

UNIVERSITY OF LONDON

**V2D: Virtual Two-Dimensional
Capillary Electrophoresis for Protein
Separation and Identification**

by

Andrea Renata Laine

A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Imperial College London
Department of Chemistry

January 2010

Declaration of Authorship

I, Andrea Renata Laine, declare that this thesis titled, 'V2D: Virtual Two-Dimensional Capillary Electrophoresis for Protein Separation and Identification' and the work presented in it are my own. I confirm that no part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution.

Signed:

Date:

Abstract

V2D: Virtual Two-Dimensional Capillary Electrophoresis for Protein Separation and Identification

Andrea Renata Laine

Virtual two-dimensional (V2D) separation is an unprecedented approach to protein separation, offering superior speed, resolution, and quantification compared to conventional techniques. Instead of performing two separations orthogonally, virtual two-dimensional electrophoresis combines two collinear separations. Time- and space-resolved measurement are employed to track and reconstruct protein migration. Unlike in conventional two-dimensional electrophoresis, the resolution is limited by the resolution of the analytical reconstruction. A microfluidic separation format is implemented, which minimises user intervention and allows fast separations. Label-free detection and quantification take place on-line, simplifying the workflow considerably.

A first part of this thesis is dedicated toward an error analysis. This establishes which parameters are critical for performance and gives an idea of performance limits. Three significant factors contributing to reconstruction inaccuracy due to non-linear migration are identified. Non-linear migration caused by non-constant separation temperature, background buffer migration, and UV-mediated protein fragmentation are studied. While these factors are of crucial importance for virtual two-dimensional separations, they represent previously unreported migration behaviour of proteins in capillary gel electrophoresis. By developing models and estimates for the non-linear migration of proteins, this could be accounted for in the reconstruction of the migration trajectory, resulting in a ten-fold improvement of the accuracy of track reconstruction.

A second part of this thesis is dedicated toward applying V2D separations. The concept of protein track reconstruction is validated in a two-buffer separation system, where proteins migrate through sections with different chemical conditions. For the first time, the concept of migration trajectory reconstruction is applied to capillary electrophoresis separations of proteins under non-constant buffer conditions. To implement V2D separations, a microfluidic system with integrated components for the two separations was developed. A protocol for rapid prototype fabrication was adapted, including surface preparation and functionalisation. Isoelectric focusing of a protein sample in an immobilised pH gradient on-chip was demonstrated.

Als ich zu Bruno sagte: "Ach Bruno, würdest du mir fünfhundert Blatt unschuldiges Papier kaufen?", antwortete Bruno, zur Zimmerdecke blickend und seinen Zeigefinger, einen Vergleich herausfordernd, in die gleiche Richtung schickend: "Sie meinen weißes Papier, Herr Oskar."

Günter Grass, *Die Blechtrommel*



Acknowledgements

It all would not have been possible without some help. My most heartfelt gratitude to:

... my supervisors John Hassard, Andrew de Mello, and Tony Cass for providing the opportunity. Their unerring dedication and wisdom pulled me through. John for giving me total freedom of research, and for those splinters of brilliance when I despaired. Andrew for always cheering me up and having confidence, and for all the admin and form-signing. Tony for some truly insightful comments, and for never missing a meeting.

... my unofficial supervisor, Paul Tjossem. For a million things, including the enthusiasm for my results, designing wacky ‘Friday afternoon’ experiments, finding countless grammar mistakes, feeding my fish, and teaching me about cask ale.

... all the good people over in deltaDOT. Phil Lewis for letting me lobotomise his code and ideas, Gary Taylor for the invaluable, sage advice regarding error analysis, and Reese Stoltzfus for giving me a hand when the code wouldn’t work. Akash Lal and Tom Witham for not minding when I put the hex keys in the wrong place. Nicola Marlin, Vaksha Patel, Deepika Devanur, and Mei-An Sung for the biochemistry advice. Fiona Pereira for teaching me how to run a separation. Paula Pino for keeping track of everything. And of course Stuart Hassard, Martin and Bruce Powell, and Paul James.

... all the good people who are no longer at deltaDOT. Matt Noy, Dimitrios Sideris, Warren Thornton, Mark Richards, and Andy Vine for interesting and helpful conversations. Lesley Caldwell for getting me started, Jeff Harford for letting me canabalise his rig. Gareth Jenkins and Alicia Yang for yellow-room support and chocolate muffins.

... the good people of the de Mello group. My dear office mates Yasemin Koç, Monpichar Srisa-Art, Alexa, FRET, Madame de Pompadour, and Rex. Kalpana Vijayakumar for showing me how to measure refractive indices. Shelly Gulati for the help with Comsol. Guillaume Chansin for getting me connected to IBje5sun. Everyone else for being nice people all around.

... Judit Nagy and Nik Zhukovsky of the IBE for showing me how to run a 2D gel. Kostis Michelakis and John Latigo for letting me use the DekTak and the spectrometer occasionally. Amir Eftekhar for letting me use the CNC PCB router.

... the friendly folks over in HEP—Dave Colling, Julia Sedgbeer, Gavin Davies, Bill Cameron, Vera Kasey, Paula Brown, and everybody else. Special thanks for the best Christmas parties of Imperial, and paint-o-clock!

... the Chemical Biology Centre: Wing-Chau Tung, Alison Stubbings, and Yvonne Tatchley. EPSRC for providing the funding. The CBC students and supervisors for being a jolly bunch.

... the wonderful people who helped me revise and proof-read this thesis: Mom, Gary, Akash, Xevi, and Matthew—thank you very, very much!

... my interns Kara Bakhshi, Abel Assefa, and Felix Quaschning, for a pinch of chaos.

... my friends and family, who helped me stay sane.

... Matthew. ♡♡♡

Contents

Declaration of Authorship	1
Abstract	2
Acknowledgements	4
Table of Contents	5
List of Figures	8
List of Tables	13
Abbreviations	14
Symbols	17
1 Introduction	21
1.1 Protein science overview	22
1.1.1 Biomedicine	22
1.1.2 The Omics	22
1.1.3 Proteomics	24
1.2 Separation science	29
1.2.1 Quantifying separations	29
1.2.2 What is needed	31
1.2.3 Protein separation techniques	32
1.2.4 Hyphenated and multidimensional separations	39
1.3 Protein electrophoresis	43
1.3.1 Background	43
1.3.2 Capillary electrophoresis	44
1.3.3 Dimensions and scaling laws	49
1.3.4 Label Free Intrinsic Imaging	50
1.4 Microfluidics and electrophoresis	51
1.4.1 One-dimensional electrophoresis on-chip	52
1.4.2 Two-dimensional electrophoresis on-chip	53
1.5 Virtual two-dimensional separations	55
2 Error Analysis	60

2.1	Introduction	61
2.1.1	Errors	61
2.1.2	Estimators	62
2.1.3	Combination of errors	66
2.1.4	Error sources in capillary gel electrophoresis	68
2.2	Experimental methods	69
2.2.1	Multiple injections as a model system for V2D error analysis	69
2.2.2	Instrumentation	70
2.2.3	Data analysis: hit-finding and track-fitting	73
2.2.4	Experimental parameters	77
2.3	Results	79
2.3.1	Characterisation of errors for standard separations	79
2.3.1.1	Vertex analysis	79
2.3.1.2	Track errors and starting position errors	82
2.3.1.3	Error correlation	85
2.3.1.4	Goodness of fit analysis	88
2.3.1.5	Shape of migration trajectories	90
2.3.1.6	Peak shape analysis	94
2.3.2	Effect of hit-finding and tracking parameters	96
2.3.2.1	Hit-finding threshold	96
2.3.2.2	Pixel subsets: length of subset	99
2.3.2.3	Pixel subsets: position of subset	101
2.3.3	Effect of different separation conditions	104
2.3.3.1	Sample dilution	104
2.3.3.2	Separation length	106
2.3.3.3	Separation voltage	108
2.3.3.4	Capillary diameter	111
2.3.4	Effect of gel formulation	114
2.3.5	Effect of band crossings and multiple injections	118
2.3.5.1	Track errors	121
2.3.5.2	Starting position errors	125
2.4	Discussion	129
2.4.1	Errors due to thermal effects	130
2.4.2	Errors due to background effects	138
2.4.3	Errors due to UV effects	141
2.4.4	Error optimisation and estimation	149
2.4.4.1	Estimate of thermal errors	150
2.4.4.2	Estimate of errors from background effects	151
2.4.4.3	Estimate of errors from UV effects	153
2.4.5	Expected performance of V2D separations	161
3	V2D Separations	166
3.1	V2D separations in a two-buffer system	166
3.1.1	Introduction	166
3.1.2	Experimental methods	169
3.1.3	Results	172
3.1.4	Discussion	181

3.2 Microfluidics for V2D separations	182
3.2.1 Introduction	182
3.2.2 Microfabrication	183
3.2.3 Fabrication of an immobilised pH gradient on-chip	186
3.2.4 Discussion and chip outlook	189
4 Conclusion	193
4.1 Summary	193
4.2 Outlook	196
Bibliography	198
Appendix	215
A Derivation of thermal model	216
A.1 Differential equations for temperature-time function	216
A.2 Coefficients for temperature-time function	220
B COMSOL conditions and parameters	223
C Cyclic olefin copolymer microfabrication process	224
D Photopolymerisation of polyacrylamide on-chip	227

