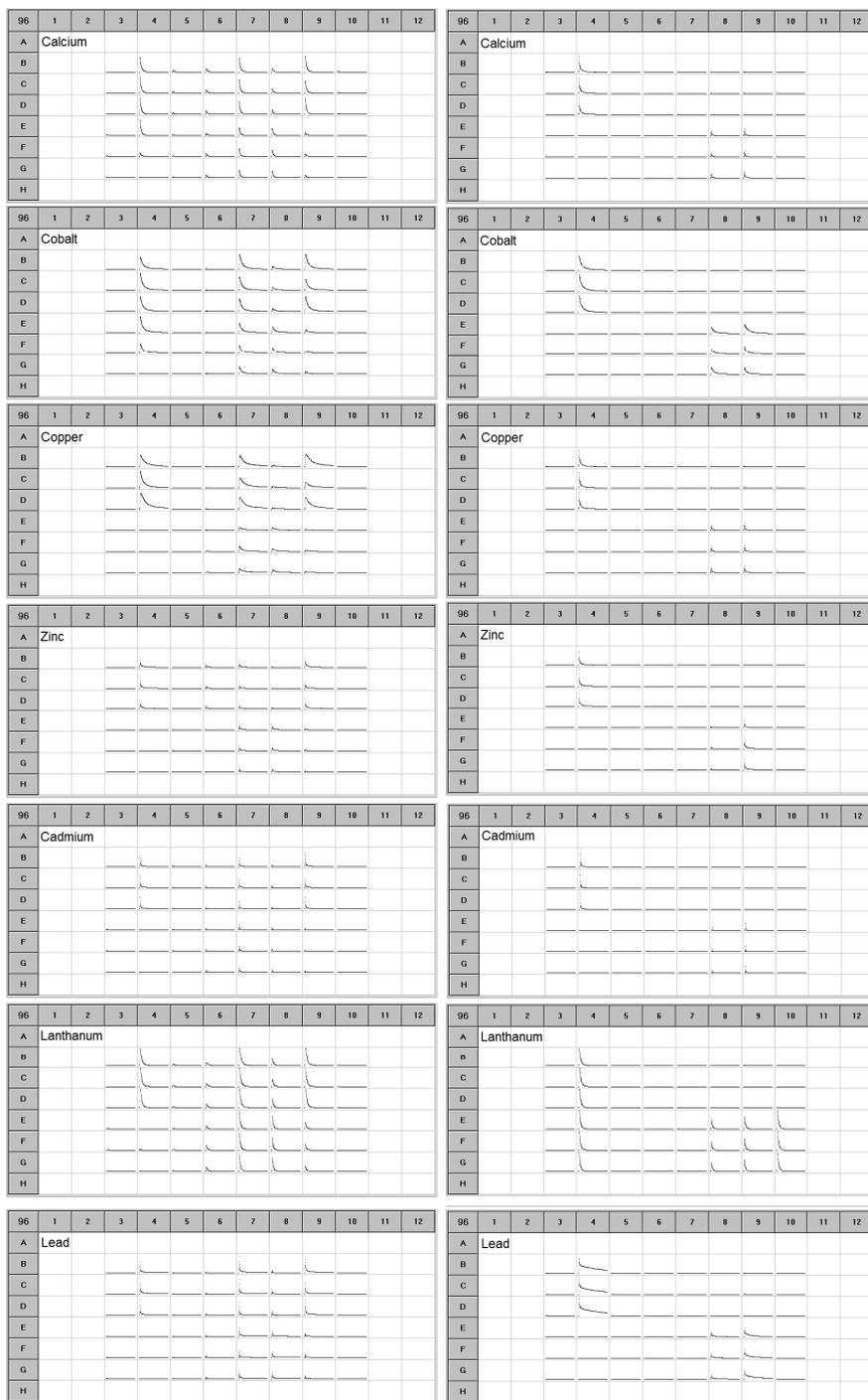


plate-3

plate-4

Figure 6-7 Visualisation of raw screening data, plates 3 and 4 of aequorin library. Continued from Figure 6-6.

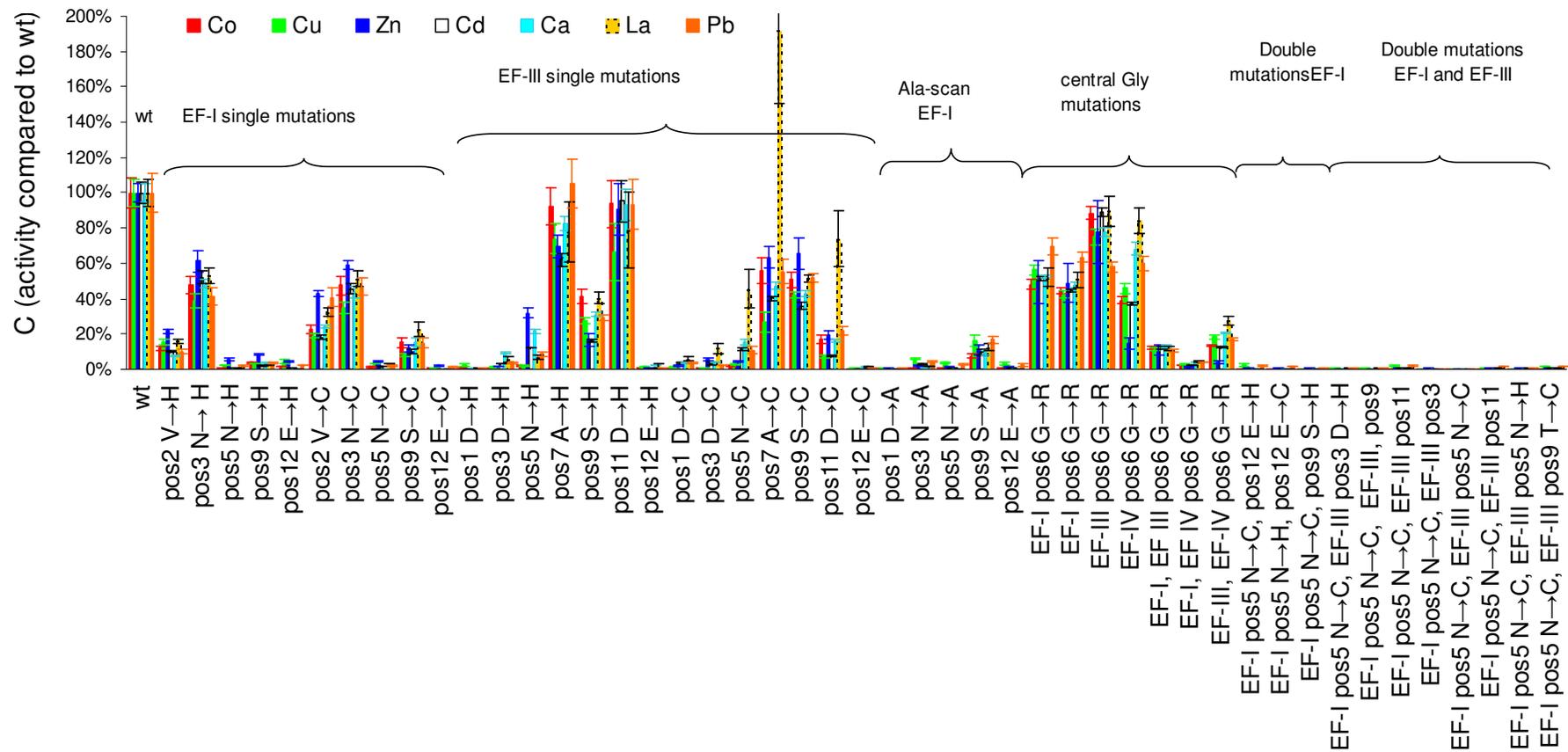


Figure 6-8 % Activity (C) compared to wild type against all ions tested, for the entire library.

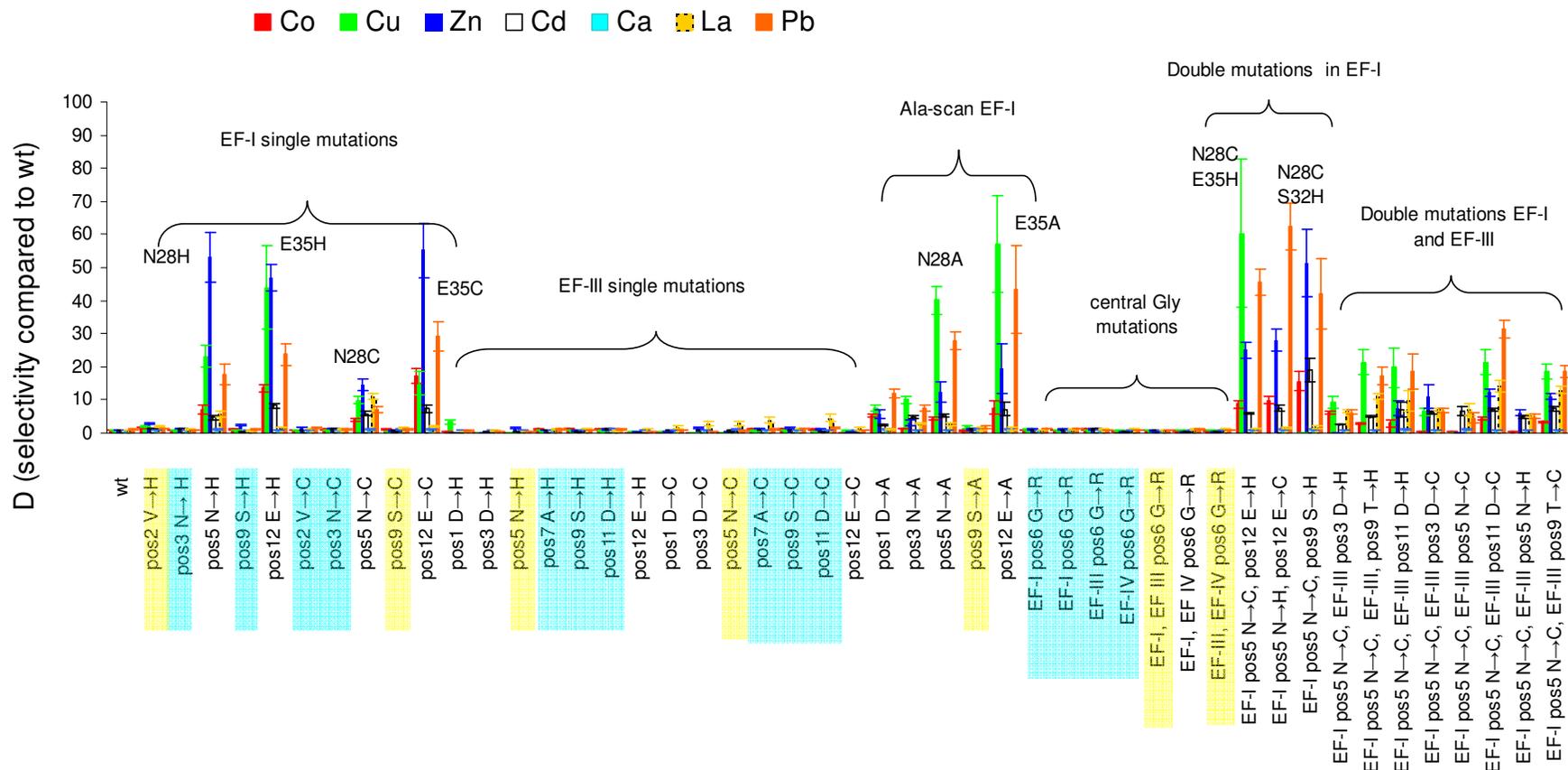


Figure 6-9 D as metric for ion selectivity shift of mutants. Highlighted labels: light blue for mutants exhibiting $C \geq 50\%$ and yellow for $10\% \leq C \leq 50\%$.

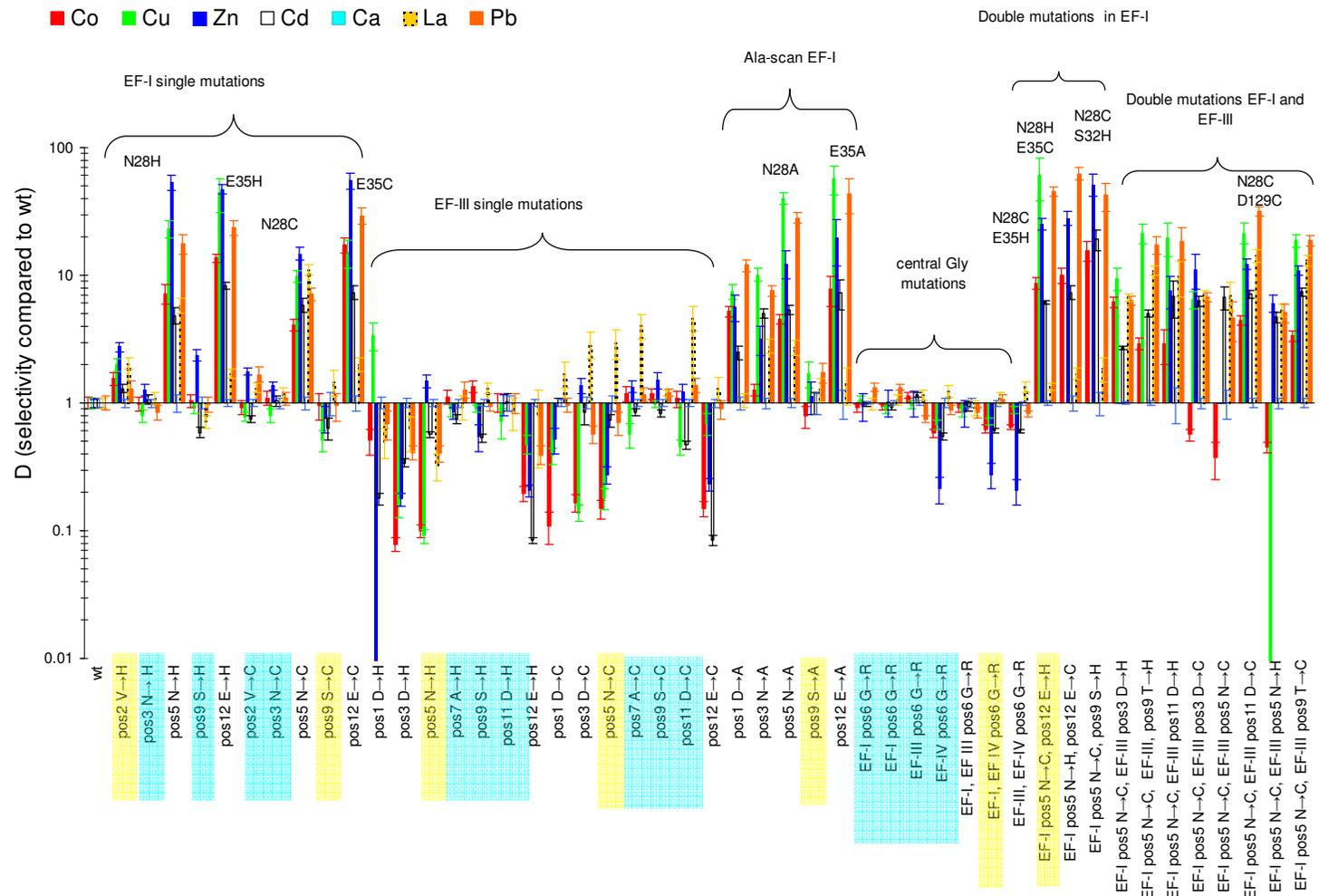


Figure 6-10 D on logarithmic scale. An alternative representation of Figure 6-9, clearly sets apart the cases where $D > 1$. D for wild type = 1. $D > 1$ indicates a stronger preference than wild type toward ions other than calcium. light blue for mutants exhibiting $C \geq 50\%$ and yellow for $10\% \leq C \leq 50\%$

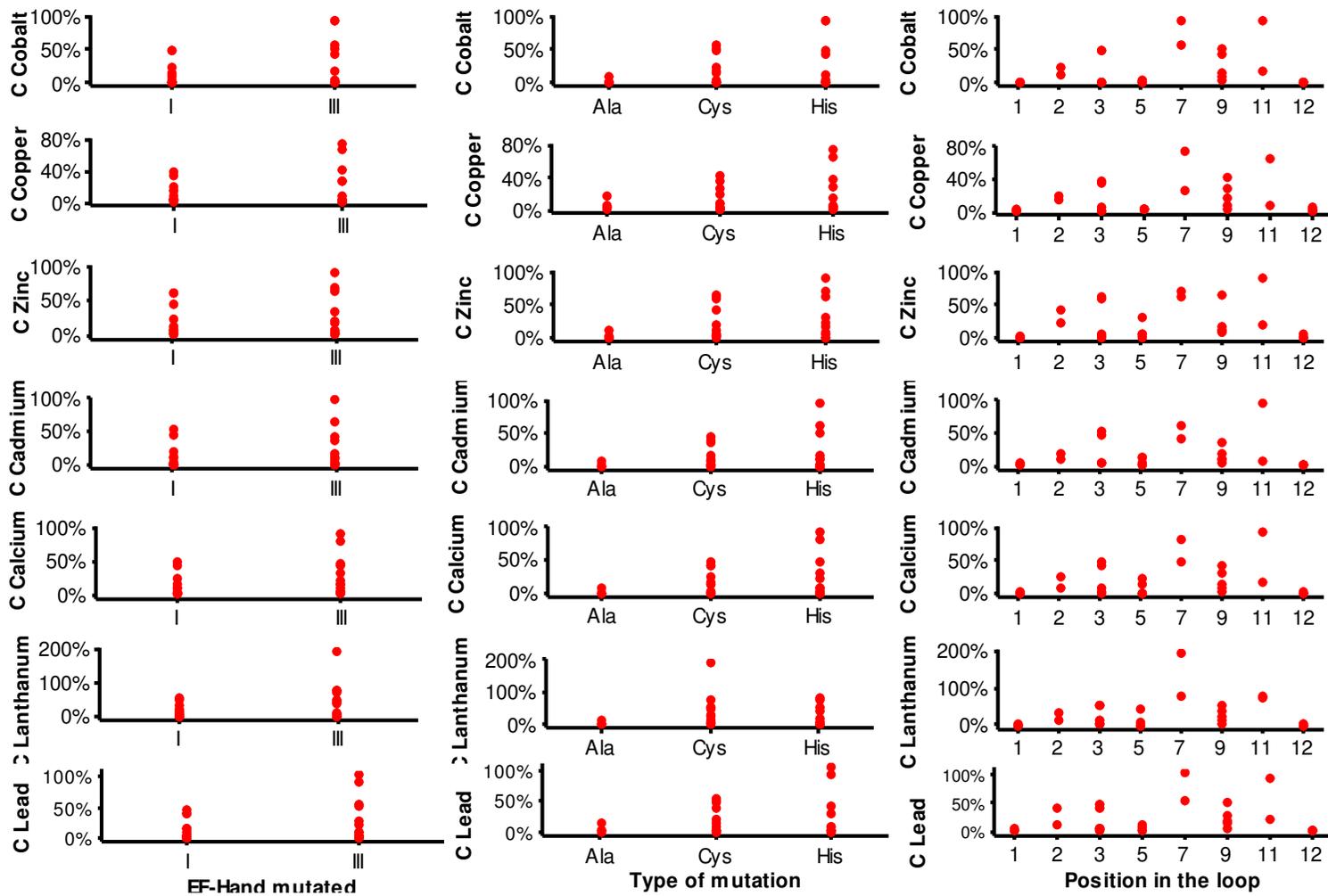


Figure 6-11 C (activity compared to wt) of single mutations. Arranged by site (EF-hand and position in the loop) and type (Ala, Cys, His) of mutation.

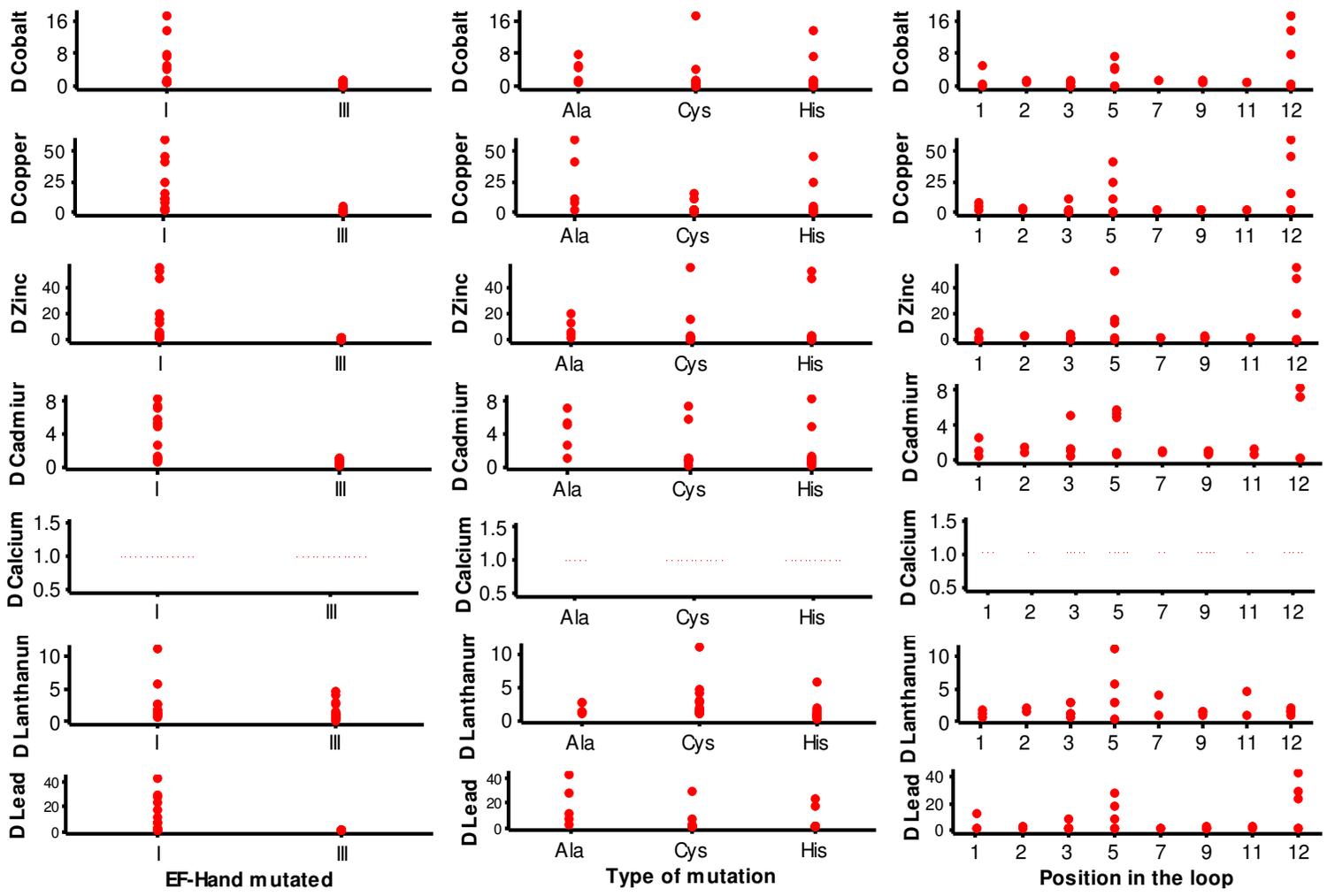


Figure 6-12 Selectivity (D) of single mutations. Arranged by site of mutation (EF-hand and position in the loop) and type.

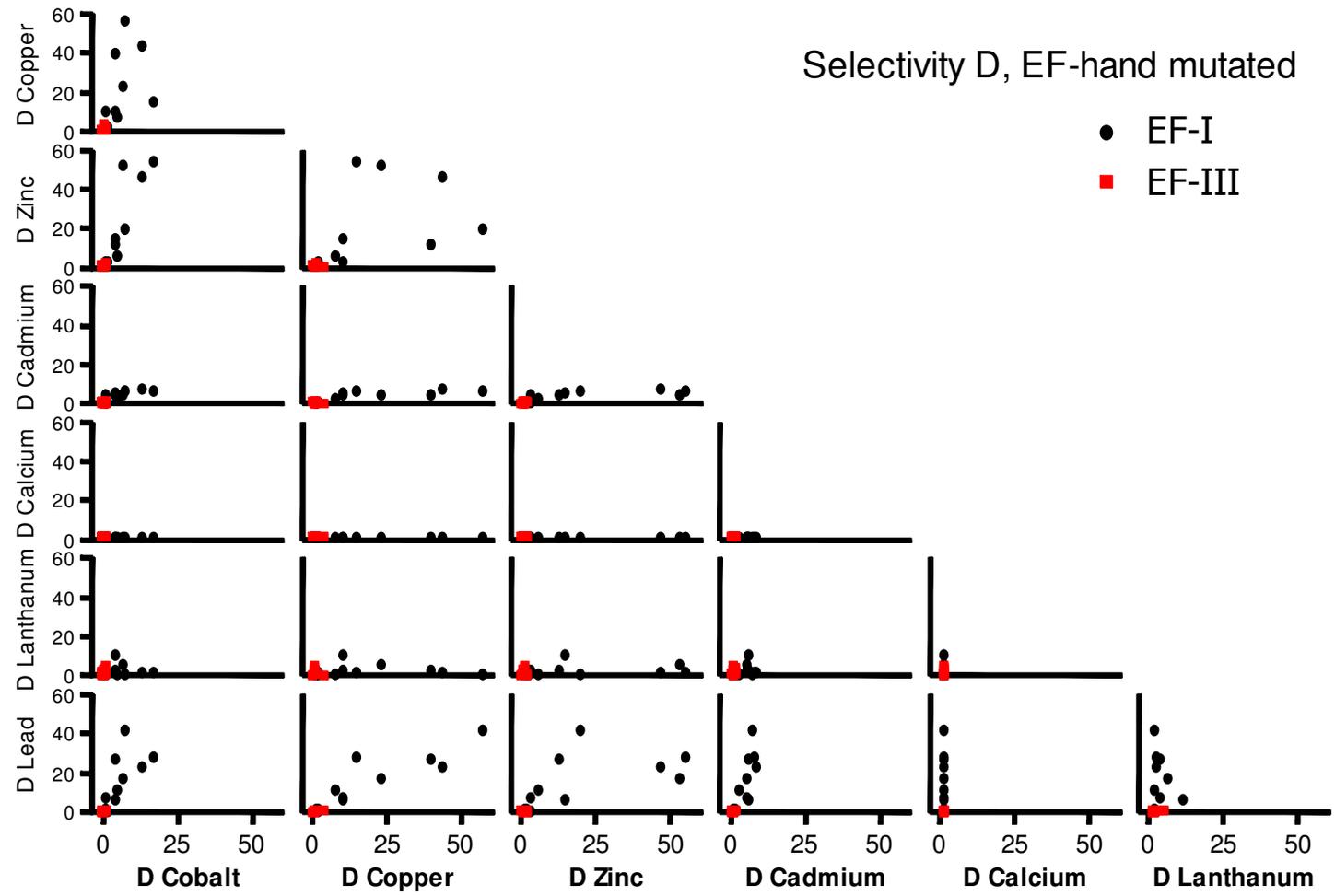


Figure 6-13 Matrix plot of D, arranged by number of EF-hand mutated. Metric D is a dimensionless number.

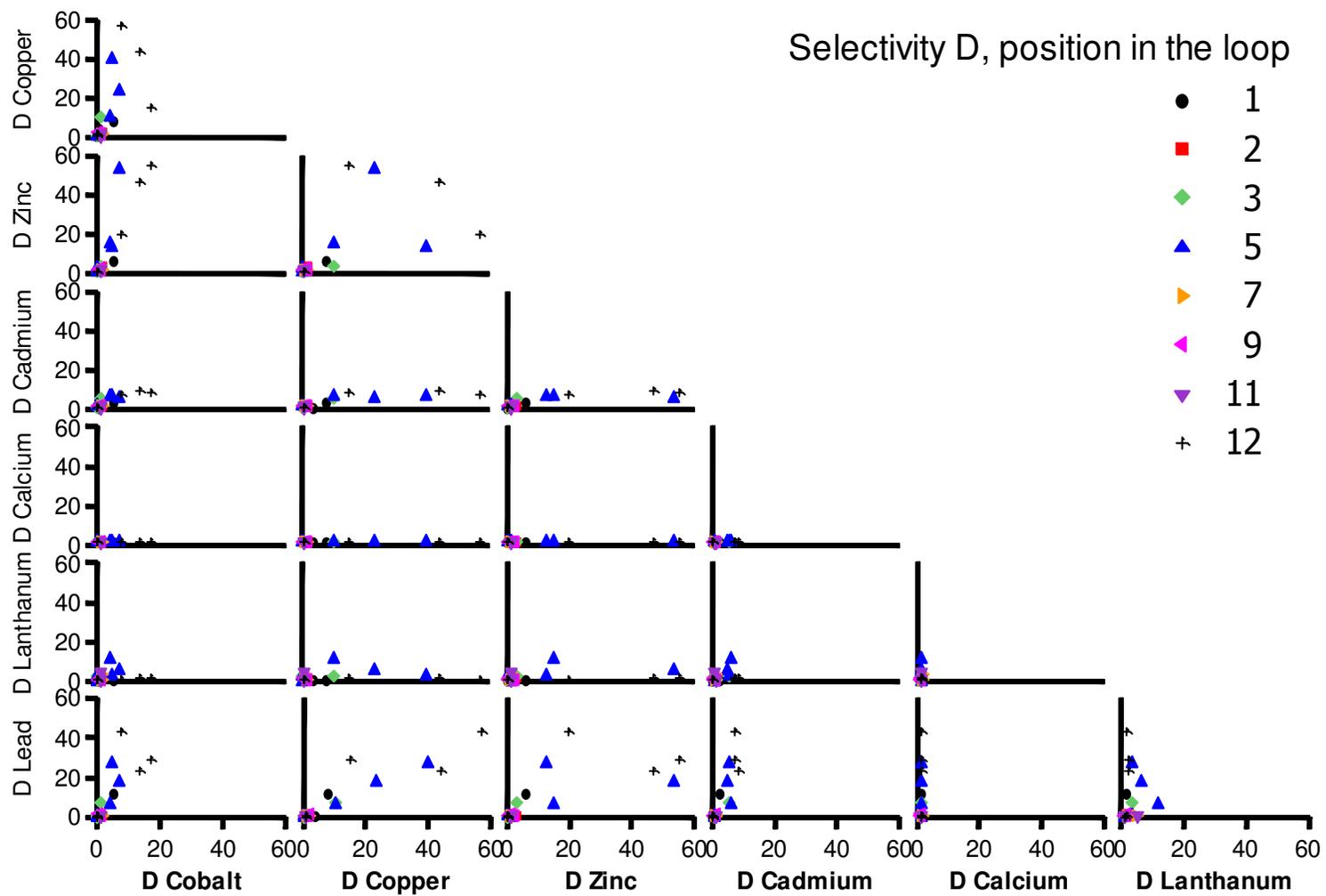


Figure 6-14 Matrix plot of D, arranged by position mutated in the EF-hand loop.

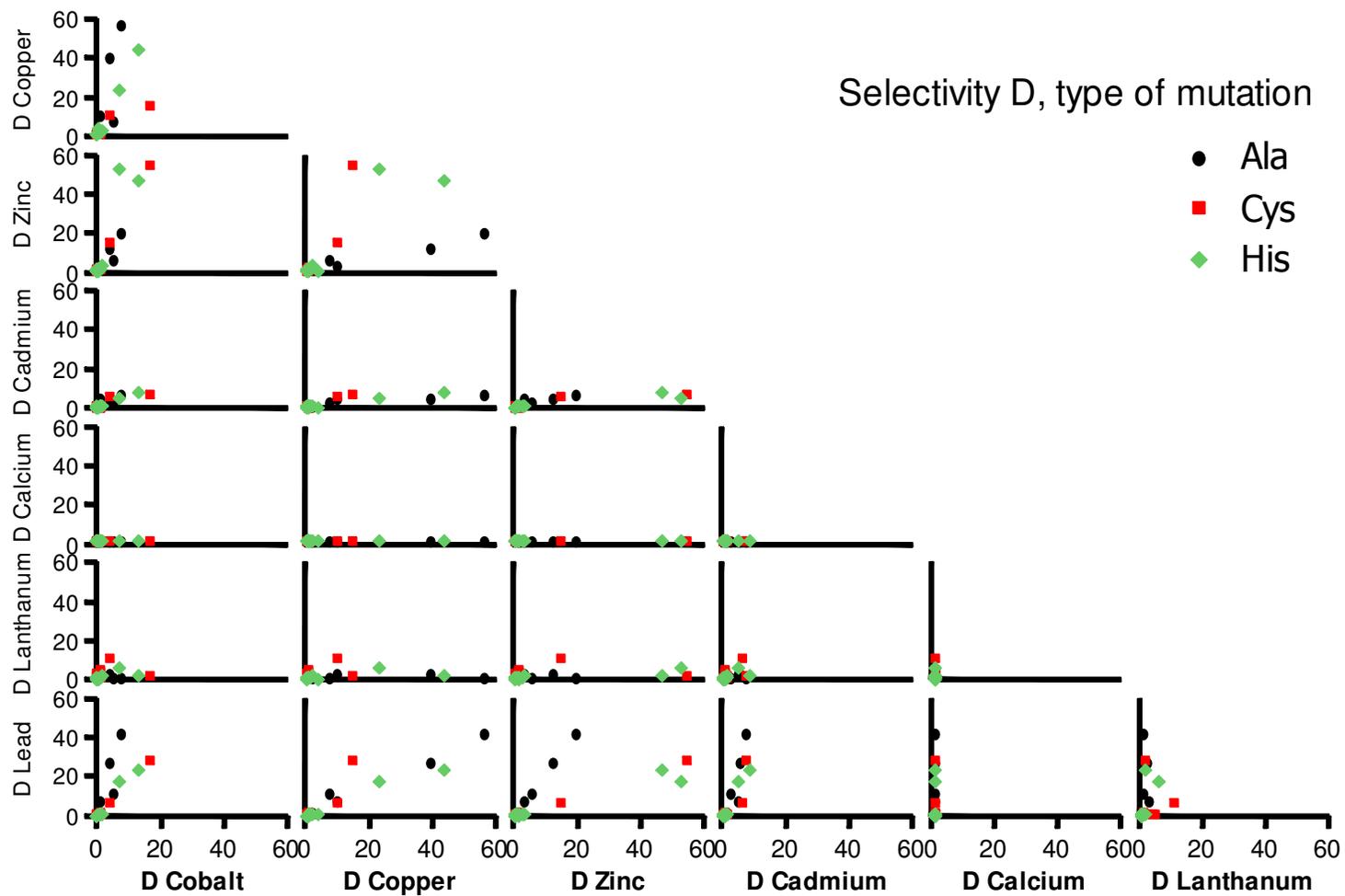


Figure 6-15 Matrix plot of D, arranged by type of substitution in the EF-hand loop.

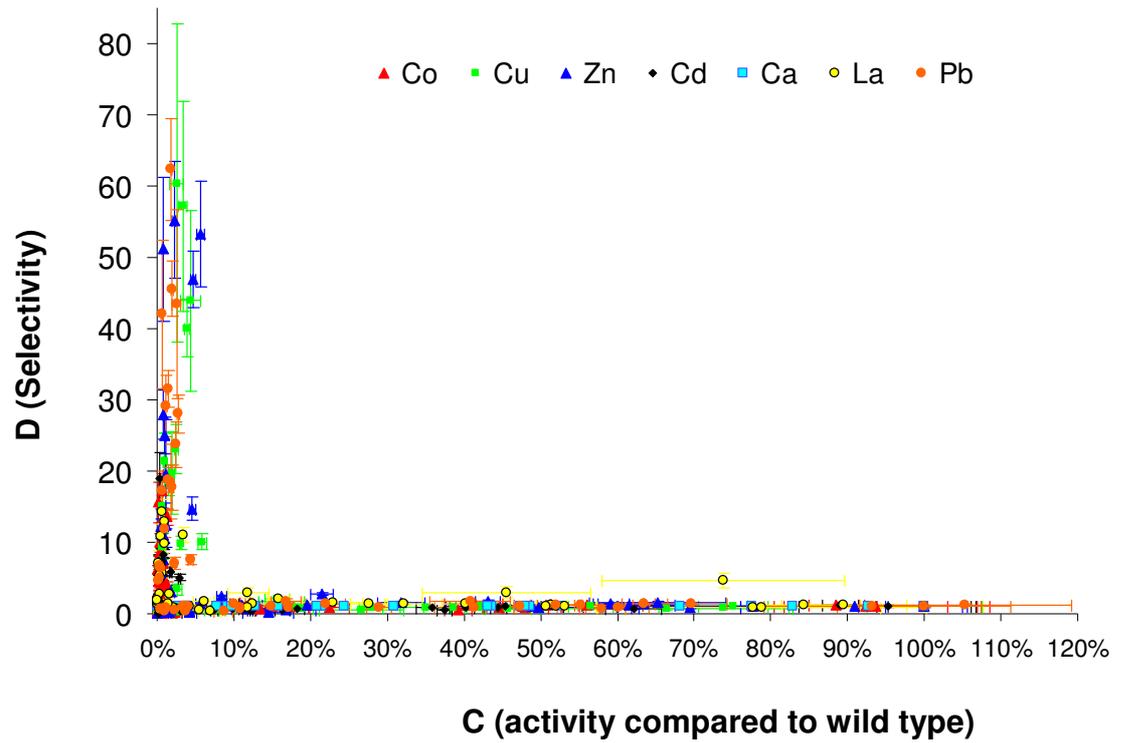


Figure 6-16 D versus C for all the mutants. Any shift of ion selectivity is associated with loss of activity for the entire mutant library.

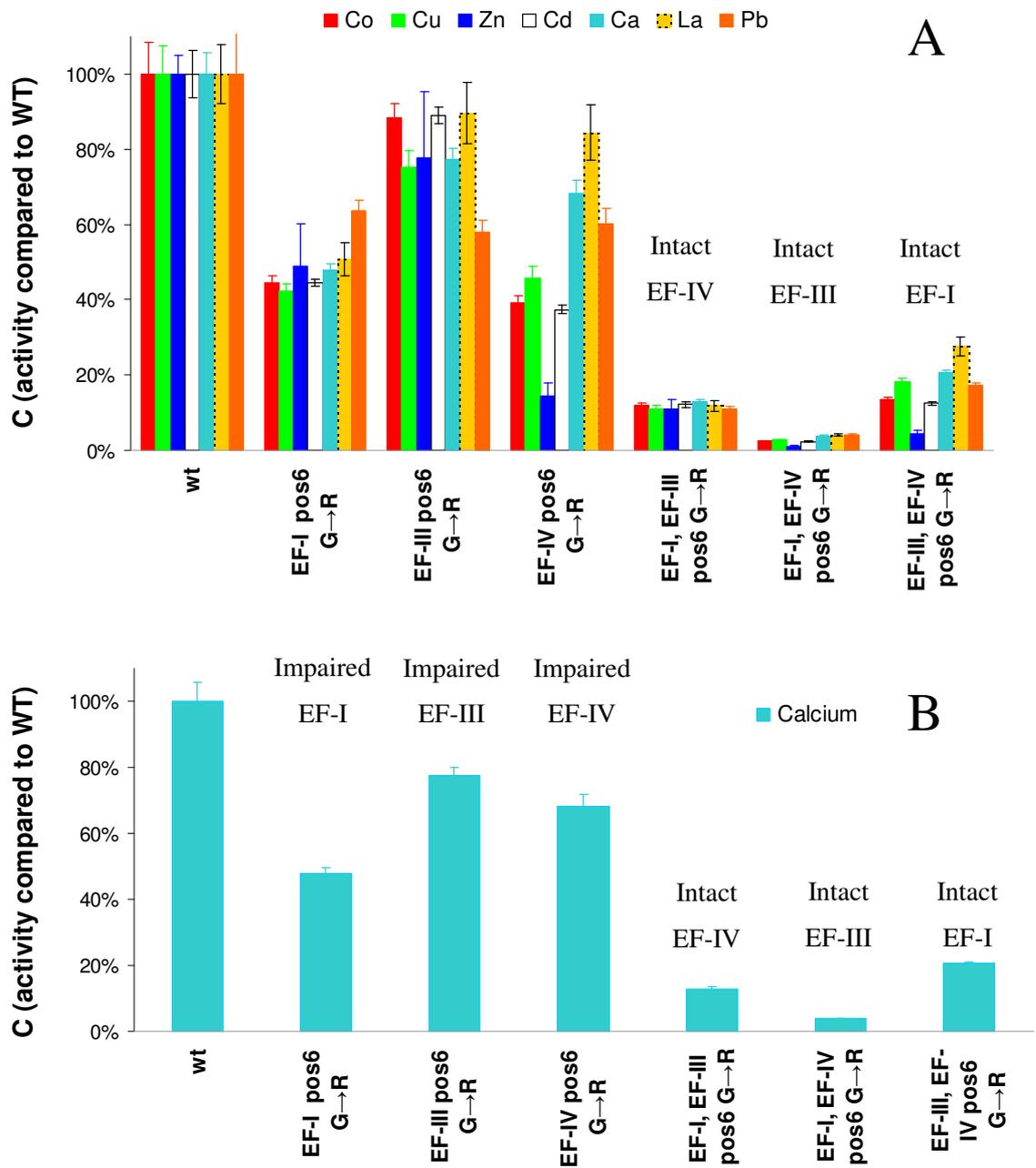


Figure 6-17 Activity of central Gly mutants (impaired EF-hands). Top: All ions. Bottom: Calcium only. Figure 6-17 zooms into the central Gly→Arg set of mutants from Figure 6-8 in order to aid visualisation.

6.4 Discussion

In Chapters 1 and 5 aequorin was shown to be a complicated allosteric and cooperative system. Calcium ions bind in the solvent-exposed EF-hands and the effect of this binding is transduced to the hydrophobic core of the protein where organic compound coelenterazine is bound as an oxidation intermediate. Destabilisation of the oxidation intermediate leads to its decarboxylation and emission of blue light. Upon ion binding the EF-hands undergo only a slight conformational change but an intricate network of hydrogen bonds connects the sites of ion binding with the sites of coelenterazine binding.

The effect of mutations placed in the EF-hand loops will be discussed in terms of aequorin's response to its natural activator (calcium and substitute lanthanum) and candidate activators (zinc, copper, cadmium, cobalt and lead).

The discussion of the library screening results will focus on salient points followed by some more detailed observations on the effect of position and type of mutation on the metrics of protein's yield (C) and ion selectivity (D).

6.4.1 Overall effect on yield

The first striking observation is that the vast majority of mutations lead to significant loss of yield compared to wild type. This applies to all ions tested. The only exception includes one mutant's response to lanthanum and will be discussed. Loss of protein function is not surprising for two reasons: (1) canonical EF-hands are finely-tuned to be highly selective for calcium and (2) the mutants of this library contain only single or double mutations out of a total of eighteen ion coordinating positions in the entire protein. These mutations are not expected to completely transform the ion binding sites in favour of new ions whilst fully retaining the intricate network of allosteric activation of aequorin.

6.4.2 Selectivity (D) versus retained activity (C)

The second major observation is that there is trade off between yield (C) and shift of selectivity (D). This effect is best illustrated in Figure 6-17. Increase of D comes at the cost of C and vice versa, highly retained activity C is associated with low or no change in selectivity. This again is not a surprising outcome in protein engineering, especially as a result of an initial round of mutation.

6.4.3 EF-I versus EF-III

Retained activity

At first glance of Figures 6-8 and 6-11 it would appear that EF-III is overall more tolerant to mutations than EF-I. Upon careful examination it can be seen that substitutions at all coordinating positions lead to 50–100% loss of luminescence in both EF-I and EF-III, for all ions tested. Mutants in EF-III exhibiting higher yields carry substitutions at positions not mutated in EF-I: position 7 (coordinating through backbone oxygen and non-conserved) and position 11 (non-coordinating and non-conserved).

Selectivity

Shift of selectivity (D) is markedly different in the two studied EF-hands (Figure 6-9, 6-10 and 6-12), in contrast to the comparable loss of retained activity (C) which is more uniform across the library. Some mutants in EF-I exhibited significant increase of D, whilst D remained very low in all EF-III mutants. This leads to the conclusion that EF-I is responsible for the selectivity of aequorin against various metal ions, whereas EF-III does not appear to be as suitable for modulating ion selectivity.

The contribution of EF-I and EF-IV is discussed in Section 5.1.3: the exiting helix of EF-I contains residues that directly hydrogen-bond to coelenterazine or are within close proximity (Figures 5-1 to 5-6). Aequorin's C-terminal tail is flanked by the entering helices of both EF-I and EF-IV. Movement of EF-I or EF-IV helices can disrupt the hydrogen bond network around the C-terminal tail which would cause relocation of Tyr184 side chain and disrupt hydrogen bonds to His169 and the peroxide. A possible sequence of events following binding of calcium was proposed by Head and

co-workers (2000), Deng and co-workers (2004) and Vysotski and Lee (2004) and summarised in Section 5.1.3.

6.4.4 Ala-scan mutants and importance of side chains

The alanine mutants (designed in Section 5.6.3) illustrate the effect of removal of the side-chain oxygens from the calcium-coordinating positions 1, 3, 5, 9 and 12 of EF-I. These results serve as references for the effect of the side-chain substitutions of histidine and cysteine at the respective positions of EF-I. Whereas the alanine substitutions completely remove the ability to coordinate ions through a side chain oxygen, the histidine and cysteine substitutions introduce the possibility to coordinate ions using an imidazole nitrogen or a sulfur of a sulfhydryl group, respectively.

Salient observations

All alanine mutants result in sharp loss of activity (C) compared to wild type (Figure 6.9) and increase of the selectivity measure D. Every single one of the conserved coordinating amino acids is very important to activity. Increase of D can occur from either reduction of affinity for calcium or increase of affinity for other ions, or both. As activity triggered by calcium is reflected in the denominator of the fraction defining D (Equation 6.2), a reduced response to calcium is sufficient to increase the value of D.

Increase of D occurs in positions 5 and 12 of the EF-I loop. Values of D among alanine, histidine and cysteine substitutions of EF-I are comparable, which suggests that the elimination of the coordinating side chains of aspartate, asparagine or glutamate is in fact the highest contributor to the shift of selectivity away from calcium. However, there is some differentiation in D values for different ions depending on the type of substitution, suggesting that histidine and cysteine do indeed fulfil some requirement for triggering luminescence. These subtle differences will be presented in Section 6.4.6.

6.4.5 Importance of loop positions

Out of the coordinating positions, the entry (position 1) and exit (position 12) of the loops are least tolerant to mutations in both EF-hands. The other coordinating

positions of the loop exhibit different levels of tolerance to mutation depending on the EF-hand they belong to. For example, position 5 in EF-I does not tolerate mutations whereas in EF-III it appears to be more tolerant. Position 3 is more tolerant to mutations in EF-I whereas position 9 is more tolerant in EF-III. All mutants that stand out in terms of selectivity contain substitutions at positions 5 and 12. As a conclusion, positions 1 and 12 must be essential for activity, followed by positions 3, 5 and 9 which may also be endowed with modulator roles in ion binding.

6.4.6 Type of substitution

The type of substitution (histidine or cysteine) contributes to the selectivity towards various ions. Whereas the alanine mutants served as reference in the coordinating positions of EF-I, the matrix plot of Figure 6-15 presents a general overview of the selectivity of mutants according to the type of substitution for all of the ions and positions tested. Some preferences can be extracted. For example, copper prefers histidine over cysteine while zinc prefers cysteine as well as histidine ligands. There is a preference between lead and alanine and between lanthanum and cysteine. However any interpretations of preference must take into account the context of the protein environment. Figures 6-12 to 6-14 capture the positioning and type of mutation but not the effect of steric, electrostatic effects, van der Waals and hydrogen bonds in the extended environment.

Mutants' increase of D for lead can partly be explained with the help of Figure 5-13: lead may require only four (IV coordinate binding) or six (VI coordinate binding) out of the eight oxygens that coordinate calcium in the native EF-hand of aquorin, whereas the loss of one coordinating oxygen is detrimental to correct calcium coordination.

Substitution of glutamate at position 12 removes two Ca^{2+} -coordinating oxygens and introduces nitrogen or sulfur instead. Substitution of asparagine at position 5 removes one side chain oxygen. Asn→His at position 5 of EF-I loop favours zinc and lead, Asn→Ala favours copper and lead and Asn→Cys exhibits the least discrimination among metals. At position 12 of EF-I Glu→His favours zinc, copper and lead while Glu→Cys favours zinc and lanthanum and Glu→Ala favours copper and lead. At position 9 Ser→His favours zinc, which does not occur for Ser→Cys or Ser→Ala.

The matrix plots of Figures 6-13 to 6-15 can aid in identifying correlations in the shift of selectivity for groups of ions. Ideally the envisaged mutant would exhibit high D values for one ion and very low values for all the remaining ions. Clusters of points or points falling onto a $y=x$ line would suggest that increase in selectivity is not evolving separately for different ions. For example, preference for zinc and copper is linked which suggests certain mutants would not make good biosensors for one in the presence of the other. However this study cannot determine dissociation constants which would be the required indicator as to whether one ion would be selected from a background of potential competitors. In conclusion, it can be seen that selectivity can potentially be tweaked through combined mutations in the metal binding loop of EF-I.

The exception: one high performer towards lanthanum

The mutation of position 7 of EF-III from alanine to cysteine (Ala123Cys) resulted in the only aequorin variant that responds to another ion (lanthanum) more efficiently than it does to calcium (Figure 6–9). The nucleophilic sulfur of cysteine attracts the increased (trivalent) positive charge of lanthanum and its small size increases the cavity size of the EF-hand loop to accommodate the slightly larger lanthanum ion. Cysteine is not a ligand normally found in calcium-binding sites and its shorter side chain could make it more difficult to coordinate to calcium ion, although calcium is coordinated to the backbone oxygen of position 7.

Lanthanum is large enough to reach the cysteine and can still use the remaining intact coordinating amino acids for successful coordination. The high tolerance of mutations in this position, as well as in position 11, suggests that these could be used as a means to alter the ion binding properties of the EF-hand without destroying the mechanism of allosteric activation.

The principle of ion discrimination through modulation of cavity size was shown in Section 5.3.1. In particular, Le Clainche and co-workers (2003) abolished calcium affinity and enhanced terbium affinity by replacing Glu with Asp in position 12 of EF-hand-like peptides.

6.4.7 Impairment of EF-hands: central Gly mutants

The central Gly mutants (Gly→Arg at position 6 of each loop) were created in order to compare the significance of the three EF-hands in the bioluminescent reaction. Two previously published strategies towards impairment of the EF-hands are discussed in Section 5.2.2. The mutational strategy by Tsuji and co-workers (1986) (Gly→Arg substitution of the highly conserved central glycine of each loop) was chosen to “disable” each of the EF-hands of aequorin (Section 5.2.2). Arginine is meant to introduce rigidity in the ion binding loop and prevent the mutated EF-hand from assuming the required conformation upon ion binding (section 5.2.2). It is possible however that an EF-hand may already be in the right order before being populated by the ion, requiring minimal adjustment upon ion binding (Section 5.2.2).

An alternative strategy for impairing EF-hands by Tricoire and co-workers (2006) consisted of replacing the position 12 glutamate with glycine, thus removing one essential amino acid that carries two coordinating side chain oxygens. These mutants were still able to carry out cooperative ion binding, as indicated by the sigmoidal shape of activity versus Ca^{2+} concentration curves (Tricoire *et al*, 2006). Although the strategy by Tricoire and co-workers was not applied for EF-hand impairment, positions 1 and 12 of the loop were targeted with His, Cys and Ala substitutions. Tricoire and co-workers (2006) showed that impairment of EF-I resulted in higher calcium sensitivity whereas impairment of EF-III or EF-IV resulted to lower calcium sensitivity. As a result, low calcium affinity was assigned to EF-I and high affinity to the C-terminal EF-hand pair, EF-III and EF-IV.

The results obtained in this work show that the substitution central Gly→Arg is less damaging than removal of a single coordinating side chain, as shown by the groupings of mutants in Figure 6-8. Amongst the single central Gly mutations (Figure 6-17A,B), impairment of EF-I results in greater loss of activity towards calcium, followed by impairment of EF-IV and EF-III. Amongst the double central Gly mutants, the one with only EF-I intact exhibits the highest activity, followed by the mutant with intact EF-IV. Both these findings suggest that EF-I is more important to activity than EF-III and EF-IV and hence more likely to dictate selectivity.

The results agree qualitatively with the findings by Tsuji and co-workers (1986) in terms of the importance of EF-I, but not regarding EF-III and EF-IV. In their work, Gly→Arg in EF-I resulted in the loss of almost 100% activity, in EF-III ~50% loss and

in EF-IV almost no loss of activity. Based on this results, the authors suggested that EF-IV is not necessary for activation of aequorin by calcium. The differences may be attributed to using a different experimental system (conditions and aequorin wild type).

EF-III is more tolerant to central Gly→Arg substitution than EF-I and EF-IV, probably due to being already preformed for the binding of calcium (Sections 5.5.2 and 5.6.1) and flexibility is not critical. Population of calcium in the loop may already have been in the right order, despite the particular mutation.

Interestingly, impairing EF-IV reduces C and D for zinc, more so than it does for the other ions, an observation made on both single and double central Gly mutants (Figure 6-17A). This suggests that EF-IV is more actively implicated in the response to zinc than the other EF-hands and that the flexibility offered by the central glycine in the loop is comparatively more important for zinc binding. Considering the low ionic radius of zinc in its preferred geometries (refer to Section 1.4 and Table 9-4 to Table 9-6), the dependency on loop flexibility is also consistent with the fact that ligands of the EF-hand would have to reach further towards the metal in order to coordinate it.

Mutants with impaired EF-I exhibit the lowest activity. Impairment of EF-III results in the least loss of activity amongst the single central Gly mutants. Out of the single intact EF-hands (double central Gly mutants) the variant carrying only EF-III intact exhibits the maximum loss of activity. EF-I is more susceptible to change of selectivity towards most of the metal ions tested (Figure 6-9), which agrees with the proposition in published work that EF-I is endowed with lower affinity towards calcium compared to the other EF-hands (Tricoire *et al.*, 2006).

Positions 1 and 12 of the loop are important in its rearrangement and position 12 is believed to be important for cooperativity (Gifford *et al.*, 2007). In wild-type aequorin an ion as small as zinc or copper cannot be coordinated by both residues as it is not possible to be within required distance from both.

The EF-I calcium-binding loop's preference for ions other than calcium can be further manipulated. In order for a smaller ion – e.g. zinc – to be tightly bound in the EF-hand and exert the required transduction of the binding effect to the protein core, the metal will have to be within a range of distances from the coordinating atoms, as observed in existing protein metal binding sites (Section 1.4, Table 1-8).

6.4.8 Double mutations

Mutations in this group were directed at a limited set of positions of loop EF-I and EF-III: position 5, 9 and 12. In the variants featuring double mutations, new ion selectivity is much higher when both mutations are in EF-I than when one mutation is in EF-I and one in EF-III. All of the double mutants tested exhibit increase of D with lead. Asn28Cys/Glu35His favours copper while Asn28Cys/Ser32His favours zinc.

6.4.9 Software for prediction of mutation effects

The substitutions will have steric and electrostatic effects, influence the hydrogen bond network, the hydrophobic and van den Waals interactions in the local environment and where applicable, formation of disulfide bridges. One possible means to predict the effects of mutations would be to use molecular modelling software such as ModLoop (Fiser and Sali, 2003). ModLoop predicts conformation of loops in protein structures by satisfaction of spatial restraints. However, a control simulation of the loop of native Ca²⁺-bound apoaequorin (PDB ID: 1SL8) resulted in an arrangement of the loop resembling that of calcium-free aequorin, PDB ID: 1EJ3. This was an expected outcome as the software cannot take into account the presence of bound metals. The coordination of calcium overcomes considerable steric clashes which would not be allowed in a local energy minimisation calculation but which are overcome by the electrostatic attraction between the oxygen ligands and calcium. Due to lack of predictive power in this case, molecular modelling software were not used to simulate mutated, ion bound loops. The effects of mutations were discussed based on properties of the respective amino acids, their position in the protein and previous structural knowledge of aequorin and other EF-hand like structures.

6.5 Conclusions

This chapter was divided in two parts:

(1) A library of forty-eight mutant variants of the apoaequorin gene was successfully created using QuikChange[®] Site-directed mutagenesis. Expression of the apoaequorin variants and reconstitution of aequorin mutants was carried out in microplates. The choice of mutations is explained in Chapter 5, Section 5.6.3.

(2) High-throughput screening of the aequorin mutant library was carried out against seven different metal ions (calcium, lanthanum, cadmium, zinc, copper, cobalt, lead). The results were discussed in terms of two assigned metrics: one for yield (C) and one for ion selectivity (D). The main conclusions from the library screening follow.

In aequorin, EF-I is the gatekeeper for ion selectivity (Section 6.2.3 for definition of metrics for retained activity C and ion selectivity D). EF-IV has significant potential as determinant of ion selectivity but was not investigated to the same extent as were EF-I and EF-III. It is potentially more important to the binding of zinc, a finding which arose from impaired EF-hand mutants (central Gly→Arg). In Chapter 5, analysis of the crystallographic structure of the EF-hands with and without calcium bound revealed that among the three calcium-binding EF-hands, EF-IV undergoes the most extensive spatial rearrangement in order to bind calcium in the optimal conformation (Section 5.5.2). The helices flanking the loops of EF-I and EF-IV contain amino acids that directly and indirectly stabilise the hydroperoxy-coelenterazine (Section 5.1.3). Conformational shifts in the loops of EF-I and EF-IV can move their respective helices and destabilise the peroxide. Future work on shifting ion selectivity of aequorin should include mutations in EF-IV.

EF-III does not seem to be endowed with ion selectivity properties in aequorin. Every mutant in EF-III exhibited very low values for the metric of selectivity D. Results from impaired EF-hand mutants (central Gly→Arg) suggest that flexibility is least required in EF-III, which agrees with the structural observation that the loop of EF-III is preformed to bind calcium (Section 5.1.3 and 6.4.3).

Selectivity towards new ions increases (and shifts away from calcium) at the cost of activity. All conserved coordinating positions are very important to calcium binding and loss of a single coordinating side chain oxygen resulted in dramatic loss of activity in both EF-I and EF-III.

As expected, the coordinating and conserved positions 1 and 12 (entry and exit) of EF-hand loops are the least tolerant to mutations. They are also 100% conserved amongst the entire EF-hand superfamily, as shown in Table 1–4, Section 1.3.4. Non-coordinating (position 11) or coordinating but not conserved (position 7) residues are generally more tolerant to mutations and may be used to fine tune ion selectivity. Within EF-I, positions 5 and 12 are most important for shifts of ion selectivity. Positions 3 and 9 are not prominent in shifting of selectivity (Section 6.4.5) but could be important in modulating activity. The importance of positions 3 and 5 is consistent with the extensive structural rearrangement of the N-terminus (positions 1–5) of the EF-hand loops upon calcium binding, as seen in the comparison between aequorin structures with (1SL8) and without calcium (1EJ3) (Section 5.5.2). The ion selectivity of the loops and the transduction of the signal to the protein core could potentially be modulated through a combination of mutations involving coordinating and non-coordinating residues. This agrees with the conclusions by Le Clainche and co-workers (2003 and 2006).

The effect of mutations is environment-dependent: location in the protein (e.g. EF-hand), position in the binding loop, type of substitution and type of metal tested all play a role in the effect of the mutation on selectivity and activity. Effects of the same substitutions are not necessarily the same across different EF-hands of aequorin (Section 6.4.5). Mutation Asp→Asn causes complete loss of calcium affinity in short EF-hand like peptides (Le Clainche *et al*, 2003 and Le Clainche *et al*, 2006) but in aequorin position 3 is conserved Asn. Oishi and co-workers (1992) working on the loops of EF-I, EF-III and EF-IV of aequorin as 22 amino acid fragments found different order of calcium binding affinity (EF-III>EF-I>EF-IV) than Tricoire and co-workers (2006) in aequorin (EF-III and EF-IV>EF-I), while and Liu and co-workers (2003) showed that in crystals of obelin (photoprotein homologous to aequorin) EF-I loop was the first to be populated by a calcium ion.

In this mutant library, selectivity is not shifted toward one new ion independently from the rest (Section 6.4.6). Additional studies would be required in order to achieve high affinity for one ion against a background of other metals in a mixture.

This experimental setup measures activity (light output) only; it does not capture structural changes upon ion binding nor does it measure free concentration of metal ions, which would then allow direct estimation of binding affinity. Selectivity (D) is inferred by comparing activity for new ions to activity for calcium.

6.5.1 Choice of mutant for further study

A number of mutants stood out as potential candidates for further characterisation, based on increased selectivity for new ions over calcium, compared to the wild type. All the mutants carried substitutions in EF-I only. Amongst the single mutants the ones that stood out were Asn28His, Asn28Cys, (at position 5 of the loop), Glu35His and Glu35Cys (at position 12 of the loop). Amongst the double mutants the ones that stood out were Asn28His/Glu35Cys, Asn28Cys/Ser32His and Asn28Cys/Glu35His. Asn28Cys/Ser32His exhibited higher selectivity for zinc and lead, but not copper; Asn28His/Glu35Cys exhibited higher selectivity for lead; Asn28Cys/Glu35His exhibited higher selectivity for copper and lead.

A variant with improved lanthanum activity, Ala123Cys was not deemed as a priority for further testing. An aequorin mutant with improved lanthanum activity had already been reported by Le Clainche and co-workers (2003), hence focus was drawn to mutants that had the potential for improved response to other metals.

Due to time constraints, only one mutant could be purified and further characterised (Chapters 7 and 8). Asn28Cys/Ser32His was chosen for its high selectivity metric D towards zinc, for the fact that two essential side chain oxygens were replaced and finally due to high cell growth in 500-mL cultures after transformation into high expression *E. coli* (Lucigen) cells.

7 Protein purification

This chapter describes the production and purification of wild-type aequorin and mutant Asn28Cys/Ser32His so that their activity may be further assessed (in Chapter 8). Mutant Asn28Cys/Ser32His emerged as an interesting candidate for further study amongst other aequorin variants, as explained in Chapter 6. The size and purity of recombinant aequorin obtained was verified by combination of luminescence activity measurements, SDS-PAGE protein electrophoresis, spectrometry, Bradford protein assay and Mass Spectrometry. The steps involved in this chapter are summarised in Figure 7-15 and further described in the respective subsections. Commercially available pure recombinant aequorin was also purified under the same conditions as wild-type and mutant Asn28Cys/Ser32His in order to subject it to the same conditions – exposure to the same chemicals, light and duration of process.

7.1 Introduction

A range of methods for the purification of apoaequorin and reconstitution of fully charged aequorin have been described in literature. Shimomura and Inouye (1999) described a method for preparing 40–60 µg/mL purified recombinant aequorin from *E. coli* cultures. Apoaequorin was expressed in the periplasmic space of *E. coli*, extracted from the cells and regenerated into aequorin in one step, by incubation with a buffer solution containing coelenterazine. A signal peptide sequence from the *Escherichia coli* outer membrane protein A (*ompA*) was used to express and translocate apoaequorin into the periplasmic space of *E. coli* cells. The signal peptide sequence is cleaved by proteases as it enters the periplasm. Apoaequorin is expected to leak from the periplasm, driven by a concentration gradient and bind coelenterazine in the medium. Some of the apoaequorin chain may bind to coelenterazine to form a holoprotein while still in the periplasm, as coelenterazine readily permeates cell membranes from the media. Shimomura and Shimomura (1981) had previously proposed that conversion of apoaequorin into aequorin increases protein stability and resistance to proteases and presumed that aequorin's more rigid conformation facilitates its extraction from the periplasm into the medium (Shimomura and Inouye, 1999). Due to the mild *in situ* extraction conditions, impurities in the extract were minimal. After

obtaining the supernatant, purification was carried out in two chromatographic steps: anion exchange (AEX) and hydrophobic interaction chromatography (HIC).

A fusion of apoaequorin cDNA to *ompA* peptide sequence was previously used by Inouye and co-workers (1989). When the apoaequorin cDNA was expressed in *E. coli*, a large excess of the recombinant protein was produced and released into the culture medium. Purification of the protein was accomplished by acid precipitation and anion exchange (DEAE-cellulose) chromatography to 95% purity (Inouye *et al.*, 1989).

An alternative method made use of the 6×His-tag sequence fused at the C-terminus of apoaequorin (Chapter 3). Glynou and co-workers (2003) reported a one-step purification and refolding of recombinant photoprotein aequorin by immobilized metal-ion affinity chromatography, utilising a hexahistidine tag fused to the N-terminus of apoaequorin (Glynou *et al.*, 2003). The yield of this method was 14–20 µg/mL of purified aequorin from an *E. coli* culture (Glynou *et al.*, 2003).

In this Chapter, purification of aequorin made use of the periplasmic translocation of apoaequorin and one-step extraction and reconstitution with coelenterazine into the medium according to Shimomura and Inouye (1999). Size exclusion chromatography was employed post aequorin reconstitution (instead of AEX and HIC) in the interest of minimising purification development time.

7.2 Materials

Commercial recombinant aequorin (AquaLite[®]) was purchased from Molecular Probes in units of 25 µg as lyophilized powder of a 1 mg/mL solution in 5 mM HEPES, 0.1 M KCl, 30 mM glucose and 5 µM EDTA, pH 7.1. Coelenterazine was purchased as 1 mg powder from Sigma-Aldrich. It was resuspended in 1 mL of ice cold ethanol to a 1 mg/mL solution immediately before use.

All other chemicals were of the highest grade (Sigma Aldrich Ltd.) and the theoretical maximum calcium contamination of stock solutions was calculated to be less than 2 µM (Section 2.5). Low retention plastics, RNase and DNase-free tips and eppendorfs were from Simport and Molecular BioProducts Inc.

7.3 Methods

7.3.1 Culture

7.3.1.1 Host strain and plasmid

E. coli host strain *E. coli*[®] BL21(DE3) (Lucigen Corp.) was transformed with the expression plasmid pETAQ. Plasmid pETAQ is a derivative of pET26b plasmid (Novagen) containing the apoaequorin cDNA and flanking DNA sequences that fuse a *pelB* signal peptide sequence at the N-terminus and a 6×His-tag at the C-terminus of the expressed apoaequorin (Section 3.3.2). This system is inducible by isopropyl-β-D-thiogalactopyranoside (IPTG) (Novagen, 2003).

7.3.1.2 Cell growth

A single colony was grown overnight in 7 mL of Luria-Bertani (LB) medium containing 50 µg/mL kanamycin at 37 °C (seed culture), as described in Sections 2.2.2–2.2.3. Two consecutive 5% inoculations were performed into fresh medium in a 500-mL flask (100 mL culture) and in a 2-L flask (500 mL culture) at 37 °C for 8 hours at 200 rpm. At the end of the 500 mL incubation, the bacterial suspension was divided into 50-mL falcon tubes and centrifuged at 4,500 rpm for 30 min. The supernatant contained negligible quantity of total apoaequorin (Shimomura and Inouye, (1999) and Section 4.4.2.2) and was decanted; any remaining supernatant was carefully removed from the pellets by pipetting. The cell paste was stored at –80 °C.

7.3.2 Regeneration and extraction of aequorin

7.3.2.1 One-step extraction and charging of active aequorin

The falcon tubes containing frozen cell paste were removed from –80°C storage and left to thaw on ice for approximately 20 min. To extract apoaequorin from the cells and reconstitute active aequorin the following procedure was followed: the thawed cell pellet in each falcon tube was dispersed in 17.5 mL of ice-cold buffer containing 50 mM Tris-HCl 10 mM EDTA and 2-mercaptoethanol or DTT (Appendix Chapter 7), pH 7.5 at 23 °C. The suspension was transferred into a 50-mL plastic beaker and stirred at ~250 rpm on ice using a magnetic stirrer. 700 µL of 1 mg/mL coelenterazine solution

(Section 7.2) were added to the stirred suspension in a dropwise manner, resulting in 92 μM coelenterazine concentration in the final mix. The beaker was covered with foil (coelenterazine is light-sensitive), placed in 4°C storage and left to stir overnight while allowing diffusion of oxygen to occur. Choice of the beaker geometry was such that the depth of the suspension was less than 3 cm to facilitate the diffusion of oxygen (Shimomura and Inouye, 1999). Following the overnight incubation the mixture was centrifuged at 13,000 rpm for 30 min at 4°C. A clear supernatant was produced that contained fully functional aequorin and the cell pellets were discarded.

Reconstitution with coelenterazine was initially performed in microwells (Section 7.4.1.1) where it was used to screen conditions for shake flask cultures; culture samples collected at different time points were measured for OD and activity post-reconstitution with coelenterazine. 100 μL of the collected samples were transferred into the corresponding wells of a new DSW plate, each containing 950 μL of cold coelenterazine solution 4 μM , 50 mM Tris-HCl, 10 mM EDTA and 5 mM 2-mercaptoethanol, pH 7.5. The DSW was covered with an inverted shallow 96-well plate (Sarstedt Inc), placed on a table top thermomixer (Eppendorf Inc) with intermittent stirring at 1,000 rpm and 1 °C; this was placed overnight in a cold cabinet (4–8 °C). The protocol was based on Shimomura and Inouye (1999) who had used a lower protein expression strain; in retrospect it was understood that 4 μM of coelenterazine was not sufficient for the apoaequorin produced in shake flasks in this work. The coelenterazine content was consequently increased in the finalised protocol for production of aequorin in 500 mL cultures (Figure 7–15). After the end of the cold incubation with coelenterazine solution, 100 μL of each well were transferred into the corresponding wells of a Lumitrac™ shallow well plate using automated liquid handling (Tecan). The activity of the regenerated aequorin was then assayed as described in Section 7.3.4.

7.3.2.2 Periplasmic fraction

Periplasmic extraction of apoaequorin from *E. coli* cultures by osmotic shock was performed using the protocol by Ausubel (1989) found in the Novagen pET system manual (2003). This preparation did not include coelenterazine for the reconstitution of active aequorin, but was an investigatory experiment to determine the localisation of apoaequorin in the cell. Cell pellet from 40 mL of *E. coli* culture was harvested by centrifugation and resuspended thoroughly in 30 mL of 30 mM Tris-HCl pH 8.0, 20% sucrose. EDTA solution (0.5 M pH 8.0) was added to final concentration of 1 mM.

PMSF was added to final concentration of 0.2 mM for protection against proteases and the mix was stirred slowly at room temperature for 10 min. The cells were collected by centrifugation at 10,000×g at 4°C for 10 min, the supernatant was removed and discarded. The pellet was resuspended in 30 mL of ice-cold 5 mM MgSO₄ and the cell suspension was stirred slowly for 10 min on ice. Periplasmic proteins should be released into the buffer during this step. The mix was centrifuged at 4°C for 10 min at 10,000×g and the shocked cells were separated from the periplasmic proteins (supernatant). One milliliter samples were collected for SDS-PAGE analysis and activity assays. The cell pellet was kept on ice for SDS-PAGE analysis of the soluble and insoluble cytoplasmic fractions. The periplasmic fraction (supernatant) was concentrated through spin filtration. Samples collected for SDS analysis were mixed with an equal volume of SDS loading buffer, immediately heated for 3 min at 85 °C to denature the proteins and stored at -20 °C.

7.3.3 Concentration of aequorin

7.3.3.1 Spin-filtration

Spin filters were used for the concentration and desalting of aequorin samples prior to size exclusion chromatography. Vivaspin[®] spin-filters (Vivascience, Germany) and Amicon Centriplus[®] (Milipore, Germany) are centrifugally operated, disposable ultrafiltration devices. Proteins are separated by size. They enable concentration of samples up to 30-fold and handle sample volumes of 0.5–15 mL (minimum final volume 200 µL). For maximum recovery the molecular weight cut-off (MWCO) should be at least 50% smaller than the molecular size of the species of interest (rule of thumb: 1/5 to 1/3 of the protein size). Spin filters of 10,000 30,000 and 100,000 kDa nominal MWCO were tested for the preparation of aequorin solutions past the coelenterazine incubation step.

Centrifugation was performed in 2–3 repeats at 4,000×g, 4°C for 15 min according to manufacturer's instructions (Vivascience, 2002) until the desired concentrate volume was reached. The spin columns were initially loaded with RO water to remove glycerine and sodium azide present in the membranes, followed by 50 mM Tris-HCl, 10 mM EDTA solution to remove cations which might contaminate the protein sample. At the end of centrifugation a yellow coloured liquid containing

concentrated aequorin (yellow indicates coelenterazine) was recovered from the concentrate pocket (retentate). 10,000 MWCO spin filters were used for desalting of aequorin samples prior to mass spectroscopy analysis or concentration prior to SDS-PAGE analysis.

7.3.4 Luminescence assay

Luminescence assays were performed on the size exclusion chromatography fractions in order to determine the active aequorin fractions. Volumes of 25–75 μL of sample were pipetted in microwells of a Lumitrac™ 96-well plate and injected with 100 μL of 10 mM CaCl_2 50 mM Tris-HCl pH 7.5 buffer solution, using the auto-injector of the BMG platereader at 23 °C. Platereader gain settings are specified in Supplementary Material. Luminescence assays were also performed in order to associate cell growth and activity with bacterial incubation time. These assays used crude bacterial sample (refer to Section 7.3.2.1 for methods and Section 7.4.1.1 for results). One hundred microliters of the regenerated aequorin mixtures were assayed in the FluoStar platereader, by measuring the light emission of the sample when an equal volume of 100 mM CaCl_2 solution in 20 mM Tris-HCl, pH 7.5 was injected.

7.3.5 Size exclusion chromatography

A 10mm internal diameter Pharmacia HR column was used, previously packed to a bed height of 27.5 cm (total gel volume 21 mL) with resin Pharmacia Superdex 75 Prep Grade (Amersham Biosciences, Freiburg, Germany). This resin is suitable for optimum molecular weight separation range of 3,000 to 70,000 (globular proteins). The column was operated with an AKTAprime™ purification system and Unicorn™ software (GE Healthcare, Uppsala, Sweden).

The quality of the column packing was checked by performing an acetone pulse test: 200 μL of 20 mg/mL acetone solution was pumped into the column through a 200 μL sample loop at 1.2 mL/min (equivalent to 60 cm/h). Efficiency and peak symmetry were calculated according to the manufacturer's protocol (GE Healthcare, 2006). $N=12,612$ plates (acceptable values $N \geq 10,000$) per meter and Peak symmetry $A_f=1.07$ (acceptable range: $0.7 \leq A_f \leq 1.3$) were found to be satisfactory.

Due to the temperature-sensitive nature of aequorin, the column and all buffers used during the purification process were kept at 0–4°C. The column was washed with 50 mM Tris-HCl solutions of decreasing EDTA concentrations (10 mM – 100 µM EDTA, 2–3 column volumes per wash) to remove any residual calcium and finally equilibrated with 50 mM Tris-HCl, 100 µM EDTA buffer, pH 7.5 (5–10 column volumes) until the baseline was stable.

The flowrate was adjusted in order to keep within the maximum pressure for the column (5 bar, 70 psi). During purification the flowrate was set at 0.2 mL/min (manufacturer's recommended rate: 10–15 cm/h) and for cleaning and equilibration it was raised to 0.5 mL/min (GE Healthcare, 2006). Lower flowrates allow for higher resolution of the purified fractions. Wild type and mutant samples were loaded at 250 µL (within 0.5–3% of the total gel volume of 21 mL). Commercial aequorin (AquaLite[®]) was loaded at 15 µg (15 µL of 1 µg/µL stock solution) through a partially filled 200 µL loop (remaining volume was equilibration buffer).

Sanitisation, storage and cleaning-in-place (CIP) were performed according to manufacturer's instructions (GE Healthcare, 2006). Cleaning-in-place was performed in order to remove precipitated proteins and non-specifically bound proteins and lipoproteins. The column was washed at 0.4 mL/min (corresponding to 20 cm/h) and reversed flow direction with 0.5 M NaOH solution. This step was followed by a washing step with 70% ethanol at the same flowrate and flow direction. The column was finally washed with two bed volumes of 20% ethanol and stored at 4 °C. Before further use, the column was equilibrated with five bed volumes of buffer.

7.3.5.1 Protein size by mass spectrometry

Mass Spectrometry analysis was performed on a KRATOS AXIMA CFR V2.3.4 mass spectrometer by the scientific support services at the Wolfson Institute for Biomedical Research, UCL. KRATOS AXIMA is a matrix-assisted laser desorption - ionisation (MALDI) time-of-flight instrument which gives mono-isotopic resolution of 3,000 up to a mass of 6,000. It can analyse masses up to approximately 150,000 Da both in positive and negative mode, which is very useful for determination of intact proteins.

The analysis was performed on aequorin fractions with luminescent activity eluted from the size exclusion chromatography step in order to verify protein. The

purified aequorin samples (recombinant wild type, mutant Asn28Cys/Ser32His and AquaLite[®]) were desalted using spin filtration columns of 10,000 MWCO membrane filters (Section 7.2.1.2). Proteins equine apomyoglobin (16.9 kDa) and rabbit aldolase (39.21 kDa) were used as size standards in 5 μ L samples of 10 pmol/ μ L.

7.3.6 Protein handling considerations

Calcium is a nearly ubiquitous ion, which even in trace concentrations could trigger the bioluminescence reaction of aequorin. Due to aequorin's high sensitivity to calcium, unprotected aequorin must be handled with extreme care and kept from contact with ordinary glass and other substances capable of liberating calcium ions (Blinks *et al.*, 2000, Hastings *et al.*, 1969). Considering native aequorin's (low) sensitivity and mutants' possible increased sensitivity toward other (than calcium) metal ions, EDTA was the cation chelator of choice for all preparatory work; EGTA would be better suited for the chelation of calcium but EDTA chelates a wider range of divalent ions more effectively. Glass and metal surfaces were avoided where possible. The glass syringe pump and metal syringe nozzle of the platereader were rinsed with decreasing concentrations of EDTA in Tris-HCl buffer (first pass: 10 mM EDTA, final pass: 10 μ M EDTA). Plasticware were also rinsed with a 50 mM Tris-HCl, 10 mM EDTA solution prior to contact with apoaequorin or aequorin.

Aequorin's sensitivity to light and temperature is due to its prosthetic group, coelenterazine. Samples containing coelenterazine or active aequorin were kept on ice, centrifugation was performed at 4 °C and chromatographic separation was performed at 4–6 °C (for handling during chromatography refer to Section 7.3.5). Samples were protected from light where possible and tubes were covered with foil.

7.4 Results and Discussion

The purification protocol was based on Shimomura and Inouye (1999) due to the simplicity of the process and the use of mild operation conditions. Two options for active protein assays emerge from the published work: (1) to obtain aequorin in a crude bacterial suspension through minimal processing (2) to isolate cell paste and subject it to further purification steps. The first option was deemed acceptable for a crude, initial HTS method amenable to automation (Chapter 4 and Chapter 6) but it contains impurities and requires high concentrations of a metal chelator (e.g. 10 mM EDTA or EGTA) to ensure aequorin is not discharged by calcium or other cations present in the mixture. The second option yields highly purified protein in a buffer of choice with minimal chelator concentrations required, all necessary conditions for detailed kinetics analysis. In the HTS protocol (Chapter 4 and Chapter 6), the purity of aequorin was 78% in a solution containing 20 mM Tris-HCl, 10 mM EDTA, 5 mM 2-mercaptoethanol, pH 7.7. In this preparation the desired protein purity would be at least 95% in a 50 mM Tris-HCl, 100 μ M EDTA buffer, pH 7.5. EDTA concentration would be further decreased by dilution for the kinetics experiments of Chapter 8.

7.4.1 Bacterial cell culture

7.4.1.1 OD and activity versus incubation time

In this section the optimum incubation time before harvesting of cells was determined. The activity of the produced aequorin is plotted in association with optical density (OD) of the cell culture versus incubation time.

500 mL cultures were grown in 2-L shake flasks as described in Section 7.3.1.2, with the exception of conducting a 10% inoculation from the 100 mL culture into the 500 mL culture. The culture was induced at 2 hours (cell culture OD \approx 1.0) with IPTG at a final concentration of 0.6 mM. One milliliter samples of the culture were collected at various time points for OD (immediate measurement) and activity measurement (samples kept on ice for further processing and assayed in one microplate). Samples were transferred into a DSW plate containing buffered coelenterazine and incubated at 1 °C overnight with occasional agitation to fully equilibrate the protein into its holo

form. Luminescence activity of each sample was assayed in the BMG FluoStar platereader.

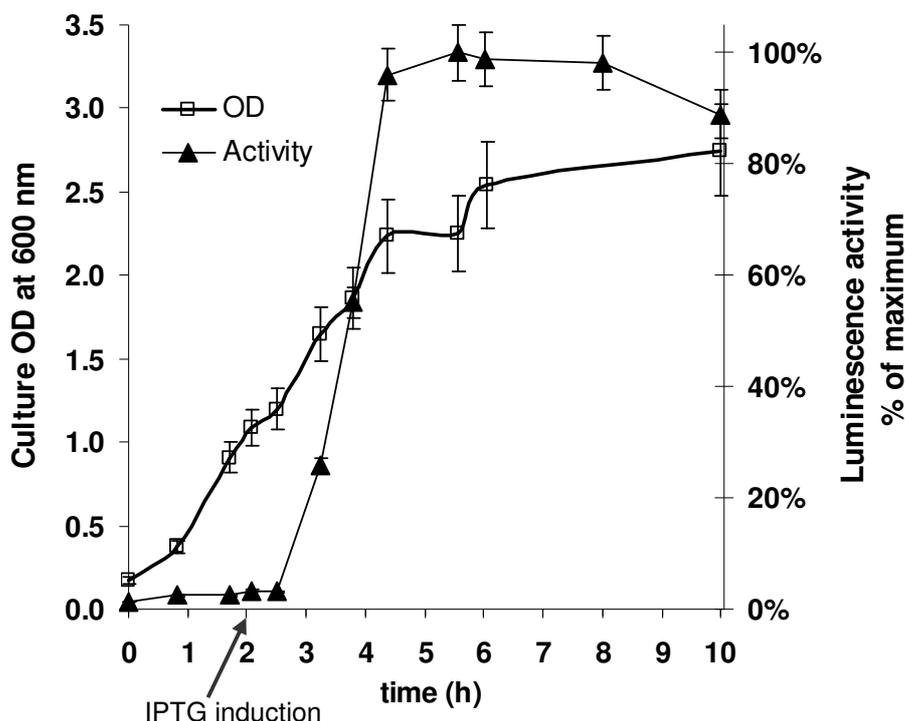


Figure 7-1 Luminescence activity and cell density of 500 mL cultures versus time. Two replicates per time point (N=2). Time zero corresponds to 10% inoculation from the overnight 100 mL culture. IPTG was added to a final concentration of 0.6 mM. Experimental setup and discussion in Section 7.3.1.2 and 7.4.1.1.

In Figure 7-1 it is shown that little apoaequorin (hence aequorin) is formed until after induction. The production of apoaequorin rapidly increases between 3–4 hours of incubation and approximately 30 min after induction. Then the total luminescence activity measured appears to plateau even though cell density continues to increase slightly from 4–10 h. This would normally suggest that it would have been possible to harvest at approximately four hours after induction, had it not been for the limiting amount of coelenterazine (4 μ M) added in this initial test (Section 7.3.2.1). Later measurements of total protein concentration and aequorin purity in shake flasks led to the usage of much higher coelenterazine concentration (92 μ M) during the cold incubation step. Figure 7-1 still provides dependable association between OD and activity until 5 h of incubation when coelenterazine became limiting. Harvesting at later time points was retained when using the higher coelenterazine concentration as Figure 7-1 showed no evidence that the aequorin was being significantly degraded.

7.4.2 Localisation of protein

7.4.2.1 Localisation of protein in the cell

In Section 4.4.2.2 it was demonstrated through activity studies that only a small quantity of apoaquorin is secreted in the culture medium during incubation. This section now examines the localisation and availability/solubility of apoaquorin in the cell pellet of the bacterial culture. The cells were retrieved by centrifugation and the pellet was re-solubilised with Tris-HCl buffer. The suspension was then sonicated to break open the cells and re-centrifuged to separate the soluble proteins in the supernatant from the insoluble proteins attached to the cell debris.

SDS-PAGE results from Figure 7-2 and Figure 7-3 clearly indicate that a substantial portion of apoaquorin produced in the cell is in the form of insoluble aggregates. These aggregates will be further discussed in Section 7.4.2.2.

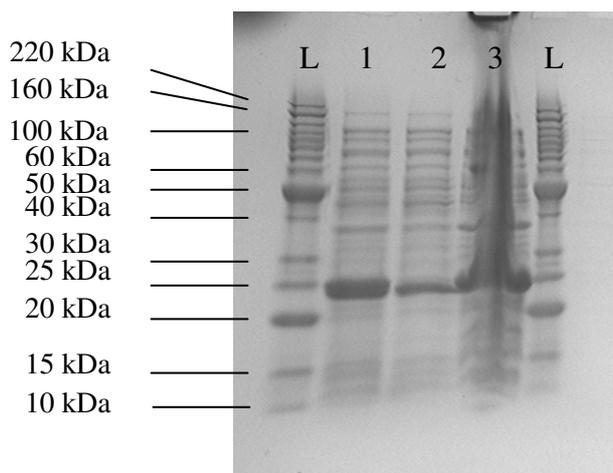


Figure 7-2 Localisation of apoaquorin in the cell pellet – (a). L: Benchmark Protein Ladder™ (Invitrogen, Figure 2-2); Lane 1: Bacterial pellet, isolated by centrifugation of cell culture; the cells were resuspended in volume of Tris-HCl buffer equal to the volume of the cell culture centrifuged. This lane represents total protein in the cell. Lane 2: The resuspended pellet represented in Lane 1 was sonicated and centrifuged. The supernatant of the sonicated mixture represents soluble protein, loaded in Lane 2. The precipitated fraction of the sonicated and centrifuged sample represents the insoluble protein in the cells, shown in Lane 3. Lanes 1, 2 and 3 were loaded with 10 μ L total volume, consisting of sample pre-mixed 1:1 with Biorad SDS loading buffer and heated to 95 $^{\circ}$ C for 5 min (Section 2.4.2). Lane L was loaded with 10 μ L of protein ladder.

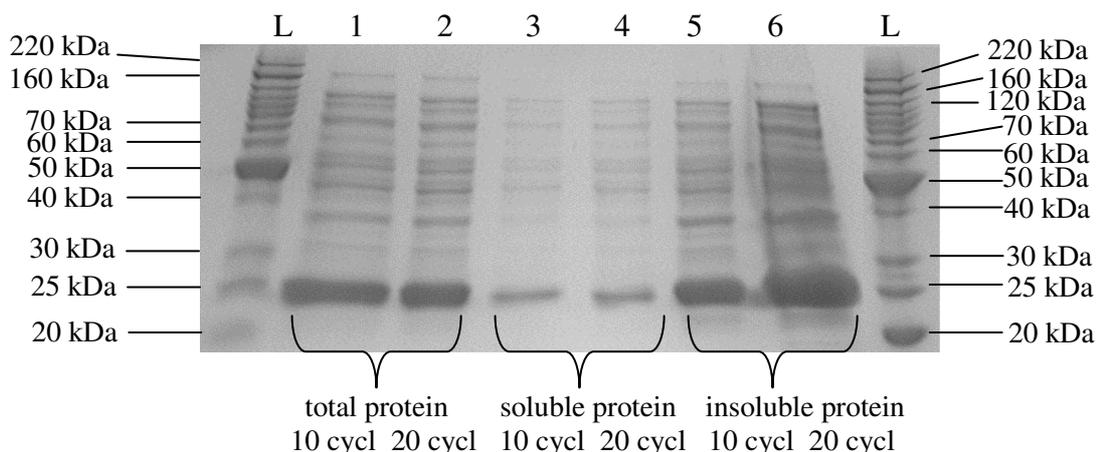


Figure 7-3 Localisation of apoaequorin in the cell pellet – (b). Corresponding to 10 and 20 cycles of sonication. Total protein (Lanes 1 and 2) corresponds to Lane 1 of Figure 7-2. Soluble protein (Lanes 3 and 4) corresponds to Lane 2 of Figure 7-2. Insoluble protein (Lanes 5 and 6) corresponds to Lane 3 of Figure 7-2. The samples were prepared as described in Figure 7-2. 10 μ L loaded per lane.

7.4.2.2 Localisation of protein after coelenterazine incubation

This section examines the localisation of apoaequorin post incubation with coelenterazine. Incubation with cold coelenterazine solution was performed as described in Section 7.3.2.1. The suspension (cells in coelenterazine buffer) contains two proteins of sizes between 20 kDa and 25 kDa, visualised as two clearly distinguishable bands in Figure 7-4. The smaller size protein is secreted in the coelenterazine buffer during the cold incubation (Lane 5). The bigger band is retained in the cell pellet (Lane 4). The smaller size band is soluble (Lane 2) as it appears in the soluble supernatant after sonication and centrifugation of the suspension.

The two bands between 20 kDa and 25 kDa both contain apoaequorin. This is verified in Section 7.4.2.3 where it is proven that bands of this size and intensity are only encoded in pETAQ and not the bacterial strain or pET26b. Taking into account that the size difference between the two bands is \sim 1 kDa (size of *pelB* leader peptide), that the bigger size band remains as insoluble within the cell and that the lower size band is secreted in the medium, it is reasonable to assume that the lower size band is apoaequorin that can be reconstituted into aequorin and the bigger band is apoaequorin still attached to the *pelB* leader peptide. Overall, most of expressed recombinant protein remained within the cells. A reasonable explanation is that the *E. coli*[®] machinery is not sufficient to cleave such large quantities of overexpressed protein.

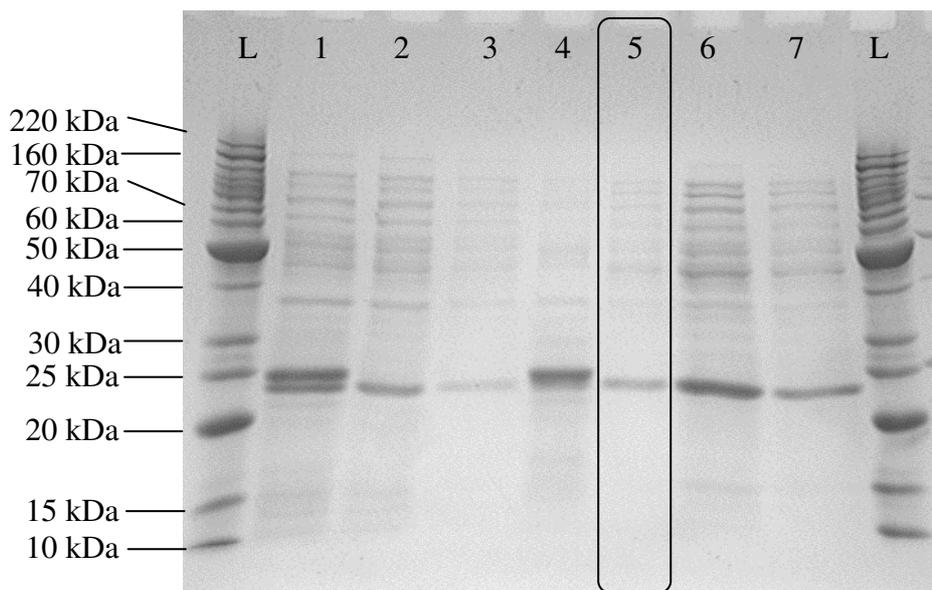


Figure 7-4 Localisation of apoaequorin after charging with coelenterazine. L: Benchmark protein ladder (Invitrogen, Figure 2-2). Lane 1: O/N suspension after cold incubation with coelenterazine (contains total protein). Lane 2: supernatant after centrifugation of sonicated O/N suspension (contains total soluble protein). Lane 3: Sample of lane 2, diluted 3-fold. Lane 4: Resuspended pellet after centrifugation of sonicated O/N suspension (contains insoluble protein attached to cells) Lane 5: Supernatant after centrifugation of O/N suspension. Contains soluble protein secreted by cells during the O/N incubation with coelenterazine solution. Apoaequorin band purity: 72%. Lane 6: Sample of lane 5 concentrated by spin centrifugation using Vivaspin6 columns (13,000 rpm, 30 min). Diluted 1:5 with Tris buffer. Lane 7: Sample of lane 5 concentrated by Amicon centriplus centrifugal filter (Milipore) (13,000 rpm, 30 min). Diluted 1:5 with Tris buffer. All lanes were loaded at 10 μ L total volume of sample (including Laemli loading buffer) or protein ladder.

The total protein concentration by Bradford assay of the protein secreted from 50 mL cell paste into 17.5 mL of coelenterazine solution was 0.275 mg/mL. This corresponds to a yield of 0.096 mg/mL of soluble apoprotein retrievable from the cell culture. Taking into account SDS-PAGE protein gel densitometry, this mixture contains 72% apoaequorin chain which if folded and reconstituted correctly can yield a maximum of 0.069 mg/mL functional aequorin or 6.9 mg per 100 mL of cell culture. The SDS-PAGE of Figure 7-4 was run under reducing conditions which would dissolve apoaequorin or aequorin aggregates and display only the linearised polypeptide bands. Size exclusion HPLC or a native SDS-PAGE assay would reveal the presence of potentially higher order aequorin/apoaequorin structures. The yield calculations in this section assume no loss of protein due to formation of aggregates post secretion from the periplasm.

7.4.2.3 Periplasmic fraction

The procedure followed is described in Section 7.3.2.2. Periplasmic extraction from cell pellets producing apoaequorin was performed in order to: (1) gain more insight into the localisation of recombinant apoaequorin the cell culture and (2) explore this method's potential as a purification step. The protocol was performed on cultures of *E. cloni*[®] cells carrying vectors pETAQ and pET26b. The latter was used as control for *E. coli* host proteins and proteins encoded in the pET26b vector without apoaequorin expression.