List of Figures

1.1	The Omics	23
1.2	Annual citation numbers for proteomics, transcriptomics, and metabolomics.	25
1.3	Relative abundance of blood plasma proteins.	26
1.4	Standard workflow in conventional electrophoresis.	27
1.5	Annual citation numbers for proteomics-related research.	27
1.6	Separation power depending on separation technique for different sizes of protein analyte.	34
1.7	SDS linearises proteins and imparts a uniform charge-to-size ratio.	35
1.8	Acidic and basic amino acids.	37
1.9	Schematic illustration of isoelectric focusing.	38
1.10	Two-dimensional gel electrophoresis.	40
1.11	Electro-osmotic flow.	45
1.12	Absorption spectra of aromatic amino acids.	47
1.13	Equiphase map.	51
1.14	Examples of previous work on two-dimensional electrophoresis in microsystems.	53
1.15	Schematic diagram of virtual 2D electrophoresis.	56
1.16	Reconstruction of isoelectric point and molecular weight with the migration trajectory of proteins.	58
1.17	Comparison of different proteomics workflows.	59
2.1	Schematic of 20 measurements with random errors.	61
2.2	Schematic of a least squares estimate for a straight line fit.	63
2.3	Distribution of detection coordinates for five protein bands.	65
2.4	χ^2 distribution for 510 degrees of freedom.	66
2.5	Schematic of protein detection data, fitted straight line, and different errors.	67
2.6	Schematic of protein bands crossing.	70
2.7	Optical train of capillary electrophoresis instrument.	71
2.8	Light ray diagram for illumination and detection.	72

2.9	Flow chart for hit-finding and track-finding algorithms.	73
2.10	Illustration of hit-finding algorithm.	74
2.11	Illustration of track-finding algorithm.	75
2.12	Electropherogram and heatmaps for different threshold values.	76
2.13	Further analysis to evaluate track data.	77
2.14	Track plot: reconstructed protein migration trajectories.	80
2.15	Track errors and starting position errors for different vertices.	81
2.16	Track error and starting position error of seven proteins, plotted against molecular weight.	82
2.17	Electropherogram of a standard separation with seven proteins.	83
2.18	Track slope and intercept for 12 repeat separations.	84
2.19	Heatmap illustrating a bubble interfering with protein tracks.	85
2.20	Starting position errors plotted against migration time and mobility.	86
2.21	Starting position errors plotted against relative migration time and mobility.	86
2.22	Starting position errors plotted against track errors.	87
2.23	Molecular weight versus mobility.	88
2.24	Goodness of fit for data from five standard separations.	89
2.25	Shape of migration trajectories.	91
2.26	Bend in migration trajectories.	91
2.27	Background disturbance: 2D projection of 3D electropherogram, and three-dimensional electropherogram of a blank run.	92
2.28	Non-linear migration trajectory in an enlarged section of the residual plot.	94
2.29	Electropherograms of pixels at different positions along the detector.	94
2.30	Peak shape as migration across detection window progresses.	95
2.31	Peak shape for five bands with and without capillary midsection covered.	96
2.32	Heatmap with tracks for different hit thresholds.	96
2.33	Track error, starting position error, and number of hits per track for different hit threshold levels.	98
2.34	Hits and tracks for different lengths of pixel subset.	99
2.35	Track errors and starting position errors for different lengths of pixel subset.	100
2.36	Goodness of fit for an increasing number of pixels considered in the track-fitting.	101
2.37	Track errors and starting position errors for 150-pixel data subsets starting at different positions along the capillary.	102
2.38	Goodness of fit for 150-pixel data subsets starting at different positions along the capillary.	103

2.39	Track errors and starting position errors at different concentrations of sample.	105
2.40	Track errors and starting position errors at different separation lengths. .	106
2.41	Residual plots for different separation lengths.	107
2.42	Track errors and starting position errors at different electric field strengths.	109
2.43	Track errors against starting position errors at different electric field strengths.	110
2.44	Residual plots for different electric field strengths.	111
2.45	Track errors and starting position errors for different capillary diameters.	112
2.46	Track errors and starting position errors for different capillary diameters under constant current conditions.	113
2.47	Electropherograms for separations with different buffers.	115
2.48	Mobility-molecular weight calibration plots for different buffers.	115
2.49	Track errors and starting position errors for different buffers.	116
2.50	Residual plots of separations with different buffers.	117
2.51	Electropherogram peak shapes along detector for different buffers.	118
2.52	Plot of tracks for a double injection separation.	119
2.53	Track fit parameters for seven repeat double injection separations.	120
2.54	Track errors plotted against molecular weight for 7 double injection runs.	122
2.55	Confidence levels for track errors differing with crossing number.	123
2.56	Starting position errors plotted against molecular weight for seven double injection runs.	125
2.57	Confidence levels for starting position errors differing by crossing number, grouped by molecular weight.	126
2.58	Starting position errors plotted against track errors for seven double injection runs.	128
2.59	Schematic of a migration trajectory with initial, average, and observed migration speeds, as well as the reconstructed starting point from the fitted migration track.	129
2.60	Calculated and measured temperature profiles.	133
2.61	Migration speed as a function of time according to time-dependent temperature.	134
2.62	Migration trajectories for different temperature-mobility coefficients B . . .	135
2.63	Migration trajectories for different initial migration speeds v_0	136
2.64	Errors for different molecular weights arising from non-linear migration due to temperature-dependent mobility ('thermal error').	137
2.65	Track shape with and without pre-flow before separation.	138

2.66	Temperature at different positions along the capillary in the course of a standard separation, with and without additional cooling.	139
2.67	Track shape with and without additional cooling.	140
2.68	Electro-osmotic flow from anode in CGE.	141
2.69	Schematic of anode directed peak asymmetry and electric field independent asymmetry.	143
2.70	Electropherograms at different positions for first and second (reverse) pass of four protein peaks.	144
2.71	HWHM plot for first and second (reverse) pass of four protein peaks. . . .	144
2.72	Sliding detection window: electropherograms of two peaks at different positions along the capillary.	145
2.73	Sliding detection window: comparison of different peaks at different capillary positions.	146
2.74	Comparison of peak asymmetry for different separation media.	147
2.75	Comparison of peaks in denaturing CZE separation with and without pre-exposure of a peak.	148
2.76	Errors for different migration times arising from thermal effects.	151
2.77	Protein migration trajectories: counterband subsections.	152
2.78	Monte Carlo simulation compared to real separation data.	154
2.79	Monte Carlo simulation of UV induced fragmentation: errors on the reconstructed starting coordinates.	155
2.80	Monte Carlo simulation of UV induced fragmentation: increasing UV exposure.	156
2.81	Monte Carlo simulation: function used to estimate errors from UV effects.	157
2.82	Mock V2D separation.	161
2.83	Correction of mock V2D separation by correlation of track errors and starting position errors.	162
2.84	Correction of mock V2D separation with error estimates.	163
2.85	Comparison of different error corrections for V2D separations.	164
3.1	Schematic illustration of a V2D separation in two different buffers. . . .	167
3.2	Schematic illustration of a V2D separation in two different buffers with a moving interface.	169
3.3	COMSOL model of the capillary with enlarged image of the buffer interface, initial conditions.	171
3.4	Mobility and EOF after addition of different polymers to CZE buffer solution.	173
3.5	Current across a capillary with different buffer fillings.	173

3.6	Reconstruction of the migration trajectory of a protein with a given interface error.	174
3.7	COMSOL simulation of interface region of two-buffer separation.	175
3.8	Repeat runs of a single protein in the two-buffer system.	176
3.9	Observed and reconstructed migration tracks of five proteins in the two-buffer system.	177
3.10	Two-buffer separation control experiment.	178
3.11	Two-buffer separation of a sample with five proteins: CZE mobility against sieving mobility.	179
3.12	Repeat runs of two-buffer separation: CZE mobility against sieving mobility.	180
3.13	Electro-osmotic flow in different sections of the two-buffer system.	181
3.14	Basic layout of a V2D chip.	183
3.15	Transmission spectra between 200 and 300 nm for a variety of polymer films.	184
3.16	Structural formula of cyclic olefin copolymer (COC).	184
3.17	Transmission spectra of COC films with residual toluene.	185
3.18	Microfabrication process: COC rapid prototyping with lamination.	186
3.19	Microfluidic chip design for generating immobilised pH gradients by diffusion.	188
3.20	Photografting of polyacrylamide gel to COC channel.	188
3.21	Formation of a dye gradient by diffusion in a microfluidic channel.	189
3.22	Isoelectric focusing of myoglobin in immobilised pH gradient on-chip.	190
3.23	Microfluidic chip design with a flush-channel adjacent to the pH gradient channel.	191
3.24	Chip design with a microfluidic channel network for forming a pH gradient.	192
B.1	Geometry, initial conditions, and boundary conditions for COMSOL model of interface region of two buffers co-migrating (not to scale).	223
C.1	Photolithographic mask design for SU-8 master for embossing microfluidic channels in COC by lamination.	225

List of Tables

1.1	Protein biomarkers for different pathologies.	28
1.2	Typical performance metrics for a variety of protein separation techniques.	42
2.1	Optimal experimental conditions for minimal errors on the reconstructed starting position.	149
2.2	Estimates of errors for different types of non-linear migration.	159
2.3	Comparison between observed and estimated starting position errors for sample separation data.	160
2.4	Estimates for overall relative errors of a V2D separation.	165
3.1	Standard sample for two-buffer V2D separations.	170
4.1	Comparison of performance metrics for V2D and conventional protein separation techniques.	195

Abbreviations

2D	2-Dimensional
2D SDS-PAGE	2 -Dimensional Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis
2DE	2 -Dimensional Electrophoresis
2DGE	2 -Dimensional Gel Electrophoresis
2DLC	2 -Dimensional Liquid Chromatography
µTAS	Micro Total Analysis System
CA	Carrier Ampholyte
CE	Capillary Electrophoresis
CEC	Capillary ElectroChromatography
CGE	Capillary Gel Electrophoresis
CHAPS	3-[(3-CHolAmidopropyl)dimethylammonio]-1-PropaneSulfonate
COC	Cyclo-Olefin Copolymer
CZE	Capillary Zone Electrophoresis
DIGE	Differential Gel Electrophoresis
DNA	DeoxyriboNucleic Acid
ELISA	Enzyme Linked ImmunoSorbent Assay
EOF	Electro-Osmotic Flow
ESI	ElectroSpray Ionisation
EVA	Equiphase Vertexing Algorithm
FT-ICR	Fourier Transform Ion Cyclotron Resonance (analyser)
GE	Gel Electrophoresis
GST	Generalised Separation Transform
HETP	Height Equivalent to one Theoretical Plate
HPLC	High Pressure Liquid Chromatography

HPLC	High Performance Liquid Chromatography
HWHM	Half Width at Half Maximum
ID	Inner Diameter
IEF	IsoElectric Focusing
IPG	Immobilised pH Gradient (strip)
ITP	Isotachophoresis
LC	Liquid Chromatography
LFII	Label Free Intrinsic Imaging
LPA	Linear PolyAcrylamide
MALDI	Matrix Assisted Laser Desorption/Ionisation
MECC	Micellar Electrokinetic Capillary Chromatography
mRNA	Messenger RiboNucleic Acid
MS	Mass Spectrometry
MuDPIT	MultiDimensional Protein Identification Technology
MW	Molecular Weight
OD	Outer Diameter
PAGE	PolyAcrylamide Gel Electrophoresis
PDA	PhotoDiode Array
PDMS	Poly(DiMethyl Siloxane)
PEO	Poly(Ethylene Oxide)
pI	Isoelectric point
PMHS	Poly(MethylHydroSiloxane)
PMMA	Poly(Methyl MethAcrylate)
PTM	Post-Translational Modification
RNA	RiboNucleic Acid
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis
SELDI	Surface Enhanced Laser Desorption/Ionisation
TEMED	TetraMethylEthylDiamine
TOF	Time-Of-Flight (detector)
UV	Ultra Violet
V2D	Virtual 2-Dimensional (electrophoresis)

Symbols

a	intercept	s (*)
<i>A</i>	cross section area	m ²
b	slope	s/mm (*)
<i>c</i>	concentration	mol/m ³
<i>c</i>	specific heat capacity (per unit mass)	J/(kgK)
<i>d</i>	diameter	m
<i>d_i</i>	inner capillary diameter	m
<i>d_a</i>	outer capillary diameter	m
<i>D</i>	diffusion coefficient	m ² /s
\vec{E}	electric field	V/m
<i>E</i>	electric field strength	V/m
<i>f(x)</i>	function <i>f</i> of <i>x</i>	
\vec{F}	force	N
<i>h</i>	temperature-increase coefficient	K/s
<i>H</i>	plate height	m
<i>I</i>	light intensity	W/m ²
<i>I</i>	electric current	A
<i>k</i>	heat transfer coefficient	W/(m ² K)
<i>k'</i>	partition ratio or column capacity factor	
<i>K</i>	phase equilibrium ratio	
<i>L</i>	length	m
<i>L_c</i>	capillary length	m
<i>L_p</i>	path length	m
<i>L_s</i>	separation length	m

*in time-position coordinate system

m	mass	kg
n	peak capacity	
n	number of measurements	
N	theoretical plate number	
p	position coordinate	mm
P	power	W
q	charge	C
Q	energy	J
r	radius	m
R	resistance	Ω
R	retardation factor	
R_s	resolution	
S	surface area	m^2
t	time	s
t	time coordinate	s
T	temperature	K
$T(t)$	temperature in dependence of time	K
T_0	initial temperature	K
T_∞	environment temperature	K
v	velocity	m/s
V	voltage	V
V	variance	
V	volume	m^3
x_i	i-th measurement value	
\bar{x}	mean, or average	
\hat{x}	estimate	
x_i, y_i	paired measurement values	
α	separation factor or selectivity factor	
χ^2	'chi square' value (goodness of fit)	
δ^*	critical exposure threshold	s
δ^{total}	total exposure time	s
ϵ_Θ	thermal error	mm

ϵ_{CB}	error from counterband interference	mm
ϵ_{UV}	error from UV effects	mm
ϵ	dielectric constant	Farad/m
ϵ	molar absorption coefficient	m^2/mol
η	viscosity	sPa
λ	wavelength	nm
μ	electrophoretic mobility	$\text{m}^2/(\text{sV})$
π	3.14159	
ρ	correlation coefficient	
ρ	density	kg/m^3
ρ	electrical resistivity, or specific resistance	Ωm
σ	electrical conductivity	S/m
σ	standard deviation, ‘error’	
σ_D^2	diffusion zone variance	m^2
σ_D	diffusion length	m
σ_a	error on intercept	mm
σ_b	error on slope	s/mm
σ_p	error on position	mm
σ_t	error on time	s
τ	migration time	s
ξ	zeta potential	V

*Für mis Grossmami. Wil si min Held isch.
For my grandmother. Because she is my hero.*

Chapter 1

Introduction

Scientific advance is largely technology driven, as exemplified by the sequencing of the human genome. Without the technique of polymerase chain reaction, vast amounts of biological sample would have been necessary to generate enough DNA to analyse. Without sophisticated electrophoretic techniques, sequencing one organism would have taken a global effort of decades of research. The same holds true for other disciplines, e.g. the Large Hadron Collider experiments at CERN, which would be unthinkable without advances in grid computing, silicon detector technology, vacuum technology, and particle accelerator technology. Countless examples can be found for the genesis of scientific insight through superior technology, with benefits for society such as improved healthcare and convenience in our daily routines. These are strong motivators for investing resources in developing and advancing the current stand of a technology.

Biological organisms are extraordinarily complex systems of molecules, built of not much more than carbon, hydrogen, oxygen, and nitrogen (these constitute 95 % of biomass; phosphorus and calcium an additional > 4 %, and the remaining < 1 % consists of over 20 different elements such as K, S, Na, Mg, I, Fe, and Cl).^[1] An average human is composed of some 10^{28} atoms, the most abundant molecule being water (65 % of total mass, 99 % of total number of molecules). There are something around 10^{14} cells in a human body, in each an estimated 1.76×10^{14} molecules.^[2]

Given the complexity of biosystems it comes as no surprise that they are susceptible to failure in manifold different ways. The World Health Organization classifies 12,420 different disease types,^[3] which are caused by the combined impact of external factors (for instance diet, sunlight exposure, competitive stress) and personal predisposition. Personal disposition can be characterised with the genome, but this alone is often not sufficient to explain or predict disease, and much less to cure. Here lies the significance of the proteome, which is the actual currency of organism vicissitude.

1.1 Protein science overview

1.1.1 Biomedicine

One of the core tasks of civilisation is to relieve suffering. In our society, medicine is one of the fundaments of wellbeing. Access to medical care is included in the universal declaration of human rights (Article 25: “Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including [...] medical care [...].”)[⁴]

In an effort to improve health care, medicine has increasingly moved from faith-based (pre-19th century) to an experience- or phenomena-based medicine (up to around 1950). Since then, medical science has developed toward a scientific approach, requiring understanding on a molecular basis. This is exemplified by how penicillin has evolved from a ‘magic bullet’ for bacterial infections, to a family of antibiotics that lyse bacteria by inhibiting various steps in the synthesis of the peptidoglycan layer of bacterial cell walls. Molecular biology itself has experienced a shift in focus over the last 100 years from identifying and characterising individual agents, such as haemoglobin in blood cells, toward exploring whole networks of interactions where many agents influence how a system performs.