Comparing the expression profile from cultures carrying each vector, it is clear that the high intensity bands between 20 kDa and 25 kDa are indeed related to apoaequorin expression and not host cell proteins (Figure 7-5 and Figure 7-6). The periplasmic extraction process leaves a considerable amount of apoaequorin within the cells. Most of the apoaequorin produced does not reach the periplasm and is retained in the cytoplasm (Lane 6 of Figure 7-5). This conclusion agrees well with the findings of Section 7.4.2.2.

Periplasmic extraction can be used as an intermediate stage in the purification of apoaequorin from an *E. coli* culture. Further dialysis can lead to apoaequorin of high purity (Lanes 7,8 of Figure 7-5) although purity and yield were not estimated. The periplasmic extraction method by Ausubel (1980) using sucrose and magnesium sulphate was not used going forward in this work. The final purification protocol used the simpler one-step extraction and reconstitution with coelenterazine by Shimomura and Inouye (1999).

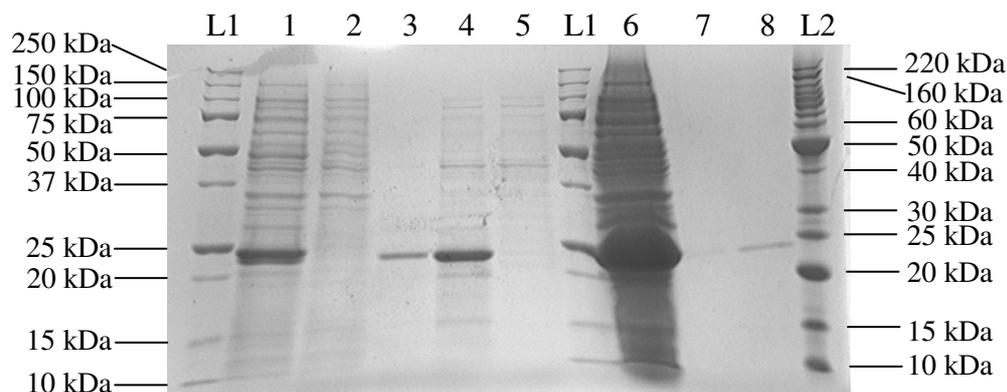


Figure 7-5 Periplasmic extraction SDS gel – (a) of cultures carrying pETAQ and pET26b. L1: Protein ladder: PrecisionPlus™ Protein Standards (Invitrogen) Lane 1: *E. coli* cells expressing pETAQ (referred to as wild type in the text), resuspended in sucrose solution. Lane 2: *E. coli* cells expressing pET26b, resuspended in sucrose solution. Lane 3: Periplasmic fraction of Lane 1 sample. Lane 4: Spin-concentrated sample of Lane 3. Lane 5: Periplasmic fraction of Lane 2 sample, spin-concentrated. Lane 6: pellet of *E. coli* cells expressing pETAQ, post periplasmic extraction. Lane 7: Dialysed sample from Lane 4. Lane 8: Spin-concentrated sample of Lane 7. L2: Protein ladder: Benchmark™ Protein ladder (Bio-Rad Laboratories). All lanes were loaded at 10 μ L total volume of sample (including Laemli loading buffer) or protein ladder.

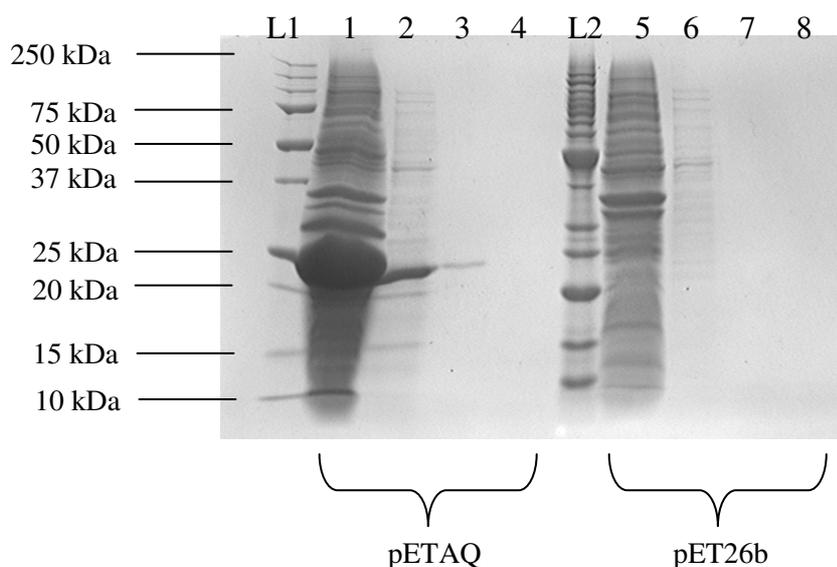


Figure 7-6 Periplasmic extraction SDS gel – (b), of cultures carrying pETAQ and pET26b. L1: Protein ladder: PrecisionPlus™ Protein Standards (Invitrogen) Lane 1: pellet of *E. coli* cells expressing pETAQ (referred to as wild type in the text) after the periplasmic extraction procedure. Same as Lane 6 in Figure 7-5. Lane 2: Periplasmic fraction of Lane 1, spin-concentrated. Same as Lane 4 in Figure 7-5. Lane 3: wash 1 of periplasmic extraction from *E. coli* cells expressing pETAQ. Lane 4: elution step of periplasmic extraction from *E. coli* cells expressing pETAQ. L2: Protein ladder: Benchmark™ Protein ladder (Bio-rad Laboratories). Lane 5: pellet of *E. coli* cells expressing pET26b (not containing the apoaequorin gene) post periplasmic extraction. Lane 6: Spin-concentrated periplasmic extraction from *E. coli* cells expressing pET26b. Same as Lane 5 of Figure 7-5. Lane 7: wash 1 of periplasmic extraction from *E. coli* cells expressing pET26b. Lane 8: elution step of periplasmic extraction from *E. coli* cells expressing pET26b. All lanes were loaded at 10 μ L total volume.

7.4.3 Clarification of protein sample

Centrifugation and filtration were tested as means of clarification of protein sample from particulate matter (such as cell debris and lipids) that may clog chromatography columns. This section compares clarification of the aequorin-coelenterazine suspension by (1) two consecutive centrifugation steps or (2) filtration through a low protein-binding 0.2 μM polyvinylidene fluoride (PVDF) filter. The activity of the protein sample was measured pre- and post-clarification (Figure 7-7). Using filtration, activity was reduced by approximately 78%; this may be due to aequorin being retained within the filter or activity lost due to cations leaching from the filter membrane and discharging aequorin. Bradford and SDS-PAGE assays would be appropriate analytics to further investigate this, but as the aim was to maximise activity of the retrieved sample, the filtration step was not used.

For all further protein sample clarifications, two consecutive runs of centrifugation at a maximum speed of 13,000 rpm at 4 °C were used, where the supernatant was retrieved at both steps (Figure 7-15).

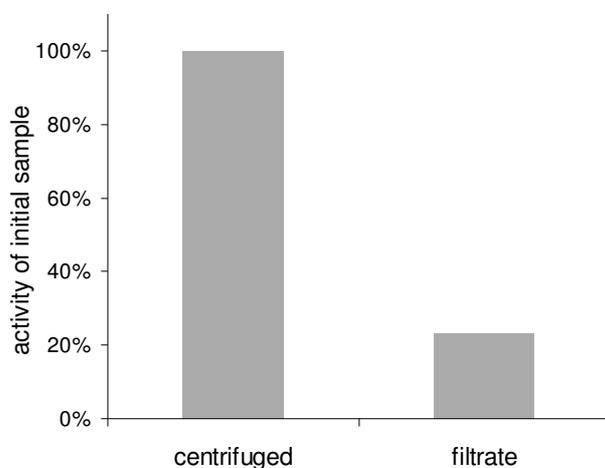


Figure 7-7 Clarification of aequorin suspension after coelenterazine incubation.

7.4.4 Enrichment of aequorin prior to chromatography

7.4.4.1 Spin-filtration

Limited sample volume can be loaded onto a size exclusion column, in this case approximately 250 μL . Prior to size exclusion chromatography, aequorin was concentrated using 30,000 MWCO spin columns (Section 7.3.3.1). An additional spin-filtration step was considered in order to test the hypothesis that aequorin could be separated from the impurities to >95% purity based on size, by employing spin columns of 30,000 and 100,000 MWCO. The retentate of the 30,000 MWCO centrifugation was filtered through a 100,000 MWCO spin column. Aequorin and a wide range of impurities were found both in the retentate, as seen in Figure 7-8. Aequorin was not separated from higher molecular weight impurities successfully using different molecular weight cut-off filters; the MWCO values are nominal, with pore sizes of wide distribution and the outcome depends on shape of the proteins, speed, duration and temperature of centrifugation. Only the 30,000 MWCO filtration/concentration step was incorporated in the final purification protocol (Section 7.4.2.2).

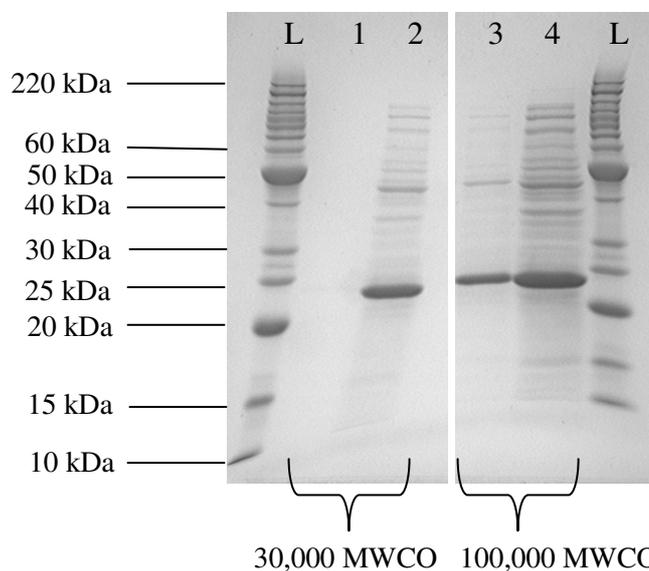


Figure 7-8 Aequorin suspension concentrated using Minisart[®] filters. The suspension contained reconstituted aequorin following incubation with coelenterazine. L: Benchmark[™] protein ladder (10 μL). Lane 1: Flowthrough of 30,000 MWCO. Lane 2: retentate of 30,000 MWCO (aequorin purity: 78.2%). Lane 3: Sample of Lane 2 filtered through 100,000 MWCO. This lane is the flowthrough (aequorin purity: 88.6%, host cell protein ~50 kDa: 7.7%). Lane 4: retentate of 100,000 MWCO (aequorin purity: 52%). Spin-centrifugation was for 2 \times 15 min, at 4,000 \times g, volume reduction factor was 17. Protein standards were loaded at 10 μL per lane and protein samples were loaded at 15 μL per lane (including 1:1 mixing with SDS loading dye).

7.4.4.2 Mass spectrometry of spin-concentrated sample

The concentrated protein sample (Section 7.4.4.1) was analysed using mass spectrometry (Section 7.3.5.1) to verify protein mass. The sample of Figure 7-9 corresponds to Lane 3 of Figure 7-8. The peak at 23,563 Da is aequorin monomer (expected value: 23,535 Da) and the peak at 47,263 Da is possibly a host cell protein, also appearing in Lane 3 of Figure 7-8. According to densitometry results these two bands account for $\geq 95\%$ of the protein content of the sample. The actual sample put through size exclusion chromatography was prepared in the same way as the sample in Lane 2 of Figure 7-8; this features aequorin of lower purity by 10% but was chosen in order to keep the processing steps and hence protein loss to a minimum.

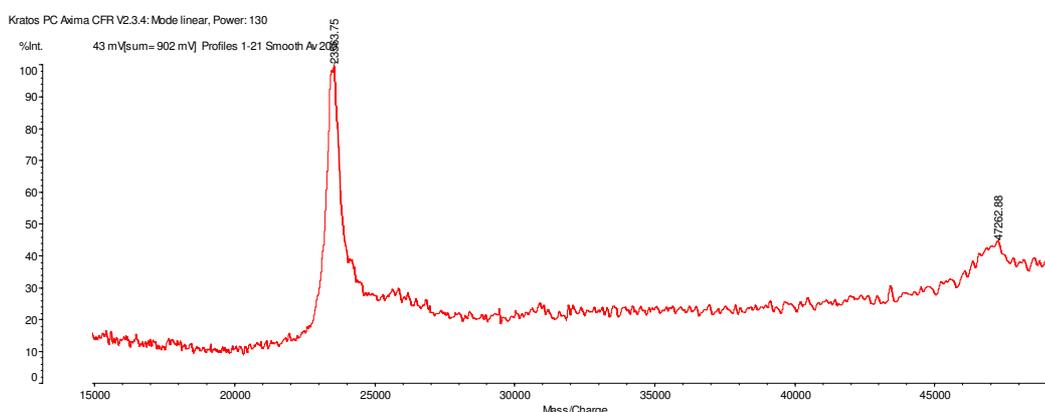


Figure 7-9 Mass spectrometry of aequorin sample prior to chromatographic purification.

7.4.5 Storage of purified fractions

Protein fractions were kept on ice throughout the duration of the wet lab work. After the active fractions were identified through activity assays, these were subjected to flash-freezing using acetone and dry ice and stored at $-80\text{ }^{\circ}\text{C}$.

7.4.6 Size exclusion chromatography of aequorin: wild-type, mutant and commercial aequorin

Size exclusion purification was performed as described in Section 7.3.5. Figure 7-10 features the chromatograms of absorbance at 280 nm matched with activity assays of the collected fractions (Section 7.3.4) for recombinant wild-type aequorin, mutant

Asn28Cys/Ser32His and commercial aequorin (AquaLite®). The UV traces of recombinant wild type and mutant Asn28Cys/Ser32His display three peaks: Peak 1, Peak 2 and Peak 3 eluting at 8, 10 and 12 mL respectively. These peaks are not well resolved. All chromatograms and raw data are appended in Supplementary Material.

Only one peak (Peak 1) at 12 mL is observed in the chromatogram of pure AquaLite®. Peak 1 exhibits the highest activity within each chromatogram and is expected to be the highly active monomer of aequorin. In a normal plot (Figure 7-10) both Peak 2 and Peak 3 appear completely inactive but on a logarithmic scale (Supplementary Material) it can be seen that there is some little activity in both. SDS-PAGE analysis of Peak 2 in both wild type and mutant reveals multiple high molecular weight bands (Supplementary Material) and displays the least luminescence activity. This minimal activity may be attributed to the activity of the preceding unresolved Peak 2 and succeeding Peak 1.

Interestingly, Peak 3 is proportionally more intense in the plot of mutant aequorin and it appears to have increased at the cost of Peak 1, compared to the UV trace of the wild type. It is possible that this particular mutant was expressed at very low levels (not analysed in this work). Mutant Asn28Cys/Ser32His has a solvent-exposed cysteine in EF-I loop which may encourage the formation of higher order aggregates of little or no activity, though disulfide bridges. It is possible that such aggregates of mutant aequorin may have formed after secretion from the cells (due to insufficient reducing agent in the solutions), thus decreasing the Asn28Cys/Ser32His monomer available for purification. Formation of aequorin/apoaequorin aggregates would explain the decrease of Peak 1 and increase of Peak 3 in the mutant plot and partially explain the low luminescence activity of Peak 3. However, SDS-PAGE analysis of the Peak 3 showed no discernible band at the size of aequorin. The difference in UV intensity of Peaks 2 and 3 between wild type and mutant could be attributed to variations in protein expression levels in the *E. coli* cultures.

Due to the poor resolution of the peaks, only partial fractions of Peak 1 of wild type and mutant were used for the kinetics experiments of Chapter 8, in order to avoid contamination of the descending Peak 2 and maintain high sample purity. The samples used in kinetics work (Chapter 8) are analysed in Sections 7.4.7 and 7.4.8.

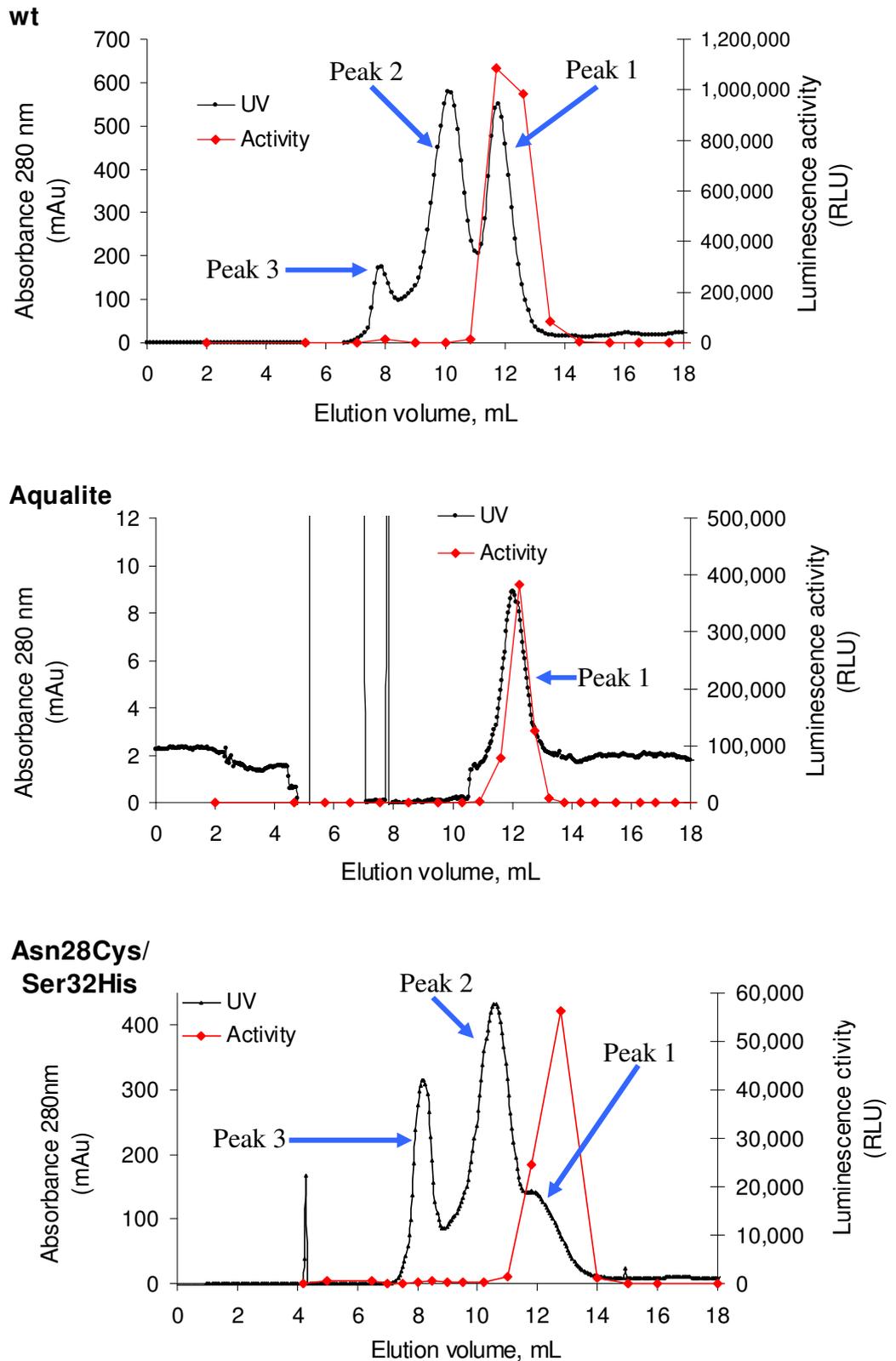


Figure 7-10 Elution profiles from size exclusion chromatography matched with activity of the eluted fractions. The midpoints of fraction volumes were used as x-axis for the activity plots. Luminescence activity RLU (relative luminescence units) are not comparable across the three plots as they were performed at different gain settings (1,100 900 and 2,000 from top to bottom) to avoid out of range readings.

7.4.7 SDS-PAGE of purified samples

Figure 7-11 presents SDS-PAGE analysis of purified wild type, mutant and commercial aequorin from size exclusion chromatography, to be used in the work of Chapter 8.

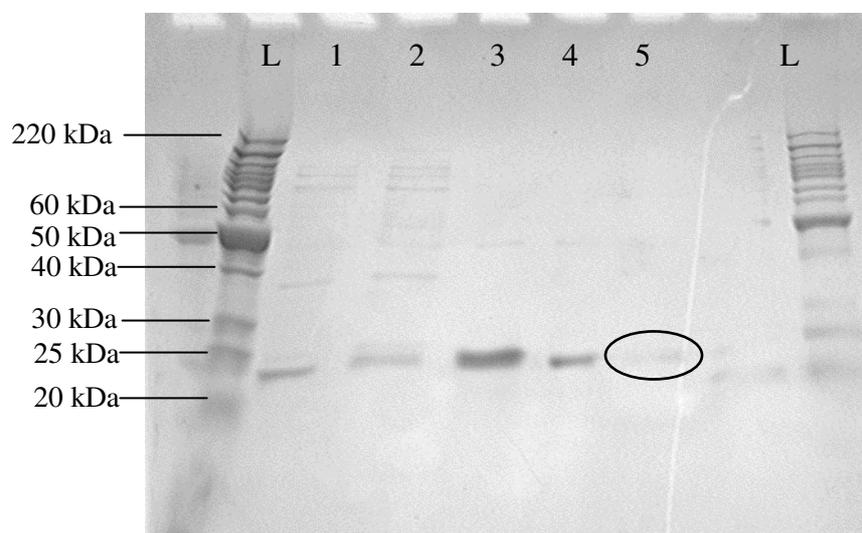


Figure 7-11 SDS-PAGE visualisation of purified aequorin samples. Lane L: Standard Benchmark™ protein Ladder (Invitrogen). Lane 1: supernatant of the aequorin suspension from high-throughput screening wells, concentrated 35-fold (aequorin purity: 78%). Lane 2: Total soluble protein of the aequorin suspension from high-throughput screening wells. The suspension was sonicated, centrifuged and the supernatant was concentrated 26-fold (aequorin purity: 66%). Lane 3: purified wild-type aequorin from size exclusion chromatography, concentrated 3-fold. Aequorin purity: 98%. Lane 4: purified mutant Asn28Cys/Ser32His, concentrated 3.5-fold. Aequorin purity: 97%. Lane 5: AquaLite® (commercially available aequorin from Molecular Probes). Aequorin purity: 100%. All samples (except Lane 5) were concentrated by spin centrifugation using a 10,000 MWCO spin column (Sartorius). Total loading volume per lane (including Laemli loading buffer): Ladder: 10 µL, Lanes 1-3: 15 µL, Lane 5 (AquaLite®): 10 µL.

Gel image densitometry was used to obtain a rough estimate of purity. Lanes 1 and 2 of Figure 7-11 show the aequorin obtained from the supernatant and total soluble protein from high-throughput screening wells, for which the purity of aequorin was 79% and 66% respectively. Lanes 3 and 4 show the aequorin obtained size exclusion purification for wild type and the Asn28Cys/Ser32His mutant, for which 98% and 97% aequorin purity was obtained respectively.

7.4.8 Mass spectrometry of purified samples

The molecular mass of aequorin was estimated through mass spectrometry at 22,927 Da (Figure 7-12) with potentially a second peak at 23,350 Da. Considering the estimation of the commercial standard (discussed below), these are both reasonably close to the estimate of 23,535 Da (including the 6×His tag), calculated theoretically using the Compute pI/Mw tool of the ExPASy Proteomics server (Gasteiger *et al.*, 2005). The coelenterazine (MW 423 Da) is presumably dissociated from the protein during the ionisation process, although the second smaller peak present at approximately 23,350 Da indicates a small residual population of the holo-aequorin.

The Asn28Cys/Ser32His mutant (Figure 7-13) gave a mass of 22,996 Da and a second peak of 23,392 Da. The two mutations were expected to increase the mass by 41 Da, whereas an increase of 69 Da was measured when compared to the experimentally determined wild type mass of 22,927 Da. This is an acceptable error in MALDI mass determination and this change at least qualitatively confirms the purification of the mutant. The second peak, with approximately 400 Da greater mass again indicates a population retained the coelenterazine during analysis.

The commercial AquaLite[®] sample gave a mass of 21,332 Da. The expected mass is 22,000 Da according to the supplier (Molecular Probes, 2001).

Given the above reduction of ~600 Da in mass determination for the commercial standard, the same apparent losses from the main peaks of wild type and mutant compared to their expected values can be attributed to assay-related calibration error (all three samples were estimated against the same calibration run). Any additional, non-assay related errors would be attributed to truncation at the C- or N-terminus. The mass difference between the first and second peaks in both wild type and mutant are consistent with retention of the chromophore during analysis. Such a peak may also be present in the trace of Figure 7-9, however the reduced resolution and peak broadening of that trace would potentially mask it.

From the mass spectrometry it is reasonable to infer that the wild type and mutant samples obtained are very pure as few alternative peaks were observed, although accurate quantitative measurements cannot be obtained from MALDI data.

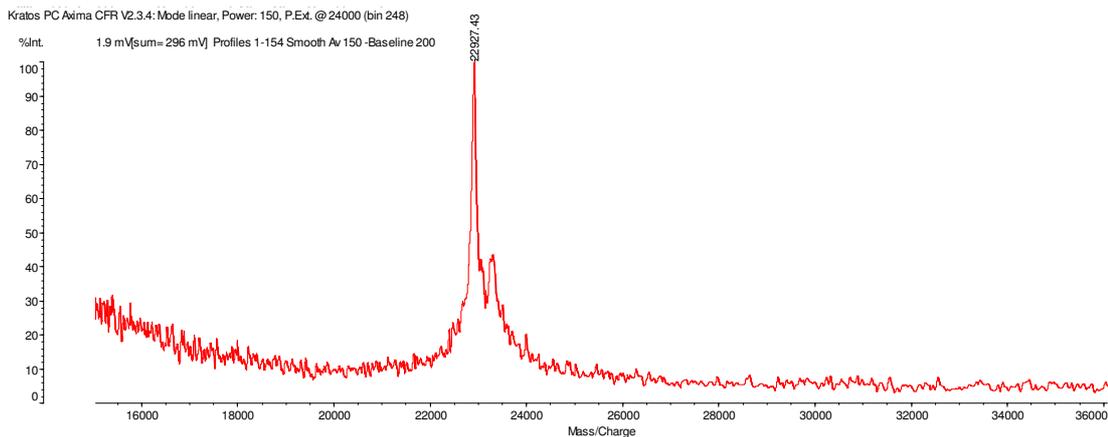


Figure 7-12 Mass spectrometry analysis of purified wild-type aequorin. Molecular weights are in Dalton. Protein concentration in the sample was approximately 157 $\mu\text{g/mL}$.

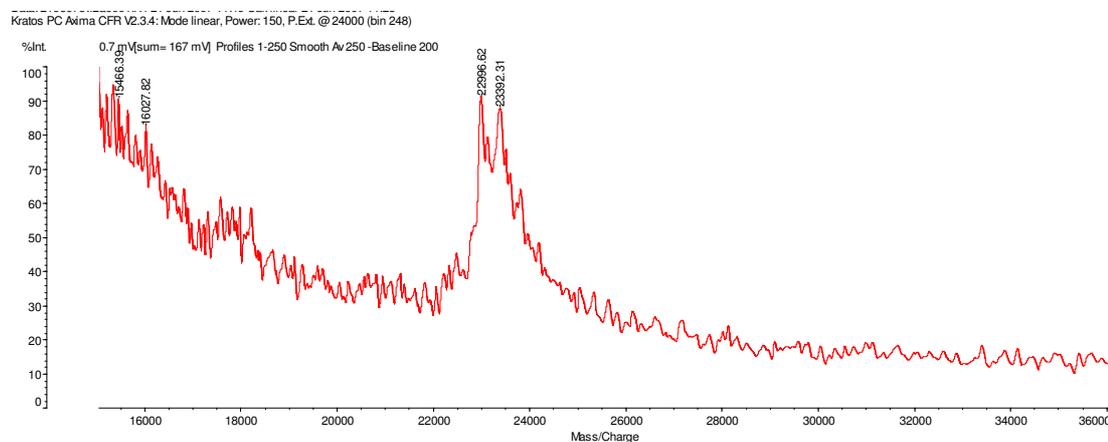


Figure 7-13 Mass spectrometry analysis of mutant Asn28Cys/Ser32His. Protein concentration in the sample was approximately 107 $\mu\text{g/mL}$.

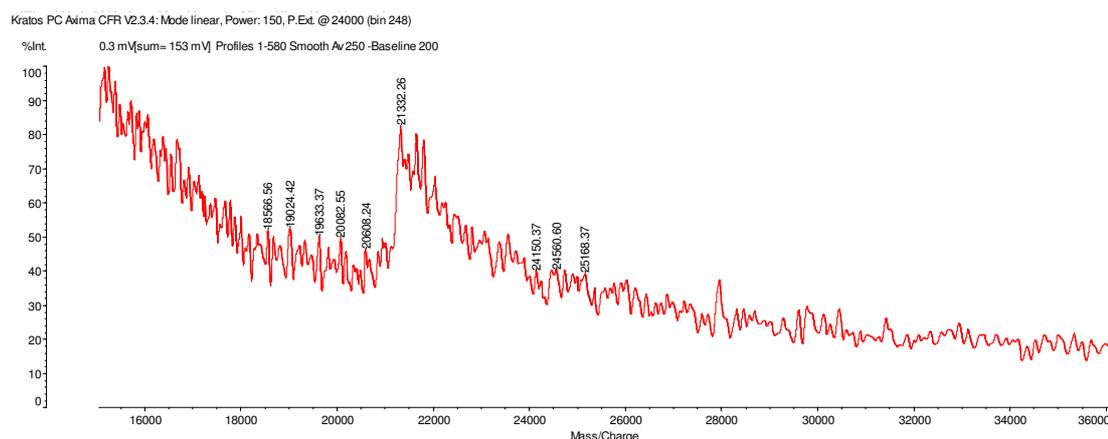


Figure 7-14 Mass spectrometry analysis of commercial aequorin AquaLite[®] (Molecular Probes) purified using the same size exclusion chromatography protocol as wild type (Figure 7-12) and mutant (Figure 7-13). Protein concentration in the sample was approximately 44.17 $\mu\text{g/mL}$.

7.5 Conclusions

A protocol for the purification of active aequorin and variants in bench scale was developed. The method yields target protein of high purity (Table 7-1) using mild extraction conditions and limited number of purification steps. The finalised procedure for the purification of recombinant aequorin and operating parameters are summarised in Figure 7-15.

This approach was sufficient to obtain small amounts of protein for the work performed in Chapter 8 and the use of size exclusion chromatography required minimal development time compared to separations based on ion exchange, hydrophobic or mixed-mode interactions. The histidine tag fused to apoaequorin that was originally envisaged for metal affinity purification (Chapter 3) was not used in this work, although an exploratory run with Ni-NTA chromatography was performed (not shown in this work).

A considerable amount (approximately half) of the apoaequorin produced in the bacterial culture was not utilised for two reasons: (1) significant portion of apoaequorin remained in the form of insoluble aggregates within the cytoplasm, possibly as the cell machinery failed to cleave the *pelB* leader peptide from the entire population of expressed protein (Section 7.4.2.2); to re-solubilise apoaequorin aggregates would require a different path of purification, use of denaturants and omission of the *pelB* leader from the DNA construct (2) only limited amount of protein can be processed through a size exclusion column at one time. Mass balance of aequorin size band between the coelenterazine extraction mix and the purified fractions show minimal to zero loss of protein (Supplementary material) during the size exclusion step, which was expected.

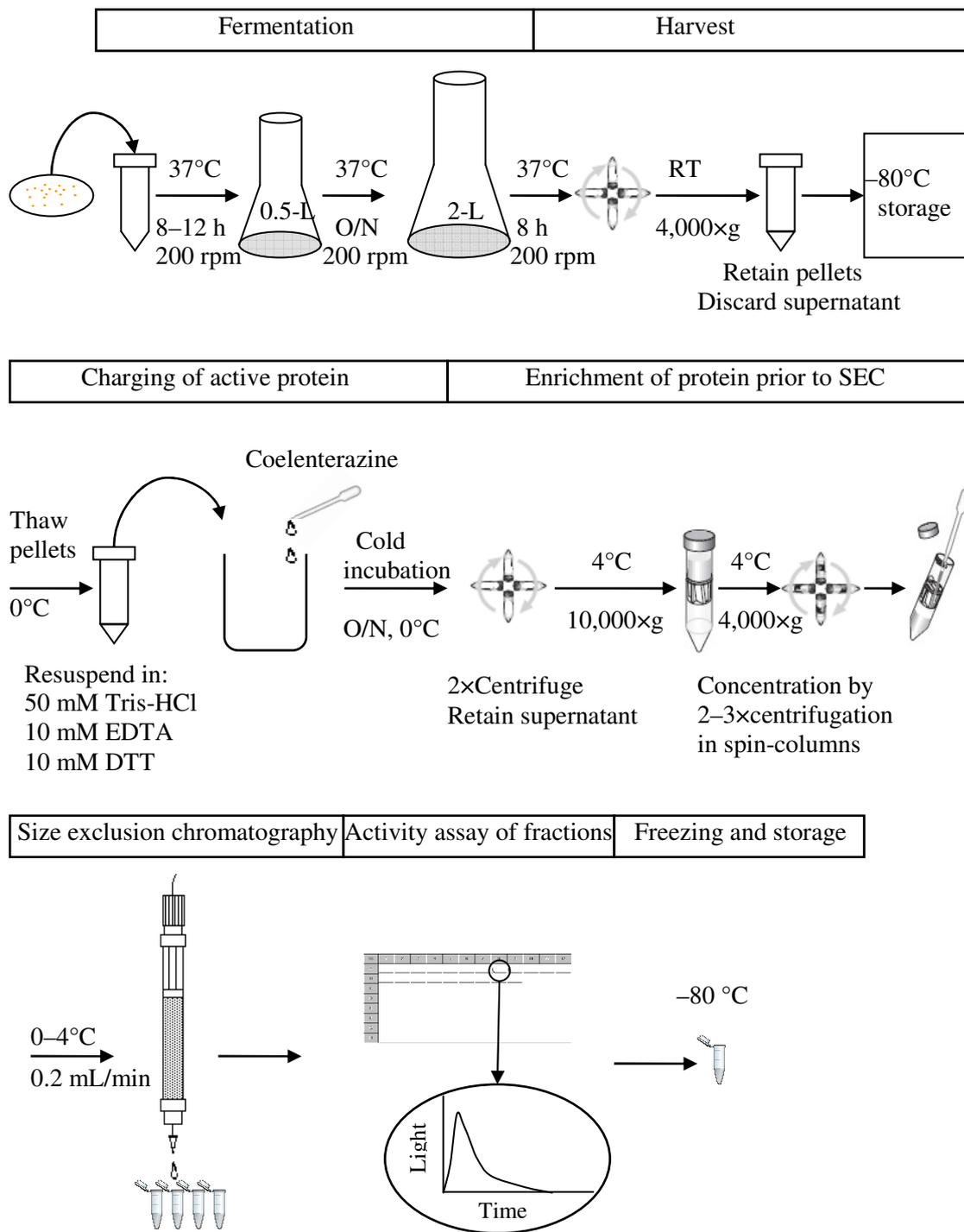


Figure 7-15 Final protocol for the purification of aequorin.

Table 7-1 Yield and purity of purified aequorin.

	Yield (mg per 100 mL culture)	Purity
Wild type	2.60	98%
Asn28Cys/Ser32His	1.04	97%

7.5.1 Improvements to the process

The resolution of size exclusion chromatography separation can be improved by increasing the length of the column, reducing the volume of sample loaded to stricter specifications of <1% of bed volume and reducing the flowrate to 0.1 mL/min from 0.2 mL/min used in this study. Superdex 75 prep grade has a fractionation range 3,000–70,000 Da but molecular markers would be tested first on the SEC column to ensure the desired size separation is achieved before processing the valuable aequorin sample.

In terms of analytics, GP-HPLC would be used for accurate estimation of protein purity instead of SDS-PAGE densitometry employed in this work. Both native and reducing SDS-PAGE would be employed for analysis of the purification fractions to identify aequorin aggregates. Silver staining method would be employed if higher detection sensitivity was needed.

If higher quantities of aequorin were required, available options would be to purify the apoaequorin from the periplasmic extract (without coelenterazine) using anion exchange chromatography as a capture step – aequorin has a low pI of 4.82 (Nanolight, 2010) – followed by hydrophobic chromatography if required. In such a case, reconstitution with coelenterazine would be performed on purified apoaequorin, and dialysis or buffer exchange step using appropriate chromatographic media (e.g. Superdex 30) would also be required.

The activity assays would all be performed at one common gain setting or more material would have to be sacrificed to perform assays at multiple gain settings in order to avoid out of range readings. This would allow specific activity values to be calculated and compared across all samples.

Published literature on aequorin reports activity in RLU's in various luminometers of each laboratory and these values are not directly comparable. Shimomura and co-workers (1990) calculated that 1 mg of pure aequorin emitted 4.80×10^{15} quanta at room temperature by calibration with the Cypridina bioluminescence reaction (Shimomura and Johnson, 1970). A more modern method for associating luminescence units to a standardised light source is the use of radioactive standards, e.g. Glowell™ luminometric standards by LUX Biotechnology which emit stable light and can readily bolt into the wells of a standard 96-well microplate. This would allow comparison of specific activity of aequorin samples across in-house experiments and published work alike.

8 Response of wild-type aequorin and mutant Asn28Cys/Ser32His to seven metal ions

In this chapter the purified wild-type aequorin and mutant Asn28Cys/Ser32His were tested for their response against seven candidate activator ions calcium, lanthanum, cadmium, lead, zinc, copper and cobalt. The work described in the previous chapters compiled literature review and theoretical structural study of aequorin (Chapters 1 and 5), designed an expression system for production of aequorin in high-throughput microplate format (Chapters 3 and 4) and created and screened a library of mutant aequorin variants against seven metal ions in high-throughput mode (Chapter 6); one variant was chosen for further investigation, to be purified (Chapter 7) and studied alongside wild-type aequorin (Chapter 8) without the interference of contaminants present in the crude screen of Chapter 4.

The purified wild-type and mutant aequorin were tested against each of the seven metal ion solutions at a wide range of metal concentrations. The light response against time was recorded.

Section 8.1 provides theoretical background on binding and kinetics useful for data interpretation. Section 8.2 compiles experimental methods and data processing specific to this chapter. Luminescence curves (light versus time) of aequorin wild type and mutant were plotted for each candidate activator ion in Section 8.3.1. Response curves (light versus metal ion concentration) that were compiled using the luminescence curve dataset were plotted in Section 8.3.2. The shapes of the luminescence curves are associated with the speed of reaction and fall under one of two types (sharp or glow-type). The shapes of the response curves indicate whether the ions serve as activators or inhibitors and allow an estimate of sensitivity of each protein to each ion.

In Section 8.3.3 a comparison is drawn between the seven ions as activators of wild-type and mutant aequorin. In Section 8.3.5 aequorin was injected with calcium and zinc (Double ion experiments) in order to test the premise that the mutated EF-hand can be triggered by the new ion (zinc) while the intact EF-hands are still triggered by calcium, thus achieving improved mutant aequorin response to zinc compared to wild type.

The kinetics of the sharp curves are discussed in Section 8.3.4.1 and the kinetics of glow curves in Section 8.3.4.2. The shapes of the luminescence, concentration-

response curves, yield and sensitivity are discussed in relation to ionic size and coordination of metal binding in Section 8.3.6.

8.1 Introduction

Work on aequorin to date has focused on the effect of calcium as its native activator. Rare earth metals (e.g. lanthanum, strontium) were found to act as calcium substitutes in triggering aequorin luminescence, however these ions are less biologically relevant. The novelty of the work described in this chapter lies in testing unusual candidates as triggers for aequorin luminescence: cadmium, zinc, copper, cobalt and lead, as well as the known activators calcium and lanthanum.

Asn28Cys/Ser32His was one of the interesting mutants emerging from the high-throughput screening of Chapter 6, chosen for its potential to exhibit altered selectivity towards new metal ions. It contains two mutations in the calcium-binding loop of EF-hand I which replace two conserved ligands for calcium ions with ligands usually found in zinc and copper binding sites; its other EF-hands remain intact. It is envisaged to have decreased affinity for calcium and increased affinity for new metals compared to wild type. A mutant with such features has not been designed and studied before, neither has it emerged from screening of randomised aequorin mutant libraries, as its low luminescence yield would have excluded it from a selection process based on high yields alone.

In the crude screen of Chapter 4, only one concentration of each metal ion was tested on each of the variants of the aequorin library. The concentration for each ion was decided based on optimal activity yielded from wild-type aequorin in cell suspension. Work in this chapter will show that the same metal ion concentration that is optimal for wild type can in fact be inhibitory for a mutant variant. This finding was taken into account in the experimental design of this Chapter.

8.1.1 Binding studies

Sections 8.1.1 and 8.1.4 discuss the differences between binding and concentration-response curves, as these concepts will be useful in data analysis. To study binding reactions quantitatively, the complex must be monitored separately from

the free biomolecules. This can be done via measuring changes in the solution in which the reaction occurs (e.g. temperature, absorbance and light emission) or by employing a secondary solution assay to probe changes in one of the biomolecules (Goodrich and Kugel, 2006). A number of fluorescence techniques (fluorescence resonance energy transfer, fluorescence anisotropy, induced fluorescence and fluorescence quenching) allow binding to be studied by monitoring changes in fluorescence due to association between a protein and a ligand. Such assays can monitor changes in a natural fluorophore (e.g. a tryptophan) present in a molecule or a label by fluorescent dye. Another way is to separate complexes from the free biomolecules based on differences in size, charge or conformation (Goodrich and Kugel, 2006).

8.1.2 Affinity and apparent affinity

Affinity describes the strength of an interaction between two molecules. To understand affinity constants (dissociation constant K_d and binding constant K_b) it is helpful to consider the rate constants governing the association and dissociation of the two molecules (Goodrich and Kugel, 2006). In a bimolecular reaction:



The dissociation constant K_d is equal to the reverse rate constant divided by the forward rate constant (Equation 8.2) and is expressed in units of M.

$$K_d = \frac{k_{-1}}{k_1} \quad \text{Equation 8.2}$$

Binding constant (K_b) is the reciprocal of K_d :

$$K_b = \frac{1}{K_d} \quad \text{Equation 8.3}$$

An interaction with a low K_d (high affinity) will likely have a large k_1 and small k_{-1} meaning that A and B will associate rapidly and the AB complex will dissociate slowly. K_d can be related to the concentrations of the three components that are present at equilibrium (A, B and AB):

$$K_d = \frac{[A] \cdot [B]}{[AB]} \quad \text{Equation 8.4}$$

where: $[A]$, $[B]$ and $[AB]$ are the molar concentrations of the reaction components at equilibrium.

At equilibrium, the population of A molecules will consist of free A and A bound in AB complexes:

$$[A_{Total}] = [A] + [AB] \quad \text{Equation 8.5}$$

$$[B_{Total}] = [B] + [AB] \quad \text{Equation 8.6}$$

The forward and reverse rate constants do not need to be known to measure the affinity constant; K_d can be measured directly (under equilibrium conditions) as follows.

Equation 8.2 is transformed into:

$$\frac{[AB]}{[A_{Total}]} = \frac{[B]}{K_d + [B]} \quad \text{Equation 8.7}$$

where $[AB]/[A_{Total}]$ is the fraction of A put into the reaction that is in the AB complex at equilibrium. If the right half of Equation 8.7 is divided by K_d it can be transformed into:

$$\frac{[AB]}{[A_{Total}]} = \frac{\frac{[B]}{K_d}}{\frac{[B]}{K_d} + 1} \quad \text{Equation 8.8}$$

When $[AB]/[A_{Total}]$ is plotted on the y-axis and $[B]$ is plotted on the x-axis, K_d is equal to the concentration of free B ($[B]$) at which the fraction bound ($[AB]/[A_{Total}]$) is half the maximum value. Equation 8.7 derives a hyperbola when x-axis is linear and one sigmoidal when a logarithmic axis is used (Figure 8-1). In both cases K_d is the concentration of B at which the fraction bound is 0.5, best expressed in Equation 8.8.

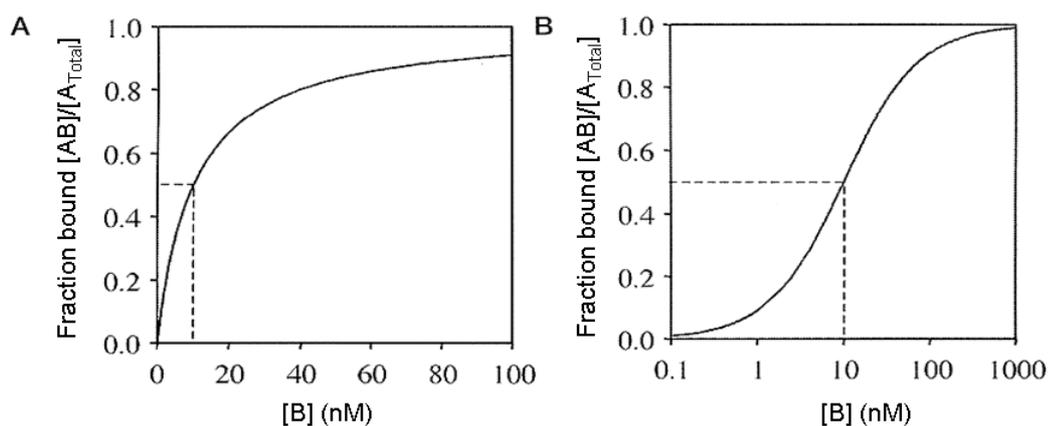


Figure 8-1 Bimolecular binding curves plotted on a linear x-axis in (A) and a logarithmic x-axis (B). The dissociation constant K_d is 10 nM in both plots. Figure reproduced from Goodrich and Kugel (2006). This is a non-cooperative system. In a cooperative system both the linear and logarithmic axis plots would be sigmoidal.

To experimentally measure the K_d using Equation 8.7 or Equation 8.8 a series of reactions are performed in which the concentration of A is kept constant and the concentration of B is varied, and the concentration of AB complex – and possibly B unbound – is determined.

In principle, any concentration of A can be used. In practise it is simplest to determine a K_d when the concentration of A in the reactions is significantly lower than the K_d for the interaction $[A_{Total}] \ll K_d$ (ideally 100-fold below the K_d). When $[A_{Total}] \ll K_d$, the amount of B in the AB complex is only a small fraction of total B, hence free [B] approximates $[B_{Total}]$ ($[B] \approx [B_{Total}]$). This simplifies performing experiments, as free [B] does not have to be measured at equilibrium and $[B_{Total}]$ (the total concentration of B in the reaction mix) can be plotted on the x-axis instead of free [B]. The K_d can then be determined using Equation 8.7. If $[A_{Total}]$ is not significantly lower than K_d , free [B] must be measured at equilibrium in each reaction as [B] does not approximate $[B]_{Total}$ under these conditions (Figure 8-2).