With the aim of understanding biochemical processes on a biomolecular basis, research is focusing increasingly on biomarkers. These are molecules that may differ between health and disease, and can be used to diagnose a pathology, to evaluate progress of a disease, or measure efficacy of a drug. Under certain circumstances biomarkers can be used as a drug target, which leads to the secondary focus of biomedical research: the corrective intervention in the form of therapy or drug development. This is the setting of the -omic disciplines, which share the study of a system of biomolecular agents.

1.1.2 The Omics

Genomics is usually considered the founding discipline of the -omics, although arguably the -ome suffix predates genomics.^[5] As early as 1924 ‘genome’ was used to refer to the haploid chromosome set of a species.^[6] The modern definition of the term ‘genome’ refers to a complete genetic sequence on one set of chromosomes—this includes the genes as well as the non-coding sequences. The first sequenced genome was that of the virus bacteriophage MS2 (with a length of over 5,300 base pairs) in 1976 by W. Fiers *et al.*^[7] The human genome with over 3×10^9 base pairs was sequenced by 2000 as a result

of international collaboration on the Human Genome Project.^[8] Today the genomes of over 25 species including mammals, fish, insects and plants have been sequenced.

In the post-genomic era, something of a paradigm shift has occurred. This is illustrated by the following observation: in the human genome there are about 25,000 protein-encoding genes. At the same time, an estimated 1,000,000 different proteins occur in humans, all encoded in those 25,000 genes. Clearly there is another level of complexity beyond the genome. This is where the other -omic disciplines come in.

The next -omic discipline that emerged was proteomics. The term was first introduced in 1995, and refers to the entire complement of proteins in a given organism or system at a given time. Proteomics is the logical complement of genomics, and is much more intimately involved in discerning function from malfunction. The proteome reflects the state of the cell, tissue, or organism. Function, health, age, and environment can influence the proteome, whereas they do not necessarily reflect on the genome. The genome alone does not determine the proteome.

Beside the genome and the proteome, something over 200 -omic terms have been coined, many of which are not in common usage, such as *Introme: the whole set of introns*.^[9] Beside genomics and proteomics, some of the more widely accepted terms are transcriptomics and metabolomics. These four prevalent -omic disciplines are illustrated in Figure 1.1.

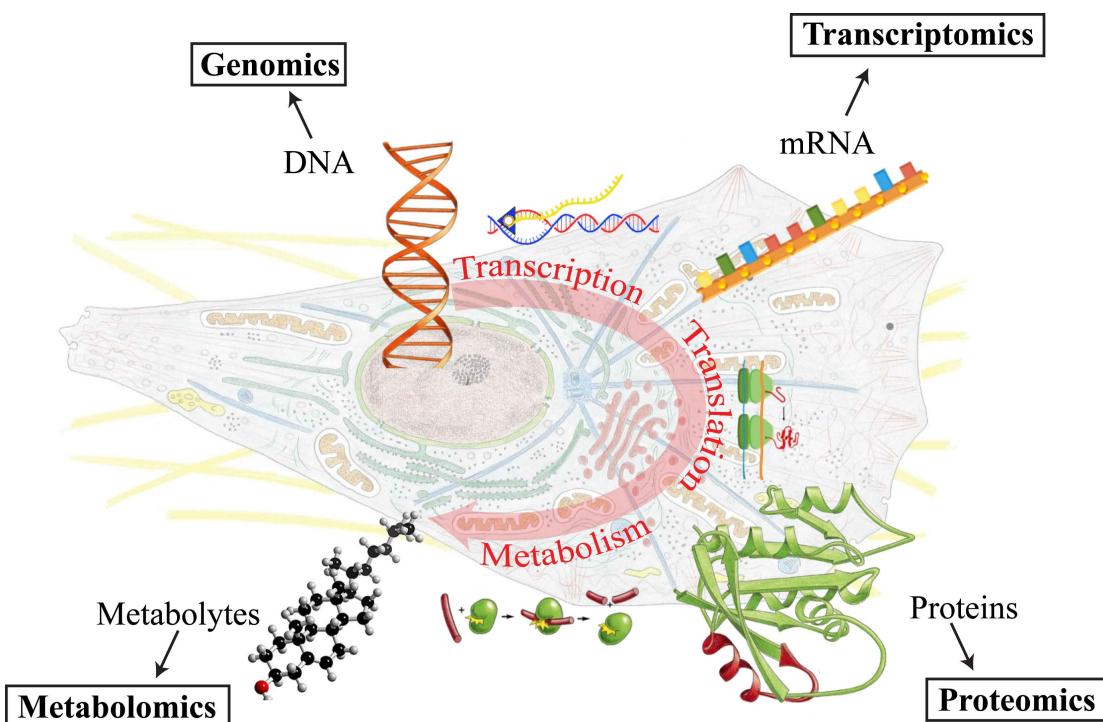


Figure 1.1: The Omics. Artwork adapted from Molecular Biology of the Cell^[10].

Transcriptomics studies the entire complement of messenger RNA present in a given single cell or a population of cells. This can be considered a linking element between genomics and proteomics, where the tools and techniques of genomics are used to gain knowledge concerning the proteome. While transcriptomics can achieve excellent results with highly advanced techniques such as DNA chips, it is still one level removed from the phenotype. Some mRNA can be non-encoding, alternative splicing can occur, and the mRNA quantity does not necessarily correlate to the protein quantity. Furthermore, proteins can undergo numerous forms of post-translational modifications (PTM). Many PTMs have a regulatory function on proteins. Common PTMs include phosphorylation and glycosylation, but many more are possible such as ubiquitination, formation of disulfide bonds, or deamination. These are differences which feature in the proteome, but not in the transcriptome.

Metabolomics studies the entire complement of small molecules and metabolites (such as metabolic intermediates, hormones and other signalling molecules) of a cell or organism. Proteins are very often participants in biochemical pathways that can involve numerous cofactors, substrates, intermediates, and products. The genome, transcriptome, proteome and metabolome are all linked in a dynamic network: the metabolism. Studying the metabolome is useful since numerous diseases such as atherosclerosis, diabetes, or osteoporosis are metabolic diseases, and often a metabolite serves well as a biomarker—for instance cholesterol for cardiovascular diseases.^[11]

1.1.3 Proteomics

Proteomics : the systematic study of amount, modifications, interactions, localisation and function of proteins in whole organisms, tissue, the cell, or subcellular components.^[12]

Since proteins are responsible for structure, transport, storage, regulation, signalling and metabolism, this definition opens a large field for research. Some sub-sets of the proteome are occasionally specified, such as the structural proteome, or enzymes. On the other hand, a global perspective is necessary given the complex behaviour of some illnesses such as cancer. One of the central questions of proteomics is whether the proteome changes in a characteristic way for a given change of state. Given this is a reasonable assumption for many illnesses, studying the protein expression profile and how it changes allows identification of a ‘fingerprint’ for that illness, indicating biomarkers and therapy possibilities.

Following the genetic revolution, proteomics has been heralded as the ‘Next Big Thing’. This seems fitting, since the word ‘protein’, derived from Greek, means ‘of first rank’.

In the past 10 years publication numbers have exploded from fewer than 100 to a couple of thousand yearly, with numerous new journals dedicated exclusively to proteomics (*PROTEOMICS*, *Journal of Proteomics*, and *Journal of Proteome Research*, to name but three). The current trend indicates a saturation of the qualifier ‘proteome’ with tendencies toward diversification and specification of proteome research. The focus is increasingly on a particular organism, or on a particular subsystem. This might be explained by a lack of techniques able to deal with the extraordinary complexity of the entire proteome. It would appear that the technology has moved beyond a one-protein-at-a-time approach, but struggles to achieve all-encompassing proteomics (see Figure 1.2).

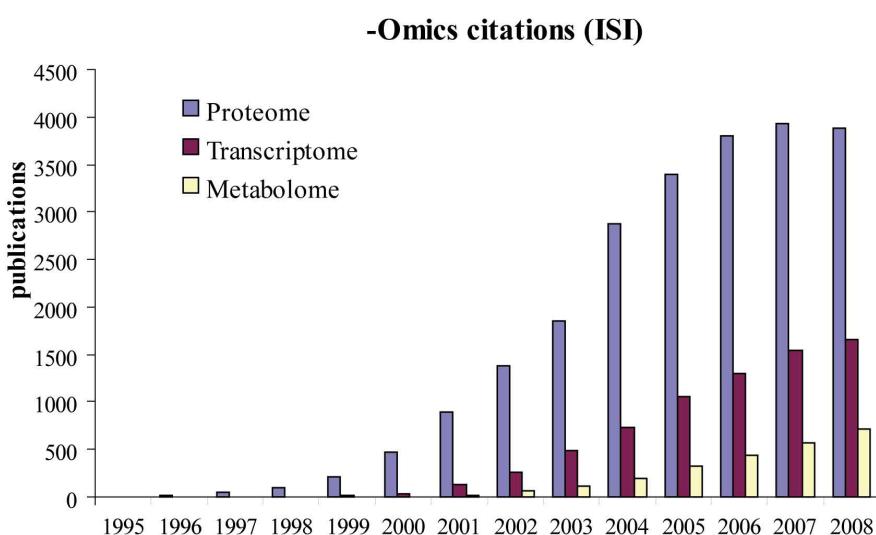


Figure 1.2: Annual citation numbers for -omics. The citation numbers for the terms ‘proteome’ and ‘proteomics’ have levelled off, whereas increasing numbers of publications relate to the transcriptome or the metabolome. Source: ISI Web of Knowledge (PubMed survey similar, data not shown).