In summary, the simplest way to measure a K_d is to set $[A_{Total}] \ll K_d$, plot $[AB]/[A_{Total}]$ versus $[B_{Total}]$ and fit the data with Equation 8.7. If these conditions cannot be met, free [B] must be measured for each $[B_{Total}]$.

Experiments to measure affinity constants must be performed under equilibrium conditions, where there is no overall change in the concentrations of the free and bound species in the reaction over time. Another experimental consideration is that in most assays one biomolecule is labelled (fluorescently or radioactively) to facilitate detecting

that molecule and the AB complex. The K_d is affected by assay conditions such as pH, temperature, salt concentration and the presence of divalent ions.

If one of the biomolecules in a bimolecular reaction is an enzyme, it might be possible to use the enzymatic reaction to measure the binding affinity between the enzyme and a second biomolecule (e.g. a regulator). A K_d measured in this way is referred to as an apparent K_d because the assay used to make the measurement is indirect (i.e. it does not directly monitor binding).

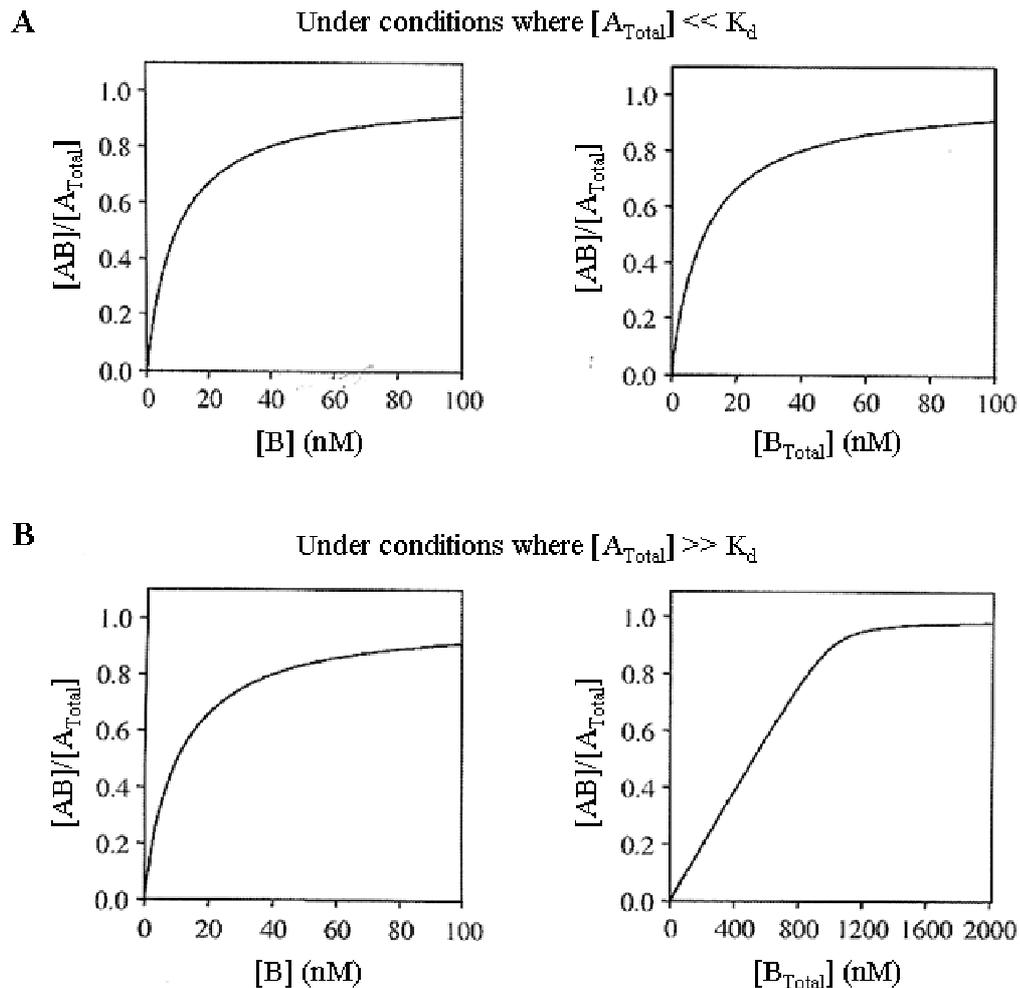


Figure 8-2 Bimolecular binding curves illustrating two experimental extremes: $[A_{Total}] \ll K_d$ and $[A_{Total}] \gg K_d$. In all plots, the K_d is 10 nM. $[A_{Total}]$ is 0.1 nM in (A) and 1000 nM in (B). Figure reproduced from Goodrich and Kugel (2006).

8.1.3 Allostery and cooperativity

In allostery a regulatory substance binds to a subunit of a mutli-subunit protein at a site other than the protein's active site. The binding alters the conformation and

functional properties of the molecule (Dorland, 2010). Cooperative binding is a special case of allostery that occurs in macromolecules with more than one binding sites. Binding to one site changes the binding affinity of the other subunits by induction of conformational changes at the other binding sites. Cooperative enzymes typically display a sigmoidal (S-shaped) plot of reaction rate versus substrate concentration” (IUPAC, 1997).

The Hill equation was formulated in 1910 to describe the sigmoidal O₂ binding curve of haemoglobin (Hill, 1910). It is a simple, empirical equation that predicts the fraction of the macromolecule saturated by ligand as a function of ligand concentration.

For the association between enzyme A and *x* identical molecules of ligand B the equilibrium between A and B can be described:



The equilibrium between A and B can be described:

$$\frac{[AB_x]}{[A_{total}]} = \frac{[A]^n}{K_d + [A]^n} \quad \text{Equation 8.10}$$

where: n=Hill coefficient and K_d is the dissociation constant.

In the case of positive cooperativity, the Hill coefficient is the slope of the ascending sigmoidal curve in a semi-log plot. The Hill equation provides useful means of assessing cooperativity in systems containing multiple B molecules binding to A, as long as its limitations are understood.

One common misuse of n is to estimate the number of B molecules bound to each molecule of A (Weiss, 1997). Another common misuse is to consider n=1 as indicator of completely independent binding (non-cooperative) and n<1 as indicator of negative cooperativity; these assumptions are not universally correct. The Hill equation can only provide clear evidence for positive cooperativity when n>1, however it cannot be used to determine negative or non-cooperative systems without also taking into account the relative affinities of each different site on A for binding B (Goodrich and Kugel, 2006). It cannot provide accurate measurements of the K_d for the interaction of B at any site of A or the exact number of B molecules that bind to A. Coefficient n is only a minimum estimate of the number of binding sites involved (Weiss, 1997).

8.1.4 Concentration-response curves

Concentration-response (or dose-response) curves are graphs that help characterise the effect of various compounds on enzymes/receptors. The x-axis plots concentration of a substance (e.g. a drug or hormone) and the y-axis plots response. Response could be almost any measurable biological function, such as enzyme activity, contraction of a muscle, change in heart rate etc (Motulsky and Christopoulos, 2003). A standard concentration-response curve (Figure 8-3) is defined by four parameters: the baseline response (Bottom), the maximum response (Top), the slope, and the concentration that provokes a response halfway between baseline and maximum (half maximal effective concentration EC_{50}).

Some of the standard shapes of many systems concentration-response curves are shown in Figure 8-3 to Figure 8-5. Stimulation is indicated when the curve goes uphill from low to high concentration of regulator (Figure 8-3). The substance may produce a full or only a partial response. Inhibition is indicated by a downward curve (Figure 8-4). Some substances may stimulate response at low concentrations and inhibit response at high concentrations, or vice-versa. This is expressed in bell-shaped concentration-response curves (Figure 8-5). Such a model is used qualitatively as a lot of data would be needed to determine all its parameters (Motulsky and Christopoulos, 2003).

A hyperbola is descriptive of many binding phenomena. Although this dependence appears sigmoidal in semi-log plots, real cooperative systems exhibit sigmoidal behaviour even in linear plots (Section 8.1.1, Figure 8-1) (Bisswanger, 2008). Since the linkage between binding and response can be very complex, any shape is possible for concentration-response curves. However, many concentration-response curves have shapes “*almost identical to hyperbolic/sigmoidal binding curves, even when multiple steps intervene between binding and measured response*”. The shape of concentration-response curves will often be similar to that of a single-step binding curve (Motulsky and Christopoulos, 2003).

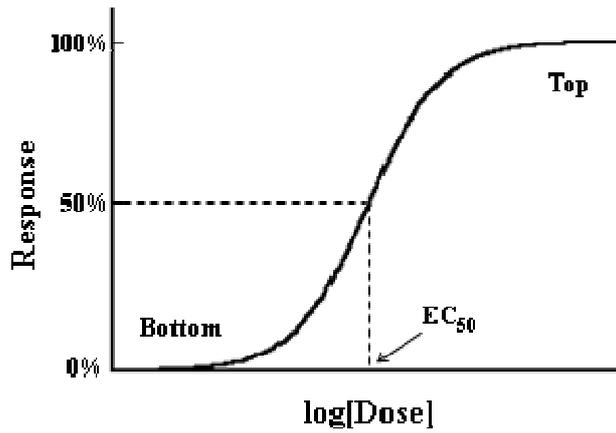


Figure 8-3 Sigmoidal concentration-response curve – stimulation. Figure reproduced from Motulski and Christopoulos (2003).

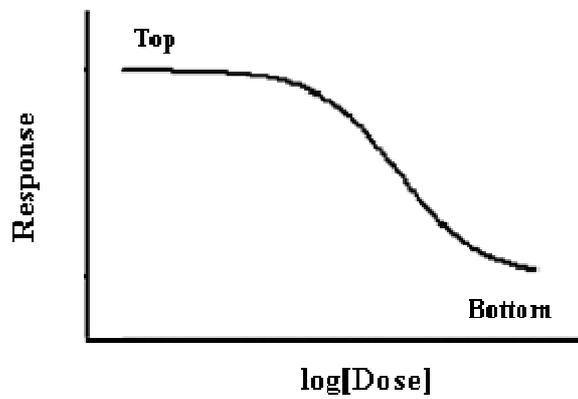


Figure 8-4 Concentration-response curve – inhibition. Figure reproduced from Motulski and Christopoulos (2003).

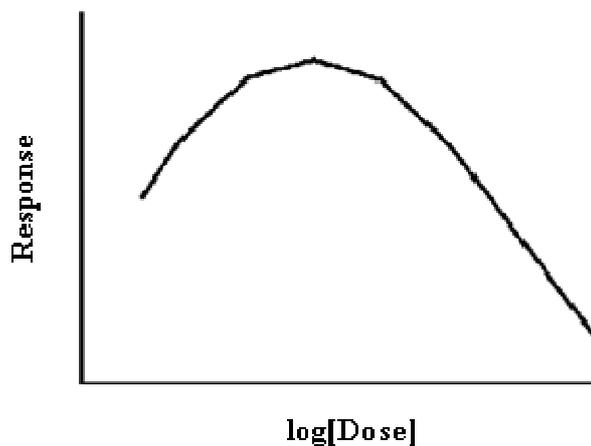


Figure 8-5 Bell-shaped concentration-response curve. Figure reproduced from Motulski and Christopoulos (2003).

The EC_{50} is determined by two properties of the regulator molecule: (1) its affinity (how tightly it binds to the enzyme) and (2) its efficacy (intensity of the response once bound). A single concentration-response experiment will encompass the effects of both affinity and efficacy. If a substance binds with high affinity and low efficacy it will produce the same concentration-response curve as one that binds with low affinity and high efficacy.

The EC_{50} is not a direct measure of drug affinity and usually not necessarily the same as the K_d for the binding of a substance to its receptor. The EC_{50} is merely the concentration of substance required to “*provoke a response halfway between the baseline and maximum responses*” (Motulsky and Christopoulos, 2003). In the cases where the concentration of an inhibitor is varied (concentration-response curve moves downhill), IC_{50} is the concentration that causes 50% inhibition (IUPAC, 2010).

8.1.5 The aequorin system

Enzymes are biomolecules that catalyse chemical reactions and are not consumed by the reactions they catalyse (Berg *et al.*, 2002). Aequorin is not a typical enzyme, as the catalysed substrate is bound within the protein chain (hydroperoxy-coelenterazine) and when the reaction is completed the spent substrate remains loosely bound to the protein chain (Section 1.3.1). One molecule of aequorin can only process one molecule of coelenterazine. The spent substrate (coelenteramide) can be separated from the protein chain by size exclusion chromatography and calcium (activator ion) bound in the EF-hands is removed by chelating agents such as EGTA or EDTA. In order to achieve regeneration of fully active aequorin fresh coelenterazine must be added (Section 1.3.8).

The activators for the bioluminescent reaction are calcium ions which act allosterically: in aequorin the activation points are the solvent accessible calcium binding EF-hands, whereas the reaction of coelenterazine oxidation occurs in the core of the protein (Section 1.3.1). Two to three calcium ions bind with positive cooperativity, i.e. the binding event of the first increases the affinity of aequorin for the consecutive calcium ions; it is also known that two calcium ions are required to complete the reaction and that even one can trigger the reaction, although with considerably lower rate (Section 1.3.7). Aequorin function does not conform to Michaelis-Menten kinetics

but produces a characteristic sigmoidal curve that declares positive cooperativity between the calcium binding sites (Section 1.3.1).

8.1.5.1 Fundamental work on the kinetics of aequorin

Hastings and co-workers (1969) performed some fundamental experiments to explore the mechanism of aequorin luminescence using rapid mixing techniques (double-stopped flow apparatus). The onset of light emission was found to be fairly rapid with a pseudo-first order rate constant of about 100 s^{-1} (half rise time of about 6 ms) and virtually independent of calcium concentration over the range examined. The decay of light intensity was exponential over at least one order of magnitude with rate constants between $1\text{--}1.2\text{ s}^{-1}$. The rate of decay was found to increase with increasing calcium concentrations from a very low value below 10^{-7} M to a maximum reached between 10^{-5} and 10^{-4} M . This rate represented the rate of aequorin utilisation. Both the rate of rise and rate of decay were found to be influenced by temperature, as shown in Table 8–1 (Hastings *et al.*, 1969). The ultimate photon yield was found to be constant over a wide range of calcium concentrations tested (Hastings *et al.*, 1969).

Table 8-1 Aequorin kinetic constants in published literature.

Kinetic constants of aequorin	at 5 °C	at 25 °C	at 50 °C
Rate constant for rise (s^{-1})	46^1	$100\text{--}300^2$	690^1
Rate constant for decay (s^{-1})	0.5^1	$1.0\text{--}1.2^2$	7^1

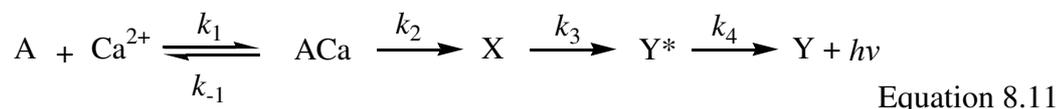
1: Hastings *et al.*, 1969

2: Loschen and Chance (1971) in Shimomura (2006)

The authors conducted experiments to determine whether the relatively long duration of light emission in a great excess of calcium should be attributed to the continued presence of free calcium or to the decay of some rapidly formed reaction intermediate that could yield light without calcium. These experiments used mixed solutions containing aequorin and calcium in the presence and absence of chelating agents. EDTA was added to bioluminescence reactions already in progress in order to deplete free calcium from the solution and calcium already bound to aequorin. The addition of excess of EDTA resulted in continuation of luminescence but with a very rapid exponential extinction, indicating that the light emission normally extending over a second or more requires the continued presence of free calcium (Hastings *et al.*, 1969).

The authors also concluded that aequorin forms an intermediate in the luminescence reaction that, once formed, irreversibly leads to the emission of light without the continued presence of free calcium. The lifetime of this intermediate was found to be $t_{1/2} \approx 7$ ms at 19 °C.

A hypothetical reaction scheme (stoichiometry unspecified) that would accommodate these results was constructed (Equation 8.11):



The first step of the proposed mechanism is the very rapid and reversible binding of calcium to aequorin. The authors proposed that – provided maximally effective calcium is present – the rate of the overall reaction is limited by the rate of the second step (k_2 approximately 1 s^{-1}) in which the intermediate X is formed. X gives rise to an excited singlet state Y^* (such states typically have lifetimes of nanoseconds) at a rate governed by k_3 , which is in the order of 100 s^{-1} . Calcium ions initiate the cascade of events, but $X \rightarrow Y^*$ does not require the presence of calcium (Hastings *et al.*, 1969).

In later publications it was proven that the excited singlet state Y^* is the excited hydroperoxy-coelenterazine (Shimomura and Johnson, 1972, Shimomura and Johnson, 1975a, Head *et al.*, 2000, Shimomura and Teranishi, 2000). If one did not assume the existence of the intermediate X, the rate of decay of light intensity found on quenching with EDTA might be attributed to the rate of dissociation of ACa (governed by k_{-1}). In this case, however, one would expect the rise-time of the luminescence reaction to be dependent on calcium concentration. Thus k_{-1} was assumed to be considerably greater than k_3 (Hastings *et al.*, 1969).

8.1.5.2 A model of double exponential decay

It was previously known that for wild type aequorin the decay rate of the flash-type luminescence curve (see Section 1.1.4 for distinction between flash- and glow-type luminescence) increases with calcium concentration, while the total light emitted (light integral) remains relatively constant (Section 8.1.5.1). In a recently published study Tricoire and co-workers (2006) focused only on the decay of the bioluminescent reaction. This was the first published proposal of a double exponential decay model of aequorin luminescence. The authors deemed that a single exponential component was not enough to describe aequorin luminescence decay because the semi-log plot of luminescence decay versus time was not linear.

The authors used random mutagenesis and functional screening to isolate mutants exhibiting flash-type luminescence with slow decay rate. They proposed that the decay of aequorin's flash-type luminescence reaction comprises of two exponential components: one fast and one slow. The authors proposed that variations of these two components dictate the peak of the luminescence curve. The decay traces of wild type and mutant aequorins were fitted with two exponentials according to Equation 8.12:

$$y = I_F \times e^{-t/\tau_F} + I_S \times e^{-t/\tau_S} \quad \text{Equation 8.12}$$

where τ_F and τ_S are the time constants of the fast and slow component respectively and I_F and I_S are the initial maximum intensities of the fast and slow component. The light integrals Σ_F and Σ_S of these components (where total light Σ_T is $\Sigma_T = \Sigma_F + \Sigma_S$) were also determined

$$\Sigma_S = I_S \times \tau_S \quad \text{Equation 8.13}$$

$$\Sigma_F = I_F \times \tau_F \quad \text{Equation 8.14}$$

Tricoire and co-workers proposed a model of aequorin calcium dependence taking into account the variations of decay rate with calcium concentration as a key determinant of light intensity. They suggested that the fast and slow components coexist across different aequorin variants and calcium concentrations. The authors also propose that the concept of parallel slow and fast decay agrees with the increase of light intensity and decay rate with increasing calcium concentrations, given that the light integral is constant at all calcium concentrations. The kinetics of slow decay mutants for a given EF-hand were found to be similar and deemed to result from the disruption of a link between an EF-hand domain and coelenterazine binding residues and not from the reduction of its calcium affinity (Tricoire *et al.*, 2006).

8.2 Materials and methods

8.2.1 Aequorin and coelenterazine

Purified fractions of wild-type aequorin and purified mutant Asn28Cys/Ser32His were prepared as described in Chapter 7. Commercially available recombinant aequorin AquaLite[®] was obtained in lyophilised form from Molecular Probes and purified with the same chromatographic methods used for wild-type and mutant Asn28Cys/Ser32His in Chapter 7.

8.2.2 Protein concentration

The concentration of purified aliquots of wild-type and mutant aequorin was determined by absorbance at 280 nm on the NanoDrop[®] ND-1000 (NanoDrop Technologies, Rockland, US).

Shimomura and Shimomura (1981) determined that in order to maintain constant specific activity aequorin must be used at concentrations <10 µg/mL. Appropriate serial dilutions of the purified aequorin fractions were made in 50 mM Tris-HCl 10 µM EDTA pH 7.5 for a final concentration of 5.1 µg/mL (approximately 0.21 µM). After 1:1 mixing with metal ion solutions the final concentration of aequorin was 2.55 µg/mL (approximately 0.11 µM). This concentration of protein in the final mixture is sufficiently low compared to previously reported values of aequorin calcium affinity – 13 µM reported by Kendall and co-workers (1992) – to allow estimation of an apparent affinity from the concentration-response curves of the aequorins versus the different metals.

Protein samples were handled as described in Section 7.3.6 in order to avoid contact with contaminating calcium in vessels as well as inactivation due to prolonged exposure at room temperature. The 5.1 µg/mL protein solutions were kept on ice in a universal tube that had been previously rinsed with 50 mM Tris-HCl, 10 µM EDTA buffer so that any metal ions which may be present on the tube surface would be chelated and removed.

8.2.3 Concentration range of metal ions

Salts of calcium, lanthanum, cadmium, zinc, copper, cobalt and lead Table 2–9 were dissolved in 50 mM Tris-HCl buffer, pH 7.5 and their actual metal ion content was determined by elemental analysis (Section 2.5.2). The concentrations of all metal ion solutions would include the concentration range from 10^{-7} M to 10^{-3} M that is relevant for biomedical and environmental studies (Pinton *et al.*, 2007, Shimomura, 1991, Shimomura, 2006, Tricoire *et al.*, 2006). Nominal dilutions were made over a range of 10^{-7} M to 10^{-2} M for all metal ions studied. The values of the nominal concentrations were corrected using total ion determination data (Section 2.5.2). As EDTA would be present in the final bioluminescence mixture of protein and metal ions, only free (unchelated by EDTA) metal ions would be available for aequorin binding.

The WinMAXC32 v2.51 Chelator program (Patton, 2002) was used to calculate the free concentration of each metal solution in the final (luminescence measurement) mix taking into account the total metal ion concentration, 2.3×10^{-7} M total contaminating calcium (traces of which are present in all chemical preparations) and the program in-built affinity constants of EDTA for all metal ions used. The original range of evenly spaced nominal ion concentrations (10^{-7} to 10^{-2} M) resulted in a much wider range of calculated free (unchelated) metal ion concentrations (10^{-16} M to 10^{-2} M) due to the chelating effect of EDTA (up to 10 μ M) in the final mix.

According to studies by Shimomura and Shimomura (1984) aequorin's detection limit of calcium ions is 1 nM in low ionic strength buffers and 100 nM in 0.15 KCl solutions. In this work 0.25 nM of free calcium in 50 mM Tris-HCl 5 μ M EDTA pH 7.5 (negative control buffer) did not trigger luminescence above noise level. 6.4 nM of free calcium in 50 mM Tris-HCl 5 μ M EDTA pH 7.5 was found to trigger measurable luminescence in wild-type aequorin.

8.2.4 Experimental setup and platereader settings

Aequorin activity was measured in a BMG FluoStar Optima platereader (Section 2.7.1) by injecting 50 μ L of protein solution (5.1 μ g/mL wild-type or mutant aequorin in 50 μ M Tris-HCl 10 μ M EDTA) into microwells containing 50 μ L of variable metal ion

concentrations in 50 mM Tris pH 7.5 (Section 8.2.3). The measurements were performed in Lumitrac™ 96-well plates at 21–23 °C.

Measurements were performed in triplicate (unless stated otherwise). Injection was set to maximum speed (420 $\mu\text{L/s}$) for rapid mixing and luminescence was measured for a total duration of 15 s. Recording of luminescence started at time 0 s and the injection of protein into the wells started at 0.3 s in order to capture any background signal before and during injection. Data were collected with 0.1 s integration time. Where background noise was detected, it was subtracted from light readings before further data analysis.

Due to significant differences in light intensity across different combinations of aequorins and metal ions, a uniform gain setting across all experiments would either cause over-range readings (e.g. wild type plus calcium) or no detection in others (e.g. mutant plus zinc). The gain settings of the photomultiplier tube were altered in order to adjust the sensitivity of the light readings through trial and error. The gain settings of the photomultiplier ranged from gain 2,000 to 4,000. At low gain settings no background signal was detected. Low levels of noise were captured at gain $\geq 3,500$. The absolute luminescence output values are not directly comparable across the entire dataset, but only within each plate or across plates processed at the same gain settings.

In order to protect the protein administered through the syringe from traces of contaminating calcium, the syringe pump and tubing were rinsed with 50 mM Tris-HCl buffer solutions with decreasing concentrations of EDTA (first wash: 10 mM, final wash: 10 μM EDTA) for the removal of metal ions and finally primed with 1 mL of sacrificial protein sample.

Addition of high calcium concentrations to protein solutions containing EDTA can cause a drop in pH. It has been reported that aequorin's stability is maximal within the pH range 6.0–8.0 (Shimomura, 2006, Shimomura and Johnson, 1969, Shimomura and Johnson, 1970). Changes in pH were measured for representative samples of each set of experiment.

Solutions containing low and high concentrations of each of the metal ions used in this chapter were mixed 1:1 with the buffer aequorin was diluted in (50 mM Tris-HCl 10 μM EDTA, pH 7.5) in order to assess the effect of EDTA in the final mix on pH. The pH of all the solutions remained between 7.1–7.4 post-mixing, with the exception of solutions containing lead. As lead precipitated in Tris-HCl, it was diluted in R.O. water instead (Section 2.5.1) and the pH of the solutions was 5.0–5.2 prior to mixing.

Post-mixing pH was 7.5 at the lowest and 2.96 at the highest concentration of lead tested. The lack of pH buffering in the lead system is expected to affect the protein's stability and activity. This will be taken into consideration in discussion of results.

8.2.5 Method for double ion experiment

Forty microliters of increasing zinc concentration solutions (1.5 μ M to 5.5 mM before mixing) were pipetted into microwells. Forty microliters of 5.1 μ g/mL purified aequorin (wild-type or mutant) in 50 mM Tris-HCl EDTA 10 μ M EDTA pH 7.5, were injected into the microwells at time 0.3 s and the reaction was allowed to proceed until time 10 s. At 10 s, 40 μ L of 50 mM Tris-HCl 50 μ M CaCl_2 (shown in Section 8.3.2 to be sufficient amount of calcium to complete the luminescence reaction and reach plateau in the dose response curves) were injected into the microwells in order to provide aequorin with enough calcium to successfully activate the luminescence reaction.

8.2.6 Data processing

8.2.6.1 Luminescence curves and concentration-response curves

Light versus time data were originally processed in Microsoft Excel and the peak and sum of luminescence counts during measurement time (15 s) were obtained. The term "Peak luminescence" was assigned to the single maximum luminescence measurement of each well. This term is relevant to the flash-type luminescent reactions and perhaps less relevant for glow-type reactions (Section 1.1.5, Figure 1.4).

Concentration-response curves were populated with the peak of the luminescence trace curve for every concentration of each ion tested versus ion concentration. As luminescence counts vary across different instruments and across different gain settings of the same instrument, the values of luminescence peaks were normalised to the maximum peak observed for that particular ion and the y-axes of the concentration-response curves were presented as percentage of maximum for the particular set of data (aequorin molecule and metal ion). The concentration-response curves were plotted in SigmaPlot software v.11.0 (Systat Software Inc., Chicago, USA) and EC_{50} values (and IC_{50} where applicable) were determined.

8.2.6.2 Exponential model fitting

Luminescence exponential decays were analysed with SigmaPlot software. Fast and slow components of double exponentials are described by their fast and slow time constants (τ_F and τ_S respectively), and their light integrals (Σ_F and Σ_S respectively), while the total light integral $\Sigma_T = \Sigma_F + \Sigma_S$ is defined in Section 8.1.5.2.

The model of double exponential decay published by Tricoire and co-workers (2006) (Section 8.1.5.2) is appropriate only to the reactions exhibiting flash luminescence. Based on this work, Equation 8.15 was fitted to the time resolved data points past the peak of the luminescence curve using SigmaPlot. The data points of each individual dataset (i.e. corresponding to each ion concentration for wild type or mutant) were normalised by the luminescence peak of the dataset. Equation 8.15 is an adaptation of Equation 8.12 to using normalised data points:

$$y = a \times e^{-t/\tau_F} + c \times e^{-t/\tau_S} \quad \text{Equation 8.15}$$

where τ_F and τ_S are the time constants of the fast and slow component respectively and a and c are the fractions of the initial maximum intensities of the fast and slow component, where $a+c=1$ (or 100% of total light contribution).

8.2.6.3 Estimation of glow luminescence kinetics

In the case of glow luminescence traces, a useful metric of the speed of decay would normally be the time required for light to fall to half of the maximum value recorded ($t_{1/2}$). In this work, measuring $t_{1/2}$ was not possible because – for most experiments – the decay was too slow to extract this parameter within the 15 s window of each experiment. Instead time was fixed at 15 s and the ratios between luminescence at 15 s over maximum luminescence of the glow curve were compared. Thus, by quantifying the progress of the decaying curve, an indirect measure of the progression of bioluminescence reaction (or reaction rate) was obtained, which was named $P_{15s/Max}$ and defined as shown in Equation 8.9. Specifically, from each glow luminescence curve $P_{15s/Max}$ was calculated:

$$P_{15s/Max} = 1 - \frac{L_{15s}}{L_{Max}} \quad \text{Equation 8.16}$$

where:

L_{15s} is the luminescence at 15 s and

L_{Max} is the maximum value of luminescence

Subtracting the luminescence ratio from 1 served to produce a metric of the progress of the reaction: higher values for $P_{15s/\text{Max}}$ correspond to higher reaction kinetics. This metric by its definition can give information of reaction kinetics only and cannot be used to describe processes related with yield of light.

Upon implementation, this proved to be a robust metric in that the variance among triplicate experiments was very low. At very low and very high ion concentrations low levels of luminescence and high noise prevented accurate estimation of L_{15s} and L_{Max} and such data points were discarded.

The glow curves are associated with low yield production. For this reason high gain settings were applied in the platereader which in turn resulted in noisy data in several cases. To mitigate this effect, averaging was used both for the estimation of both L_{Max} and of L_{15s} . Specifically, to determine the maximum response observed, the raw data was processed using a moving average filter. A moving average is commonly used with time series data to smooth out short-term fluctuations and highlight longer-term trends or cycles. For each time point the average of the two preceding points, the current point and the two following points was calculated. To minimise the effect of noise on L_{15s} value, this was calculated as the mean of the five last data points (for time 14.5 to 14.9 s).

8.2.7 Artefacts from syringe injection

Artefacts are common occurrences in luminescence and fluorescence readings and they are often associated with injection events. They are different to background noise as to their timing and duration: background noise is present in blank samples or samples where a reaction has not been initiated and its values are fairly stable; the experimental results can be corrected for noise by means of a simple subtraction. Artefacts are different in that they can occur in one or more specific instances during a experiment, e.g. more often associated with injection or other events. Frictional forces due to a very rapid injection can excite the probe and produce an injection artefact (Van Dyke *et al.*, 2002). In this work, luminescence artefacts associated with the event of injection were identified and removed from the data set.

8.2.8 Assay objectives and limitations

The purpose of the assay was to assess the response of aequorin wild-type and its mutant variant Asn28Cys/Ser32His to various metal ions and compare the different ions in their ability to induce luminescence in aequorin.

The experimental setup consists of mixing protein with metal ion solutions and recording the light output. It is important to note the assay's limitations. There is no direct or indirect measurement of the amount of ions free in solution or bound by the protein, hence binding curves are not attainable. Structural/conformational changes occurring in the protein as a result addition of ions were not monitored. The assay did not record any potential spectral changes in light emission but measured all visible light as produced by the luminescence reaction.

With these limitations in mind, the following types of information can be elicited:

- qualitative and quantitative analysis on the effect (activating or inhibitory) of ions on luminescence activity of aequorin wild-type and mutant by plotting concentration-response curves
- the rate of luminescence can be quantified into a (set of) kinetic constants or characteristic times
- a measure for sensitivity for each ion can be introduced, by identifying the lowest ion concentration at which bioluminescence proceeds or by means of estimating EC50.

These metrics are fairly easy to estimate and are quite relevant in the applications of aequorin as a metal biosensor.

8.3 Results and discussion

Seven metal ions – calcium, lanthanum, cadmium, lead, zinc, copper and cobalt – were tested as potential activators of wild type and mutant aequorin. Luminescence response (activity) versus time was monitored.

Section 8.3.1 presents one typical shape of the luminescence traces for each combination of protein-metal tested. The shape of the curves will be discussed and maximum values from each luminescence curve at every ion concentration will be used to compile concentration-response curves in Section 8.3.2. Concentration-response curves show sensitivity of the protein to each metal and the effect of each metal as an activator or inhibitor at the ranges of concentrations tested. In Section 8.3.3 the ions are compared to each other for their ability to trigger light production in wild-type and mutant aequorin.

Kinetic parameters for the experimental results were derived using two methods of analysis depending on the shape of the luminescence traces in Sections 8.3.4.1 and 8.3.4.2.

Double ion experiments in Section 8.3.5 show the effect of two ions on aequorin (zinc and calcium), with zinc envisaged to bind to and possibly activate the mutated EF-hand loop and calcium to bind to the non-mutated EF-hands.

The experimental results regarding shape of luminescence curves, concentration-response curves, sensitivity and yield are discussed in relation to structural parameters derived in Chapter 5.

8.3.1 Luminescence curves: flash and glow-type

The light intensity (luminescence) was recorded as described in Section 8.2.4. Luminescence was plotted against time for each metal ion concentration. Both wild type and mutant were tested with twelve metal ion concentrations of seven different metal ions and each experimental condition was performed in triplicate. For simplicity one representative luminescence trace from each aequorin-metal set was chosen to illustrate the shape of the graphs in Figure 8-6 (wild type) and Figure 8-7 (mutant).

From the entire dataset of time course experiments with wild-type aequorin and all concentrations of calcium (Supplementary Material), it was interesting to notice that the rise time was independent of calcium concentration (luminescence peak occurred at 1 s) but the rate of decay was affected by calcium concentration. Both observations are in agreement with published literature (Hastings *et al.*, 1969).

Although luminescence traces may vary slightly in sharpness between different ion concentrations for each aequorin-metal combination, the curves overall fall under two distinct shapes. Luminescence triggered by calcium, lanthanum and lead is flash-type, characterised by a sharp onset and sharp decay of light (Figure 8-6) Luminescence traces triggered by cobalt and zinc are glow-type, characterised by a slow rise to a fairly broad peak followed by a slow decay (Figure 8-7).

Luminescence curves triggered by copper were classified as glow-type because the shapes of the traces varied from a non-sharp peak to a “classic” glow-type shape depending on the concentration of copper (Supplementary Material). Cadmium presents a special case in that it triggers luminescence traces of a unique shape that seems to be a contribution of a sharp response followed by a glow-type luminescence curve with a slow decay.

The classification of the shape of luminescence curves is the same in wild-type aequorin and variant Asn28Cys/Ser32His, but overall the kinetics of decay in the mutant are slower compared to wild type. In wild-type aequorin, luminescence reached a sharp peak at 1 to 1.1 s, with the injection occurring between 0.3 and 0.4 s. Gain settings varied across different ions and different proteins (Appendix Chapter 8, Table 10-14) therefore the output of luminescence units on the y-axes cannot be compared across the different plots.

Table 8-2 Luminescence curve shapes.

	Wild-type aequorin	Mutant Asn28Cys/Ser32His
Calcium	Flash-type	Flash-type
Lanthanum	Flash-type	Flash-type
Lead	Flash-type	Flash-type
Cadmium	combination	combination
Cobalt	Glow-type*	Glow-type*
Zinc	Glow-type	Glow-type
Copper	Glow-type	Glow-type

*Borderline flash-type

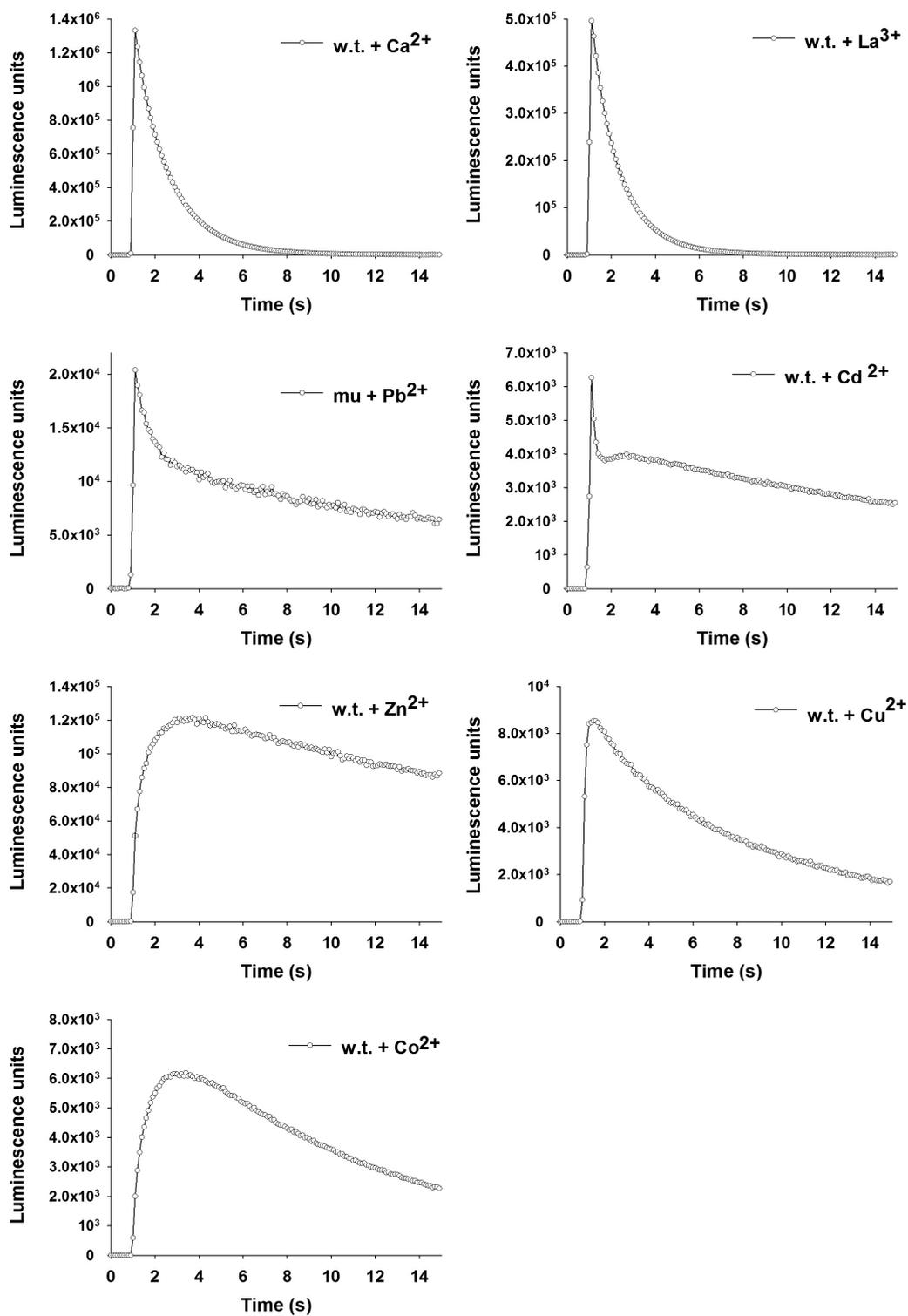


Figure 8-6 Light versus time – Wild-type aequorin. Tested with Ca²⁺, La³⁺, Pb²⁺, Cd²⁺, Cu²⁺, Zn²⁺, Co²⁺ at the respective ion concentrations that trigger maximal response of activity and correspond to the peak of the concentration-response curves in Figure 8-8. 2.55 μg/mL of protein post mix, pH 7.45 (lead at pH ~5), T=23 °C.

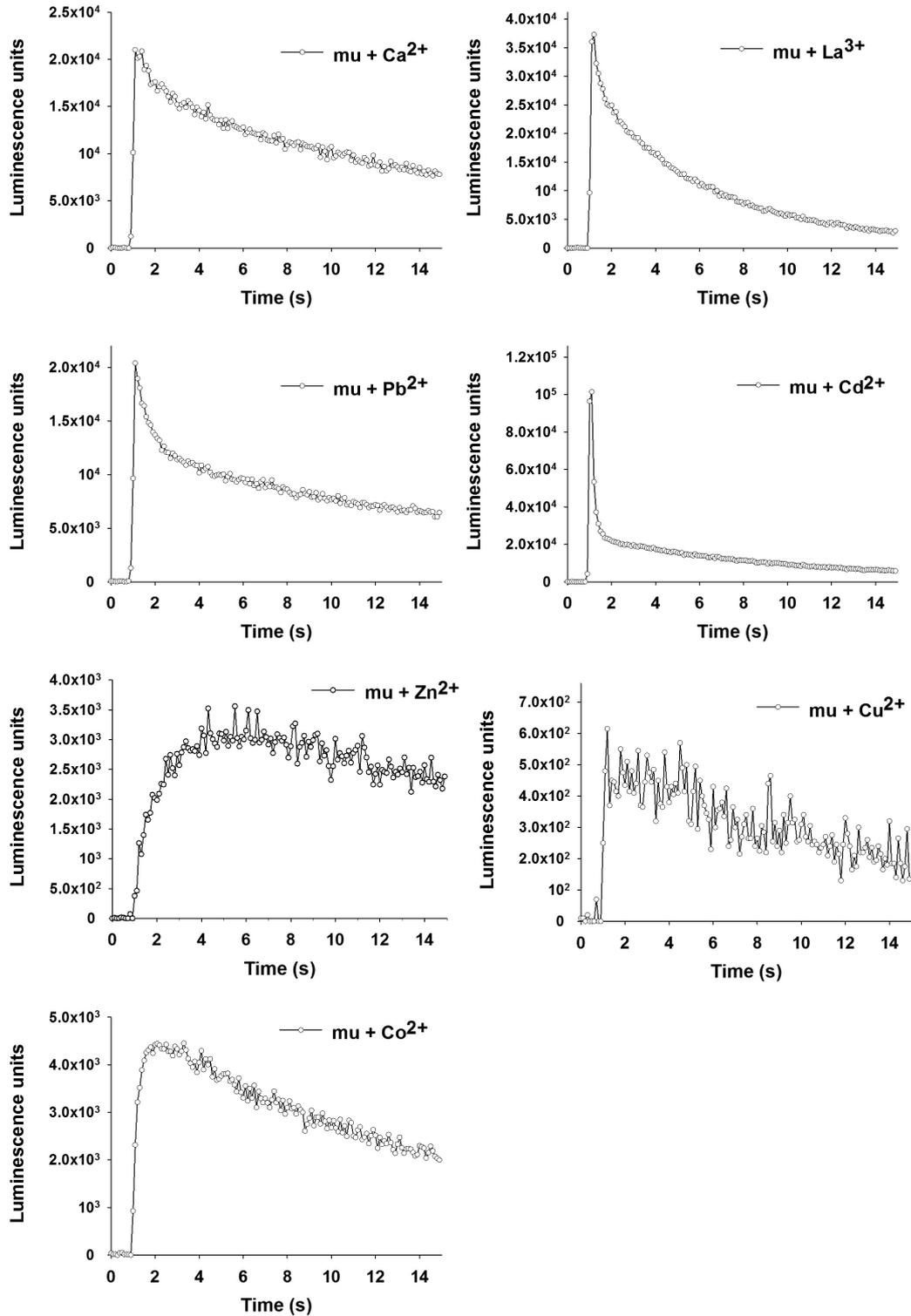


Figure 8-7 Light versus time – Mutant Asn28Cys/Ser32His. Tested with Ca²⁺, La³⁺, Pb²⁺, Cd²⁺, Cu²⁺, Zn²⁺, Co²⁺ at the respective ion concentrations that trigger maximal response of activity and correspond to the peak of the concentration-response curves in Figure 8-9. 2.55 μg/mL of protein post mix, pH 7.45 (lead at pH ~5), T=23 °C.

8.3.2 Concentration-response curves

Concentration-response (or concentration-response) curves were drawn to study the effect of varying concentrations of metal on wild-type and mutant aequorin. Triplicates of a range of ion concentrations per aequorin-metal combination were used. To simplify comparisons all results were normalised by the highest measured activity of each set so that the y-axis of response is the percentile of maximum response achieved within each set of aequorin-metal combination (e.g. wild type and calcium, mutant and calcium etc).

A visual inspection of the concentration-response curves reveals that the different ions exert different effect on wild-type and mutant aequorin. The sigmoidal curves of wild type triggered by each of calcium, lanthanum and cadmium suggest cooperative binding (both linear and semi-logarithmic plots were sigmoidal). The sigmoidal curve of wild type plus calcium agrees well with the sigmoidal curve of commercial aequorin AquaLite[®] (Figure 8-8), which indicates that the preparation of aequorin in this work and the presence of His-tag and two additional amino acids at the C-terminus (Section 3.3.4, Figure 3-8 and Figure 3-9) did not alter its properties in regard to shape of the bioluminescence trace and calcium sensitivity. Additionally, it will be shown in Section 8.3.4.1 that kinetics of wild type and mutant aequorin are very similar.

The profiles of mutant aequorin triggered by each of calcium, cadmium and lanthanum resemble a sigmoidal curve but are not as concrete as the profiles of wild type with the same ions. In the case of lanthanum, it hints toward a sigmoidal shape but very few non-zero points were available. For biomedical systems however, concentrations of free lanthanum higher than 0.1 M would not be applicable. It is noteworthy that mutant aequorin shows dramatically decreased sensitivity to lanthanum compared to wild type.

The concentration-response curves of lead, cobalt, zinc and copper with wild type and mutant are all bell shaped (Section 8.1.4). This indicates a departure from the cooperative binding model. At low concentrations these ions act as partial activators of aequorin until a maximum is reached. At higher concentrations they have an inhibitory effect on aequorin luminescence for both wild type and mutant. In terms of ion sensitivity, wild-type was found to be more sensitive to all other ions than its native activator calcium, as expressed by the maximum values (and consequently the EC₅₀)

occurring at significantly lower concentrations of metals. However the activity produced with other ions as activators was significantly lower than the wild-type plus calcium combination, summarised in Figure 8-10, Section 8.3.3. Characteristic values and qualitative observations of the concentration-response curves are summarised in Table 8-3. Due to lack of some data points in the interesting “ascending range” of some of the concentration-response curves, it was not possible to extract exact values but only ranges for EC_{50} and IC_{50} (where applicable).

Overall, mutant aequorin exhibited increase of sensitivity to calcium and zinc (reduction of EC_{50} of approximately one order of magnitude) compared to wild type. The opposite effect was seen with lanthanum, where mutant aequorin almost lost its sensitivity as shown by an estimated increase of EC_{50} of between four and seven orders of magnitude compared to wild type (Table 8-3). The sensitivity to cadmium remained relatively unchanged. Comparisons of sensitivity to copper, cobalt and lead were not possible due to lack of data in the very low ranges where the respective EC_{50} values resided.

These results can be further discussed in terms of physical processes. In general, increase of sensitivity (decrease of EC_{50}) means either faster or stronger binding, or faster light production by the activated intermediate.

The increase of mutant sensitivity to calcium agrees with findings by Tricoire and co-workers (2006) in regard to effect of mutations in EF-I. The authors reported the “seemingly paradoxical” increase of calcium sensitivity in mutants that contained altered EF-I but a decrease of sensitivity in mutants that contained altered EF-III or EF-IV. As a result they assumed that EF-I must have lower calcium affinity than EF-III and EF-IV; the exact increase in sensitivity was not quantified as the ascending range in the mutant concentration-response curve was outside the calcium concentration range tested by the authors. The work in this section only included mutated EF-I with different type of amino acid substitutions from Tricoire’s work (EF-I mutations in this work: Asn29Cys/Ser32His and in Tricoire: Asn26Asp, Glu35Gly and Val44Ala) but the agreement in respect to increase of calcium sensitivity is still interesting.

The dramatic decrease of lanthanum sensitivity in the mutant alludes to EF-I being crucial in the binding of lanthanum. Indeed, in the analysis of the mutant library in Chapter 6 the metric of selectivity shift (D) for lanthanum and mutant Asn28Cys/Ser32His was very low compared to that for zinc and copper.

From the compilation of Table 8-3 an association emerges between the shape of the luminescence curves with the shape of the concentration-response curves. Flash-type luminescence curves contribute to sigmoidal concentration-response curves whereas glow-type luminescence curves contribute to bell-shaped concentration-response curves. Some seeming variations from this rule are cadmium, lead and cobalt. Cadmium triggers a combination of flash and glow luminescence and its flash-type contribution most likely leads to a sigmoidal concentration-response curve. Luminescence triggered by cobalt was classified as glow-type, however its luminescence versus time traces are sharper than the glow-type traces produced by zinc and copper and more blunt than the classical flash-type curves of calcium, lanthanum and lead (traces of the entire dataset are provided in Supplementary Material). The relative sharpness of some of the traces by cobalt would contribute to a sigmoidal concentration-response curve.

Lead triggers flash-type luminescence but at higher concentrations an inhibition effect is observed that leads to a bell-shaped concentration-response curve.

On enzyme inhibition due to metal ions

It is possible for heavy metals, such as lead, silver and mercury, to act as enzyme inhibitors by bonding with side groups such as -SH, -COO⁻, -OH (Hovde, 2011). ‘Sulfur-seeking’ metals, such as cadmium and copper, would be expected to bind mainly to thiols and zinc would be expected to bind mainly to carboxylic acids at cytoplasmic pH (Martell and Smith, 1982).

Binding to the active site can cause competitive reversible or irreversible inhibition and binding to a site other than the active site causing non-competitive reversible inhibition. In biological systems, heavy metal toxicity is caused by tight binding of a metal such as mercury, lead, aluminum, or iron, to a functional group at the active site of an enzyme. At high concentrations heavy metals are relatively non-specific for the enzymes they inhibit and inhibit a large number of enzymes.

In the case of aequorin it is expected that metal ions binding to the EF-hands causing partial or no activation, by definition also cause at least partial inactivation of aequorin. It is entirely possible that at higher concentrations they also bind to sites other than the ion-binding loops causing additional inhibition with the cumulative inhibition effect manifested in the downslope of the bell-shaped curves and reduction at the high concentrations of lead’s sigmoidal concentration-response curve.

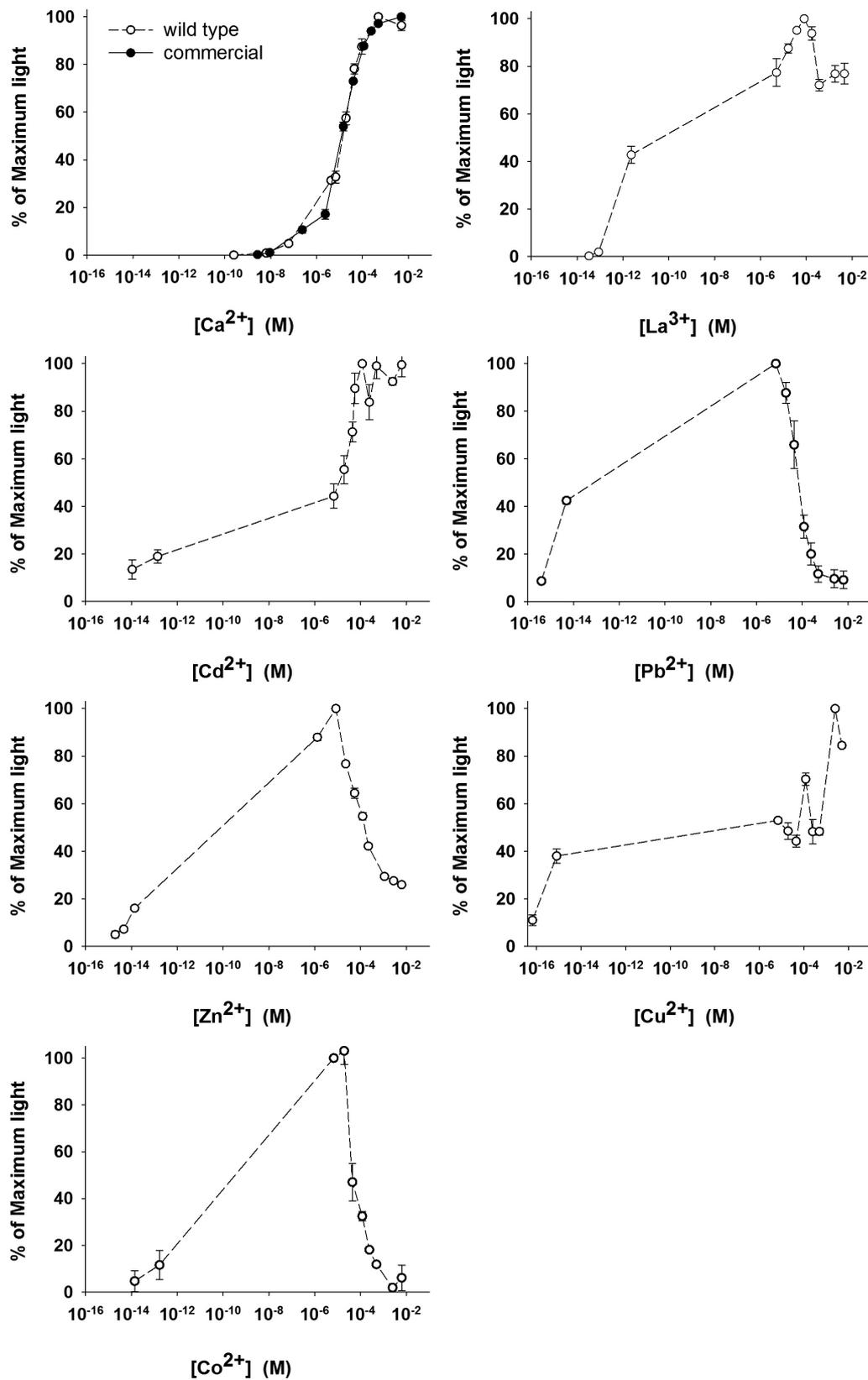


Figure 8-8 Concentration-response curves of wild-type aequorin. Tested with Ca^{2+} , La^{3+} , Pb^{2+} , Cd^{2+} , Cu^{2+} , Zn^{2+} , Co^{2+} . 2.55 μ g/mL protein post mixing, pH 7.45 (lead at pH \sim 5.0), T=23 $^{\circ}$ C.

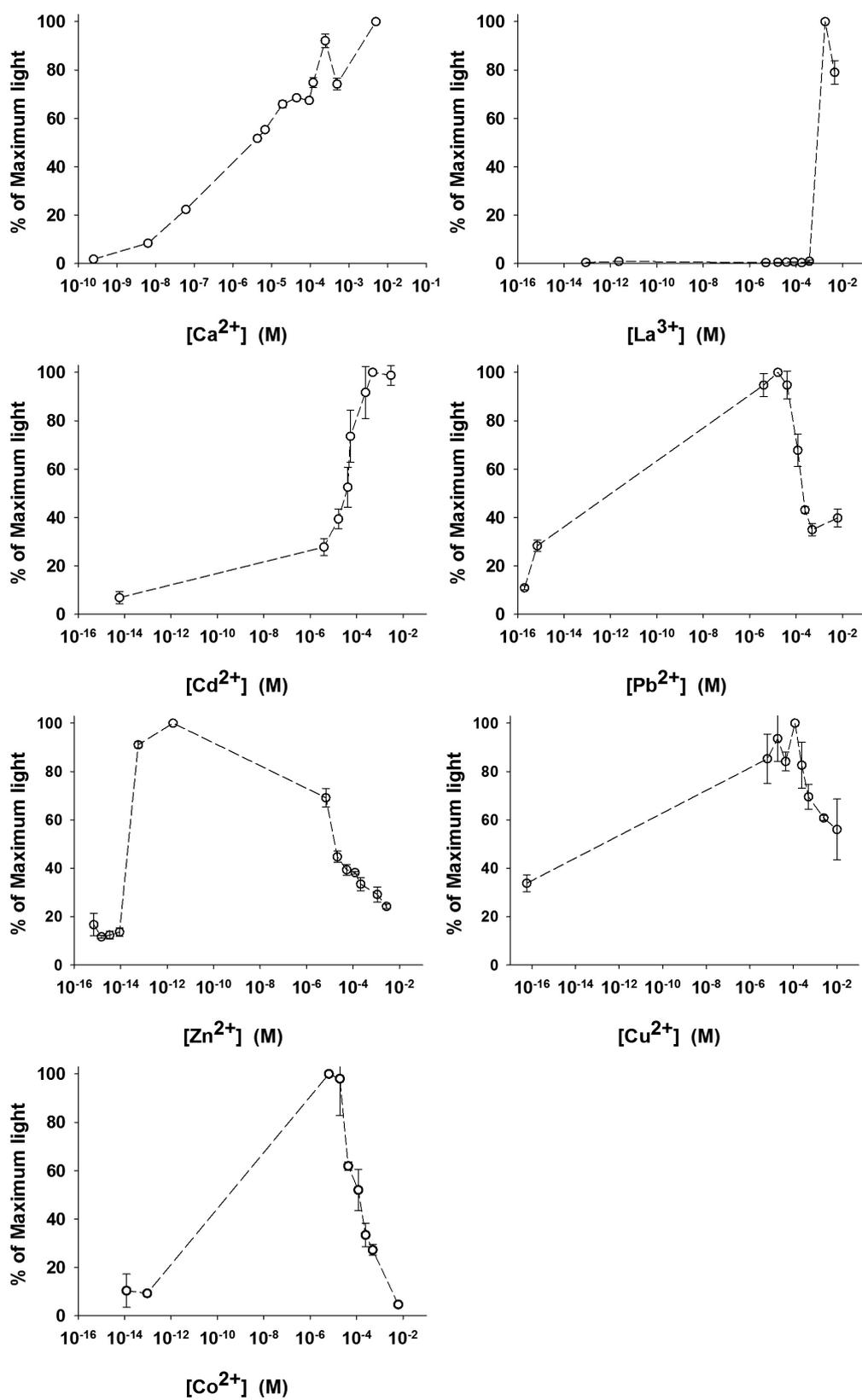


Figure 8-9 Concentration-response curves of mutant Asn28Cys/Ser32His. Tested with Ca²⁺, La³⁺, Pb²⁺, Cd²⁺, Cu²⁺, Zn²⁺, Co²⁺. 2.55 μ g/mL protein post mixing, pH 7.45 (lead at pH ~5.0), T=23 °C.

Table 8-3 Summary of characteristics of luminescence versus time and concentration-response curves. When values for EC₅₀ and IC₅₀ could not be drawn from the plots due to lack of data points, a range of potential values is provided instead.

wild-type aequorin				mutant Asn28Cys/Ser32His				
	EC ₅₀ (M)	IC ₅₀ (M)	luminescence curve	concentration-response curve	EC ₅₀ (M)	IC ₅₀ (M)	luminescence curve	concentration-response curve
Ca ²⁺	1.2×10 ⁻⁵ –1.6×10 ⁻⁵	n.a.	flash-type	sigmoidal	2×10 ⁻⁶	n.a.	flash-type	sigmoidal
Zn ²⁺	10 ⁻¹³ –10 ⁻¹⁰	1.3×10 ⁻⁴	glow-type	bell-shaped	2×10 ⁻¹⁴	2×10 ⁻⁵	glow-type	bell-shaped
Cu ²⁺	3×10 ⁻¹⁵ –3×10 ⁻⁸	n.d.	glow-type	bell-shaped	10 ⁻¹⁴ –10 ⁻⁸	3×10 ⁻²	glow-type	bell-shaped
Co ²⁺	10 ⁻¹² –2×10 ⁻¹⁰	5×10 ⁻⁵	glow-type	bell-shaped	5×10 ⁻¹³ –2×10 ⁻¹⁰	10 ⁻⁴	glow-type	bell-shaped
Cd ²⁺	10 ⁻⁵	n.a.	combination	sigmoidal	3×10 ⁻⁵	n.a.	combination	sigmoidal
Pb ²⁺	10 ⁻¹⁴ –10 ⁻¹³	9×10 ⁻⁵	flash-type	bell-shaped	2×10 ⁻¹⁵ –10 ⁻¹²	2×10 ⁻⁴	flash-type	bell-shaped
La ³⁺	5×10 ⁻¹¹ –10 ⁻⁸	n.a.	flash-type	sigmoidal	9×10 ⁻⁴	n.a.	flash-type	sigmoidal (?)

n.d.: not determined

n.a.: not applicable

IC₅₀ values are only applicable in bell-shaped concentration-response curves

8.3.3 Light yield comparison across all metals with wild-type and mutant aequorin

In biomedical research or environmental monitoring the minimum measurable output of a luminescent biosensor would be the peak or sum (cumulative light) of emitted light during the duration of the experiment. In this Section, the peak and sum of luminescence for the experiment duration (15 s) are compared between:

- purified wild-type aequorin with each of the metal ions tested (Figure 8-10)
- purified mutant Asn28Cys/Ser32His with each of the metal ions tested (Figure 8-11)

As discussed in Section 8.1.4, the maximum response occurs with different concentrations of different ions and varies between wild-type and mutant aequorin. For simplicity, the comparison was made at each ion's optimum concentration with wild type and mutant respectively. This decreased the number of experiments required. The metal ion concentrations used in each case are presented in Table 8-4.

Photomultiplier gain settings were an additional consideration. In contrast to work in Sections 8.3.1 and 8.3.2 the gain settings within wild type experiments were kept constant (Gain 2,300) in order to allow direct comparison between luminescence outputs. The respective gain settings used for mutant comparisons were 2,300 and 4,000. Gain of 2,300 allowed direct comparison across wild type and mutant aequorin with calcium but failed to detect light from some of the mutant-ion sets. Gain of 4,000 allowed luminescence from low producers (mutant variant with zinc, cobalt, copper) to be detected that would otherwise have been undetectable at the low gain setting of 2,000.

As illustrated in Figure 8-10 and Figure 8-11, the order in which metal ions exert the highest peak of luminescence in aequorin ranked from highest to lowest is:

- For wild type:
Calcium > Lanthanum > Lead > Cobalt > Cadmium > Zinc > Copper
- For mutant Asn28Cys/Ser32His
Lanthanum > Cadmium > Calcium > Lead > Zinc > Cobalt > Copper

Due to the shape and kinetics of the luminescence curves, the ranking may vary slightly where peak or sum of luminescence is taken into account for the entire duration of the experiment (15 s). For example, the peak response of wild type to lead is significantly higher than response to cobalt, whereas the sum of light triggered by cobalt is slightly higher than that triggered by lead for the experiment duration. This is associated with the different speed at which reactions proceed.

The peak of light is the most relevant metric for two reasons: a high intensity peak would allow lower limits of detection, hence increased sensitivity of the sensor; a sharp peak with a rapidly descending tail would mean it is possible to study rapid fluctuations of the analyte of interest. In this system, high luminescence intensity seems to be linked to increased sharpness of peak, hence indirectly suggests fast response kinetics.

Table 8-4 Metal ion concentrations used for overall yield comparisons. The available metal ion concentrations in the final mix (metal ion plus protein solution) that triggered the highest activity. These were used for the overall comparison of wild type and mutant performance with all ions. Free ion concentration post-mixing was calculated as described in Section 8.2.3.

Ion	Most effective ion concentration for wild type	Most effective ion concentration for mutant
Calcium	10^{-2}	10^{-2}
Lanthanum	8.1×10^{-5}	1.8×10^{-3}
Cadmium	1.9×10^{-5}	2.4×10^{-4}
Lead	7×10^{-6}	4×10^{-6}
Zinc	8.4×10^{-6}	1.8×10^{-12}
Copper	1.2×10^{-4}	3.9×10^{-6}
Cobalt	1.9×10^{-5}	1.9×10^{-5}

Luminescence activity of mutant plus calcium in comparison with wild-type plus calcium at 10^{-2} M and gain setting 2,300 was $0.032 \% \pm 0.001 \%$ for Peak and $0.128 \% \pm 0.003 \%$ for Sum.

In theory the total light produced by the luminescence reaction is constant (Hastings *et al.*, 1969). This should apply at all experimental conditions given infinite time. In practice, infinite time is not applicable. The sum of light readings for the duration of each experiment was used as it was a concrete set of results and relevant to the kinetics of the reaction (slower kinetics → lower sum of light in a set timeframe).

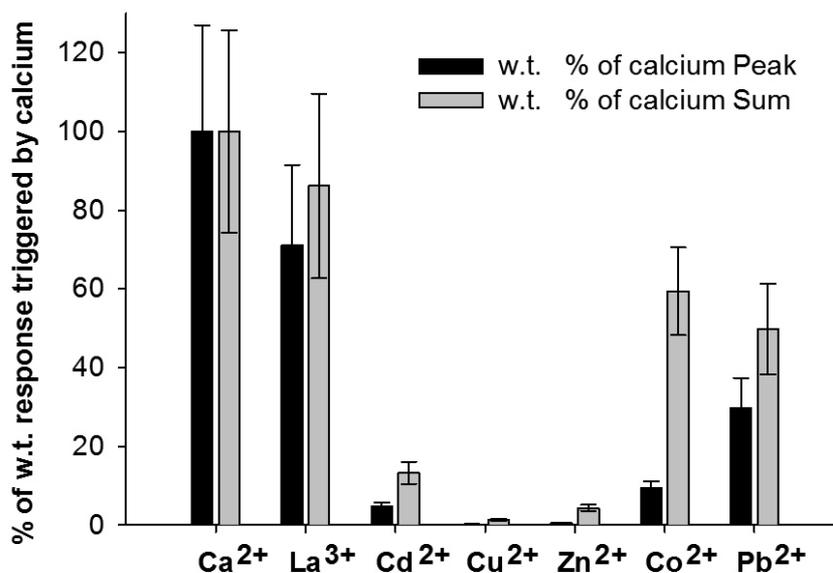


Figure 8-10 Wild type against each metal ion. 2.55 $\mu\text{g/mL}$ aequorin, pH 7–7.5 at 23 $^{\circ}\text{C}$ (with the exception of lead that was diluted in R.O. water and post-mix pH was 5.0). Error bars are standard error of the mean. Each metal ion injected was at the optimal concentration derived from the concentration-response curves in Section 8.3.2 and summarised in Table 8-4.

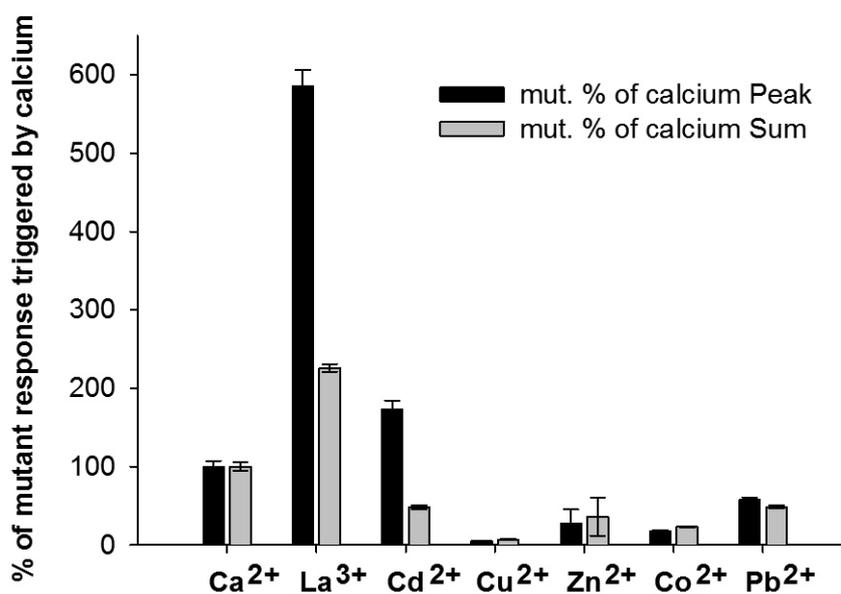


Figure 8-11 Mutant Asn28Cys/Ser32His against each metal ion. 2.55 $\mu\text{g/mL}$ of protein, pH 7–7.5 at 23 $^{\circ}\text{C}$ (with the exception of lead that was diluted in R.O. water and post-mix pH was 5.0). Error bars are standard error of the mean. Each metal ion injected was at the optimal concentration derived from the concentration-response curves in Section 8.3.2 and summarised in Table 8-4.

8.3.4 Analysis of bioluminescence kinetics

In Section 8.3.1 it was shown that experimental results fall under two main groups in terms of the shape of luminescence curves: flash-type and glow-type. The shape of luminescence curves makes them amenable to different methods of analysis in terms of kinetics.

Flash-type luminescence comprises of the data sets of wild-type and mutant aequorin with calcium, lanthanum and lead. Whereas the ascent from zero to the peak value is the same amongst the data set, the luminescence decay differs with aequorin type (wild type or mutant), ion type and ion concentration and were fitted to a model of exponential decay in Section 8.3.4.1.

Glow-type data were processed in Section 8.3.4.2 on the basis of progression of the bioluminescence reaction within the time course of the experiment. Due to the combination of flash- and glow-type luminescence present in the cadmium-triggered curves, neither method was deemed suitable.

8.3.4.1 Flash-type luminescence: Model of double exponential decay

The decay luminescence of wild-type and mutant aequorin with each of calcium, lanthanum and lead (flash-type luminescence curves) was fitted with the double exponential decay model proposed by Tricoire and co-workers (2006) and introduced in Section 8.1.5.2.

Time constants (τ) express of the duration of a phenomenon whereas kinetic constants – the inverse of time constants (k) – express the speed of a phenomenon. High values of τ indicate a slow luminescence branch and low values indicate a fast luminescence branch. Salient observations will be discussed in regards to the values of the time constants and the contribution of the fast and slow component to total light.

The slow and fast time constants (τ_S and τ_F respectively) as well as contribution of the slow and fast component to the integral luminescence (Σ_S/Σ_T and Σ_F/Σ_T respectively) were derived using SigmaPlot as described in Section 8.2.6.2 and plotted against a wide range of ion concentrations (Figure 8-12 and Figure 8-13). Commercial aequorin Aqualite[®] was also tested with calcium in order to allow comparison with wild-type aequorin produced in this work and the results of the exponential model fitting are presented in Figure 8-14.

Tricoire and co-workers (2006) fitted each trace of luminescence decay to a sum of two exponentials and found that it gives better description of the data compared to fitting a single exponential component (Section 8.2.6.2). This is something that was also observed in this work.

Kinetics of commercial versus wild-type aequorin

In order to verify that the His-tag and additional amino acids (resulting from pET cloning of Chapter 3) in aequorin did not alter the bioluminescence kinetics, kinetic parameters were derived for both wild-type and commercial aequorin (Aqualite[®], Section 8.2.1).

Both proteins exhibited the same trend of τ_F increasing slightly with increasing calcium concentrations and τ_S decreasing with increasing calcium concentrations until they both reached a plateau (Figure 8-12 and Figure 8-14). The ranges of time constants for both proteins were approximately the same: 1.2–14 s (τ_S) for the slow and 0.2–1.3 s (τ_F) for the fast components. A similar trend is observed in the contribution of the slow and fast branch to integral light (Σ_S/Σ_T and Σ_F/Σ_T respectively), with the fast branch increasing its contribution at higher calcium concentrations. Since no significant differences were observed, the remaining analysis in this section focused on comparisons between wild-type and mutant aequorin.

Effect of mutation on the fast reaction time constants

Contrasting wild-type to mutant aequorin, no significant differences in the numeric values were observed among the ions causing flash luminescence (calcium, lanthanum and lead) (Table 8-5). Examination of Figure 8-12 and Figure 8-13 shows that although some variation exists, there is a general trend. This trend consists of a gradual increase of τ_F from 0.2 s between low and middle ion concentrations, reaching a plateau range around 1.2 s at high concentrations of calcium and lanthanum. The τ_F values for wild type plus lead fluctuated between 0.1 and 1.2 s without exhibiting a clear trend. The saturation of τ_F values at mid and high calcium concentrations was consistent between this work and Tricoire (2006).

Effect of mutation on the slow reaction time constants

In contrast to their fast counterparts, the slow time constants τ_S differ between wild-type and mutant aequorin. As shown in Table 8-6 τ_S values can increase up to one

order of magnitude for the mutant protein. This is clearly seen in the cases of calcium and lead and might possibly apply to lanthanum. However, due to low sensitivity of mutant aequorin to lanthanum, the data set available for decay fitting was too limited (Figure 8-12 and Figure 8-13) and data were fitted at only two lanthanum concentrations ($\sim 10^{-3}$ and 10^{-2} M). At these lanthanum concentrations time constants for wild type and mutant were comparable.

The curves depicting τ_s follow a similar trend across different ions and type of protein (wild type or mutant) used. Specifically, all curves show a slow smooth reduction on the value of τ_s with increasing metal ion concentration, which is consistent with the notion that by increasing ion concentrations one should expect to see faster kinetics. This trend is also consistent with previously published work on wild-type aequorin and different variants with calcium.

Wild type: comparison between different ions

Contrasting the behaviour of wild-type aequorin in the presence of calcium and lanthanum where the branches of τ_s and τ_f tend to converge, in lead the slow and the fast branches move in parallel to each other (Figure 8-12, Table 8-6). The slow time constants of wild type plus lead are lower (faster kinetics) than those triggered by calcium and lanthanum, while the fast time constants are in the same range. However, the contribution of the fast branch to the total light (Σ_F/Σ_T) as well as the actual yields of light produced with lead were much less than their calcium and lanthanum counterparts during the time course of the experiment, which suggests reduced efficiency in light production. The actual light yield comparisons were derived from the raw data used for the luminescence curves (Supplementary Material) and summarised in the bar chart of overall yield comparisons (Section 8.3.3). Such high contribution of the slow reaction branch is not a desirable feature for a biosensor envisaged to capture fast metal ion fluctuations at real time.

Mutant aequorin versus different ions

In this case both the lanthanum and lead systems exhibit faster kinetics (both slow and fast branch) than the calcium system (Figure 8-13). This is useful in that it can promote the development of a biosensor which shows a specific fast response for these metals. However it should be noted that high concentrations of lead result in

inhibition (refer to concentration-response curves for lead in Figure 8-8, Figure 8-9 and Section 8.3.3).

Fast and slow rate contributions to bioluminescence

The most important and consistent observation regarding the contribution of the fast and slow branches to total light produced is that the contribution of the slow reaction is dominant in mutant aequorin, reaching approximately 99% for all three ions (Figure 8-12 and Figure 8-13). One way to read this is that the fast branch is essentially blocked by the mutations.

Extracting the fast and slow components of bioluminescence as well as integrating to calculate photon yield for each component is a noisy process. When the time constants tend to converge to values of the same order of magnitude, the pre-exponential components can vary significantly, as expressed by high error bars in some graphs. Computationally, the slow component is expected to produce more light since light production persists for a longer time interval, which is reflected in the results. The calculation of Σ_S/Σ_T and Σ_F/Σ_T takes into account both the time constants and the pre-exponential terms of the fittings (Section 8.2.6.2). Due to the nature of the calculation of Σ_S/Σ_T and Σ_F/Σ_T , ($\Sigma_S/\Sigma_T + \Sigma_F/\Sigma_T = 100\%$) the branches of these components are mirror images of each other on the plots. In terms of contributions to total light, it is not easy to compare with previously published work (Tricoire *et al.*, 2006) since the authors plotted Σ_S and Σ_F normalised by the maximum point in their curves and not by the total Σ_T for both curves.

Following the assumptions that the fast branch of the reaction is blocked by the mutations (while the slow path remains almost intact), it is reasonable to assume that the two branches operate independently.

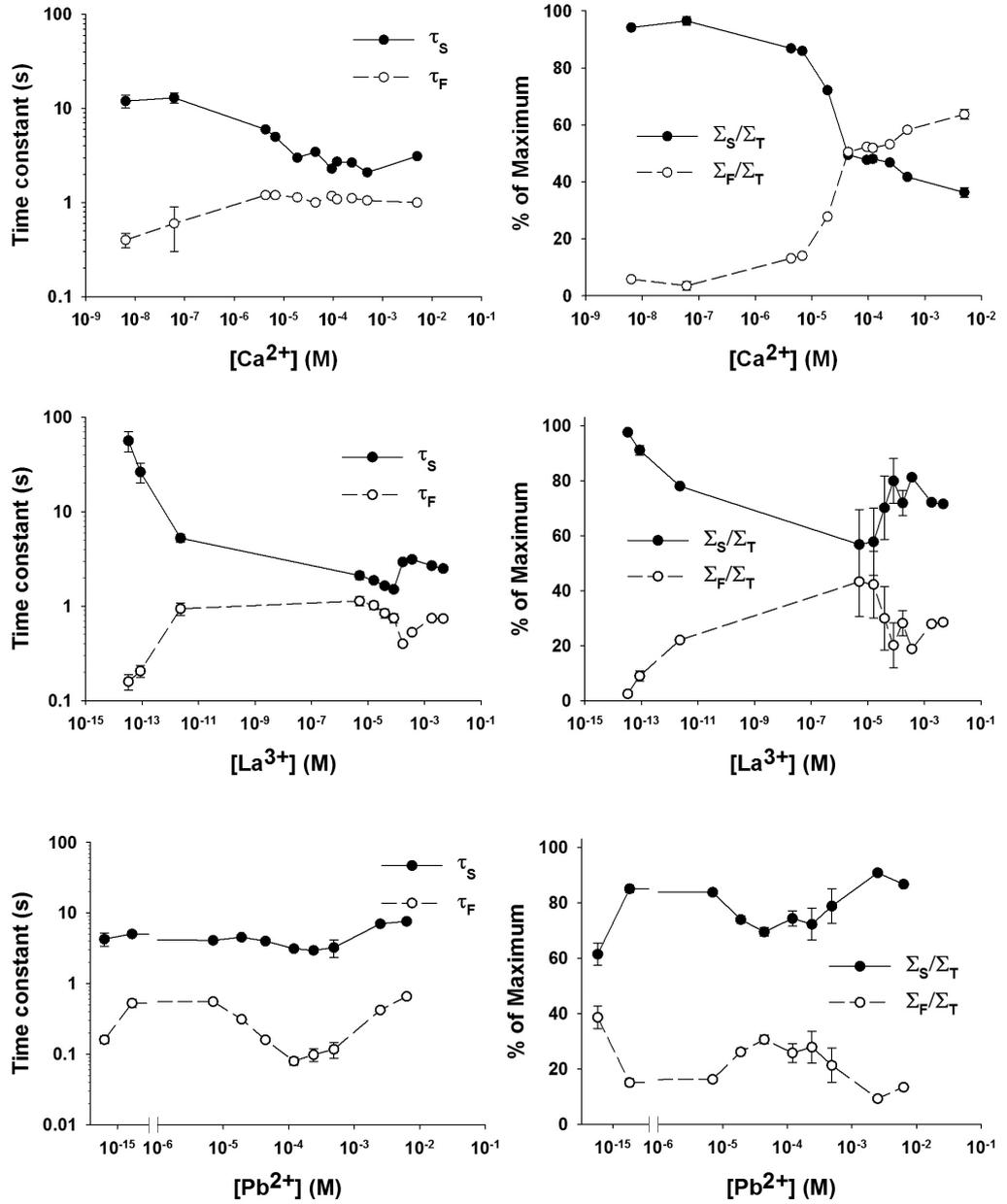


Figure 8-12 Decay kinetics of wild-type aequorin. Left: slow and fast decay time constants (τ_S and τ_F respectively) plotted versus ion concentration (calcium, lanthanum and lead). Right: Contribution of the slow (Σ_S/Σ_T) and fast (Σ_F/Σ_T) light integrals to total light versus ion concentration. Error bars are standard error of the mean.

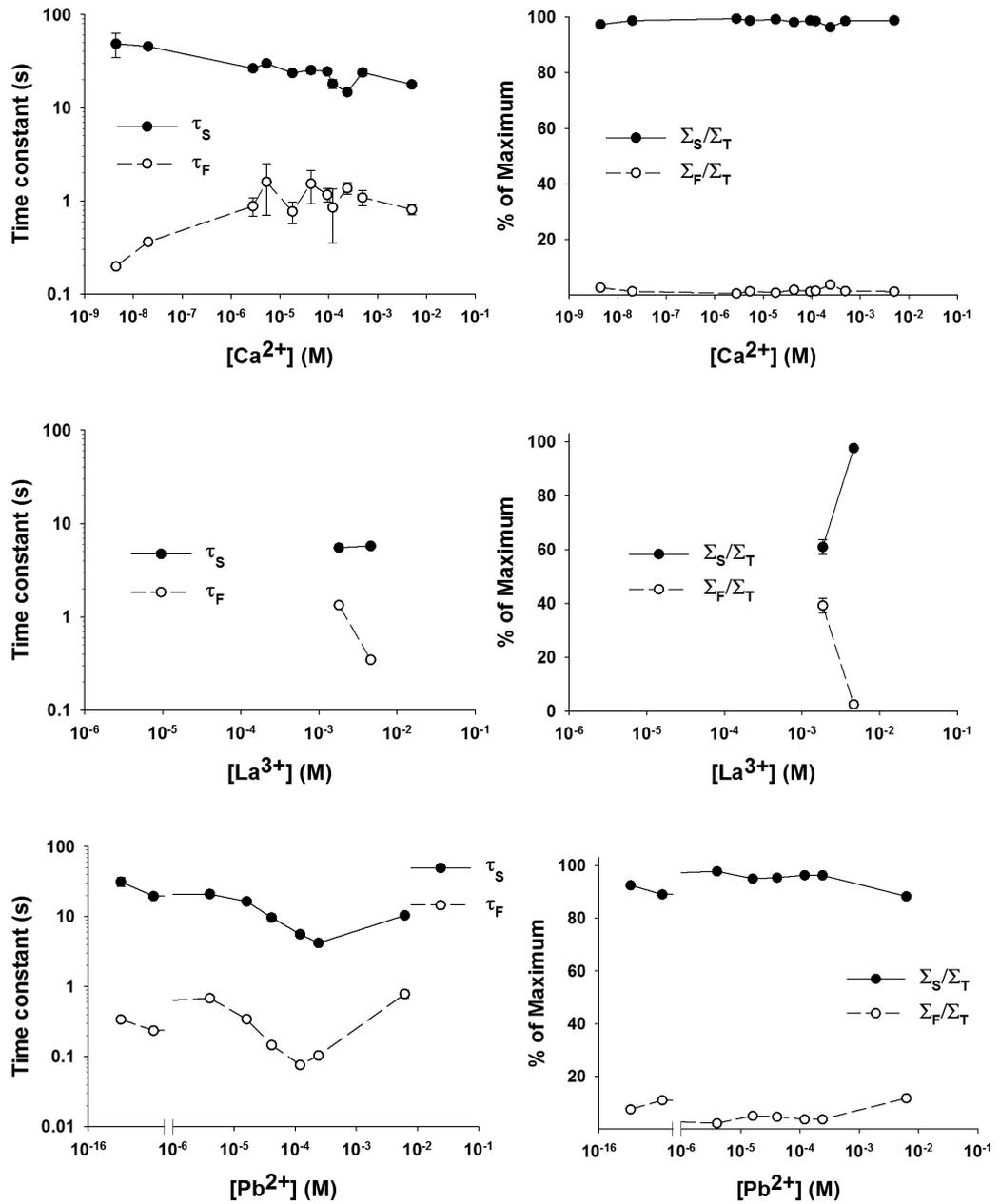


Figure 8-13 Decay kinetics of mutant Asn28Cys/Ser32His aequorin variant. Left: slow and fast decay time constants (τ_S and τ_F respectively) plotted against ion concentration (calcium, lanthanum and lead). Right: Contribution of the slow (Σ_S/Σ_T) and fast (Σ_F/Σ_T) light integrals to total light against ion concentration. Error bars are standard error of the mean.

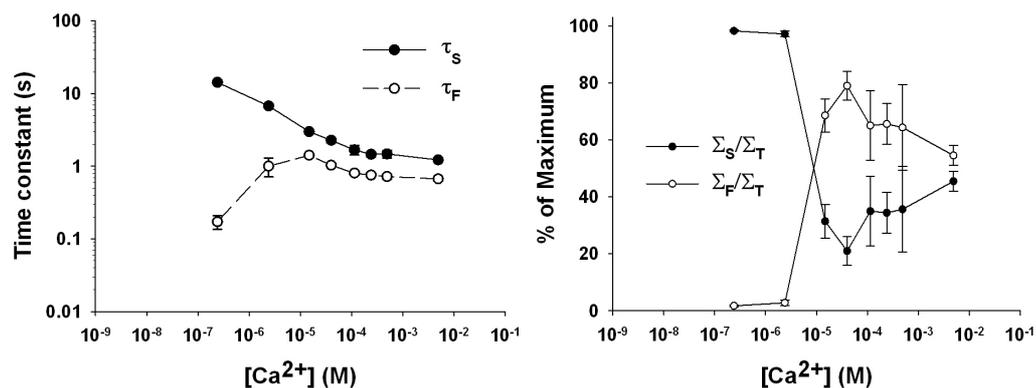


Figure 8-14 Decay kinetics of commercial aequorin AquaLite[®]. Left: slow and fast decay time constants (τ_S and τ_F respectively) plotted against calcium concentration. Right: Contribution of the slow (Σ_S/Σ_T) and fast (Σ_F/Σ_T) light integrals to total light against calcium concentration.

Table 8-5 Fast time constant (τ_F) ranges for flash-type curves

	AquaLite [®]	Wild-type aequorin	mutant aequorin
Calcium	0.17–1.3	0.4–1.2 s	0.19–1.6 s
Lanthanum	n.a.	0.15–1.13 s	0.3–1.3 s
Lead	n.a.	0.07–0.66 s	0.07–0.78 s

Table 8-6 Slow time constant (τ_S) ranges for flash-type curves

	AquaLite [®]	wild-type aequorin	mutant aequorin
Calcium	1.22–14.3	2.09–12 s	14.7–48.7 s
Lanthanum	n.a.	1.5–56.5 s	5.5–5.7 s *
Lead	n.a.	2.9–7.6 s	4.2–31.4 s

*limited dataset for mutant plus lanthanum (lanthanum concentrations 10^{-3} – 10^{-2} M)

8.3.4.2 Glow-type luminescence kinetics

In the previous section flash-type reactions were dissected in two exponential components for the light decay. The very rapid ascend of the luminescence curves was not analysed. In the case of glow-type luminescence triggered by the metals zinc, copper and cobalt, the double exponential decay model was not applicable. The ascending and decay part of luminescence would require a far more intricate model that is beyond the scope of this work.

This section looks at the overall kinetics of the reaction. A simple metric, $P_{15s/Max}$, was devised as an appropriate measure of the decay rate of the glow curves during the experiment duration of 15 s (Section 8.2.6.3). High values of $P_{15s/Max}$ indicate that at 15 s the bioluminescence reaction is almost complete and therefore $P_{15s/Max}$ can be used as an indirect measure of the speed of the bioluminescence reaction. Values for $P_{15s/Max}$ were plotted for each ion for wild type and mutant in Figure 8-15. Table 8-7 summarises the reaction progress metric $P_{15s/Max}$ for all the glow-type experiments.

Out of the three metals examined, zinc was found to trigger the slowest kinetics. The mutant plus zinc system was faster than the wild type plus zinc over the range of concentrations studied. The same applies to copper at all concentrations with the exception of the highest concentration of 10^{-2} M where $P_{15s/Max}$ for wild type and mutant converged.

In contrast to zinc and copper, the mutations did not affect the rate of bioluminescence for cobalt. This difference can be attributed to the nature of the mutations (substitutions of Asp and Ser by His and Cys respectively) which in theory is conducive to enhance binding of zinc and copper (Section 1.4, Section 5.6).

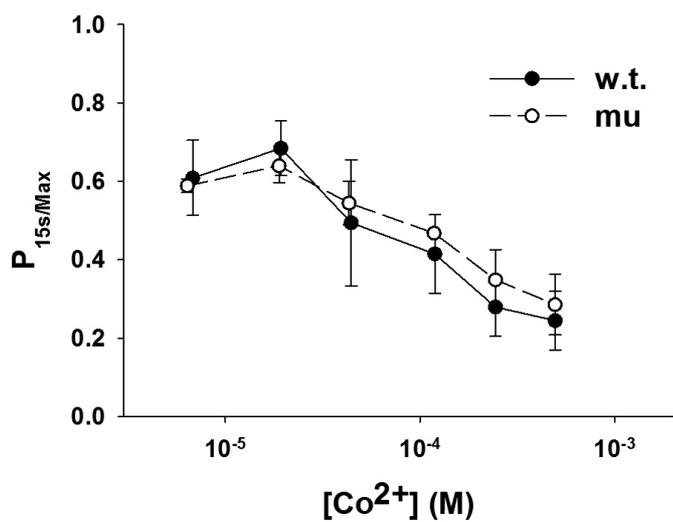
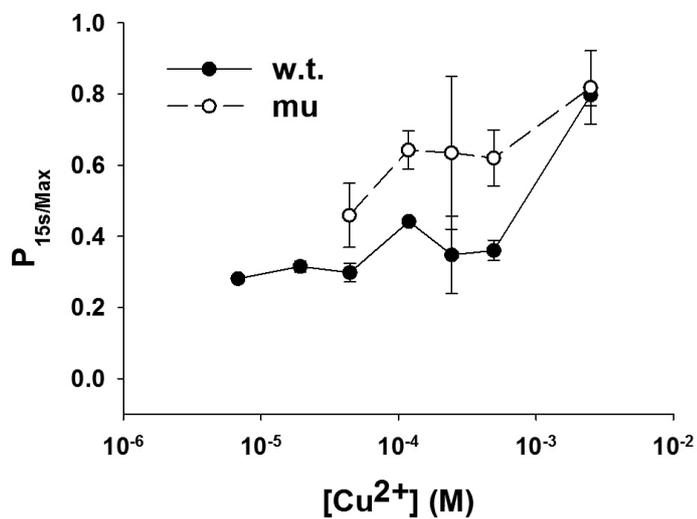
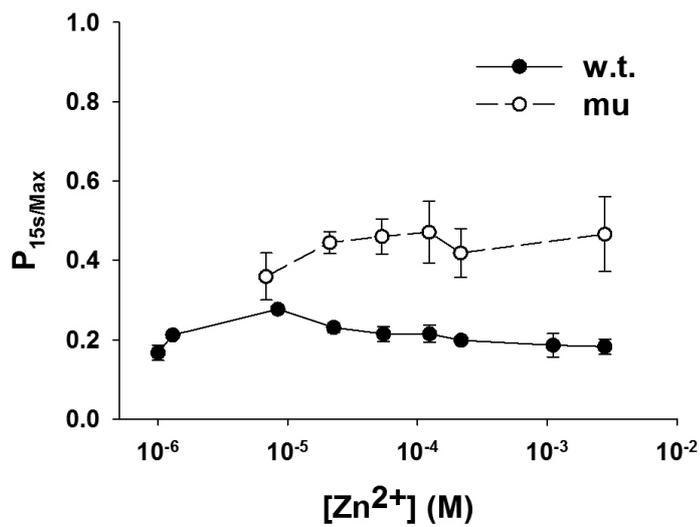


Figure 8-15 Metric of reaction progress $P_{15s/Max}$ for glow-type curves. Reaction progress of wild-type and mutant aequorin versus concentration of each of zinc, copper and cobalt.

Table 8-7 Ranges of $P_{15/Max}$ for glow-type curves. $P_{15/Max}$ is a dimensionless number.

Metal	Wild-type aequorin	Mutant aequorin
Zinc	0.17–0.28	0.36–0.47
Copper	0.28–0.82	0.46–0.82
Cobalt	0.24–0.68	0.26–0.59

8.3.5 Double ion experiments: wild type plus Zn^{2+} and Ca^{2+}

A potentially promising result of increased sensitivity of mutant aequorin towards zinc (Section 8.3.2) inspired a new experiment where aequorin was injected with a combination of zinc and calcium.

In the experiments described thus far, each of wild type or mutant was only mixed with one type of metal ion. Mutant Asn28Cys/Ser32His was designed to provide a more favourable environment for zinc or copper binding but the mutations were in only one out of the three calcium-binding loops. The remaining (native) EF-hands however retain their affinity for calcium and will still require calcium for their successful contribution to luminescence. In this light, no single ion could alone evoke the highest possible result, as inevitably one or more EF-hands would be lacking their preferred activator ion.

The premise of this experimental design was that the loop of the mutated EF-hand (EF-I) can be populated by zinc whilst the remaining intact calcium-binding loops of EF-III and EF-IV can be populated by their native activator, calcium. The double ion experiment setup is described in Section 8.2.5. In summary, aequorin was mixed with zinc at a range of concentrations. After a few seconds the mix was enhanced with sufficient amount of calcium to complete the luminescence reaction. This amount was determined from the concentration-response curves of wild-type and mutant plus calcium in Section 8.3.2.

The two metal ions were introduced at separate times in order to allow monitoring of the zinc effect before addition of calcium that was likely to occupy the binding sites (and exclude zinc from binding) from the beginning of the experiment. An

additional reason arose from results of previous experiments that showed that light yield produced by zinc binding was very low compared to calcium. Mixing aequorin and zinc first allowed recording of zinc-induced luminescence that would be masked by the intensity of calcium-activated luminescence.

If increased selectivity for zinc has been successfully introduced in EF-I of the mutant, then by providing both zinc and calcium EF-I would be primed with zinc and EF-III and EF-IV would be primed with calcium. Zinc and calcium in combination should then trigger a higher luminescence yield in the mutant than calcium alone, and certainly higher than zinc alone.

Both assumptions were found to be true. Figure 8-16 shows that in wild type, calcium alone triggers the same luminescence yield as calcium plus low concentrations of zinc ($\sim 10^{-14}$ M). At increasing concentrations of zinc the light yield decreases, possibly following a sigmoidal function with negative slope. These findings suggest that zinc does not contribute in the activation of any EF-hand at low concentrations and that in the presence of the abundant activator calcium, zinc cannot compete for the binding sites. As higher concentration of zinc is added, zinc competes with calcium for at least one binding site and at ever increasing concentrations it blocks calcium's access to the calcium-binding loops. Previous work in Section 8.3.2, Figure 8-8, showed that the concentration-response curves of zinc are bell-shaped with a maximum at 10^{-6} M. A generic detrimental effect on aequorin at higher concentrations was assumed.