Of the organic molecules of a cell, proteins are one of the more abundant and diverse. In a liver cell, an estimated 7.9×10^9 protein molecules are present.^[12] Estimates for how many different protein types are present in a single cell vary between 10,000 and 100,000 depending on type of cell and whether post translational modifications are considered.^[12,13] The number of protein molecules per protein species, the dynamic range, varies between 10 copies and 1 million copies in any given cell. Figure 1.3 illustrates the abundance of different protein species present in blood serum. Often it is the least abundant proteins that are indicative of the processes and changes an organism is undergoing.

One of the most powerful resolving techniques for proteins, two-dimensional (2D) gel electrophoresis, can resolve up to 10,000 proteins. Under most circumstances 5,000 resolved proteins may be considered a success. To analyse organisms such as *Escherichia*

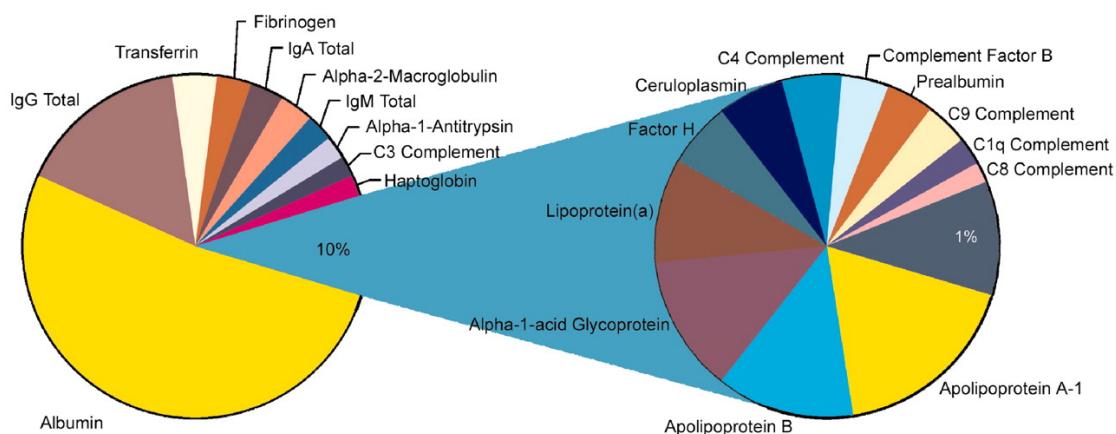


Figure 1.3: Pie chart representing the relative contribution of proteins within plasma. Only 22 different proteins constitute 99 % of the total protein mass, with albumin alone contributing 50 %. From Tirumalai *et al*^[14].

coli with no more than 1,000 proteins present at any one time, 2D gel electrophoresis is an ideal tool, but for many applications it quickly reaches its limits. This will be discussed in further detail in Section 1.2.

To study the proteome, the experimental cornerstones are separation, identification, and quantification. Separation is necessary to reduce the complexity, allowing selection of subfractions of interest. Quantification is often equally important, because the increase or decrease of a particular protein species may be meaningful. Once protein groups of interest have been determined, they are characterised, with an outlook toward identification.

The standard workflow is sample preparation followed by 2D gel electrophoresis for separation and quantification, and finally mass spectrometry (MS) for identification. This is illustrated in Figure 1.4. Although other techniques exist and are in use, 2D gel electrophoresis combined with MS is very common. Liquid chromatography based separations are also widely used and will be discussed in Section 1.2.3.

Sample preparation includes removal of non-protein components of tissue (DNA, small molecules, lipids), and possibly pre-fractionation of the protein bulk (e.g. cytoplasmic, cytoskeletal, etc.). 2D gel electrophoresis separates proteins by two independent properties, molecular weight and isoelectric point. This gives a more powerful separation than either one on its own (as discussed in detail in Section 1.2.4). The separated (and localised) proteins are visualised and quantified, either with fluorescent protein derivatisation or by staining. By comparing the quantities of the separate protein groups to a reference, groups of interest are determined. Mass spectrometry is the technique of choice for identification. To this end proteins are usually cleaved into fragments in a

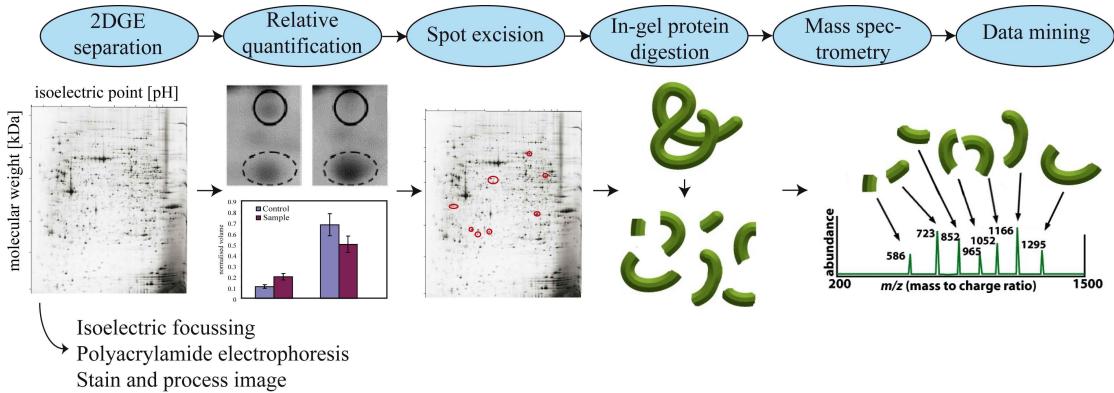


Figure 1.4: Standard workflow in conventional electrophoresis. The prepared sample is separated by two-dimensional electrophoresis. Relative quantification is achieved by comparison with a control. Once spots of interest have been determined, they are cut out of the gel for identification. Proteins are digested and cleaned before they are introduced into a mass spectrometer. The mass spectrum of a protein is matched to protein database information for identification. Artwork adapted from Molecular Biology of the Cell^[10].

controlled way, for instance by trypsin digestion. The resulting set of peptides is characteristic of the parent protein. The mass spectrum of the fragment mixture is determined, and matched to information in protein databases. After a protein has been identified, further studies on structure and function may follow. For elucidating protein structure X-ray crystallography or NMR spectroscopy may be used. For studying interactions, affinity-based assays such as western blotting, enzyme linked immunosorbent assays (ELISA), or protein chips may be used. *In vivo* interaction can be studied with a yeast two-hybrid system.^[15,16]

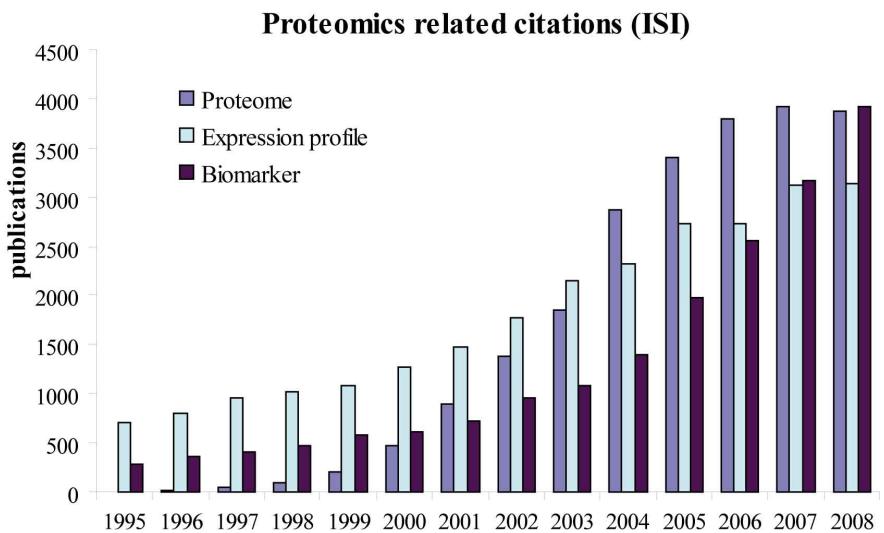


Figure 1.5: Annual citation numbers for proteomics-related research. While the citation numbers of proteome-related research have levelled off, biomarker-related research is on the increase. Source: ISI Web of Knowledge.