In the mutant, the shape of the concentration-response curve of the double ion experiment is different (Figure 8-17). Calcium alone is not performing as highly as calcium plus a small addition of zinc. This observation may suggest that traces of zinc incurred a positive effect in the mutant that calcium alone cannot offer. The reason for this change is likely to be the increase of selectivity of EF-I towards zinc and decrease towards calcium in mutant Asn28Cys/Ser32His. These statements make sense considering the types of mutations introduced in EF-I and discussed in Chapters 5 and 6.

The shape of light versus time plot (Figure 8-16 A and Figure 8-17 A) though suggests that despite the positive shift toward zinc selectivity, the mechanism of bioluminescence was overall impaired, as the yield was significantly reduced and the characteristic sharp peak of flash luminescence of wild type was blunted in the mutant experiment when calcium was added. A possible explanation is that the type of

mutation introduced was disruptive to the propagation of signal between ion binding and emission of light.

In many ways this experiment is more practical and more meaningful than the addition of single ions to aequorin, due to the importance of the cooperative function of the EF-hands in aequorin. Future screening for incremental shifts of selectivity toward new ions could use a mix of calcium and the new ion, rather than the new ion alone.

The main finding of this section was that in wild type 100% of maximum activity was achieved by using calcium alone whereas in the mutant 100% of maximum activity required the addition of a small amount of zinc ions. The conclusion from this section is that the mutations engineered in Asn28Cys/Ser32His resulted in successful localised binding of zinc in EF-I with only partial activation of aequorin bioluminescence.

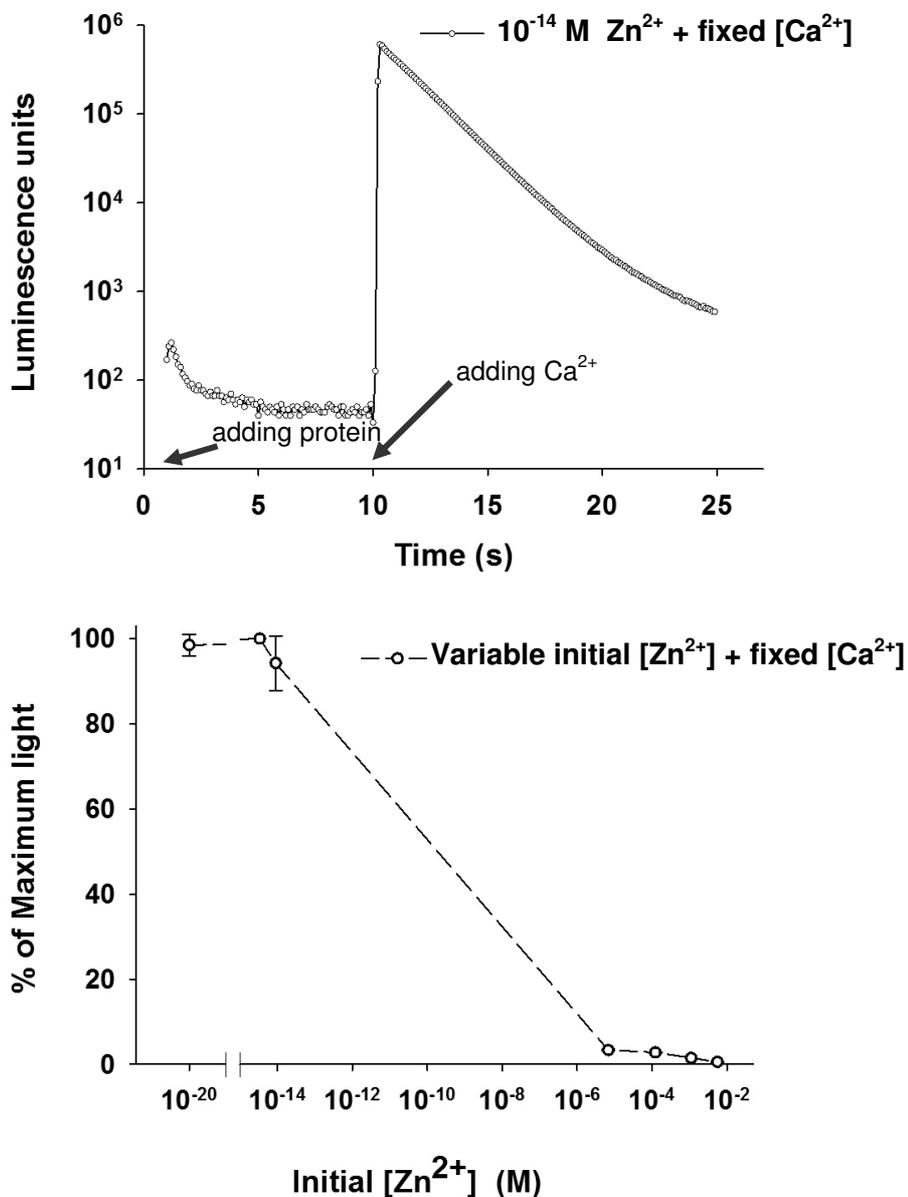


Figure 8-16 Wild type triggered by zinc and calcium. 40 μL of zinc were injected into 40 μL of purified wild type aequorin solution at 0.3 s at a wide range of concentrations, followed by injection of 40 μL of 50 μM calcium at 10 s (16.7 μM final calcium concentration in the well). Top: Typical trace of luminescence versus time from the resulting data set. The y-axis was set to logarithmic. In a linear scale the effect of zinc would be barely visible due to the very low light yield produced compared to the light produced upon addition of calcium in wild-type aequorin. Bottom: Concentration-response curve comprised from the Peak value of all the luminescence curves of the data set. The same trend was observed when plotting the Sum of the light readings for the duration of the experiment. A similar curve was produced when injecting 10 mM calcium (fixed) at 10 s (3.33 mM final calcium concentration in the well).

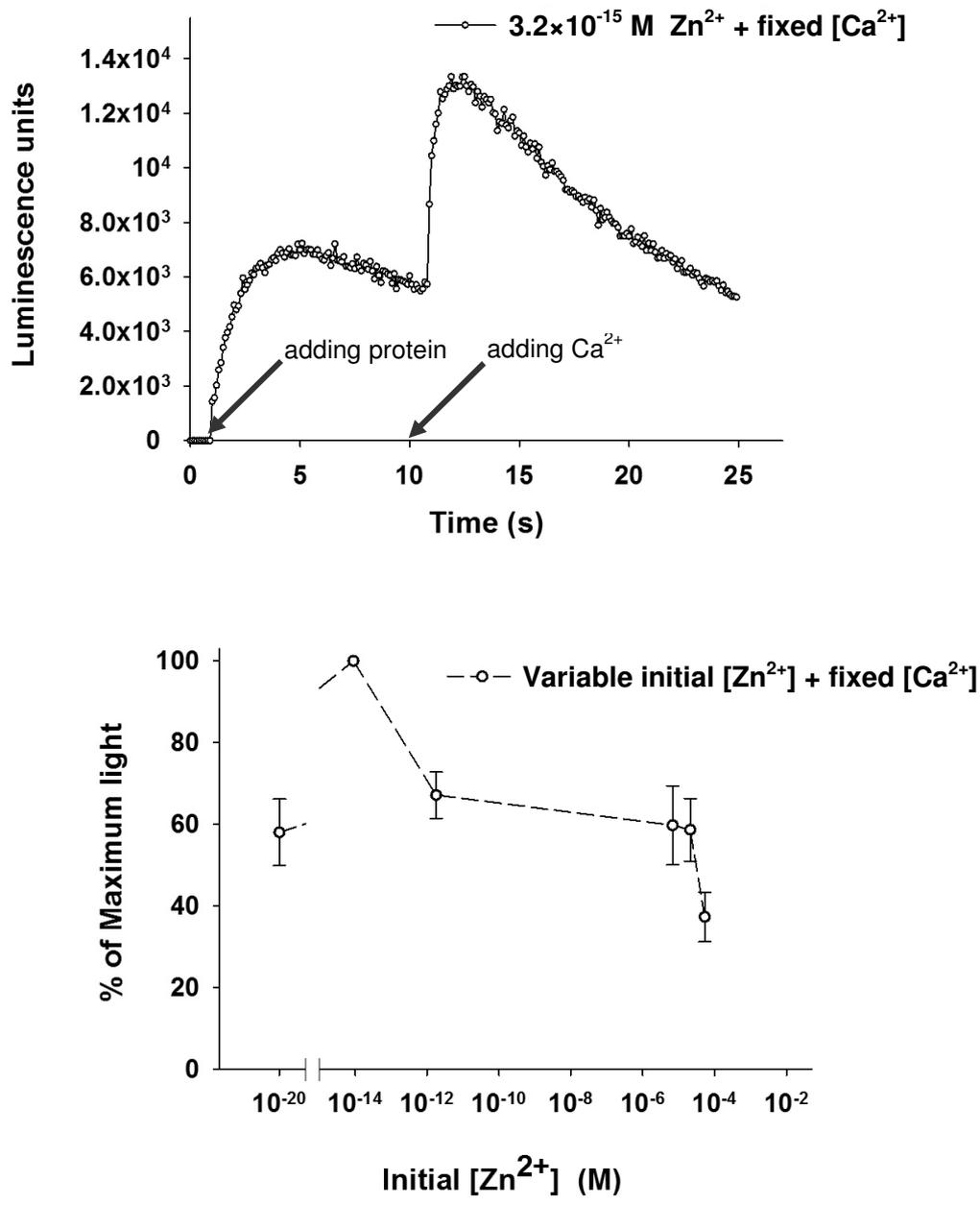


Figure 8-17 Mutant Asn28Cys/Ser32His aequorin triggered by zinc and calcium. 40 μL of zinc were injected into 40 μL of purified wild type aequorin solution at 0.3 s at a wide range of concentrations, followed by injection of 40 μL of 50 μM calcium at 10 s (16.7 μM final calcium concentration in the well). Top: Typical trace of luminescence versus time from the resulting data set. In contrast with Figure 8-16 the y-axis was set to linear scale, as the effect of zinc was comparable to that of calcium in mutant aequorin. Bottom: Concentration-response curve comprised from the Peak value of all the luminescence curves of the data set. The same trend was observed when plotting the Sum of the light readings for the duration of the experiment. A similar curve was produced when injecting 10 mM calcium (fixed) at 10 s (3.33 mM final calcium concentration in the well).

8.3.6 Correlation of ion size and shape of luminescence curve

In this section experimental findings of this chapter - light yield, shape of luminescence, shape of concentration-response curves, sensitivity (EC_{50}) - are discussed in relation with structural information drawn from Chapters 1 and 5.

In Figure 8-18 the shape of the luminescence curves of aequorin was plotted against the coordination numbers of ions within metal-ligand complexes and its potential for interatomic overlap between the metal ion and the coordinating ligands in the loop of EF-I of aequorin. The metric of interatomic overlap ($D-r-R$) was introduced in Section 5.4.2 as the interatomic distance (D) between calcium and each of the oxygen ligands in the pentagonal bipyramidal coordination (PDB ID: 1SL8) subtracted by the atomic radii of each pair of atoms, where r is the radius of the metal ion and R is the radius of the coordinating oxygen atom. The overlap calculations were obtained by “fixing” each ion in the position of calcium in the crystallographic structure of calcium-bound apoaequorin (PDB ID: 1SL8). This positioning of the different ions is a simplistic assumption in order to study their potential effect within the local environment of the EF-I ion binding loop. With the reasonable assumption that the seven-coordinate binding of calcium is optimum for triggering aequorin luminescence, ions that are able to bind in a similar manner should also be effective.

In Figure 8-18 $MIN < 0$ (green bar) means that there is interatomic overlap between at least one metal-ligand ($M-L$) pair. $MAX < 0$ (blue bar) indicates that interatomic overlap occurs between all of the metal and each of the ligands in the crystallographic structure. SUM (red bar) is the sum of the overlap between the metal ion and every coordinating ligand within the loop of EF-I.

In calcium, lanthanum, cadmium and lead there is overlap in every metal-ligand combination, as denoted by both minimum and maximum values being negative. Cobalt would appear to be an exception as it produces higher yields than cadmium. One would expect all ions with $D-r-R > 0$ to achieve lower yields than ions with $D-r-R < 0$. However cobalt triggers more light than cadmium and triggers luminescence curves that are overall sharper than those of zinc and copper. Despite its small size, cobalt stands out amongst equally small metals and indeed yields higher luminescence than the larger cadmium.

In Figure 8-19 the metric of sensitivity EC_{50} of wild type to each metal ion was plotted against the sharpness of the luminescence curves, the ionic size and the metric of interatomic overlap D-r-R.

Sharpness of luminescence reflected in the flash-type curves is associated with efficiency of the reaction. However the cobalt-triggered glow-type luminescence was the sharpest (and therefore fastest) compared to that of zinc and copper (Section 8.3.1).

It was not possible to derive specific EC_{50} values for every combination of aequorin type and metal ion and a range was derived instead (Table 8-3). A causal correlation between sensitivity and ionic size was not derived: native activator calcium has a lower sensitivity (higher values of EC_{50}) than its most successful replacement lanthanum but similar to that of cadmium which is a lesser performer than lanthanum. The range of potential EC_{50} values for the poorer performers zinc, copper and cobalt makes aequorin more sensitive to them than calcium. For a bioreporter, the sensitivity of the molecule should lie within a relevant concentration range, i.e. where fluctuations of the measured target molecule (metal ions in this case) occur and are of interest. If studying a new metal ion that is present in very low concentrations (e.g. lower than those of calcium), then the envisaged biosensor must be able to detect these concentrations and their fluctuations if applicable, hence the EC_{50} value must be lower than that of calcium.

What differentiates cobalt from zinc and copper is its ability to bind in octahedral coordination geometry, as is calcium, whereas zinc is mostly found in tetrahedral coordination and copper in square pyramidal and square planar coordinations in metalloproteins (Section 1.4, Table 1-7). In coordination geometries that include a higher number of participating atoms, higher radii allow atoms to interact from a distance that is still not sterically prohibitive. In aequorin calcium coordinates with the oxygen of a water molecule additionally to the six coordinating atoms of the EF-hand loop, thus making its coordination geometry pentagonal bipyramidal. Geometry of binding is of great importance in achieving an efficient bioluminescence reaction. The higher coordination number a new ion can bind in, the easier it will fit the native binding sites of aequorin.

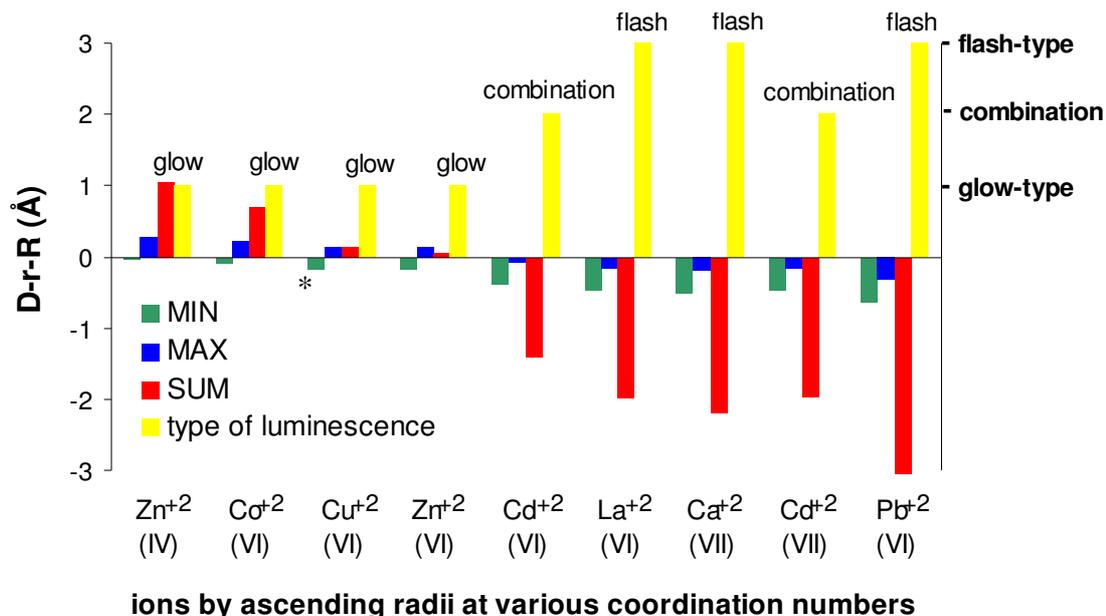


Figure 8-18 Is radii overlap associated with sharpness of luminescence curves? For some ions, like Zn²⁺ and Cd²⁺, more than one possible radius is plotted due to strong preference for coordinations with lower coordination number (CN) in metal binding sites. The asterisk at the glow classification for cobalt luminescence denotes the relative sharpness of the cobalt curves compared to the other glow-type curves of zinc and copper. Coordination numbers that were considered are shown in parentheses.

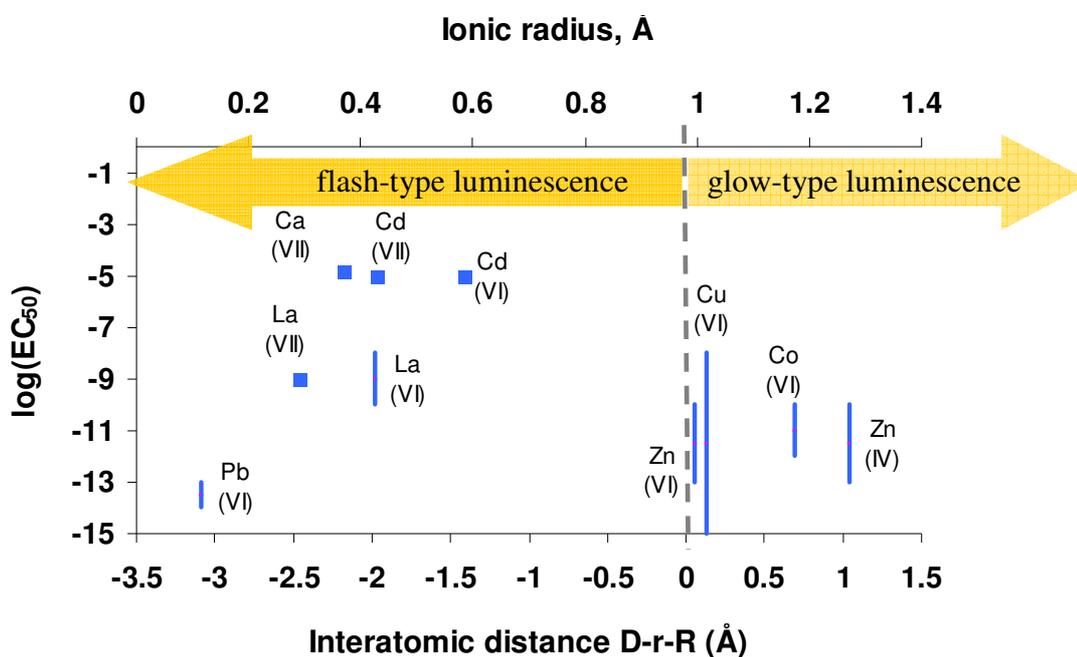


Figure 8-19 Type of luminescence, metal ion sensitivity and interatomic distance in the EF-I loop of wild-type aequorin. Long blue bars indicate a range of possible logEC₅₀, as the actual value could not be estimated from the respective concentration-response graphs. LogEC₅₀ values for wild type plus metal ions were taken from Table 8-3. As in Figure 8-18, more than one possible ionic radius (and hence coordination numbers) was considered. Possible coordination numbers (CN) are shown in parenthesis.

8.4 Conclusion

This chapter aims to describe the effect of the seven metal ions – calcium, lanthanum, cadmium, zinc, copper, cobalt and lead - on aequorin and its mutant variant Asn28Cys/Ser32His. It does not aspire to construct a detailed mechanistic description of the processes involved. Different experiments and data analysis would be required for this purpose.

The experimental setup measured output of luminescence versus time (Section 8.2.8). The light response itself is the result of a complex set of events starting from ion binding and resulting in subtle conformational changes that lead to the destabilisation of the hydroperoxy-coelenterazine in the centre of the protein and concomitant production of light (Section 5.1 and 5.6.2). Interpretation regarding the specific effect of the mutations or different metals on the function of the molecule is difficult, especially in absence of structural and binding data (Section 8.2.8). Therefore general conclusions were drawn in this work.

Ions as partial activators

All ions other than the native activator calcium, act as partial activators for aequorin, each to a different extent. Significant loss of luminescence was observed in mutant Asn28Cys/Ser32His with all metal and in the wild type with metals other than calcium and lanthanum. The levels of activation range from <1% to 70% of native activity (Section 8.3.3). Lanthanum was a successful substitute for calcium, with relatively high yields of luminescence and concentration-response curves being the same shape as that of calcium. Low yields triggered by other ions suggest that cooperativity links are broken and that some of the ions act as inhibitors at higher concentrations (Section 8.3.2).

In this chapter comparisons were made at each ion's optimal concentration in order to assess the extent of activation/inactivation under the conditions tested. An alternative way to conduct comparisons would be to focus at specific ranges of metal ion concentrations, as only certain ranges of metal concentration are relevant to different biological systems or environmental samples. This type of comparison would be essential when considering the suitability of a metal ion biosensor in a relevant environment.

Size, coordination of metal binding and efficiency of reaction

The traces of luminescence versus time fell under two major categories (Table 8-2). Traces of the first category – which is also the case of the “classic” aequorin bioluminescence with native activator calcium – are characterised by a sharp increase in luminescence producing a distinct peak followed by a fast decay. This type of luminescence was termed “flash-type” luminescence. The second major category is characterised by a much smoother ascent, substitution of the distinct peak with a wide maximum and followed by a slow descent. This behaviour was termed “glow-type” luminescence.

Flash-type luminescence is associated with higher yields as manifestation of an efficient overall process of ion binding and production of light. Radii overlap is necessary for flash-type and higher yields. Analysis took into account the preferred coordination geometries of these ions, as well as their ionic size. The larger ions calcium, lanthanum, cadmium and lead produce flash-type luminescence, whereas the smaller ions trigger glow-type luminescence. This reflects the size selectivity of the EF-hand. Size selectivity would be a major consideration in planning future mutational strategies.

Both the requirement of the protein environment for effective binding and the propensity of metals to bind at specific geometries will dictate the performance of aequorin variants with new ions. It will be very challenging to completely redesign the EF-hand loop whilst keeping the intrinsic network of interactions within the protein intact. This network is necessary for the appropriate conformational changes to be propagated to the centre of the molecule for effective bioluminescence production.

In terms of the effect of ions on EC_{50} and yield, it is not possible to uncouple the causal effect of binding in the EF-hands or at other sites of the protein causing inhibition.

A mutant with positive response to zinc

Experimental results showed that EF-I was successfully altered to develop a higher affinity towards zinc as well as to contribute to a mutant-specific improvement in luminescence production compared to wild type. In the double ion experiment (Section 8.3.5) wild-type and mutant aequorin were exposed to a combination of zinc and

calcium; this approach allowed zinc the opportunity to activate the mutated EF-hand (EF-I) and calcium to activate the remaining native hands.

The mutant achieved higher activity with a specific combination of zinc and calcium than with either zinc alone or calcium alone. However it was acknowledged that this was still low level activity and that although the mutation caused a successful increase of selectivity towards zinc, it also caused impairment of the bioluminescence reaction mechanism.

Sensitivity to metal ions

In addition to studying the time course of the luminescence reaction, the response of aequorin to several concentrations of each metal ion was studied. This process resulted in metrics for the sensitivity of aequorin (wild-type and mutant) to each ion. Concentration-response curves were drawn to study the effect of a wide range of concentrations of each ion on the proteins.

In terms of the shape of concentration-response curves, two main groups emerge: sigmoidal and bell-shaped. From the dataset in this work, an overall association emerges between shape of luminescence and shape of concentration-response curves. Specifically, the datasets that agree with the flash-type luminescence produce sigmoidal concentration-response curves and those that produce glow-type luminescence produce bell-shaped concentration-response curves. At first glance, lead and cobalt appeared to be exceptions. Lead triggers flash-type luminescence but its concentration-response behaviour is bell-shaped both for wild-type and mutant aequorin. However this could be due to non-specific enzymatic inhibition (Section 8.3.2), lack of buffering in the lead nitrate solutions (Section 8.2.3) and consequent loss of aequorin stability at low pH conditions (Section 1.3.3). Cobalt produced sigmoidal concentration-response curve. The luminescence was classified as glow-type but the luminescence curves triggered by cobalt were sharper than those triggered by zinc and copper.

Despite the lack of structural data, a reasonable assumption is that new ions bind to the EF-hands but fail to exert the full scale effect of the native activator calcium.

Kinetics of the reactions

For wild type, the results for the flash-type luminescence were consistent with previous work on the independence of the rate of ascent eluding to first order kinetics

and the dependence of descend (decay) of the luminescence curves on calcium concentrations (Hastings *et al.*, 1969).

For the flash luminescence reactions a previously published double exponential decay model (Sections 8.1.5.2 and 8.3.4.1) was applied. The model proposed the existence of two branches of luminescence – one fast and one slow. Time constants and estimates of contribution to total light were derived for both the postulated slow and fast branch of the reaction. The results were consistent with published literature in terms of the values of the time constants for wild type and calcium.

The combinations of (1) wild-type with metals other than calcium or lanthanum and (2) mutant aequorin with every metal tested, result in the slow branch of the reaction becoming even slower (increase of time constant = slower reaction) and in increase of the slow branch's contribution to the total light produced. Conversely, the fast branch of the reaction contributed less than 1% of total light in these cases.

The glow-type dataset (cobalt, zinc, copper) was analysed using a simple metric of reaction progression for the duration of the experiments (Section 8.2.6.3 and 8.3.4.2). It was found that the kinetics of zinc and copper were faster in the mutant than the wild type, whereas in cobalt no differences were observed. This suggests that from a kinetics point of view, the mutant responded to the presence of zinc and copper more positively than wild-type aequorin did. Cadmium did not conform with the standard models (flash-type or glow-type) in that its flash-type curve was followed by a distinct glow-type shape and as a result its kinetic behaviour was left for future analysis.

The study of mutant Asn28Cys/Ser32His in comparison to wild-type aequorin yielded some step improvements on certain aspects and allowed insight on some basic requirements and consequences of redesigning ion selectivity in the EF-hands of aequorin. Additionally, Asn28Cys/Ser32His included mutations in only one EF-hand out of the three of aequorin. Further work would be needed to approach a mutant with shifted selectivity toward new ions that will retain high luminescence yields.

9 Epilogue

9.1 Overall summary of this project

Aequorin is a bioluminescent photoprotein that emits a flash of blue light upon binding of calcium ions. Due to this property, along with its non-invasive nature and its response to calcium in a biologically relevant dynamic range, aequorin has been used as a real-time intracellular bioreporter for more than forty years. The active protein comprises of a polypeptide chain (apoaequorin) that non-covalently stabilises a chromophore (coelenterazine) in an intermediate oxidative state (hydroperoxy-coelenterazine) within the polypeptide's hydrophobic core. Coelenterazine is the source of bioluminescence; the decarboxylation of hydroperoxy-coelenterazine leads to release of energy. Aequorin has three highly conserved calcium-binding sites (EF-hand loops) that are exposed to the solvent and bind calcium ions with positive cooperativity. The binding of calcium results in local and global conformational changes that ultimately lead to the decarboxylation of the hydroperoxy-coelenterazine.

In the early days of the protein's discovery it was noted that aequorin is potentially responsive to other cations; however, aequorin was highly selective toward calcium at biologically relevant concentrations and produced very low light yield in the presence of most other ions. Members of the lanthanide group (e.g. lanthanum, terbium) constituted an exception, as they were found to be good substitutes for calcium in aequorin bioluminescence. This potential promiscuity of aequorin towards other metal ions was not pursued further; researchers might have been discouraged by the significantly low light yields originally reported.

In this work, the molecule of aequorin was used as a template in a strategy of rational mutations with the goal of creating a novel metal ion bioreporter. The envisaged molecule would have high affinity towards new metals and reduced/abolished affinity for calcium. Mutations were engineered with a bias for zinc and copper binding, as these metals are particularly interesting in the study of neurological diseases.

The structure and function of the protein was discussed in detail through compilation of theoretical analysis and published literature (Chapters 1 and 5). Previous

mutational work on aequorin and other calcium-binding proteins was used as a guide to make informed choices on a mutational strategy (Chapter 5). A library of forty eight aequorin mutants was designed in order to be tested against seven metal ions as potential activators of bioluminescence: calcium, lanthanum, zinc, copper, cadmium, cobalt and lead. The mutational approach consisted of:

(1) replacing the calcium-coordinating residues of aequorin with amino acids that are known to bind zinc and copper ions (histidine and cysteine). These residues are exotic to the specific positions in the EF-hands of aequorin, all other photoproteins or any other member of the EF-hand superfamily. Substitutions were made either one by one or in pairs of two per aequorin molecule. Mutations of this type had not been reported in any previous work.

(2) removing functional groups from the calcium-coordinating positions of the EF-hand loops by means of alanine replacements. This would indicate whether the performance of mutants is due to the loss of the calcium-coordinating side chain oxygens (carboxyl- or hydroxyl-) or due to the introduction of sulfur atoms and imidazole nitrogens (from cysteine and histidine substitutions respectively).

(3) replacing the highly conserved central glycine of the calcium-binding loops with arginine. The central glycine does not coordinate calcium but is present in all calcium-binding EF-hands and in the entire EF-hand superfamily. It is assumed to act as a “hinge”, allowing the coordinating residues to reach for the calcium ion. Arginine was thought to reduce the loop’s flexibility and hence impair binding of calcium and cooperative action between the EF-hands.

Each of the three binding sites carries six coordinating residues. It was impossible to exhaust all possibilities of the desired substitutions and their combinations within the duration of this project. Most of the mutational work was prioritised for the loop of EF-I, followed by EF-III, with the least number of mutations designed for the loop of EF-IV. This “ranking” was based on published work on the potential importance of the three EF-hands in aequorin activity. A random or random-rational approach was reserved for a later stage, once the study of specific mutations would be conducted.

The gene encoding for apoaequorin was cloned into an expression system (Chapter 3) in order to facilitate expression and screening of the mutant library. The expression vector would incorporate the following features: (1) promoter site and ribosome-binding site (RBS) for transcription and translation; (2) signal peptide

sequence for localisation of the expressed apoaequorin in the periplasmic space of the cell, e.g. *pelB* leader sequence; (3) 6×His-tag encoding sequence for the option of metal affinity purification; (4) unique restriction sites flanking the apoaequorin gene.

Following the cloning step a microplate based high-throughput screening (HTS) process was developed to include the following steps (Chapter 4): (1) microwell cultures (inoculation and induction); (2) charging of active protein by incubation with coelenterazine; (3) copies of a 96-well microplate library; (4) high-throughput screening luminescence activity assay

A library of forty eight rational aequorin variants was created using site-directed mutagenesis. The library was screened using the HTS method developed and the results were analysed based on two metrics: (1) metric of the shift of ion selectivity (D) and (2) metric of % activity compared to wild type. The results were analysed in relation to the type of mutation, its position in the protein (EF-hand and position within the EF-hand) and type of metal ion tested.

A number of mutants stood out for their shift of selectivity towards new ions. Due to time constraints only one candidate (Asn28Cys/Ser32His) was selected for further study. This mutant carried two substitutions in the loop of EF-I and in the HTS activity assay exhibited high shift of preference toward zinc.

In order to study the bioluminescence reaction without the contaminants present in the crude mix (diluted cell culture) of the HTS library, wild-type apoaequorin and Asn28Cys/Ser32His were produced and purified at lab scale (Chapter 7).

The purified wild type and mutant were tested against a wide range of concentrations for each of the seven metal ions (Chapter 8). The outputs of the analysis included: (1) shape of luminescence versus time traces (flash- or glow-type); (2) determination of ion effect (activation or inhibition) through concentration-response curves; (3) sensitivity of the molecule to each ion (determination of EC₅₀); (4) kinetics of the reactions in both mutant and wild type triggered by the different metals; (5) comparison of all ions on their ability to activate wild-type and mutant aequorin and (6) discussion of the above findings in association with the size of metal ions and structural characteristics of the protein.

Double ion experiments were also performed; in these experiments the protein sample was injected with a combination of zinc and calcium ions. In principle, if affinity for zinc increased, this experiment would allow zinc to populate – and

potentially “activate” – the mutated EF-hand, whilst calcium ions would populate the remaining native EF-hands.

9.2 Overall conclusions

9.2.1 Cloning of apoaequorin gene

Two expression vectors were successfully created: pETAQ is a pET26b derivative and TOPAQ is a pCR-Blunt II-TOPO[®] derivative, both containing the apoaequorin cDNA, the *pelB* signal peptide sequence, unique restriction sites flanking the gene and 6xHis-tag fusion downstream of the T7 promoter. The pET cloning approach used predominantly restriction digest and ligation steps whereas for the TOPO[®] route the desired features were added by sequential PCR amplifications. The TOPO[®] route was pursued as a mitigation strategy to the pET route. As expected, the pET based vector was the higher expresser (verified in Chapter 4) and was used for all subsequent work.

9.2.2 High-throughput screening process

A practical, automated microplate-based process was successfully developed and optimised to cover every stage of the library creation and screening starting from colonies of transformants: (1) cell culture, inoculation and induction; (2) charging of active protein variants in microplates; (3) replicate copies of the library and (4) bioluminescence activity screen. The method was originally developed using wild-type aequorin and can be used with large libraries (using automated colony picking) or smaller rational mutation libraries (using manual picking). The method utilised cell culture suspension (not purified protein) which still allowed for high activity signals and zero background noise. It served as an initial “crude” screen of the entire library of mutants.

9.2.3 Mutant library screening

In the majority of mutants studied, shift of selectivity towards new ions (and away from calcium) increases at the cost of activity (Chapter 6). The vast majority of mutations lead to significant loss of activity (50–100% loss) compared to wild type. This observation was common for all ions tested. In contrast to metric of retained activity (metric C) which is fairly uniform across the library, shift of selectivity (metric D) is markedly different in the EF-hands studied. Some EF-I mutations exhibit high values of D whereas every mutant in EF-III exhibited very low values of the same metric. Aequorin was able to function with only one intact EF-hand, at low but detectable levels of activity.

EF-I may be the driver of aequorin selectivity. EF-III does not seem to be endowed with ion selectivity properties in aequorin. Indeed, results from impaired EF-hand mutants (central Gly→Arg) suggest that loop flexibility is least required in EF-III. The same set of results suggests that EF-IV has significant potential as determinant of ion selectivity. These experimental findings agree with previous structural observations (Chapter 5) that EF-III is already preformed to bind calcium and the helices attached to EF-I and EF-IV are potentially very important in the stability/destabilisation of hydroperoxy-coelenterazine through hydrogen bond interactions.

Loss of even a single conserved calcium-coordinating side chain resulted in dramatic loss of activity. This became evident from the alanine substitutions as well as from the histidine and cysteine substitutions. The amino acids marking the entry and exit of the EF-hand loops (100% conserved) were the least tolerant to mutations. Within EF-I, the positions most important for modulation of ion selectivity are 5 and 12. As expected, non-coordinating or coordinating but non-conserved residues were more tolerant to mutations; such residues could be mutated to fine-tune ion selectivity with little cost on activity levels.

The mutants that stood out for further work carried histidine or cysteine substitutions in position 5 of the EF-I loop. Asn28Cys/Ser32His exhibited higher selectivity for zinc and lead, but not copper; Asn28His/Glu35Cys exhibited higher selectivity for lead; Asn28Cys/Glu35His exhibited higher selectivity for copper and lead. One variant (Ala123Cys) exhibited improved lanthanum activity compared to calcium; however an EF-hand-like peptide with improved lanthanum activity had

already been reported and was not considered to be a novel outcome to pursue in this work.

9.2.4 Purification of aequorin

Wild-type apoaequorin and its variant Asn28Cys/Ser32His were purified from shake flask *E. coli* cultures using mild extraction conditions: one-step periplasmic cell extraction and incubation with coelenterazine to form active protein, followed by size exclusion chromatography. This approach yielded 0.61 mg of active wild type and 0.24 mg of active mutant Asn28Cys/Ser32His, corresponding to purified 2.6 mg of wild type and 1.04 mg of mutant aequorin per 100 mL of *E. coli* culture. The decreased yield of mutant monomer compared to wild type is attributed to its solvent-exposed cysteine, which likely caused higher order aggregates via intermolecular disulfide bridges. The size exclusion chromatogram of the mutant indicates a larger elution peak at higher molecular weights compared to that of wild type.

9.2.5 Study of purified wild-type aequorin and mutant Asn28Cys/Ser32His

Significant loss of luminescence was observed in all experimental conditions – combinations of wild-type and mutant with all metal ions – compared to the activity of wild type aequorin with calcium. All ions other than calcium act as partial activators for wild-type aequorin, each to a different extent. Despite the lack of structural data, a reasonable assumption is that new ions bind to the EF-hands but fail to exert the full scale effect of the “native” activator. Confirming previous reports on the wild type, lanthanum was a successful substitute for calcium, producing relatively high yields of luminescence (~70% compared to the peak of luminescence triggered by calcium) and a sigmoidal concentration-response curve, suggesting cooperativity in ion binding. The levels of activation by the other ions were significantly lower (ranging from <1% up to 40% of native activity) and in some cases produced bell-shaped concentration response curves. This suggests that cooperativity links are broken and that some of the ions act as inhibitors at higher concentrations.

The most significant outcome was that EF-I was successfully altered to contribute to a mutant-specific improvement: Asn28Cys/Ser32His was able to utilise zinc for its bioluminescence reaction when wild type was not. A combination of zinc and calcium produced more light in the mutant variant than calcium alone did. This was not the case in wild type. On a theoretical basis it is reasonable to claim that the mutations in EF-I increased affinity for zinc and decreased affinity for calcium but the experimental results also proved that this affinity shift translates into protein function.

A clear association emerged between the size of the metal ions and the efficiency of the bioluminescence reaction as expressed by the yield, the luminescence curves and the concentration-response curves. The features of bioluminescence produced by wild type and calcium were set as benchmark for the most efficient bioluminescence function possible. The traces of luminescence versus time fell under two major categories: (1) flash-type luminescence, including the “classic” aequorin bioluminescence with native activator calcium and known substitute lanthanum and (2) glow-type luminescence.

The larger ions can interact electrostatically with protein atoms that are located at further positions of the EF-hand loop. They are able to bind in coordination geometries of higher numbers, for example in octahedral geometry, which is the closest to the pentagonal bipyramidal binding of calcium in aequorin. Cobalt – a small ion with similar size to zinc and copper – triggers sharper luminescence traces than zinc and copper and produces higher yields than those triggered by cadmium – a much bigger ion. In databases of protein-metal and inorganic metal complexes, cobalt is shown to have a propensity to bind in octahedral coordination geometries. These results were in line with the notion that successful binding is dictated by both the protein environment and the metal ion and that coordination geometry is more important than size of the ion alone (although size can inform coordination geometry options).

Analysis of kinetics was based on the shape of the luminescence traces; flash-type traces were fitted to exponential decay curves whilst glow-type traces were analysed with an *ad hoc* metric for reaction progression. In the flash-type sets (aequorin with calcium, lanthanum or lead), the contribution of mutant aequorin or of ions other than calcium and lanthanum resulted in slower kinetics. A previously published model/assumption of two independent luminescence decay reactions was applied to the decay data. In the glow-type sets (aequorin with cobalt, zinc or copper), the mutation seems to have increased the speed of the reaction for zinc and copper but not for cobalt,

suggesting that these mutations were a positive step toward a bioreporter for these metals. Cadmium was not analysed in this work due to the shape of its luminescence curve (combination of sharp and glow-type) but would be interesting for further study.

9.3 Future work

This work produced an experimental platform for expressing and testing aequorin mutants in a high-throughput manner; rational mutagenesis was used to gain insight on the effect of key amino acid substitutions. It also highlighted the complexity of the task at hand. In an oversimplified view, the successful function of aequorin as a bioreporter depends on two events overall: binding of metal ions and propagation of the conformational changes from the three EF-hands to the protein core. Direction for future work would involve the following considerations: (1) a strategy for simplifying the problem; (2) mutational approach; (3) an alternative protein scaffold; (4) alternative applications of the envisaged molecule.

9.3.1 Breaking down the problem

The problem can be broken into distinct projects: one pertaining to ion binding affinity, followed by one pertaining to activity. Short peptides corresponding to the EF-hand loop can be mutated and used as models for metal binding studies. In order for such peptides to acquire a native-like fold, some alterations (disulfide bonds) will need to be incorporated. It is also expected that the metal affinities between the peptides and metal ions will vary when the same motifs are incorporated in the structure of the entire aequorin molecule. However this strategy will facilitate pure binding studies and remove the additional complexity of allostery and cooperativity in aequorin. This approach was employed by other researchers in the past and described in Chapter 5. The sequences of the successful peptides can then be introduced in the aequorin structure or into protein scaffolds other than aequorin (an option also discussed later in this Chapter).

9.3.2 Mutational strategy

Size of the binding cavity

Size selectivity of the metal binding cavity was found to be decisive in aequorin's effective response to metal ions. In order for the EF-hand loop to accommodate smaller ions, a more drastic mutational strategy may examine decreasing the size of the loop by removal of amino acids. In the past the opposite was done by addition of a glycine or substitution for a smaller coordinating residue (Glu→Asp) in order to effectively increase the size of the loop. Placing at least one histidine and one cysteine as close as possible to the entry and exit of the binding loop could achieve a more prominent conformational response as a result of metal binding.

Shortening the loop may have serious implications in the transmission of signal to the core of the protein. It may also prove to be ineffective, as despite of reduction of the loop size, the distance between the two flanking α -helices which mark the entry and exit of the loop may not actually decrease sufficiently; this distance is dictated by the remaining protein structure. Finally, shortening the loop will alter the current pairing of the EF-hand loops (EF-I to EF-II and EF-III to EF-IV) and risk disrupting the conformational transduction cascade.

Rational mutations in the binding sites

The results of the mutant library screening (Chapter 6) showed that removal of any calcium-coordinating side chains is detrimental to aequorin's response to calcium and that introduction of at least one histidine or cysteine can bring about a selectivity shift toward zinc, copper or lead. New rational substitutions in the loop would be designed based on these initial findings.

Random-rational mutations in the first coordination shell

As next step after implementing and screening rational mutations in a shortened loop, a random-rational approach could be applied within and in the close vicinity of the binding site (e.g. using cassette mutagenesis or MEGAWHOP). This approach would require screening of large mutant libraries.

Random-rational mutations in the extended protein structure

In this work the mutational strategy was based on precise structural knowledge of the EF-hands and of coordination preferences of metal ions in proteins (as shown in statistical databases). Mutations targeted the EF-hand loop, which comprises the first coordination shell of metal binding. In future work, additional rounds of random mutagenesis will be required in the entirety of the molecule, exploring long range interactions for beneficial results in: (1) the potential for repair of the aequorin activity and (2) additional metal selectivity modulation through distant mutations, which has been shown to be possible in metalloproteins.

9.3.3 Repairing aequorin activity

After altering the affinity of the metal binding sites, the scope of the next stage would be to restore the light yield of aequorin with further rounds of random mutagenesis. In this work, the property that was rewarded was the shift of selectivity towards new ions and not the absolute yield. In a screen aiming for repaired activity the selection should be based on absolute yields.

9.3.4 Alternative scaffold

In order to avoid the complexities of aequorin function and even the requirement for addition of coelenterazine, an alternative scaffold can be proposed for a metal ion bioreporter. The proposed chimeric protein would incorporate at least one EF-hand and domains of a fluorescent molecule – e.g. variants of the green fluorescent protein (GFP). Conformational changes upon ion binding would translate in altering fluorescence emission characteristics of the chimera. This approach is discussed in Chapter 1 and Section 1.2.5). Omitting coelenterazine would remove light and temperature constraints in protein stability, as well as the need for introduction of coelenterazine by diffusion or microinjection. It would allow for a bioreporter that can be utilised repeatedly and not depleted after one use.

9.3.5 Alternative application

It would also be useful to consider the landscape of other options for a protein molecule that selects for specific metals from its environment. An alternative goal from that of a biosensor would be to use a variant of apoaequorin immobilised on a surface (without its cofactor coelenterazine), or use another protein with multiple EF-hands as a “sponge” for toxic metals in bioremediation projects.

9.4 Analytics for future work

The experimental setup used in this work measures luminescence (activity assay). For future work it would be advisable to complement results from the activity assay with analytics that provide additional insight into the binding event (conformational response to binding of metals) and the metal uptake by the protein.

Conformational response

Conformational responses of polypeptides upon ion binding can potentially be tracked by monitoring changes in the fluorescence spectra upon titration with metal ions. This is a microplate based, high-throughput and inexpensive method but it requires the presence of aromatic amino acids. If not already present, aromatic residues would need to be introduced in the structure at least for the screening stage. It should be noted that any substitutions can affect the function of the peptide under study and may be undesirable in the final construct of the functional protein.

Other methods that have been used to monitor changes in the structure of aequorin (or isolated EF-hand peptides) in response to presence of metal ions are nuclear magnetic resonance (NMR) spectroscopy, circular dichroism (CD) spectroscopy and electrospray mass spectroscopy. These are all low-throughput options with relatively high sample consumption requirements – higher than the contents of a microwell – and NMR in particular is the most expensive, least accessible option. To use such methods, interesting candidates would be cherry-picked based on the results of high-throughput bioluminescence (activity) and fluorescence (conformational changes) screening, produced in sufficient quantities and submitted for further analysis. The bioluminescence activity assay would be applicable only if working with the entire

aequorin molecule, whereas a fluorescence assay that tracks conformational changes could be applicable to any polypeptide – provided that its structural changes can be translated into fluorescence changes. Fit for purpose for each method would be verified before finalising the analytical strategy.

Metal uptake

The amount of metal ion bound by the protein can be deduced from the free metal ion remaining in the mixture after the luminescence assay. Metal-sensitive electrodes can be used to measure free metal ions in the test sample. Another method would be to use ultracentrifugation in order to precipitate out the protein with any metals bound and use atomic adsorption spectroscopy to measure unbound metals in the supernatant.

These are both low-throughput methods and the sample requirement for metal-sensitive electrodes exceeds the capacity of microplate formats. In this case, analysis would be performed on mutants selected from an activity screen and later produced in the milliliter (instead of microliter) scale.

Sensitivity and selectivity

The successful candidate must be fit for its application and environment. Selectivity and sensitivity of the molecule to metal ions are both very important. A tier of experiments pertaining to the intended use (e.g. *in vivo* for real-time intracellular imaging) would be required to qualify the envisaged biosensor. If the bioreporter is not suitable for its intended environment, information on selectivity and sensitivity may help identify an alternative context where the bioreporter would be of value.

Sensitivity is related to the bioavailability of the metal ion under study. The bioreporter would be tested with the relevant range of metal ion concentrations. For example, aequorin has been an excellent calcium sensor partly because its EC₅₀ is within the biologically relevant range of intracellular calcium concentrations. Zinc would require a far more sensitive sensor to detect changes in the biologically relevant range of this metal.

Selectivity is related to the bioavailability of the metal ion under study versus potentially competing ions. In this mutant library, selectivity is not shifted toward one new ion independently from the rest (Section 6.4.6). Additional work would be required in order to achieve high affinity for one ion against a background of other

metals in a mixture. If competing metals are not present in the specific environment under study (e.g. cell, cellular compartment), exclusivity of ion binding is not necessary. Using aequorin as an example, the lanthanides can be excellent replacements for calcium, however they are not present in the biological systems routinely studied.

9.5 Value and risks of the project

The search for a functional aequorin mutant with altered metal selectivity was a high risk-high gain endeavour. Further work would be required to increase understanding of the structure-function of aequorin with alternative activator ions. The canonical binding site of the EF-hand superfamily evolved to selectively bind calcium over a background of other metals. Cooperativity between the EF-hands and the allosteric activation of the protein further complicate the project.

The three ion binding sites (EF-hands) are spatially distinct from the domain where light is produced (hydrophobic core of the protein). There is no guarantee that successful shift of the binding loop affinities for a new metal ion can accommodate the necessary conformational changes for production of light. The risk was reflected in the results, whereby improved response of mutant to zinc resulted in significant loss of activity compared to aequorin's full potential.

Additional risk was in assuming promiscuity of aequorin toward other metal ions; this was originally reported in the 1960s but no follow up work was performed. There was a risk that aequorin luminescence with various ions was merely due to non-specific binding which caused destabilisation and discharge of the protein. In this work, the fact that different ions induce different types of luminescence profiles and concentration-response curves suggests that different ions indeed exert different impact on the activity of aequorin and that an attempt to study and tweak their binding was valid.

In terms of the rational mutagenesis strategy, the single and double substitutions constituted a rather crude initial approach. The work with forty eight mutants produced useful information on the importance of the three EF-hands and of specific positions and residues within the EF-hand loops in the response of aequorin to various metals.

Most mutations in a protein are expected to be deleterious. The designed mutations were not expected to completely transform the ion binding sites in favour of

new ions whilst retaining the full activity of the native protein, for the following reasons: (1) canonical EF-hands are fine-tuned to be highly selective for calcium; (2) the mutants of this library contain only single or double mutations out of a total of eighteen ion-coordinating positions across the entire molecule and (3) activity is dependent on successful binding followed by propagation of each binding event through an intricate network of cooperativity and allosteric activation. It is highly unlikely to completely redesign the EF-hand loops whilst keeping the allosteric and cooperative functions intact so that the appropriate conformational changes are propagated to the centre of the molecule for effective bioluminescence production.

In continuing to pursue the originally envisaged aequorin mutant, it should be noted that the mechanism of calcium binding and cooperativity have co-evolved in aequorin, thus the potential for a highly functional molecule with altered ion selectivity may not be achieved. The options discussed in Future work (Section 9.3) can be viable alternatives, although they too have limitations to be considered on an individual basis.

Finally, it is hoped that the present work may contribute to future methodologies of screening mutant libraries and identifying promising mutants for new applications. This itself is an interesting challenge, especially when there is a trade-off between two desired properties (e.g. selectivity and yield). To reward only high activity mutants could lead to missed opportunities. Disqualifying mutants with low activity in the first rounds of screening would exclude valuable candidates where activity could have been restored with further rounds of mutagenesis. Also, certain desired properties cannot be directly measured but must be deciphered (e.g. gradual shifts in selectivity) with appropriate *ad hoc* metrics.

Appendix

Appendix Chapter 1

Table 10-1 Intracellular metal ion sensors

Type of sensor	Name	Application	Reference
bioluminescent protein	aequorin		(Czarnik, 1995)
Fluorescent, synthetic	Fura-2	Ca ²⁺ , Zn ²⁺ , Mn ²⁺ , Pb ²⁺ , Ba ²⁺ , Cd ²⁺ , Co ²⁺ , Sr ²⁺	Kwan and Putney, 1990; Hinkle <i>et al</i> , 1992; Tomsig and Suszkiw, 1990; Atar <i>et al</i> , 1995; Tomat <i>et al</i> , 2008 in(Palmer, 2009)
Zinpyr-labeled fusion proteins		Zn ²⁺	
Derivatives of benzylguanine linked to Indo-1	BG-Indo-1, BG-Indo-2, BG-Indo-3	Ca ²⁺	(Keppler <i>et al.</i> , 2003) and Gronemeyer <i>et al</i> , 2006 in Palmer, 2009
BG-Indo-1 derivatives reacted with SNAP-tag	SNAP-Indo-1	Ca ²⁺	Bannwarth <i>et al</i> , 2009 in Palmer, 2009
FRET-based, ECFP-calmodulin-M13-cpVenus (redesigned calmodulin and M13), interaction of domains	Cameleon D3	Ca ²⁺	Palmer <i>et</i> , 2006 & Wallace <i>et al</i> , 2008 in Chudakov <i>et al</i> , 2010
FRET-based, ECFP-calmodulin-M13-cpVenus, interaction of domains	Yellow Cameleon 3.6	Ca ²⁺	Nagai <i>et al</i> , 2004 in Chudakov <i>et al</i> , 2010
FRET-based, CFP-2x(COOH-terminal lobe of troponin C)-cpYFP, structural rearrangement of domain	TN-XXL	Ca ²⁺	Mank <i>et al</i> , 2008 in Chudakov <i>et al</i> , 2010
Insertion of structurally rearranging calmodulin domain into YFP	Camgaroo-2	Ca ²⁺	Griesbeck <i>et al</i> , 2001 in Chudakov <i>et al</i> , 2010
Single cpFP fused to interacting domains, M13-cpGFP-calmodulin	Case12	Ca ²⁺	Souslova <i>et al</i> , 2007 in Chudakov <i>et al</i> , 2010

Type of sensor	Name	Application	Reference
Single cpFP fused to interacting domains, M13-cpGFP-calmodulin	GCaMP3	Ca ²⁺	Tallini <i>et al</i> , 2006 & Tian <i>et al</i> , 2009 in Chudakov <i>et al</i> , 2010
Fluo-3 dye		Ca ²⁺	Czarnik, 1995
Quin-2		Ca ²⁺	Czarnik, 1995
Fura-2		Ca ²⁺	Czarnik, 1995
SBF1		Na ⁺	Czarnik, 1995
PBF1		K ⁺	Czarnik, 1995
SPQ		Cl ⁻	Czarnik, 1995
TMAPQ		Cl ⁻	Czarnik, 1995
FURAPTRA (Mag-fura-2)		Mg ²⁺	Czarnik, 1995
FRET, Intramolecular single domain, apoK1-er, CFP YFP		Ca ²⁺	Osibow <i>et al</i> , 2006 in Li <i>et al</i> , 2006
FRET, Intramolecular multiple domain, CaM M13, CFP YFP BFP GFP		Ca ²⁺	Miyawaki <i>et al</i> , 1997 in Li <i>et al</i> , 2006
Small fluophore	Zinquin	Zn ²⁺	Colvin <i>et al</i> , 2006
Small fluophore	ZnAF-1, ZnAF-2, ZnAF-3	Zn ²⁺	Hirano <i>et al</i> , 2000; Hirano <i>et al</i> , 2002; Komatsu <i>et al</i> , 2005
Small fluophore	FluoZin-1, FluoZin-2, FluoZin-3	Zn ²⁺	Eichelsdoerfer <i>et al</i> , 2010; Domaille <i>et al</i> , 2008 ; Gee <i>et al</i> , 2002
Small fluophore	FuraZin	Zn ²⁺	(Sensi <i>et al.</i> , 1999)
Small fluophore	RhodZin-3	Zn ²⁺	(Sensi <i>et al.</i> , 2003)
Phenanthroline-based	Phen Green FL	Cu ²⁺ , Cu ⁺ , Fe ²⁺ , Hg ²⁺ , Pb ²⁺ , Cd ²⁺ , Ni ²⁺	(Colvin <i>et al.</i> , 2008, Kikuchi <i>et al.</i> , 2004) (Tougu <i>et al.</i> , 2008)
luxCDABE-based		Pb ²⁺ , Hg ²⁺ , Zn ²⁺	Corbisier <i>et al</i> , 1996; Barkay <i>et al</i> , 1998; Erbe <i>et al</i> , 1996

Table 10-2 Multiple alignment of calcium-binding proteins a) aequorin, b) mitrocomin, c) obelin, d) clytin, e) human calmodulin f) bovine calmodulin, g) human parvalbumin, h) intestinal Ca²⁺-binding protein, i) sarcoplasmic Ca²⁺-binding protein and j) troponin C using ClustalW (Thompson, 1994) (DNA and protein sequence IDs in Table 1–6).

	EF-I																																				
a:	T	P	D	F	D	N	P	K	W	I	G	R	H	K	H	M	F	N	F	L	D	V	N	H	N	G	R	I	S	L	D	E	M	V	Y		
b:	T	T	D	F	D	N	P	K	W	I	A	R	H	K	H	M	F	N	F	L	D	I	N	S	N	G	Q	I	N	L	N	E	M	V	H		
c:	K	T	D	F	D	N	P	R	W	I	K	R	H	K	H	M	F	D	F	L	D	I	N	G	N	G	K	I	T	L	D	E	I	V	S		
d:	R	P	N	F	D	N	P	K	W	V	N	R	H	K	F	M	F	N	F	L	D	I	N	G	D	G	K	I	T	L	D	E	I	V	S		
e:	M	A	D	Q	L	T	E	E	Q	I	A	E	F	K	E	A	F	S	L	F	D	K	D	G	D	G	T	I	T	T	K	E	L	G	-		
f:	-	A	D	Q	L	T	E	E	Q	I	A	E	F	K	E	A	F	S	L	F	D	K	D	G	D	G	T	I	T	T	K	E	L	G	-		
g:	M	T	D	L	L	N	A	E	D	I	K	K	A	V	G	A	F	S	A	T	D	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
h:	-	-	-	-	K	S	P	E	E	L	K	G	I	F	E	K	Y	A	A	K	E	G	D	P	N	-	-	-	-	-	-	-	-	-	-	-	
i:	L	N	D	F	Q	K	Q	K	I	K	F	T	F	D	F	F	L	D	M	N	H	D	G	S	I	Q	D	N	D	F	E	D	M	M	T		
j:	-	-	-	-	-	-	-	-	G	V	F	Y	F	N	M	N	R	V	F	V	Y	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 10-2 continued

a:	K	A	S	D	I	V	I	N	N	L	G	A	T	P	E	Q	A	K	R	H	K	D	A	V	E	A	F	F	G	G	A	G	M	K	Y	
b:	K	A	S	N	I	I	C	K	K	L	G	A	T	E	E	Q	T	K	R	H	Q	K	C	V	E	D	F	F	G	G	A	G	L	E	Y	
c:	K	A	S	D	D	I	C	A	K	L	E	A	T	P	E	Q	T	K	R	H	Q	V	C	V	E	A	F	F	R	G	C	G	M	E	Y	
d:	K	A	S	D	D	I	C	A	K	L	G	A	T	P	E	Q	T	K	R	H	Q	D	A	V	E	A	F	F	K	K	I	G	M	D	Y	
e:	-	-	-	-	T	V	M	R	S	L	G	Q	N	P	T	E	A	E	L	Q	D	M	I	N	E	V	D	A	D	D	L	-	-	P	G	
f:	-	-	-	-	T	V	M	R	S	L	G	Q	N	P	T	E	A	E	L	Q	D	M	I	N	E	V	D	A	D	-	-	-	-	-	G	
g:	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	F	D	H	K	K	F	F	Q	M	V	G	L	K	K	-	-	-	-		
h:	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
i:	R	Y	K	E	V	N	K	G	S	L	S	D	A	D	Y	K	S	M	Q	A	S	L	E	D	E	W	R	D	L	K	G	R	A	D	I	
j:	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 10-2 continued

a:	G	V	E	T	E	W	P	E	Y	I	E	G	W	K	R	L	A	S	E	E	L	K	R	Y	S	K	N	Q	I	T	L	I	R	L	W				
b:	D	K	D	T	T	W	P	E	Y	I	E	G	W	K	R	L	A	K	T	E	L	E	R	H	S	K	N	Q	V	T	L	I	R	L	W				
c:	G	K	E	I	A	F	P	Q	F	L	D	G	W	K	Q	L	A	T	S	E	L	K	K	W	A	R	N	E	P	T	L	I	R	E	W				
d:	G	K	E	V	E	F	P	A	F	V	D	G	W	K	E	L	A	N	H	D	L	K	L	W	S	Q	N	K	K	S	L	I	R	D	W				
e:	N	G	T	I	D	F	P	E	F	L	T	-	-	-	-	-	-	-	-	-	-	M	M	A	R	K	M	K	D	T	D	S	E	E	E				
f:	N	G	T	I	D	F	P	E	F	L	T	-	-	-	-	-	-	-	-	-	-	M	M	A	R	K	M	K	D	T	D	S	E	E	E				
g:	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	K	S	A	D	D
h:	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
i:	N	K	D	D	V	V	S	W	E	E	Y	L	A	M	W	E	K	T	I	A	T	C	K	S	V	A	D	L	P	A	W	C	Q	N	R				
j:	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 10-2 continued

	EF-III																																				
a:	G	D	A	L	F	D	I	I	D	K	D	Q	N	G	A	I	S	L	D	E	W	K	A	Y	T	K	S	A	G	-	-	-	I	I	Q		
b:	G	D	A	L	F	D	I	I	D	K	D	R	N	G	S	V	S	L	D	E	W	I	Q	Y	T	H	C	A	G	-	-	-	I	Q	Q		
c:	G	D	A	V	F	D	I	F	D	K	D	G	S	G	T	I	T	L	D	E	W	K	A	Y	G	K	I	S	G	-	-	-	I	S	P		
d:	G	E	A	V	F	D	I	F	D	K	D	G	S	G	S	I	S	L	D	E	W	K	A	Y	G	R	I	S	G	-	-	-	I	C	S		
e:	I	R	E	A	F	R	V	F	D	K	D	G	N	G	Y	I	S	A	A	E	L	R	H	V	M	T	N	L	G	-	-	-	E	K	L		
f:	I	R	E	A	F	R	V	F	D	K	D	G	N	G	Y	I	S	A	A	E	L	R	H	V	M	T	N	L	G	-	-	-	E	X	L		
g:	V	K	K	V	F	H	M	L	D	K	D	K	S	G	F	I	E	E	D	E	L	G	F	I	L	K	G	F	S	P	D	A	R	D	L		
h:	-	-	-	-	-	-	Q	L	S	K	E	E	L	K	L	L	L	Q	T	E	F	P	S	L	L	K	G	G	S	-	-	-	-	-	-	-	
i:	I	P	F	L	F	K	G	M	D	V	S	G	D	G	I	V	D	L	E	E	F	Q	N	Y	C	K	N	F	Q	-	-	-	-	-	-	L	
j:	-	-	-	-	-	-	-	-	-	T	L	K	R	G	Q	P	N	H	F	A	L	S	E	V	G	N	E	R	Y	T	Y	I	G	-	-	-	-

Table 10-2 continued

	EF-IV																																		
a:	S	S	E	D	C	E	E	T	F	R	V	C	D	I	D	E	S	G	Q	L	D	V	D	E	M	T	R	Q	H	L	G	F	W	Y	T
b:	S	R	G	Q	C	E	A	T	F	A	H	C	D	L	D	G	D	G	K	L	D	V	D	E	M	T	R	Q	H	L	G	F	W	Y	S
c:	S	Q	E	D	C	E	A	T	F	R	H	C	D	L	D	D	S	G	D	L	D	V	D	E	M	T	R	Q	H	L	G	F	W	Y	T
d:	S	D	E	D	A	E	K	T	F	K	H	C	D	L	D	N	S	G	K	L	D	V	D	E	M	T	R	Q	H	L	G	F	W	Y	T
e:	T	D	E	E	V	D	E	M	I	R	E	A	D	I	D	G	D	G	Q	V	N	Y	E	E	F	V	Q	M	M	T	A	K	-	-	-
f:	T	D	E	E	V	D	E	M	I	R	E	A	D	I	D	G	D	G	Q	V	N	Y	E	E	F	V	Q	M	M	T	A	K	-	-	-
g:	S	A	K	E	T	K	M	L	M	A	A	G	D	K	D	G	D	G	K	I	G	V	D	E	F	S	T	L	V	A	E	S	-	-	-
h:	-	-	-	T	L	D	E	L	F	E	E	L	D	K	N	G	D	G	E	V	S	F	E	E	F	Q	V	L	V	K	K	I	S	Q	-
i:	Q	C	A	D	V	P	A	V	Y	N	V	I	T	D	G	G	K	V	T	F	D	L	N	R	Y	K	E	L	Y	Y	R	L	L	T	S
j:	-	-	-	-	-	-	-	N	G	F	T	S	T	K	C	P	L	V	I	A	S	E	H	N	I	P	Y	L	M	D	K	E	G	N	G

Table 10-2 continued

a:				M	D	P	A	C	E	K	L	Y	G	G	A	V	P	-
b:				V	D	P	T	C	E	G	L	Y	G	G	A	V	P	Y
c:				L	D	P	E	A	D	G	L	Y	G	N	G	V	P	-
d:				L	D	P	N	A	D	G	L	Y	G	N	F	V	P	-
e:				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
f:				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
g:				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
h:				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
i:				P	A	A	D	A	G	N	T	L	M	G	Q	K	P	-
j:	Y	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 10-3 Preferred coordination geometries of metal ions in CSD structures. Table reproduced from Kuppuraj *et al*, 2009.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
IA	IIA	IIIB	IVB	VB	VIB	VIIIB		VIIIB		IB	IIIB	IIIA	IVA	VA	VIA
H															
Li ⁺	Be ²⁺											B	C	N	O
<i>T_d</i>	<i>T_d</i>														
Na ⁺	Mg ²⁺											Al ³⁺	Si	P	S
<i>O_h</i>	<i>O_h</i>											<i>T_d</i>			
K	Ca ²⁺	Sc ³⁺	Ti ²⁺	V ²⁺	Cr ²⁺	Mn ⁺	Fe ²⁺	Co ⁺	Ni ²⁺	Cu ⁺	Zn ²⁺	Ga ³⁺	Ge ⁴⁺	As	Se
<i>O_h</i>	TBP	SPI/ <i>O_h</i>	<i>T_d</i>	<i>T_d</i>	<i>T_d</i>	<i>T_d</i>									
Rb	Sr ²⁺		Ti ⁴⁺	V ³⁺	Cr ³⁺	Mn ²⁺	Fe ³⁺	Co ²⁺	Ni ³⁺	Cu ²⁺					
	SA		<i>O_h</i>	<i>O_h</i>	<i>O_h</i>	<i>O_h</i>	<i>O_h</i>	<i>O_h</i>	SPI	SP/SPI					
				V ⁴⁺	Mo ²⁺	Mn ³⁺	Ru ²⁺	Co ³⁺							
				<i>O_h</i>	cTBP	<i>O_h</i>	<i>O_h</i>	<i>O_h</i>							
		Y ³⁺	Zr ⁴⁺	Nb ⁵⁺	Mo ³⁺	Tc ³⁺	Ru ³⁺	Rh ⁺	Pd ²⁺	Ag ⁺	Cd ²⁺	In ³⁺	Sn ⁴⁺	Sb	Te
		cTBP	<i>O_h</i>	<i>O_h</i>	<i>O_h</i>	<i>O_h</i>	<i>O_h</i>	SPI	SPI	<i>T_d</i>	<i>T_d</i> / <i>O_h</i>	<i>T_d</i>	<i>T_d</i>	<i>O_h</i>	
					Mo ⁴⁺			Rh ³⁺	Pd ⁴⁺						
					<i>O_h</i>			<i>O_h</i>	<i>O_h</i>						
					Mo ⁵⁺		Os ²⁺	Ir ⁺	Pt ²⁺	Au ⁺	Hg ²⁺	Tl ³⁺	Pb ²⁺	Bi ³⁺	Po
					<i>O_h</i>		<i>O_h</i>	SPI	SPI	L _n	L _n	<i>T_d</i>	<i>T_d</i>	TBP/ <i>T_d</i>	
					Mo ⁶⁺		Os ³⁺	Ir ³⁺	Pt ⁴⁺	Au ³⁺					
					TBPr		<i>O_h</i>	<i>O_h</i>	<i>O_h</i>	SPI					
Cs	Ba ²⁺		Hf ⁴⁺	Ta ⁵⁺	W ²⁺	Re ⁺	Os ⁴⁺								
	cTBP/SA		cTBP	<i>O_h</i>	cTBP	<i>O_h</i>	<i>O_h</i>								
					W ⁴⁺	Re ³⁺	Os ⁶⁺								
					<i>O_h</i>	<i>O_h</i>	<i>O_h</i>								
					W ⁶⁺	Re ⁴⁺									
					<i>O_h</i>	<i>O_h</i>									
						Re ⁵⁺									
						<i>O_h</i>									
						Re ⁷⁺									
						TBP									

^a CN = 2: L_n = linear. CN = 4: *T_d* = tetrahedral; *~T_d* = semi-tetrahedral; SPI = square planar. CN = 5: TBP = trigonal bipyramidal; SP = square pyramidal. CN = 6: *O_h* = octahedral; TBPr = trigonal bipyramidal. CN = 7: cTBP = capped trigonal bipyramidal. CN = 8: SA = square antiprismatic.

Table 10-4 Metal-ligand distances in the CSD and PDB. The numbers of observations and mean distances in structures in the PDB determined at near atomic resolution and in the CSD with R factor < 0.065. Sample standard deviations and metal-coordination numbers from all coordination geometries are included. Table reproduced from Harding, 2006.

	Ca	Mg	Mn	Fe	Co	Cu†	Zn	Na	K
<i>M</i> –H ₂ O									
PDB									
<i>N</i> _{obs}	302	269	17	8	2	4	31	133	24
Mean distance (Å)	2.40 (10)	2.09 (8)	2.22 (6)	2.17 (8)	—	—	2.06 (13)	2.42 (19)	2.82 (14)
CSD									
<i>N</i> _{obs}	169	326	289	121	552	379	270	334	104
Mean distance (Å)	2.39 (5)	2.07 (3)	2.19 (4)	2.09 (5)	2.09 (3)	2.13 (22)†	2.09 (5)	2.41 (10)	2.80 (19)
<i>M</i> –O monodentate carboxylate									
PDB									
<i>N</i> _{obs}	105	43	19	12	1	1	16	4	1
Mean distance (Å)	2.33 (7)	2.08 (8)	2.12 (5)	2.10 (6)	2.01	1.96	2.01 (9)	2.3 (3)	2.8
CSD									
<i>N</i> _{obs}	170	4	5	8	33	95	84	931	1049
Mean distance (Å)	2.38 (7)	2.05 (5)	2.15 (1)	2.03 (2)	2.05 (6)	1.96 (4)	1.99 (5)	2.41 (11)	2.82 (13)
<i>M</i> –O main-chain carbonyl									
PDB									
<i>N</i> _{obs}	130	12	—	—	—	—	5	44	25
Mean distance (Å)	2.36 (10)	2.26 (23)	—	—	—	—	—	2.46 (24)	2.80 (15)
CSD									
<i>N</i> _{obs}	6	4	8	26	30	137	12	15	11
Mean distance (Å)	2.39 (11)	—	2.19 (5)	2.04 (6)	2.08 (5)	2.04 (14)†	2.07 (5)‡	2.37 (6)	2.67 (10)
<i>M</i> –N of imidazole (for His)									
PDB									
<i>N</i> _{obs}	2	3	22	24	7	19	62	3	—
Mean distance (Å)	—	—	2.16 (5)	2.03 (8)	2.04 (9)	2.02 (4)	2.04 (4)	—	—
CSD									
<i>N</i> _{obs}	1	—	10	7	47	110	34	—	—
Mean distance (Å)	—	—	2.25 (3)	2.17 (1)	2.14 (5)§	2.02 (9)†	2.01 (4)	—	—
<i>M</i> –S of thiolate (for Cys)									
PDB									
<i>N</i> _{obs}	—	—	—	239	—	10	59	—	—
Mean distance (Å)	—	—	—	2.30 (3)	—	2.15 (9)	2.34 (5)	—	—
CSD									
<i>N</i> _{obs}	—	—	43	47	46	3	28	10	—
Mean distance (Å)	—	—	2.35 (4)	2.28 (4)	2.25 (4)§	—	2.28 (4)	2.88 (8)	—

Table 10-5 Average metal-ligand distances for metal ions as a function of metal charge, coordination number (CN), donor atom's charge and coordination number of ligand (number of atoms connected to the donor atom of the ligand). Table reproduced from Kuppuraj *et al.*, 2009.

metal ion	metal charge	metal CN	donor atom	ligand charge	ligand CN	M-L (#) ^a (Å)
Li	1	4	O ^{metal}	0	3	1.92 ± 0.03 (55)
			O	0	3	1.93 ± 0.04 (218)
			O	0	2	1.91 ± 0.04 (35)
			O	-1	2	-
			N	0	3	2.03 ± 0.03 (30)
Na	1	6	S	-1	2	-
			O ^{metal}	0	3	-
			O	0	3	2.30 ± 0.13 (38)
			O	0	2	-
			O	-1	2	-
Mg	2	6	N	0	3	2.46 ± 0.02 (18)
			S	-1	2	-
			O ^{metal}	0	3	2.06 ± 0.02 (308)
			O	0	3	2.10 ± 0.04 (78)
			O	0	2	2.05 ± 0.04 (92)
Ca	2	6	O	-1	2	2.07 ± 0.04 (28)
			N	0	3	2.19 ± 0.06 (56)
			S	-1	2	-
			O ^{metal}	0	3	2.33 ± 0.01 (30)
			O	0	3	2.34 ± 0.01 (40)
Ca	2	7	O	0	2	2.30 ± 0.02 (52)
			O	-1	2	-
			N	0	3	2.47 ± 0.03 (24)
			S	-1	2	-
			O ^{metal}	0	3	2.39 ± 0.04 (62)
Ca	2	7	O	0	3	2.40 ± 0.03 (24)
			O	0	2	2.39 ± 0.03 (11)
			O	-1	2	-
			N	0	3	2.54 ± 0.03 (16)
			S	-1	2	-
Mn	2	6	O ^{metal}	0	3	2.18 ± 0.03 (243)
			O	0	3	2.20 ± 0.04 (24)
			O	0	2	2.14 ± 0.07 (112)
			O	-1	2	2.15 ± 0.05 (72)
			N	0	3	2.27 ± 0.04 (173)
Fe	2	6	S	-1	2	2.60 ± 0.04 (14)
			O ^{metal}	0	3	2.11 ± 0.03 (103)
			O	0	3	2.15 ± 0.03 (23)
			O	0	2	2.08 ± 0.04 (89)
			O	-1	2	2.07 ± 0.05 (32)
Fe	3	6	N	0	3	2.06 ± 0.11 (507)
			S	-1	2	2.46 ± 0.13 (13)
			O ^{metal}	0	3	2.01 ± 0.05 (42)
			O	0	3	-
			O	0	2	1.99 ± 0.05 (159)
Co	2	6	O	-1	2	1.99 ± 0.02 (39)
			N	0	3	2.06 ± 0.10 (218)
			S	-1	2	2.34 ± 0.10 (42)
			O ^{metal}	0	3	2.09 ± 0.03 (482)
			O	0	3	2.11 ± 0.05 (58)
Ni	2	6	O	0	2	2.07 ± 0.04 (289)
			O	-1	2	2.07 ± 0.05 (167)
			N	0	3	2.13 ± 0.07 (597)
			S	-1	2	2.42 ± 0.01 (22)
			O ^{metal}	0	3	2.07 ± 0.03 (634)
Cu	2	4	O	0	3	2.09 ± 0.04 (119)
			O	0	2	2.05 ± 0.04 (424)
			O	-1	2	2.06 ± 0.06 (236)
			N	0	3	2.10 ± 0.04 (1034)
			S	-1	2	2.44 ± 0.06 (169)
Zn	2	4	O ^{metal}	0	3	1.95 ± 0.02 (53)
			O	0	3	1.96 ± 0.02 (16)
			O	0	2	1.91 ± 0.01 (542)
			O	-1	2	1.94 ± 0.02 (334)
			N	0	3	1.97 ± 0.03 (855)
Zn	2	4	S	-1	2	2.26 ± 0.03 (95)
			O ^{metal}	0	3	1.99 ± 0.04 (21)
			O	0	3	2.15 ± 0.10 (23)
			O	0	2	1.96 ± 0.05 (237)
			O	-1	2	1.97 ± 0.03 (109)
			N	0	3	2.03 ± 0.04 (603)
			S	-1	2	2.32 ± 0.05 (246)

Appendix Chapter 5

Table 10-6 Primers for mutants Nr 1-8

#	Numbering according to 1EJ3	Numbering according to 1SL8	position in EF-I	position in EF-III	position in EF-IV	Primer sequence 5'-3'
1	Asp24Ala	Asp26Ala	1			For CACAAGCACATGTTTAATTTTCTTGCGGTCAACCACAATGGAAGGATCTCTC Rev GAGAGATCCTTCCATTGTGGTTGACCGCAAGAAAATTAACATGTGCTTGTG
2	Asn26Ala	Asn28Ala	3			For GCACATGTTTAATTTTCTTGATGTGCGGCACAATGGAAGGATCTCTCTTGACG Rev CGTCAAGAGAGATCCTTCCATTGTGCGCGACATCAAGAAAATTAACATGTGC
3	Asn28Ala	Asn30Ala	5			For GTTTAATTTTCTTGATGTCAACCACGCGGGAAGGATCTCTCTTGACGAGATGG Rev CCATCTCGTCAAGAGAGATCCTTCCCGGTGGTTGACATCAAGAAAATTAAC
4	Ser32Ala	Ser34Ala	9			For CTTGATGTCAACCACAATGGAAGGATCGCGCTTGACGAGATGGTCTACAAGG Rev CCTTGTAGACCATCTCGTCAAGCGGATCCTTCCATTGTGGTTGACATCAAG
5	Glu35Ala	Glu37Ala	12			For GGAAGGATCTCTCTTGACGCGATGGTCTACAAGGCGTCCGATATTG Rev CAATATCGGACGCCTTGTAGACCATCGCGTCAAGAGAGATCCTTCC
6	Gly29Arg	Gly31Arg	6			For ACCACAATCGTAGGATCTCTC Rev AGATCCTACGATTGTGGTTG
7	Gly122Arg	Gly124Arg		6		For ACCAAAATCGTGCTATTTTAC Rev TGAAATAGCACGATTTTGGTC
8	Gly158Arg	Gly160Arg			6	For ATGAAAGTCGTCAGCTCGATG Rev ATCGAGCTGACGACTTTCATC

Table 10-7 Primers for mutants Nr 9-17

#	Numbering according to 1EJ3	Numbering according to 1SL8	position in EF-I	position in EF-III	position in EF-IV	Primer sequence 5'-3'
9	Gly29Arg/Gly122Arg	Gly31Arg/Gly124Arg	6	6		For Template Gly29Arg, primers of #7 Rev Template Gly29Arg, primers of #7
10	Gly29Arg/Gly158Arg	Gly31Arg/Gly160Arg	6		6	For Template Gly29Arg, primers of #8 Rev Template Gly29Arg, primers of #8
11	Gly122Arg/Gly158Arg	Gly124Arg/Gly160Arg		6	6	For Template Gly122Arg, primers of #8 Rev Template Gly122Arg, primers of #8
12	Asp24His	Asp26His	1			For AAGCACATGTTTAATTTTCTTCATGTCAACCACAATGGAAGG Rev TTCCATTGTGGTTGACATGAAGAAAATTAACATGTGCTTG
13	Val25His	Val27His	2			For CATGTTTAATTTTCTTGATCATAACCACAATGGAAGGATC Rev GATCCTTCCATTGTGGTTATGATCAAGAAAATTAACATG
14	Asn26His	Asn28His	3			For ATGTTTAATTTTCTTGATGTCCATCACAATGGAAGGATCTC Rev AGATCCTTCCATTGTGATGGACATCAAGAAAATTAACATG
15	Asn28His	Asn30His	5			For TTCTTGATGTCAACCACCATGGAAGGATCTCTCTTGACG Rev TCGTCAAGAGAGATCCTTCCATGGTGGTTGACATCAAG
16	Ser32His	Ser34His	9			For GTCAACCACAATGGAAGGATCCATCTTGACGAGATGGTCTACAAGG Rev CCTTGTAGACCATCTCGTCAAGATGGATCCTTCCATTGTGGTTGAC
17	Glu35His	Glu37His	12			For ACAATGGAAGGATCTCTCTTGACCATATGGTCTACAAGGCG Rev GCCTTGTAGACCATATGGTCAAGAGAGATCCTTCCATTGTG

Table 10-8 Primers for mutants Nr 18-23

#	Numbering according to 1EJ3	Numbering according to 1SL8	position in EF-I	position in EF-III	position in EF-IV	Primer sequence 5'-3'
18	Asp24Cys	Asp26Cys	1			For AGCACATGTTTAATTTTCTTTGCGTCAACCACAATGGAAGG Rev CCTTCCATTGTGGTTGACGCAAAGAAAATTAACATGTGC
19	Val25Cys	Val27Cys	2			For CATGTTTAATTTTCTTGATCATAACCACAATGGAAGGATC Rev GATCCTTCCATTGTGGTTATGATCAAGAAAATTAACATG
20	Asn26Cys	Asn28Cys	3			For ATGTTTAATTTTCTTGATGTCTGCCACAATGGAAGGATCTC Rev GAGATCCTTCCATTGTGGCAGACATCAAGAAAATTAACATG
21	Asn28Cys	Asn30Cys	5			For TTCTTGATGTCAACCACTGCGGAAGGATCTCTCTTGACG Rev CTCGTCAAGAGAGATCCTTCCGCAGTGGTTGACATCAAG
22	Ser32Cys	Ser34Cys	9			For GTCAACCACAATGGAAGGATCTGCCTTGACGAGATGGTCTACAAGG Rev CTTGTAGACCATCTCGTCAAGGCAGATCCTTCCATTGTGGTTGAC
23	Glu35Cys	Glu37Cys	12			For ACAATGGAAGGATCTCTCTTGACTGCATGGTCTACAAGGCG Rev CTTGTAGACCATGCAGTCAAGAGAGATCCTTCCATTGTG

Table 10-9 Primers for mutants Nr 24-32

#	Numbering according to 1EJ3	Numbering according to 1SL8	position in EF-I	position in EF-III	position in EF-IV	Primer sequence 5'-3'
24	Asp117His	Asp119His		1		For GGTGATGCATTGTTCGATATCATTTCATAAAGACCAAAAATGGAGCTATTTCACTGG Rev CCAGTGAAATAGCTCCATTTTGGTCTTTATGAATGATATCGAACAATGCATCACC
25	Asp119His	Asp121His		3		For CATTGTTCGATATCATTGACAAACATCAAAAATGGAGCTATTTCACTGG Rev CCAGTGAAATAGCTCCATTTTGGTCTTTGTCAATGATATCGAACAATG
26	Asp121His	Asp123His		5		For CGATATCATTGACAAAGACCAACATGGAGCTATTTCACTGGATGAATGG Rev CCATTCATCCAGTCAAATAGCTCCATGTTGGTCTTTGTCAATGATATCG
27	Ala123His	Ala125His		7		For CGATATCATTGACAAAGACCAAAAATGGACATATTTCACTGGATGAATGG Rev CCATTCATCCAGTCAAATGCATCCATTTTGGTCTTTGTCAATGATATCG
28	Thr125His	Ser127His		9		For GACAAAGACCAAAAATGGAGCTATTCATCTGGATGAATGGAAAGCATAACACC Rev GGTGTATGCTTTCCATTCATCCAGATGAATAGCTCCATTTTGGTCTTTGTC
29	Asp127His	Asp129His		11		For CCAAAAATGGAGCTATTTCACTGCATGAATGGAAAGCATAACCAAAATCTGC Rev GCAGATTTGGTGTATGCTTTCCATTCATGCAGTCAAATAGCTCCATTTTGG
30	Glu128His	Glu130His		12		For CCAAAAATGGAGCTATTTCACTGGATCATTGGAAAGCATAACCAAAATCTGC Rev GCAGATTTGGTGTATGCTTTCCAATGATCCAGTCAAATAGCTCCATTTTGG
31	Asp117Cys	Asp119Cys		1		For GGTGATGCATTGTTCGATATCATTGCAAAGACCAAAAATGGAGCTATTTCACTGG Rev CCAGTGAAATAGCTCCATTTTGGTCTTTGCAAATGATATCGAACAATGCATCACC
32	Asp119Cys	Asp121Cys		3		For CATTGTTCGATATCATTGACAAATGCCAAAATGGAGCTATTTCACTGG Rev CCAGTGAAATAGCTCCATTTTGGCATTGTCAATGATATCGAACAATG

Table 10-10 Primers for mutants Nr 33-40

#	Numbering according to 1EJ3	Numbering according to 1SL8	position in EF-I	position in EF-III	position in EF-IV	Primer sequence 5'-3'
33	Asp121Cys	Asp123Cys		5		For TTCGATATCATTGACAAAGACCAATGCGGAGCTATTTCACTGGATGAATGG Rev CCATTCATCCAGTGAAATAGCTCCGCATTGGTCTTTGTCAATGATATCGAA
34	Ala123Cys	Ala125Cys		7		For CGATATCATTGACAAAGACCAAAATGGATGCATTTCACTGGATGAATGG Rev CCATTCATCCAGTGAAATGCATCCATTTTGGTCTTTGTCAATGATATCG
35	Thr125Cys	Ser127Cys		9		For GACAAAGACCAAAATGGAGCTATTTGCCTGGATGAATGGAAAGCATAACACC Rev GGTGTATGCTTTCCATTCATCCAGGCAAATAGCTCCATTTTGGTCTTTGTC
36	Asp127Cys	Asp129Cys		11		For CCAAAATGGAGCTATTTCACTGTGCGAATGGAAAGCATAACCAAATCTGC Rev GCAGATTTGGTGTATGCTTTCCATTCGCACAGTGAAATAGCTCCATTTTGG
37	Glu128Cys	Glu130Cys		12		For CCAAAATGGAGCTATTTCACTGGATTGCTGGAAAGCATAACCAAATCTGC Rev GCAGATTTGGTGTATGCTTTCCAGCAATCCAGTGAAATAGCTCCATTTTGG
38	Asn28Cys/Glu35His	Asn30Cys/Glu37His	5, 12			For GGAAGGATCTCTCTTGACCATATGGTCTACAAGGCGTC Rev GACGCCTTGTAGACCATATGGTCAAGAGAGATCCTTCC
39	Asn28His/Glu35Cys	Asn30His/Glu37Cys	5, 12			For GAAGGATCTCTCTTGACTGCATGGTCTACAAGG Rev CCTTGTAGACCATGCAGTCAAGAGAGATCCTTC
40	Asn28Cys/Ser32His	Asn30Cys/Ser34His	5, 9			For GTCAACCACTGCGAAGGATCCATCTTGACGAGATGGTCTACAAG Rev CTTGTAGACCATCTCGTCAAGATGGATCCTTCCGCAGTGGTTGAC

Table 10-11 Primers for mutants Nr 41-48

#	Numbering according to 1EJ3	Numbering according to 1SL8	position in EF-I	position in EF-III	position in EF-IV	Primers, template
41	Asn28Cys/Asp119His	Asn30Cys/Asp121His	5	3		template Asn28Cys, primers for Asp119His
42	Asn28Cys/Asp121His	Asn30Cys/Asp121His	5	5		template Asn28Cys, primers for Asp121His
43	Asn28Cys/Ser125His	Asn30Cys/Ser127His	5	9		template Asn28Cys, primers for Ser125His
44	Asn28Cys/Asp127His	Asn30Cys/Asp129His	5	11		template Asn28Cys, primers for Asp127His
45	Asn28Cys/Asp119Cys	Asn30Cys/Asp121Cys	5	3		template Asn28Cys, primers for Asn28Cys
46	Asn28Cys/Asp121Cys	Asn30Cys/Asp123Cys	5	5		template Asn28Cys, primers for Asp121Cys
47	Asn28Cys/Asp127Cys	Asn30Cys/Asp129Cys	5	11		template Asn28Cys, primers for Asp127Cys
48	Asn28Cys/Ser125Cys	Asn30Cys/Ser127Cys	5	9		template Asn28Cys, primers for Ser125Cys

Table 10-12 Ionic radii of metals according to their coordination and spin state. Original data from Shannon, 1976.

Element	Charge	Coordination	Spin State	Ionic Radius
Copper IV	2	IV		0.57
Copper IVSQ	2	IVSQ		0.57
Cobalt IV	2	IV	High spin	0.58
Zinc IV	2	IV		0.6
Cobalt VI ls	2	VI	Low spin	0.65
Copper V	2	V		0.65
Cobalt V	2	V		0.67
Zinc V	2	V		0.68
Copper VI	2	VI		0.73
Zinc VI	2	VI		0.74
Cobalt VI	2	VI	High spin	0.745
Cadmium IV	2	IV		0.78
Cadmium V	2	V		0.87
Cadmium VI	2	VI		0.95
Lead IVPY	2	IVPY		0.98
Calcium VI	2	VI		1.00
Cadmium VII	2	VII		1.03
Lanthanum VI	3	VI		1.032
Calcium VII	2	VII		1.06
Lanthanum VII	3	VII		1.1
Lead VI	2	VI		1.19
Lead VII	2	VII		1.23

Appendix Chapter 6

Table 10-13 Library OD(600 nm) for aequorin wild-type and mutants. OD values were converted from platereader into spectrometer units as described in Chapter 4, Section 4.4.1. Positions of triplicates are indicated through bold outlines and the wild type internal controls of each plate are highlighted.

plate 1												
<>	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			0.99	2.07	1.42	1.99	1.21	1.64	1.21	0.07		
C			0.90	2.01	1.53	2.02	1.13	1.75	1.23	0.05		
D			1.06	2.00	1.75	2.10	1.26	1.92	1.15	0.06		
E			1.36	2.08	1.62	1.77	1.17	1.12	0.89	0.06		
F			1.29	1.73	1.76	1.46	1.03	1.38	0.81	0.05		
G			1.23	2.10	2.06	1.94	1.02	1.60	0.77	0.09		
H												
plate 2												
<>	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			1.65	1.98	1.60	1.47	2.09	2.05	2.25	1.59		
C			1.64	1.83	1.46	1.39	2.19	2.38	2.18	1.58		
D			1.73	2.18	1.48	1.51	2.13	2.39	2.08	1.45		
E			1.30	1.76	1.04	1.03	1.84	1.61	1.36	1.34		
F			1.19	1.74	1.11	1.18	1.67	1.37	1.49	1.32		
G			1.32	2.10	1.25	1.53	1.84	1.91	1.56	1.42		
H												
plate 3												
<>	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			0.80	1.87	1.67	0.76	1.28	0.74	1.96	1.90		
C			0.86	1.86	0.90	0.95	1.31	0.74	1.90	1.79		
D			0.82	2.00	1.05	0.77	2.36	0.92	1.90	1.80		
E			1.91	1.97	1.99	1.71	1.73	1.03	1.79	1.91		
F			1.99	2.40	1.99	1.58	1.68	1.01	1.80	1.76		
G			1.92	1.93	1.97	1.60	1.94	1.03	1.93	1.81		
H												
plate 4												
<>	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			0.85	2.01	1.33	0.85	0.69	0.68	0.82	0.67		
C			0.75	1.93	0.84	0.71	0.71	0.74	0.92	2.45		
D			0.96	2.15	0.80	0.77	0.74	0.79	0.58	0.56		
E			0.56	1.94	0.66	0.64	0.76	1.79	1.56	1.44		
F			0.84	2.06	0.81	0.69	1.33	1.76	1.05	2.09		
G			0.48	2.09	0.86	0.72	0.75	1.97	1.34	2.03		
H												

```

      10      20      30      40      50      60
Aequorin WT  ATGACCAGCGAACAATACTCAGTCAAGCTTACACCAGACTTCGACAACCCAAAATGGATT
Asn28Cys     ATGACCAGCGAACAATACTCAGTCAAGCTTACACCAGACTTCGACAACCCAAAATGGATT

      70      80      90      100     110     120
Aequorin WT  GGACGACACAAGCACATGTTTAATTTTCTTGATGTCAACCACAAATGGGAGGATCTCTCTT
Asn28Cys     GGACGACACAAGCACATGTTTAATTTTCTTGATGTCAACCACAAATGGGAGGATCTCTCTT

      130     140     150     160     170     180
Aequorin WT  GACGAGATGGTCTACAAGGCGTCCGATATTGTTATAAACAATCTTGGAGCAACACCTGAA
Asn28Cys     GACGAGATGGTCTACAAGGCGTCCGATATTGTTATAAACAATCTTGGAGCAACACCTGAA

      190     200     210     220     230     240
Aequorin WT  CAAGCCAAACGTCACAAAGATGCTGTAGAAGCCTTCTTCGGAGGAGCTGGAATGAAATAT
Asn28Cys     CAAGCCAAACGTCACAAAGATGCTGTAGAAGCCTTCTTCGGAGGAGCTGGAATGAAATAT

      250     260     270     280     290     300
Aequorin WT  GGTGTAGAAACTGAATGGCCTGAATACATCGAAGGATGGAAAAGACTGGCTTCCGAGGAA
Asn28Cys     GGTGTAGAAACTGAATGGCCTGAATACATCGAAGGATGGAAAAGACTGGCTTCCGAGGAA

      310     320     330     340     350     360
Aequorin WT  TTGAAAAGGTATTCAAAAACCAAATCACACTTATTCGTTTATGGGGTGATGCATTGTTC
Asn28Cys     TTGAAAAGGTATTCAAAAACCAAATCACACTTATTCGTTTATGGGGTGATGCATTGTTC

      370     380     390     400     410     420
Aequorin WT  GATATCATTGACAAAGACCAAAATGGAGCTATTTCACTGGATGAATGGAAAGCATAACCC
Asn28Cys     GATATCATTGACAAAGACCAAAATGGAGCTATTTCACTGGATGAATGGAAAGCATAACCC

      430     440     450     460     470     480
Aequorin WT  AAATCTGCTGGCATCATCCAATCGTCAGAAGATTGCGAGGAAACATTAGAGTGTGCGAT
Asn28Cys     AAATCTGCTGGCATCATCCAATCGTCAGAAGATTGCGAGGAAACATTAGAGTGTGCGAT

      490     500     510     520     530     540
Aequorin WT  ATTGATGAAAGTGGACAGCTCGATGTTGATGAGATGACAAGACAACATTTAGGATTTTGG
Asn28Cys     ATTGATGAAAGTGGACAGCTCGATGTTGATGAGATGACAAGACAACATTTAGGATTTTGG

      550     560     570     580     590     600
Aequorin WT  TACACCATGGATCCTGCTTGCGAAAAGCTCTACGGTGGAGCTGTCCCCCTCGAGCACCAC
Asn28Cys     TACACCATGGATCCTGCTTGCGAAAAGCTCTACGGTGGAGCTGTCCCCCTCGAGCACCAC

      610
Aequorin WT  CACCACCACCACTGA
Asn28Cys     CACCACCACCACTGA

```

Figure 10-1 Sequencing results for mutant Asn28Cys. The mutation AAT→TGC, corresponding to Asn→Cys at amino acid position 5 of the loop in EF-I is highlighted.