Since proteomics has become a major focus of scientific research, the number of reports on biomarkers has made a parallel jump (see Figure 1.5). A selection of potential protein biomarkers is shown in Table 1.1. Although these biomarkers are still far from routine clinical application, and many still need validation, this is a very promising field.^[17–19] Some of the first applications of such biomarkers include cancer screening tests for prostate specific antigen, carcinoembryonic antigens, α -fetoprotein, β 2-microglobulin, neuron-specific enolase, and amylase.^[20,21]

Table 1.1: Protein biomarkers for different pathologies.^[16]

Pathology	Protein biomarkers
Ovarian cancer	RHOGDI, glyoxalase-1, FK506BP, haptoglobin, HAP1, transferrin, apolipoprotein A1, truncated form of transthyretin, inter- α -trypsin inhibitor heavy chain H4, cSHMT, Tbx3, utrophin ^[17]
Prostate cancer	Annexin-1, prohibitin, calgranulin-B, serotransferrin precursor, gelsolin, HSP-60, SBP, IgBPLP ^[22]
Breast cancer	HSP-27, HSP-60, HSP-90, PCNA, transglein, RS/DJ-1, ubiquitin, ferritin, light chain, 14-3-3- σ , cSHMT, Tbx3, utrophin, tumor protein D52, folate receptor α ^[17]
Lung cancer	PGP.5, cytokeratins
Hepatocellular carcinoma	Hcc-1, lamin B1, sarcosine dehydrogynase
Leukaemia	Op18, nm23-H1
Colorectal cancer	HSP-70, S100-A9, S100-A11
Bladder cancer	Keratins, psoriasin
Alzheimer's disease	Antithrombin III, carnosinase I, α 1-antichymotrypsin, zinc- α 2-glycoprotein ^[23]
Rheumatic arthritis	α -Enolase, citrullinated α -enolase, serum amyloid A, calgranulin B, myeloid-related protein 8 ^[24]

Despite the huge potential of proteomics research, numerous limitations exist. The two techniques with the greatest level of detail and the highest accuracy, 2D gel electrophoresis and mass spectrometry, are demanding of the operator, and time-consuming. Achieving good reproducibility poses high demands on sample preparation and handling. The instrumentation, in particular mass spectrometers, is expensive and limits widespread application. It seems that the technology is still in a state of flux, and no true standard has established itself. Until these problems are addressed, research of the proteome will struggle to realise its full potential.^[25,26]

1.2 Separation science

Separation : the act or process of moving things apart, categorising or distinguishing things, or splitting something into component parts.^[27] Separation science deals with the fundamental physical and chemical phenomena involved in the achievement of separations, as well as with the development, application, and reproducibility of various separation processes.^[28]

One of the great challenges of proteomics lies in the extraordinary complexity of the analyte. Strategies to reduce the complexity and bring order into it are essential. This is the role that separation science plays in proteomics. Over the years many different approaches have been developed. 2D gel electrophoresis has been in use for over 30 years,^[29] with many refinements since its invention. Others, such as protein chips, are more recent developments.^[30] The variety of techniques that are available indicates that there is no universal separation solution, and each approach has some severe disadvantages.

In this section, metrics of separations will be introduced. The dominant techniques of liquid chromatography, gel electrophoresis, capillary electrophoresis, and protein chips will be introduced. A brief excursion into the technique of mass spectrometry will also be included. Hyphenated separation techniques, where two different separations are combined, will be discussed and compared. This will provide an overview of what is available, and what the limitations are. Table 1.2 at the end of this section summarises the typical performance of different techniques available for protein separation.

1.2.1 Quantifying separations

Separations can be quantified in various terms such as resolution, accuracy, sensitivity, dynamic range, sample load, or peak capacity. Other characteristics include the throughput and the ability to automate a separation.

In separation operations, the phase equilibrium ratio K is defined as:^[31]

$$K = \frac{y_i}{x_i} \tag{1.1}$$

where y_i and x_i are the mole fractions of a species present in two phases. In chromatographic separations, K is termed the distribution ratio, and y_i and x_i are replaced by the analyte concentrations c in the stationary (c_s) and mobile (c_m) phases. The partition

ratio or column capacity factor k' of a chromatographic separation is:^[32]

$$k' = \frac{q_s}{q_m} \quad (1.2)$$

where q_s and q_m are the quantities of analyte in stationary and mobile phase, respectively. With $q = cV$ for a volume V , Equation 1.2 can be written as:

$$k' = K \frac{V_s}{V_m} \quad (1.3)$$

The average fraction of time an analyte molecule spends in the stationary phase is:

$$\frac{t_a^{(s)}}{t_a^{(total)}} = \frac{q_s}{q_m + q_s} = \frac{k'}{1 + k'} \quad (1.4)$$

and similarly, the average fraction of time a molecule spends in the mobile phase is:

$$\frac{t_a^{(m)}}{t_a^{(total)}} = \frac{q_m}{q_m + q_s} = \frac{1}{1 + k'} \quad (1.5)$$

If the mobile phase flows through a column at a constant velocity of v_m , then from Equation 1.5 the apparent velocity v_a of an analyte band becomes:

$$v_a = v_m \frac{1}{1 + k'} = v_m R \quad (1.6)$$

R is the retardation factor, or the fractional rate of movement of a band relative to the mobile phase. Given the analyte retention time t_r and the solvent retention time t_m are given, then R can be expressed as:

$$\frac{v_a}{v_m} = \frac{t_m}{t_r} = \frac{1}{1 + k'} = R \quad (1.7)$$

With Equations 1.1–1.7, the analyte retention volume V_r can be described in dependence of the mobile phase volume (dead volume) V_m with:

$$V_r = \frac{V_m}{R} = V_m(1 + k') = V_m + KV_s \quad (1.8)$$

Equations 1.1–1.8 allow a quantitative description of the chromatographic separation process. For quantifying the merit of a separation, the column efficiency or number of theoretical plates (also plate number) N is defined as:^[33]

$$N = \left(\frac{t_r}{\sigma_i} \right)^2 = \left(\frac{V_r}{\sigma_V} \right)^2 \quad (1.9)$$

using the variance σ_i of the band in the chromatogram. Alternatively, the baseline peak width w_b or the width at half maximum $w_{h/s}$ can be employed:

$$N = 16 \left(\frac{t_r}{w_b} \right)^2 = 5.54 \left(\frac{t_r}{w_{h/2}} \right)^2 \quad (1.10)$$

The height equivalent to one theoretical plate (HETP), or the plate height H , is a related measure of column efficiency using the column length L :

$$H = \frac{L}{N} \quad (1.11)$$

The resolution R_s of two neighboring bands is defined by the distance between the peaks divided by the peak widths:

$$R_s = \left(\frac{t_{R_2} - t_{R_1}}{w_1 + w_2} \right) \quad (1.12)$$

With Equation 1.10 the resolution can be expressed in terms of the plate height:

$$R_s = \frac{\sqrt{N}}{4} \left(1 - \frac{t_1}{t_2} \right) = \frac{\sqrt{N}}{4} \left(1 - \frac{1 + k'_1}{1 + k'_2} \right) \quad (1.13)$$

The separation factor or selectivity factor α is defined as:

$$\alpha = \frac{k'_2}{k'_1} \quad (1.14)$$

In Equation 1.13 the resolution is related to the plate number and the column capacity factor. The plate number changes with the length of the chromatographic column, the particle size, and the velocity of the mobile phase. The capacity factor depends on the interactions between stationary phase, mobile phase, and analyte.

For comparing different separation techniques, the concept of peak capacity is useful. In simple terms, this is the number of analyte species a separation can distinguish (number of peaks that can be fitted into a chromatogram, each peak being separated from its neighbor by 4σ).

1.2.2 What is needed

As discussed in Section 1.1.3, the analyte is a mixture of proteins with 10^4 – 10^5 different types of proteins. The width of the dynamic range of protein abundance is 10^5 – 10^9 , depending on sample type (e.g. tissue or serum). Because amplification of proteins, unlike DNA, is not possible, the sensitivity needs to be good enough to detect the least abundant protein.

Proteins have several properties by which they can be separated. Molecular weight, isoelectric point, hydrophobicity, charge, or affinity for a particular ligand are commonly used properties. Most proteins have a mass of less than 100 kDa, but some clinically interesting proteins (e.g. antibodies) are among the > 100 kDa molecular weight proteins. Theoretical genome calculations predict a majority of proteins with isoelectric points (pI) between 4 and 6.5, with a second group of proteins of pI 8–12.^[15] The greatest challenge for protein separations is the detection of low abundance proteins, as well as separation of proteins with extreme properties—especially low and high molecular weight, low and high pI, and high hydrophobicity. Quantification, especially absolute quantification, is not always satisfactory, and reproducibility remains problematic. Another hurdle is that many separation techniques need considerable user skill and effort, and that some of the instrumentation is expensive, highly sophisticated and not widely accessible. In the words of a recent proteomics textbook:

“It should be emphasised that [...] many of the technologies in proteomics are still prototypical and better material, instrument design and methodology are expected to improve sensitivity, resolution and repeatability in future.”^[16]

1.2.3 Protein separation techniques

Prefractionation: This is an important tool to improve sensitivity ('deep proteome mining') and reduce sample complexity. Techniques such as centrifugation, chromatography, or electrophoresis are often employed. Centrifugation can separate cell organelles. Similarly, by successively adding different solvents, various protein fractions can be washed out of a whole cell sample. A typical commercial kit separates cytoplasmic, membrane, soluble nuclear, chromatin-bound and cytoskeletal protein fractions (Subcellular Protein Fractionation Kit from Thermo Scientific Inc.). Most liquid chromatographic techniques are applicable to prefractionation. Electrophoretic techniques that are performed in free medium (i.e. non-gel based) are used, especially techniques related to isoelectric focusing.^[34] Depletion of high abundance proteins is a variant of prefractionation. Closely related, so-called ‘equaliser beads’ capture equal amounts of all the different protein species, with the help of ligand libraries.^[35] With all prefractionation techniques, it must be kept in mind that with each additional treatment step come losses, the potential for contamination and variation, the introduction of biases and systematic effects, and additional time and effort investments.

Liquid chromatography (LC), also high pressure liquid chromatography or high performance liquid chromatography (HPLC): This method of separation depends on the relative affinities of the different sample components with a mobile phase and a stationary phase. The sample mixture and a solvent (the mobile phase) are pumped through a column (the stationary phase) where interactions between the proteins, solvent, and stationary phase occur, causing the different species to spend longer or shorter time in the column. This is quantified with the chromatogram from a detector at the end of the column. Different columns separate proteins by different properties. Typically the stationary phase is a column filled with packed beads, which have unique chemical and physical properties.

In size exclusion chromatography, the stationary phase is porous but inert (for instance agarose beads, or SephadexTM). Put simply, smaller molecules reside longer in the column since they spend more time in the pores than larger proteins. This is the only mode of liquid chromatography where stationary phase interactions do not play a role in the separation.

In ion exchange chromatography, proteins interact with a stationary phase that carries ionic groups on the surface. Usually the sample is allowed to adsorb to the column before gradual elution is effected by changing the ionic strength of the buffer. As the buffer ionic strength increases, increasingly highly charged, strongly binding proteins are displaced with buffer ions and eluted. The separation mechanism is based on ionic interactions between the analyte and the mobile phase, and competing interactions of the mobile phase.

Reversed phase chromatography is also based on reversible adsorption of proteins to the column, but instead of charge, the interaction is based on hydrophobicity. Elution is achieved with an organic solvent such as acetonitrile. Hydrophobic interaction chromatography is similar to reversed phase chromatography, with hydrophobic ligands on the column, but the elution is achieved with an inverse salt gradient instead of an organic solvent gradient.

Affinity chromatography is based on the specific interaction of certain proteins with ligands that are presented on the surface of the column matrix. After the sample is applied to the column, unbound proteins are flushed out, before the bound proteins are eluted in a second step. Various types of affinity tags are used, such as antibodies or antigens, lectins for glycoprotein enrichment, and metal ions for phosphoproteins.

Figure 1.6 illustrates how typical liquid chromatographic separations perform.^[36] Plate capacities (i.e. theoretical number of species that can be resolved in a separation) are plotted for different molecular weights. Notably, chromatographic separations are most

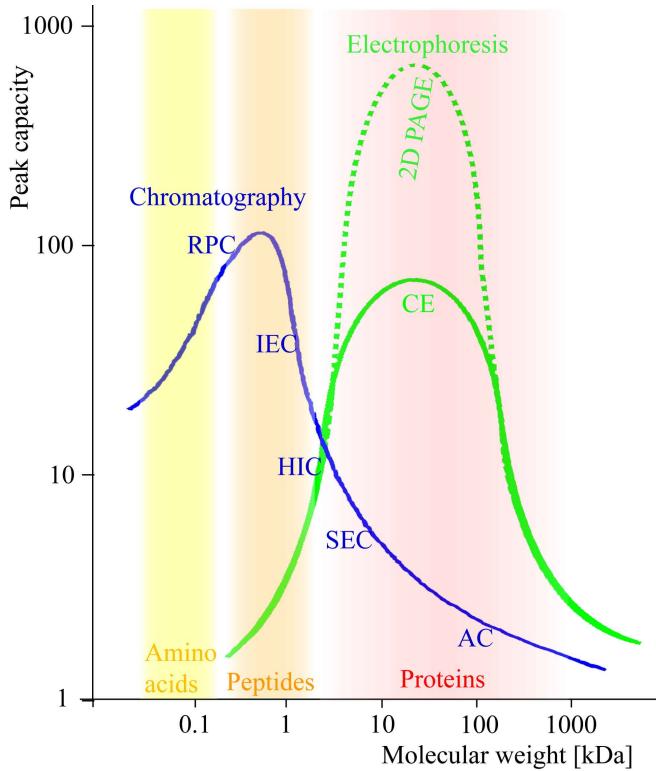


Figure 1.6: Separation power (peak capacity) for different sizes of protein analyte, depending on separation technique. RPC: reverse phase chromatography, IEC: ion exchange chromatography, HIC: hydrophobic interaction chromatography, SEC: size exclusion chromatography, AC: affinity chromatography, CE: capillary electrophoresis, 2D PAGE: two-dimensional polyacrylamide gel electrophoresis. Adapted from Westermeier and Naven^[36].

efficient for proteins smaller than 1 kDa. This is at the lower end of protein molecular weights, and more relevant for protein fragments (peptides). For analysis of unfragmented proteins, the separation power lags behind other techniques.

As a pressure driven separation, the resolution in chromatography is affected by flow dispersion, as opposed to electrophoresis, where diffusion is more relevant. Amongst the advantages of liquid chromatography are the compatibility with high sample loads, or the concentrating capabilities, and the large number of experimental parameters that can be tuned to design the protein separation. This can become critical if only a particular species is under investigation, but is less useful if a more global overview is necessary. Amongst the drawbacks of liquid chromatography are difficulties associated with quantifying the amounts of proteins. This is very important for comparing differences between samples. Another disadvantage of the technique is that the property by which the proteins are separated (e.g. charge) is not easily quantified. This could be a useful help in identifying a protein. It should also be mentioned that the visualisation of the separation is less intuitive than for instance gel electrophoresis. Overall, liquid

chromatography is powerful for separating peptides, and for prefractionation, but is not necessarily the ideal technique for protein separations.

Protein chips also protein microarrays, or biochips: This is one of the newest additions to the palette of separation techniques. Protein chips are related to chromatographic techniques in so far as they rely on similar interactions, but the hardware is similar to DNA chips. The chip consists of an array of spots that are designed to capture specific proteins or protein classes. Surface arrays are available for hydrophobic interactions, ion exchange interactions, metal affinity interactions, alongside a range of more specific baits such as antibodies.^[30] Detection can be achieved with fluorescent labels (similar to conventional enzyme-linked immunosorbent assays, ELISA). Protein chips can be combined with mass spectrometry, as is the case in surface enhanced laser desorption/ionisation (SELDI) mass spectrometry.^[37]

Although protein chips are amenable to automation and high throughput screening, obtaining capture agents for the global proteome is problematic.^[38] Protein chips are best suited for confirming the absence or presence of a known agent, but not for discovering unknown or unexpected entities. This, and the fact that many capture agents are difficult and expensive to produce and purify, makes protein chips a tool with limited applications.

Gel electrophoresis (GE), also polyacrylamide gel electrophoresis (PAGE), or sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE): In denaturing gel electrophoretic separation of proteins, the sample is usually treated with sodium dodecyl sulphate (SDS). This is an anionic detergent that denatures and binds to proteins, resulting in a uniform charge per unit length of protein, as illustrated in Figure 1.7.

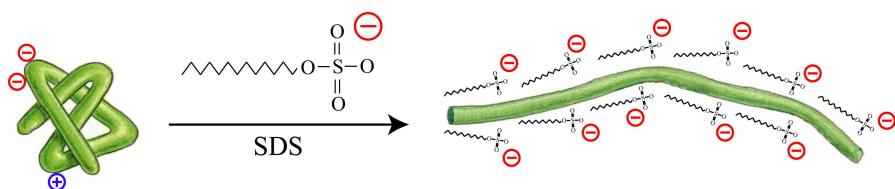


Figure 1.7: SDS linearises proteins and imparts a uniform charge-to-size ratio. Artwork adapted from Molecular Biology of the Cell^[10].

Native PAGE separations without SDS can also be performed, but are far less common. The electrophoretic force molecules experience is proportional to their charge. The drag a protein molecule experiences in free medium is proportional to its mass. If the charge-to-size ratio is uniform for all different species of a mixture, then they all move with the same velocity, regardless of size. If the mixture is made to migrate through a type of

sieve, smaller molecules travel farther in the same time than larger, more ‘cumbersome’ molecules. In PAGE the sieving matrix is crosslinked polyacrylamide, which provides a solid, porous network with a suitable mesh size.

In practice, the SDS-treated sample mixture is deposited in a reservoir at one end of a gel slab, and an electric voltage is applied over the gel, inducing protein migration through the gel toward the anode. After the separation, proteins can be visualised by staining, and the molecular weights determined by comparison to a known molecular weight marker.

SDS-PAGE separations were devised by Shapiro *et al*^[39] and have been a staple technique in biomolecular research for the last 40 years, with a variety of commercial systems available. Although the technique is relatively easy to use and does not require sophisticated instrumentation, it is not adequate for visualising the global proteome. Like liquid chromatography, gel electrophoresis does not have the capacity to resolve large numbers of proteins.