Appendix Chapter 7

Optimisation of cold incubation conditions

The type and concentration of reducing agents was improved during this experiment. Bacterial pellet containing apoaequorin and harvested from 50 mL culture was incubated overnight in two different ice-cold buffers: (1) 50 mM Tris-HCl, 10 mM EDTA, 5 mM 2-mercaptoethanol and (2) 50 mM Tris-HCl, 10 mM EDTA, 10 mM DTT.

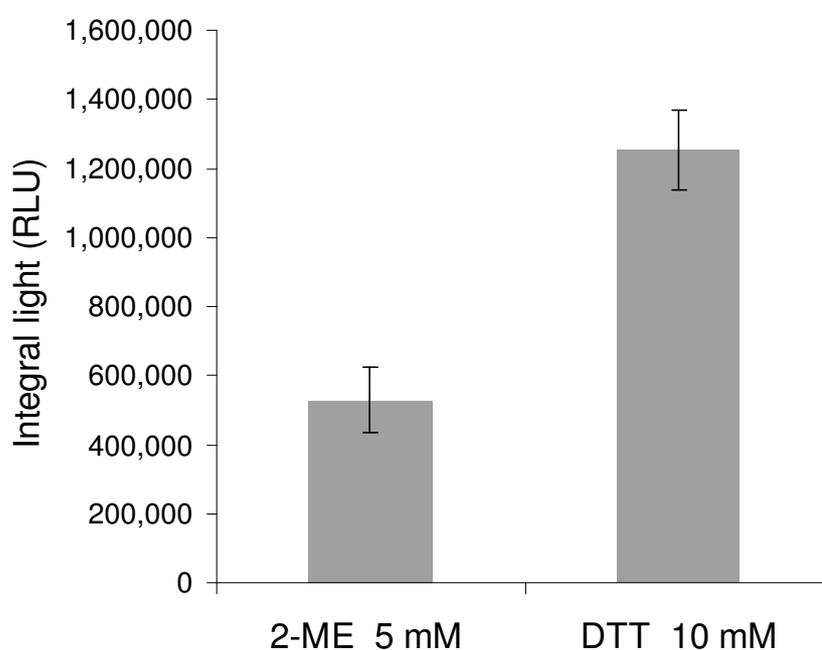


Figure 10-2 Optimisation of reducing agent in charging of aequorin.

Appendix Chapter 8

Table 10-14 Gain settings used for luminescence curves.

	Wild-type aequorin	Mutant Asn28Cys/Ser32His
Calcium	2,500	4,000
Lanthanum	2,500	4,000
Lead	2,500	4,000
Cadmium	2,500	4,000
Cobalt	2,500	4,000
Zinc	3,900	4,000
Copper	3,000	4,000
Zinc+calcium	2,500	4,000

References

- ALLEN, D. G. & BLINKS, J. R. 1978. Calcium transients in aequorin-injected frog cardiac muscle. *Nature*, 273, 509-13.
- ALLEN, D. G., BLINKS, J. R. & PRENDERGAST, F. G. 1977. Aequorin luminescence: relation of light emission to calcium concentration--a calcium-independent component. *Science*, 195, 996-8.
- ALLEN, F. H. 2002. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr B*, 58, 380-8.
- ALVAREZ, J. & MONTERO, M. 2002. Measuring [Ca²⁺] in the endoplasmic reticulum with aequorin. *Cell Calcium*, 32, 251-60.
- ARNOLD, F. H. 1993. Protein engineering for unusual environments. *Curr Opin Biotechnol*, 4, 450-5.
- ARNOLD, F. H. 1998. When blind is better: protein design by evolution. *Nat Biotechnol*, 16, 617-8.
- ARNOLD, F. H. 2001. Combinatorial and computational challenges for biocatalyst design. *Nature*, 409, 253-7.
- ARNOLD, F. H., WINTRODE, P. L., MIYAZAKI, K. & GERSHENSON, A. 2001. How enzymes adapt: lessons from directed evolution. *Trends Biochem Sci*, 26, 100-6.
- ATKINS, P. W. 2006. *Atkins' Physical chemistry*, Oxford, Oxford University Press.
- BACHMANN, T. 2003. Transforming cyanobacteria into bioreporters of biological relevance. *Trends Biotechnol*, 21, 247-9.
- BARKALIFA, R., HERSHFINKEL, M., FRIEDMAN, J. E., KOZAK, A. & SEKLER, I. 2009. The lipophilic zinc chelator DP-b99 prevents zinc induced neuronal death. *Eur J Pharmacol*, 618, 15-21.
- BELKIN, S. 2003. Microbial whole-cell sensing systems of environmental pollutants. *Curr Opin Microbiol*, 6, 206-12.
- BENSON, D. A., KARSCH-MIZRACHI, I., LIPMAN, D. J., OSTELL, J. & WHEELER, D. L. 2005. GenBank. *Nucleic Acids Res*, 33, D34-8.
- BERG, J. M., TYMOCZKO, J. L. & STRYER, L. 2002. *Biochemistry*, New York, W H Freeman.
- BERNARD, P., GABANT, P., BAHASSI, E. M. & COUTURIER, M. 1994. Positive-selection vectors using the F plasmid ccdB killer gene. *Gene*, 148, 71-4.
- BISSWANGER, H. 2008. *Enzyme Kinetics - Principles and Methods*.
- BLINKS, J. R. 1990. Use of photoproteins as intracellular calcium indicators. *Environ Health Perspect*, 84, 75-81.
- BLINKS, J. R., MATTINGLY, P. H., JEWELL, B. R., VAN LEEUWEN, M., HARRER, G. C. & ALLEN, D. G. 2000. Practical aspects of the use of Aequorin as a Calcium Indicator: Assay, Preparation, Microinjection and Interpretation of Signals. San Diego, California: Academic Press.
- BORNSCHEUER, U. F. & KAZLAUSKAS, R. J. 2009. Survey of protein engineering strategies. *Nat Chem Biol*, 5, 526-529.
- BRADFORD, M. M. 1976. A rapid and sensitive for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248-254.
- BRINI, M. 2008. Calcium-sensitive photoproteins. *Methods*, 46, 160-6.
- BRINI, M., DE GIORGI, F., MURGIA, M., MARSAULT, R., MASSIMINO, M. L., CANTINI, M., RIZZUTO, R. & POZZAN, T. 1997. Subcellular analysis of

- Ca²⁺ homeostasis in primary cultures of skeletal muscle myotubes. *Mol Biol Cell*, 8, 129-43.
- BRINI, M., MARSAULT, R., BASTIANUTTO, C., ALVAREZ, J., POZZAN, T. & RIZZUTO, R. 1995. Transfected aequorin in the measurement of cytosolic Ca²⁺ concentration ([Ca²⁺]_c). A critical evaluation. *J Biol Chem*, 270, 9896-903.
- BRODERSEN, D. E., NYBORG, J. & KJELDGAARD, M. 1999. Zinc-binding site of an S100 protein revealed. Two crystal structures of Ca²⁺-bound human psoriasin (S100A7) in the Zn²⁺-loaded and Zn²⁺-free states. *Biochemistry*, 38, 1695-704.
- BROOKS, B. R., BRUCCOLERI, R. E., OLAFSON, D. J., STATES, D. J., SWAMINATHAN, S. & KARPLUS, M. 1983. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry*, 4, 187-217.
- BRUN, M. A., TAN, K. T., NAKATA, E., HINNER, M. J. & JOHNSON, K. 2009. Semisynthetic fluorescent sensor proteins based on self-labeling protein tags. *J Am Chem Soc*, 131, 5873-84.
- CARBONE, M. N. & ARNOLD, F. H. 2007. Engineering by homologous recombination: exploring sequence and function within a conserved fold. *Current Opinion in Structural Biology*, 17, 454-459.
- CHARBONNEAU, H., WALSH, K. A., MCCANN, R. O., PRENDERGAST, F. G., CORMIER, M. J. & VANAMAN, T. C. 1985. Amino acid sequence of the calcium-dependent photoprotein aequorin. *Biochemistry*, 24, 6762-71.
- CHIESA, A., RAPIZZI, E., TOSELLO, V., PINTON, P., DE VIRGILIO, M., FOGARTY, K. E. & RIZZUTO, R. 2001. Recombinant aequorin and green fluorescent protein as valuable tools in the study of cell signalling. *Biochem J*, 355, 1-12.
- CHOPRA, S. & RANGANATHAN, A. 2003. Protein evolution by "codon shuffling": a novel method for generating highly variant mutant libraries by assembly of hexamer DNA duplexes. *Chem Biol*, 10, 917-26.
- COLVIN, R. A., BUSH, A. I., VOLITAKIS, I., FONTAINE, C. P., THOMAS, D., KIKUCHI, K. & HOLMES, W. R. 2008. Insights into Zn²⁺ homeostasis in neurons from experimental and modeling studies. *Am J Physiol Cell Physiol*, 294, C726-42.
- CORPET, F. 1988. Multiple sequence alignment with hierarchical clustering. 16, 9.
- COVALYS BIOSCIENCES, I. 2009. *Snap-tag protein-labelling products* [Online]. Covalys. Available: <http://www.covalys.com> [Accessed 09/09/09].
- CRETON, R., KREILING, J. A. & JAFFE, L. F. 1999. Calcium imaging with chemiluminescence. *Microsc Res Tech*, 46, 390-7.
- CZARNIK, A. W. 1995. Desperately seeking sensors. *Chem Biol*, 2, 423-8.
- DAL PERARO, M., SPIEGEL, K., LAMOUREUX, G., DE VIVO, M., DEGRADO, W. F. & KLEIN, M. L. 2007. Modeling the charge distribution at metal sites in proteins for molecular dynamics simulations. *J Struct Biol*, 157, 444-53.
- DALBY, P. A. 2003. Optimising enzyme function by directed evolution. *Curr Opin Struct Biol*, 13, 500-5.
- DALBY, P. A. 2011. Strategy and success for the directed evolution of enzymes *Current opinion in structural biology*, 21, 8.
- DAUNERT, S., BARRETT, G., FELICIANO, J. S., SHETTY, R. S., SHRESTHA, S. & SMITH-SPENCER, W. 2000. Genetically engineered whole-cell sensing

- systems: coupling biological recognition with reporter genes. *Chem Rev*, 100, 2705-38.
- DEGRADO, W. F. 1997. Proteins from scratch. *Science*, 278, 80-1.
- DELANO, W. L. 2002. The PyMOL Molecular Graphics System. In: SCIENTIFIC, D. (ed.). San Carlos, CA, USA.: DeLano Scientific.
- DEN BOEF, G. & HULANICKI, A. 1983. Recommendations for the usage of selective, selectivity and related terms in analytical chemistry. *Pure and Applied Chemistry*, 55, 4.
- DENG, L., MARKOVA, S. V., VYSOTSKI, E. S., LIU, Z. J., LEE, J., ROSE, J. & WANG, B. C. 2004. Crystal structure of a Ca²⁺-discharged photoprotein: implications for mechanisms of the calcium trigger and bioluminescence. *J Biol Chem*, 279, 33647-52.
- DENG, L., VYSOTSKI, E. S., MARKOVA, S. V., LIU, Z. J., LEE, J., ROSE, J. & WANG, B. C. 2005. All three Ca²⁺-binding loops of photoproteins bind calcium ions: the crystal structures of calcium-loaded apo-aequorin and apo-obelin. *Protein Sci*, 14, 663-75.
- DEO, S. K. & DAUNERT, S. 2001. Luminescent proteins from *Aequorea victoria*: applications in drug discovery and in high throughput analysis. *Fresenius J Anal Chem*, 369, 258-66.
- DEO, S. K., LEWIS, J. C. & DAUNERT, S. 2001. C-terminal and n-terminal fusions of aequorin with small peptides in immunoassay development. *Bioconjug Chem*, 12, 378-84.
- DESANTIS, G., WONG, K., FARWELL, B., CHATMAN, K., ZHU, Z., TOMLINSON, G., HUANG, H., TAN, X., BIBBS, L., CHEN, P., KRETZ, K. & BURK, M. J. 2003. Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). *J. Am Chem Soc*, 125, 11476-77.
- DIARRA, A., SHELDON, C., BRETT, C. L., BAIMBRIDGE, K. G. & CHURCH, J. 1999. Anoxia-evoked intracellular pH and Ca²⁺ concentration changes in cultured postnatal rat hippocampal neurons. *Neuroscience*, 93, 1003-16.
- DIKICI, E., QU, X., ROWE, L., MILLNER, L., LOGUE, C., DEO, S. K., ENSOR, M. & DAUNERT, S. 2009. Aequorin variants with improved bioluminescence properties. *Protein Eng Des Sel*, 22, 243-8.
- DORLAND, W. 2010. Dorland's Illustrated Medical Dictionary. In: ANDERSON, D. (ed.). <http://medical-dictionary.thefreedictionary.com>.
- DRAKE, S. K. & FALKE, J. J. 1996. Kinetic tuning of the EF-hand calcium binding motif: the gateway residue independently adjusts (i) barrier height and (ii) equilibrium. *Biochemistry*, 35, 1753-60.
- DUDEV, T. & LIM, C. 2003. Principles governing Mg, Ca, and Zn binding and selectivity in proteins. *Chem Rev*, 103, 773-88.
- DUMON, C., VARVAK, A., WALL, M. A., FLINT, J. E., LEWIS, R. J., LAKEY, J. H., MORLAND, C., LUGINBÜHL, P., HEALEY, S., TODARO, T., DESANTIS, G., SUN, M., PARRA-GESSERT, L., TAN, X., WEINER, D. P. & GILBERT, H. J. 2008. Engineering hyperthermostability into a GH11 xylanase is mediated by subtle changes to protein structure. *J Biol Chem.*, 283, 22557-64.
- DUPRIEZ VJ, M. K., LE POUL E, BURGEON E, DETHEUX M. 2002. Aequorin-based functional assays for G-protein-coupled receptors, ion channels, and tyrosine kinase receptors. *Receptors Channels*, 8, 319-330.

- FALLER, P. 2009. Copper and zinc binding to amyloid-beta: coordination, dynamics, aggregation, reactivity and metal-ion transfer. *ChemBiochem*, 10, 2837-45.
- FALLER, P. & HUREAU, C. 2009. Bioinorganic chemistry of copper and zinc ions coordinated to amyloid-beta peptide. *Dalton Trans*, 1080-94.
- FRIARD, O. 2004. AnnHyb: a tool for working with and managing nucleotide sequences in multiple formats. 4.913 ed.: Bioinformatics Organisation <http://www.bioinformatics.org/annhyb>.
- GARZON-RODRIGUEZ, W., YATSIMIRSKY, A. K. & GLABE, C. G. 1999. Binding of Zn(II), Cu(II), and Fe(II) ions to Alzheimer's A beta peptide studied by fluorescence. *Bioorg Med Chem Lett*, 9, 2243-8.
- GASTEIGER, E., GATTIKER, A., HOOGLAND, C., IVANYI, I., APPEL, R. D. & BAIROCH, A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.*, 31, 3784-8.
- GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. 2005. *Compute pI / Mw Tool* [Online]. Available: http://www.expasy.org/tools/pi_tool-doc.html [Accessed 18/09/07].
- GE HEALTHCARE, B.-S. 2006. Instructions 71-7127-00 AF: Gel Supedex Filtration Media. Uppsala, Sweden: GE Healthcare Biosciences.
- GIFFORD, J. L., WALSH, M. P. & VOGEL, H. J. 2007. Structures and metal-ion-binding properties of the Ca²⁺-binding helix-loop-helix EF-hand motifs. *Biochem J*, 405, 199-221.
- GLYNOU, K., IOANNOU, P. C. & CHRISTOPOULOS, T. K. 2003. One-step purification and refolding of recombinant photoprotein aequorin by immobilized metal-ion affinity chromatography. *Protein Expr Purif*, 27, 384-90.
- GODDARD, J. P. & REYMOND, J. L. 2004. Recent advances in enzyme assays. *Trends Biotechnol*, 22, 363-70.
- GOODRICH, J. A. & KUGEL, J. F. 2006. *Binding and Kinetics for Molecular Biologists*, New York, Cold Spring Harbor Laboratory Press.
- GREER, L. F., 3RD & SZALAY, A. A. 2002. Imaging of light emission from the expression of luciferases in living cells and organisms: a review. *Luminescence*, 17, 43-74.
- GU, C., CHEN, S., XU, X., ZHENG, L., LI, Y., WU, K., LIU, J., QI, Z., HAN, D., CHEN, G. & HUO, X. 2009. Lead and cadmium synergistically enhance the expression of divalent metal transporter 1 protein in central nervous system of developing rats. *Neurochem Res*, 34, 1150-6.
- HADDOCK, S. H., RIVERS, T. J. & ROBISON, B. H. 2001. Can coelenterates make coelenterazine? Dietary requirement for luciferin in cnidarian bioluminescence. *Proc Natl Acad Sci U S A*, 98, 11148-51.
- HADDOCK, S. H. D. M., C.M.; CASE. 2004. *The Bioluminescence Web Page* [Online]. University of California, Santa Barbara Available: <http://www.lifesci.ucsb.edu/~biolum/> [Accessed 01/03/04].
- HALL, T. 2004. BioEdit. 7.0.4.1. ed. Carlsbad, CA Ibis Biosciences
- HALL, T. 2007. BioEdit. 7.0.4.1. ed. Carlsbad, CA Ibis Biosciences
- HAN, D., WANG, H. & YANG, P. 2008. Molecular modeling of zinc and copper binding with Alzheimer's amyloid beta-peptide. *Biometals*, 21, 189-96.
- HARDING, M. M. 2004. The architecture of metal coordination groups in proteins. *Acta Crystallogr D Biol Crystallogr*, 60, 849-59.
- HARDING, M. M. 2006. Small revisions to predicted distances around metal sites in proteins. *Acta Crystallogr D Biol Crystallogr*, 62, 678-82.

- HARVEY, E. N. 1952. *Bioluminescence*, New York, New York Academy of Sciences.
- HASTINGS, J. W., MITCHELL, G., MATTINGLY, P. H., BLINKS, J. R. & VAN LEEUWEN, M. 1969. Response of aequorin bioluminescence to rapid changes in calcium concentration. *Nature*, 222, 1047-50.
- HAY, M., RICHARDS, J. H. & LU, Y. 1996. Construction and characterization of an azurin analog for the purple copper site in cytochrome c oxidase. *Proc Natl Acad Sci U S A*, 93, 461-4.
- HEAD J.F., I. S., TERANISHI K., SHIMOMURA O. 2000. PDB ID: 1EJ3 Crystal Structure of Aequorin 2000 ed.: RCSB Protein Data Bank.
- HEAD, J. F., INOUE, S., TERANISHI, K. & SHIMOMURA, O. 2000. The crystal structure of the photoprotein aequorin at 2.3 Å resolution. *Nature*, 405, 372-6.
- HERRING, P. J. 1978. *Bioluminescence in action*, London : Academic Press.
- HILL, A. V. 1910. The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *The Journal of Physiology*, 40, iv–vii.
- HOGREFE, H. H., CLINE, J., YOUNGBLOOD, G. L. & ALLEN, R. M. 2002. Creating randomized amino acid libraries with the QuikChange Multi Site-Directed Mutagenesis Kit. *Biotechniques*, 33, 1158-60, 1162, 1164-5.
- HOVDE, S. 2011. *Course notes: principles of organic and biological chemistry related to the health science field* [Online]. Hartnell University, Department of Chemistry. Available: <http://www.hartnell.edu/faculty/showde/chem23/EnzymeInhibition.htm>.
- INOUE, S., AOYAMA, S., MIYATA, T., TSUJI, F. I. & SAKAKI, Y. 1989. Overexpression and purification of the recombinant Ca²⁺-binding protein, apoaequorin. *J Biochem*, 105, 473-7.
- INOUE, S., NOGUCHI, M., SAKAKI, Y., TAKAGI, Y., MIYATA, T., IWANAGA, S. & TSUJI, F. I. 1985. Cloning and sequence analysis of cDNA for the luminescent protein aequorin. *Proc Natl Acad Sci U S A*, 82, 3154-8.
- INOUE, S., OHMIYA, Y., TOYA, Y. & TSUJI, F. I. 1992. Imaging of luciferase secretion from transformed Chinese hamster ovary cells. *Proc Natl Acad Sci U S A*, 89, 9584-7.
- INVITROGEN 2004. Instruction Manual: Zero Blunt TOPO PCR Cloning Kit. *Version M*.
- IUPAC 1997. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). created by . . . In: NIC, M., JIRAT J. & KOSATAB. (eds.) *Glossary for chemists of terms used in biotechnology* Oxford: Blackwell Scientific Publications.
- IUPAC. 2010. *IUPAC Glossary of Terms Used in Toxicology* [Online]. Bethesda. Available: <http://sis.nlm.nih.gov/enviro/iupacglossary/glossaryi.html> [Accessed 13/11/10].
- IZUTSU, K. T., FELTON, S. P., SIEGEL, I. A., NICHOLLS, J. I., CRAWFORD, J., MCGOUGH, J. & YODA, W. T. 1974. An investigation of the ionic calcium assay based on the jellyfish protein aequorin. *Anal Biochem*, 58, 479-84.
- IZUTSU, K. T., FELTON, S. P., SIEGEL, I. A., YODA, W. T. & CHEN, A. C. 1972. Aequorin: its ionic specificity. *Biochem Biophys Res Commun*, 49, 1034-9.
- KAPLAN, J. & DEGRADO, W. F. 2004. De novo design of catalytic proteins. *Proc Natl Acad Sci U S A*, 101, 11566-70.
- KARPLUS, M. & KURIYAN, J. 2005. Molecular dynamics and protein function. *Proc Natl Acad Sci U S A*, 102, 6679-85.

- KARR, J. W., KAUPP, L. J. & SZALAI, V. A. 2004. Amyloid-beta binds Cu²⁺ in a mononuclear metal ion binding site. *J Am Chem Soc*, 126, 13534-8.
- KAZLAUSKAS, R. J. 2005. Enhancing catalytic promiscuity for biocatalysis. *Current opinion in structural biology*, 9, 195-201.
- KAZLAUSKAS, R. J. & BORNSCHEUER, U. T. 2009. Finding better protein engineering strategies. *Nat Chem Biol*, 5, 526-529.
- KEANE, A., PHOENIX, P., GHOSHAL, S. & LAU, P. C. 2002. Exposing culprit organic pollutants: a review. *J Microbiol Methods*, 49, 103-19.
- KENDALL, J. M., SALA-NEWBY, G., GHALAUT, V., DORMER, R. L. & CAMPBELL, A. K. 1992. Engineering the CA(2+)-activated photoprotein aequorin with reduced affinity for calcium. *Biochem Biophys Res Commun*, 187, 1091-7.
- KEPLER, A., GENDREIZIG, S., GRONEMEYER, T., PICK, H., VOGEL, H. & JOHNSSON, K. 2003. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.*, 21, 86-89.
- KIKUCHI, K., KOMATSU, K. & NAGANO, T. 2004. Zinc sensing for cellular application. *Curr Opin Chem Biol*, 8, 182-91.
- KOHLMEIER, S., MANCUSO, M., TECON, R., HARMS, H., VAN DER MEER, J. R. & WELLS, M. 2007. Bioreporters: gfp versus lux revisited and single-cell response. *Biosens Bioelectron*, 22, 1578-85.
- KUPPURAJ, G., DUDEV, M. & LIM, C. 2009. Factors governing metal-ligand distances and coordination geometries of metal complexes. *J Phys Chem B*, 113, 2952-60.
- LAWSON, A. J., WALKER, E. A., WHITE, S. A., DAFFORN, T. R., STEWART, P. M. & RIDE, J. P. 2009. Mutations of key hydrophobic surface residues of 11 beta-hydroxysteroid dehydrogenase type 1 increase solubility and monodispersity in a bacterial expression system. *Protein Science*, 18, 1552-63.
- LE CLAINCHE, L., FIGUET, M., MONTJARDET-BAS, V., BLANCHARD, S. & VITA, C. 2006. Modulating the affinity and the selectivity of engineered calmodulin EF-Hand peptides for lanthanides. *Biotechnol Bioeng*, 95, 29-36.
- LE CLAINCHE, L., PLANCQUE, G., AMEKRAZ, B., MOULIN, C., PRADINES-LECOMTE, C., PELTIER, G. & VITA, C. 2003. Engineering new metal specificity in EF-hand peptides. *J Biol Inorg Chem*, 8, 334-40.
- LEWIS, J. C. & DAUNERT, S. 2000. Photoproteins as luminescent labels in binding assays. *Fresenius J Anal Chem*, 366, 760-8.
- LIU, Z. J., STEPANYUK, G. A., VYSOTSKI, E. S., LEE, J., MARKOVA, S. V., MALIKOVA, N. P. & WANG, B. C. 2006. Crystal structure of obelin after Ca²⁺-triggered bioluminescence suggests neutral coelenteramide as the primary excited state. *Proc Natl Acad Sci U S A*, 103, 2570-5.
- LIU, Z. J., VYSOTSKI, E. S., DENG, L., LEE, J., ROSE, J. & WANG, B. C. 2003. Atomic resolution structure of obelin: soaking with calcium enhances electron density of the second oxygen atom substituted at the C2-position of coelenterazine. *Biochem Biophys Res Commun*, 311, 433-9.
- LOSCHEN, G. & CHANCE, B. 1971. Rapid kinetic studies of the light emitting protein aequorin. *Nature: New biology*, 11, 2.
- LU, Y. & VALENTINE, J. S. 1997. Engineering metal-binding sites in proteins. *Curr Opin Struct Biol*, 7, 495-500.
- LUTZ, S., OSTERMEIER, M., MOORE, G. L., MARANAS, C. D. & BENKOVIC, S. J. 2001. Creating multiple-crossover DNA libraries independent of sequence identity. *Proc Natl Acad Sci U S A*, 98, 11248-53.

- MACKEREL JR, A. D., BROOKS III, C. L., NILSSON, L., ROUX, B., WON, Y. & KARPLUS, M. 1998. The Encyclopedia of Computational Chemistry. *CHARMM: The Energy Function and Its Parameterization with an Overview of the Program*. John Wiley & Sons: Chichester.
- MARINO, T., RUSSO, N., TOSCANO, M. & PAVELKA, M. 2010. On the metal ion (Zn(2+), Cu(2+)) coordination with beta-amyloid peptide: DFT computational study. *Interdiscip Sci*, 2, 57-69.
- MARTELL, A. E. & SMITH, R. M. (eds.) 1982. *Critical Stability Constants*, New York: Plenum Press.
- MARTIN, J.-R. 2008a. In Vivo Brain Imaging: Fluorescence or Bioluminescence, Which to Choose? *Journal of Neurogenetics*, 285 - 307.
- MARTIN, J. R. 2008b. In Vivo Brain Imaging: Fluorescence or Bioluminescence, Which to Choose? *J Neurogenet*, 1-23.
- MERRIAM-WEBSTER 2009. Biosensor. In: ONLINE, M.-W. (ed.). Springfield, MA 01102: Merriam-Webster, Incorporated.
- MILLS, C. E. 2008. *Bioluminescence of Aequorea, a hydromedusa* [Online]. Mills, C.E. Available: <http://faculty.washington.edu/cemills/Aequorea.html> [Accessed 05/11/2008].
- MIURA, T., SUZUKI, K., KOHATA, N. & TAKEUCHI, H. 2000. Metal binding modes of Alzheimer's amyloid beta-peptide in insoluble aggregates and soluble complexes. *Biochemistry*, 39, 7024-31.
- MIYAWAKI, A. 2008. Green fluorescent protein glows gold. *Cell*, 135, 987-90.
- MIYAZAKI, K. & TAKENOUCI, M. 2002. Creating random mutagenesis libraries using megaprimer PCR of whole plasmid. *Biotechniques*, 33, 1033-4, 1036-8.
- MOISESCU, D. G., ASHLEY, C. C. & CAMPBELL, A. K. 1975. Comparative aspects of the calcium-sensitive photoproteins aequorin and obelin. *Biochim Biophys Acta*, 396, 133-140.
- MOLECULAR PROBES. 2001. *AquaLite Recombinant Aequorin (A-6785)* [Online]. Available: <http://probes.invitrogen.com/media/pis/mp06785.pdf> [Accessed 05/09/03].
- MONTERO, M., BRINI, M., MARSAULT, R., ALVAREZ, J., SITIA, R., POZZAN, T. & RIZZUTO, R. 1995. Monitoring dynamic changes in free Ca²⁺ concentration in the endoplasmic reticulum of intact cells. *EMBO J*, 14, 5467-75.
- MORINAGA, Y., FRANCESCHINI, T., INOUE, S. & INOUE, M. 1984. Improvement Of Oligonucleotide-Directed Site-Specific Mutagenesis Using Double-Stranded Plasmid DNA. *Nature Biotechnology*, 2, 636-9.
- MORLEY, K. L. & KAZLAUSKAS, R. J. 2005. Improving enzyme properties: when are closer mutations better? *Trends Biotechnol*, 23, 231-7.
- MOTULSKY, H. & CHRISTOPOULOS, A. 2003. Fitting models to biological data using linear and non-linear regression. A practical guide to curve fitting.
- MYERS, P., R. ESPINOSA, C. S. PARR, T. JONES, G. S. HAMMOND, AND T. A. DEWEY. 2009. *The Animal Diversity Web* [Online]. Available: http://animaldiversity.ummz.umich.edu/site/accounts/classification/path/Aequorea_victoria.html#Aequorea%20victoria [Accessed 10/12/09].
- NANDA, V., ROSENBLATT, M. M., OSYCZKA, A., KONO, H., GETAHUN, Z., DUTTON, P. L., SAVEN, J. G. & DEGRADO, W. F. 2005. De novo design of a redox-active minimal rubredoxin mimic. *J Am Chem Soc*, 127, 5804-5.
- NANOLIGHT, T. 2010. *Product Technical Note: Aequorin products* [Online]. Available: <http://nanolight.com> [Accessed 01/03/10].

- NOMURA, M., INOUE, S., OHMIYA, Y. & TSUJI, F. I. 1991. A C-terminal proline is required for bioluminescence of the Ca(2+)-binding photoprotein, aequorin. *FEBS Lett*, 295, 63-6.
- NOVAGEN 2003. pET System Manual. 10th Rev. B 0403 ed.
- OHASHI, W., INOUE, S., YAMAZAKI, T. & HIROTA, H. 2005. NMR analysis of the Mg²⁺-binding properties of aequorin, a Ca²⁺-binding photoprotein. *J Biochem*, 138, 613-20.
- OHMIYA, Y. & HIRANO, T. 1996. Shining the light: the mechanism of the bioluminescence reaction of calcium-binding photoproteins. *Chem Biol*, 3, 337-47.
- OISHI, O., NAGATOMO, A., KOHZUMA, T., ODA, N., MIYAZIMA, T., OHNO, M. & SAKAKI, Y. 1992. Validity of putative calcium binding loops of photoprotein aequorin. *FEBS Lett*, 307, 272-4.
- OLD, C. W. 1994. *Principles of gene manipulation*
- PACE, C. N., VAJDOS, F., FEE, L., GRIMSLEY, G. & GRAY, T. 1995. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci*, 4, 2411-23.
- PALMER, A. E. 2009. Expanding the repertoire of fluorescent calcium sensors. *ACS Chem Biol*, 4, 157-9.
- PALMER, J. 2002. *Metalloprotein Program Project Overview* [Online]. The Scripps Research Institute. Available: <http://www.scripps.edu/research/metallo/> [Accessed 16/01/03].
- PATTON, C. 2002. MAXCHELATOR: a program for determining the free metal concentration in the presence of chelators. v 2.51 ed. Hopkins Marine Station, Pacific Grove, CA 93950: Stanford University.
- PERSSON, B.-A. & VESSMAN, J. 2001. The use of selectivity in analytical chemistry - some considerations. *TrAC Trends in Analytical Chemistry*, 20, 526-532.
- PINTON, P., RIMESSI, A., ROMAGNOLI, A., PRANDINI, A. & RIZZUTO, R. 2007. Biosensors for the detection of calcium and pH. *Methods Cell Biol*, 80, 297-325.
- PLÜCKTHUN, A. & MAYO, S. L. 2007. The design of evolution and the evolution of design *Current Opinion in Structural Biology*, 17, 451-453.
- PRASHER, D., MCCANN, R. O. & CORMIER, M. J. 1985. Cloning and expression of the cDNA coding for aequorin, a bioluminescent calcium-binding protein. *Biochem Biophys Res Commun*, 126, 1259-68.
- PRENDERGAST, F. G. 2000. Bioluminescence illuminated. *Nature*, 405, 291-3.
- REGAN, L. 1995. Protein design: novel metal-binding sites. *Trends Biochem Sci*, 20, 280-5.
- ROURA, S., GÁLVEZ-MONTÓN, C. & BAYES-GENIS, A. 2013. Bioluminescence imaging: a shining future for cardiac regeneration. *J Cell Mol Med.*, 17, 693-703.
- ROWE, L., COMBS, K., DEO, S., ENSOR, C., DAUNERT, S. & QU, X. 2008a. Genetically modified semisynthetic bioluminescent photoprotein variants: simultaneous dual-analyte assay in a single well employing time resolution of decay kinetics. *Anal Chem*, 80, 8470-6.
- ROWE, L., ROTHERT, A., LOGUE, C., ENSOR, C. M., DEO, S. K. & DAUNERT, S. 2008b. Spectral tuning of photoproteins by partnering site-directed mutagenesis strategies with the incorporation of chromophore analogs. *Protein Eng Des Sel*, 21, 73-81.

- RULISEK, L. & VONDRASEK, J. 1998. Coordination geometries of selected transition metal ions (Co²⁺, Ni²⁺, Cu²⁺, Zn²⁺, Cd²⁺, and Hg²⁺) in metalloproteins. *J Inorg Biochem*, 71, 115-27.
- SALA-NEWBY, G. B., BADMINTON, M. N., EVANS, W. H., GEORGE, C. H., JONES, H. E., KENDALL, J. M., RIBEIRO, A. R. & CAMPBELL, A. K. 2000. Targeted bioluminescent indicators in living cells. *Methods Enzymol*, 305, 479-98.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. 1989. *Molecular cloning: A Laboratory Manual*, New York, Cold Spring Harbor Laboratory Press.
- SCHMIDT, A. M., MULLER, H. N. & SKERRA, A. 1996. A Zn(II)-binding site engineered into retinol-binding protein exhibits metal-ion specificity and allows highly efficient affinity purification with a newly designed metal ligand. *Chem Biol*, 3, 645-53.
- SCOTT, D., DIKICI, E., ENSOR, M. & DAUNERT, S. 2011. Bioluminescence and its impact on bioanalysis. *Annu Rev Anal Chem (Palo Alto Calif)*, 4, 297-319.
- SEEBECK, F. P. & HILVERT, D. 2003. Conversion of a PLP-dependent racemase into an aldolase by a single active site mutation. *J. Am Chem Soc*, 125, 10158-9.
- SENSI, S. L., TON-THAT, D., WEISS, J. H., ROTHE, A. & GEE, K. R. 2003. A new mitochondrial fluorescent zinc sensor. *Cell Calcium*, 34, 281-4.
- SENSI, S. L., YIN, H. Z., CARRIEDO, S. G., RAO, S. S. & WEISS, J. H. 1999. Preferential Zn²⁺ influx through Ca²⁺-permeable AMPA/kainate channels triggers prolonged mitochondrial superoxide production. *Proc Natl Acad Sci U S A*, 96, 2414-9.
- SHANNON, R. D. 1976. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallographica. (1976)*, A32, 751-767.
- SHIMOMURA, O. 1986. Isolation and properties of various molecular forms of aequorin. *Biochem J*, 234, 271-7.
- SHIMOMURA, O. 1991. Preparation and handling of aequorin solutions for the measurement of cellular Ca²⁺. *Cell Calcium*, 12, 635-43.
- SHIMOMURA, O. 1995a. Luminescence of aequorin is triggered by the binding of two calcium ions. *Biochem Biophys Res Commun*, 211, 359-63.
- SHIMOMURA, O. 1995b. A short story of aequorin. *Biol Bull*, 189, 1-5.
- SHIMOMURA, O. 2005. The discovery of aequorin and green fluorescent protein. *J Microsc*, 217, 1-15.
- SHIMOMURA, O. 2006. *Bioluminescence: Chemical principles and methods*, World Scientific Pub Co Inc.
- SHIMOMURA, O. & INOUE, S. 1996. Titration of recombinant aequorin with calcium chloride. *Biochem Biophys Res Commun*, 221, 77-81.
- SHIMOMURA, O. & INOUE, S. 1999. The in situ regeneration and extraction of recombinant aequorin from Escherichia coli cells and the purification of extracted aequorin. *Protein Expr Purif*, 16, 91-5.
- SHIMOMURA, O., INOUE, S., MUSICKI, B. & KISHI, Y. 1990. Recombinant aequorin and recombinant semi-synthetic aequorins. Cellular Ca²⁺ ion indicators. *Biochem J*, 270, 309-12.
- SHIMOMURA, O. & JOHNSON, F. H. 1969. Properties of the bioluminescent protein aequorin. *Biochemistry*, 8, 3991-3997.
- SHIMOMURA, O. & JOHNSON, F. H. 1970. Calcium binding, quantum yield, and emitting molecule in aequorin bioluminescence. *Nature*, 227, 1356-7.

- SHIMOMURA, O. & JOHNSON, F. H. 1972. Structure of the light-emitting moiety of aequorin. *Biochemistry*, 11, 1602-8.
- SHIMOMURA, O. & JOHNSON, F. H. 1973. Further data on the specificity of aequorin luminescence to calcium. *Biochem Biophys Res Commun*, 53, 490-4.
- SHIMOMURA, O. & JOHNSON, F. H. 1975a. Chemical nature of bioluminescence systems in coelenterates. *Proc Natl Acad Sci U S A*, 72, 1546-9.
- SHIMOMURA, O. & JOHNSON, F. H. 1975b. Regeneration of the photoprotein aequorin. *Nature*, 256, 236-8.
- SHIMOMURA, O., JOHNSON, F. H. & SAIGA, Y. 1962. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *J Cell Comp Physiol*, 59, 223-39.
- SHIMOMURA, O., MUSICKI, B., KISHI, Y. & INOUE, S. 1993. Light-emitting properties of recombinant semi-synthetic aequorins and recombinant fluorescein-conjugated aequorin for measuring cellular calcium. *Cell Calcium*, 14, 373-8.
- SHIMOMURA, O. & SHIMOMURA, A. 1981. Resistivity to denaturation of the apoprotein of aequorin and reconstitution of the luminescent photoprotein from the partially denatured apoprotein. *Biochem J*, 199, 825-8.
- SHIMOMURA, O. & SHIMOMURA, A. 1984. Effect of calcium chelators on the Ca²⁺-dependent luminescence of aequorin. *Biochem J*, 221, 907-10.
- SHIMOMURA, O. & TERANISHI, K. 2000. Light-emitters involved in the luminescence of coelenterazine. *Luminescence*, 15, 51-8.
- SHIMOMURA, O., TOMA, S., CHONG, K. T., NAKAGAWA, A., TERANISHI, K., INOUE, S., INOUE, S., GODA, M., SHRIBAK, M., TRAN, P. T., HEAD, J. F., KISHI, Y., MUSICKI, B., CAMPBELL, A. K., SALA-NEWBY, G., ASTON, P. J., KALSHEKER, N., SHIMOMURA, A., BAKER, P. F., SCHAPIRA, A. H., JOHNSON, F. H., MORISE, H., HORI, K., CORMIER, M. J., JOHNSON, F. H., KOHAMA, Y. & SAIGA, Y. 2009. Discovery of green fluorescent protein (GFP) (Nobel Lecture). *Angew Chem Int Ed Engl*, 48, 5590-602.
- SHRESTHA, S., PAENG, I. R., DEO, S. K. & DAUNERT, S. 2002. Cysteine-free mutant of aequorin as a photolabel in immunoassay development. *Bioconjug Chem*, 13, 269-75.
- SIEBER, V., MARTINEZ, C. A. & ARNOLD, F. H. 2001. Libraries of hybrid proteins from distantly related sequences. *Nat Biotechnol*, 19, 456-60.
- SNYDER, E. E., BUOSCIO, B. W. & FALKE, J. J. 1990. Calcium(II) site specificity: effect of size and charge on metal ion binding to an EF-hand-like site. *Biochemistry*, 29, 3937-43.
- STEMMER, W. P. 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc Natl Acad Sci U S A*, 91, 10747-51.
- STRATAGENE 2004a. pBluescript II Phagemid Vectors.
- STRATAGENE 2004b. QuickChange II Site-Directed Mutagenesis Kit.
- STRATAGENE. 2005. *QuikChange® II Site-Directed Mutagenesis Kit* [Online]. Agilent Technologies. Available: <http://www.stratagene.com/products/displayProduct.aspx?pid=573> [Accessed 18/01/06].
- STRELTSOV, V. A., TITMUSS, S. J., EPA, V. C., BARNHAM, K. J., MASTERS, C. L. & VARGHESE, J. N. 2008. The structure of the amyloid-beta peptide high-affinity copper II binding site in Alzheimer disease. *Biophys J*, 95, 3447-56.

- STRYNADKA, N. C. & JAMES, M. N. 1989. Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu Rev Biochem*, 58, 951-98.
- TALMARD, C., BOUZAN, A. & FALLER, P. 2007. Zinc binding to amyloid-beta: isothermal titration calorimetry and Zn competition experiments with Zn sensors. *Biochemistry*, 46, 13658-66.
- TANAHASHI, H., ITO, T., INOUE, S., TSUJI, F. I. & SAKAKI, Y. 1990. Photoprotein aequorin: use as a reporter enzyme in studying gene expression in mammalian cells. *Gene*, 96, 249-55.
- THOMPSON, J. D., HIGGINS, D.G. AND GIBSON, T.J 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22:4673-4680.
- TOMA, S., CHONG, K. T., NAKAGAWA, A., TERANISHI, K., INOUE, S. & SHIMOMURA, O. 2005. The crystal structures of semi-synthetic aequorins. *Protein Sci*, 14, 409-16.
- TOUGU, V., KARAFIN, A. & PALUMAA, P. 2008. Binding of zinc(II) and copper(II) to the full-length Alzheimer's amyloid-beta peptide. *J Neurochem*, 104, 1249-59.
- TRAN, M. C. 1993. *Biosensors*, London, Chapman & Hall.
- TRICOIRE, L., TSUZUKI, K., COURJEAN, O., GIBELIN, N., BOUROUT, G., ROSSIER, J. & LAMBOLEZ, B. 2006. Calcium dependence of aequorin bioluminescence dissected by random mutagenesis. *Proc Natl Acad Sci U S A*, 103, 9500-5.
- TRUONG, K., SAWANO, A., MIZUNO, H., HAMA, H., TONG, K. I., MAL, T. K., MIYAWAKI, A. & IKURA, M. 2001. FRET-based in vivo Ca²⁺ imaging by a new calmodulin-GFP fusion molecule. *Nat Struct Biol*, 8, 1069-73.
- TSUJI, F. I., INOUE, S., GOTO, T. & SAKAKI, Y. 1986. Site-specific mutagenesis of the calcium-binding photoprotein aequorin. *Proc Natl Acad Sci U S A*, 83, 8107-8111.
- TSUJI, F. I., OHMIYA, Y., FAGAN, T. F., TOH, H. & INOUE, S. 1995. Molecular evolution of the Ca(2+)-binding photoproteins of the Hydrozoa. *Photochem Photobiol*, 62, 657-61.
- TSUZUKI, K., TRICOIRE, L., COURJEAN, O., GIBELIN, N., ROSSIER, J. & LAMBOLEZ, B. 2005. Thermostable mutants of the photoprotein aequorin obtained by in vitro evolution. *J Biol Chem*, 280, 34324-31.
- ULMER, K. M. 1983. Protein engineering. *Science*, 219, 666-71.
- VAN DYKE, K., VAN DYKE, C. & WOODFORK, K. 2002. *Luminescence Biotechnology. Instruments and Applications*, Boca Raton, Fla. ; London : CRC Press.
- VINKENBORG, J. L. 2010. *Fluorescent sensor proteins for intracellular metal imaging*. PhD Thesis, Eindhoven University of Technology.
- VINKENBORG, J. L., NICOLSON, T. M., BELLOMO, E. A., KOAY, M. S., RUTTER, G. A. & MERKX, M. 2009. Genetically encoded FRET sensors to monitor intracellular Zn²⁺ homeostasis. *Nature Methods*, 6, 4.
- VIPOND, I. B., MOON, B. J. & HALFORD, S. E. 1996. An isoleucine to leucine mutation that switches the cofactor requirement of the EcoRV restriction endonuclease from magnesium to manganese. *Biochemistry*, 35, 1712-21.
- VIVASCIENCE, L. 2002. Technical Data and Operating Instructions for Vivaspin 6 and 20. Hannover.

- VYSOTSKI, E. S. & LEE, J. 2004. Ca²⁺-regulated photoproteins: structural insight into the bioluminescence mechanism. *Acc Chem Res*, 37, 405-15.
- WALDRON, K. J., RUTHERFORD, J. C., FORD, D. & ROBINSON, N. J. 2009. Metalloproteins and metal sensing. *Nature*, 460, 8.
- WATKINS, N. J. & CAMPBELL, A. K. 1993. Requirement of the C-terminal proline residue for stability of the Ca(2+)-activated photoprotein aequorin. *Biochem J.*, 293, 181-5.
- WEISS, J. N. 1997. The Hill equation revisited: uses and misuse. *FASEB J*, 11, 7.
- WICKNER, W., DRIESSEN, A. J. & HARTL, F. U. 1991. The enzymology of protein translocation across the Escherichia coli plasma membrane. *Annu Rev Biochem*, 60, 101-24.
- WONG, V. 2006. *Cold Light* [Online]. London: Royal Society of Chemistry. Available: <http://www.rsc.org/education/teachers/learnnet/inspirational/resources/7.4.pdf> [Accessed 04/12/06].
- YAGI, K. 2007. Applications of whole-cell bacterial sensors in biotechnology and environmental science. *Appl Microbiol Biotechnol*, 73, 1251-8.
- YOU, L. & ARNOLD, F. H. 1996. Directed evolution of subtilisin E in Bacillus subtilis to enhance total activity in aqueous dimethylformamide. *Protein Eng*, 9, 77-83.
- ZHAO, H., GIVER, L., SHAO, Z., AFFHOLTER, J. A. & ARNOLD, F. H. 1998. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat Biotechnol*, 16, 258-61.
- ZHOU, Y. H., ZHANG, X. P. & EBRIGHT, R. H. 1991. Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. *Nucleic Acids Res*, 19, 6052.