Capillary electrophoresis (CE): This is a family of techniques that share a common capillary-based format. Representatives include micellar electrokinetic capillary chromatography (MECC), isotacheophoresis (ITP), capillary zone electrophoresis (CZE), isoelectric focusing (IEF), and also capillary gel electrophoresis (CGE). CGE is closely related to gel electrophoresis in so far as the basic separation mechanism is the same, with SDS-treated proteins being sieved to separate by molecular weight. Instead of a crosslinked polyacrylamide slab gel, the sieving matrix is an entangled polymer solution such as linear polyacrylamide (LPA), poly(ethylene oxide) (PEO), or dextran. This solution is filled into capillaries, typically fused silica capillaries with 25–100 µm inner diameter. The sample mixture is injected at one end of the capillary, and proteins are made to migrate by applying an electric voltage over the capillary. Proteins are visualised near the cathode as they migrate past a detector.

In SDS-PAGE, the separation speed is limited by the heat generated by the flow of electric current (Joule heating). In capillary electrophoresis, the volume is very small with a high surface area-to-volume ratio, which is favourable for increased heat dissipation. This allows high field strengths of 500 V/cm and more, an order of magnitude greater than the typical field strength in SDS-PAGE. CE allows faster separations for the same peak capacity. Although CE can resolve a greater number of proteins than gel electrophoresis, the peak capacity (under good circumstances 100) is insufficient for a proteome with over 1,000 species.^[40–42] Both CE and SDS-PAGE can only determine the absolute molecular weight value of a protein by calibration against a known standard.

Isoelectric focusing (IEF): This is another electrophoretic technique, which, unlike SDS gel electrophoresis, relies on the intrinsic charge of proteins.

Most amino acid residue groups are neutral, but some are acids (glutamic acid and aspartic acid), while others are bases (histidine, lysine and arginine), see Figure 1.8. The basic groups protonate if the buffer pH is below their respective pI and contribute a positive charge to the intrinsic charge of a protein. The acidic groups ionise and carry a negative charge if the buffer pH is above their respective pI.

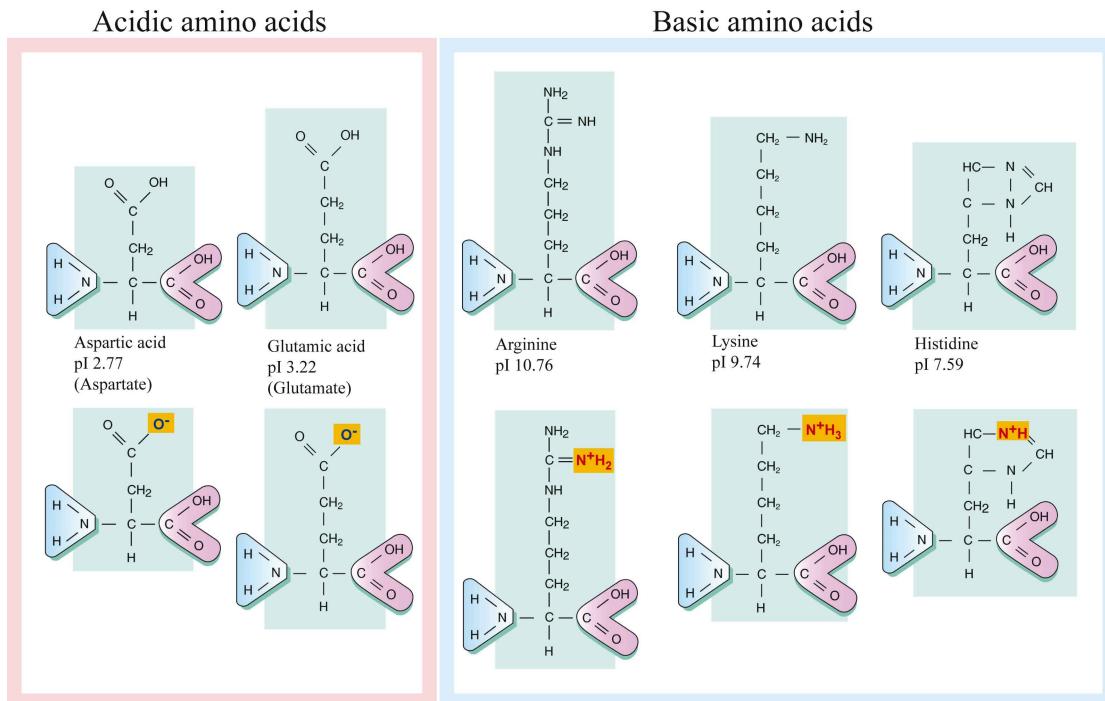


Figure 1.8: Acidic amino acids: glutamic acid and aspartic acid; and basic amino acids: histidine, lysine and arginine. Artwork from Mosby's Medical Dictionary^[43].

The relative number of basic and acidic amino acids is characteristic of a protein and is quantified in the protein's isoelectric point (pI). This is the pH at which the net charge of a protein is zero. Conversely, the further away the pH is from a protein's pI, the greater the charge on the protein.

In isoelectric focusing, the separation can take place in a slab-gel-based format, or in a capillary. Proteins are introduced into a pH gradient and subjected to an electric field. This causes the proteins to move toward their isoelectric point in the pH gradient. Once at the isoelectric point, proteins have zero net charge, and remain static. Figure 1.9 shows an illustration of isoelectric focusing.

Proteins not only separate, but actually focus at their isoelectric point, effectively concentrating the sample. This has many important implications. One is that sample injection does not affect the separation. Another is that theoretically, the separation

distance does not affect the resolution, favouring short separations. For routine isoelectric focusing, the pH gradient is generated with an immobilised pH gradient strip (IPG).

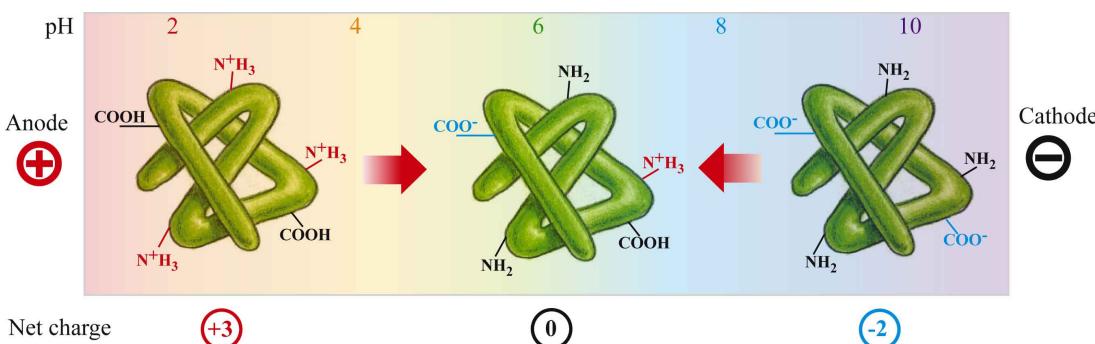


Figure 1.9: Isoelectric focusing: a protein migrates toward its isoelectric point in a pH gradient when an external electric field is applied. Artwork adapted from Molecular Biology of the Cell^[10].

Mass spectrometry (MS): This is the most common technique for protein identification, but it is in essence a separation technique, and as such a brief overview will be included here. MS separates peptides of differing mass-to-charge ratio (m/z) in the gas phase with electric or magnetic fields. Several techniques are used for ionisation, and for mass analysis. Peptides are usually ionised either by electrospray ionisation (ESI), or by matrix assisted laser desorption/ionisation (MALDI). Proteomics mass analysers include quadrupoles, ion-traps, time-of-flight (TOF) detectors, and Fourier transform ion cyclotron resonance (FT-ICR) analysers. Under some circumstances two mass analysers are coupled in series (tandem MS or MS/MS, for instance TOF-TOF), whereby further fragmentation of the peptides is induced between analysers. There are many different setups and combinations in use for MS analysis, each with strengths and weaknesses, but no real standard.

Generally the resolution of MS is very high, with excellent accuracy and sensitivity ($10^{-15}\text{--}10^{-1}$ mole quantities of sample with an precision of 0.1–0.01 %). Measurements are quick (in the range of minutes), and only small amounts of sample are required (less than 1 pmol). Some of the limitations are: larger molecules can be fragmented during ionisation (favouring peptide analysis over protein analysis); the same peptide can carry multiple charges (in particular in large molecules and ESI) which makes spectrum interpretation challenging; and quantification is difficult without internal standards and isotope labelling. For these reasons MS is usually applied not to investigate the proteome, but to identify individual proteins separated with other techniques.

For protein identification, peptide mass fingerprinting is the preferred technique. Starting with ideally a single protein, site-specific protein cleavage (typically with trypsin) is performed. The cleavage products are analysed to give a mass spectrum, which is compared to predicted mass spectra from protein databases. In the case of tandem MS, select peptides can be sequenced in addition to the mass spectrum, giving further information for identification.^[15,16,36]

1.2.4 Hyphenated and multidimensional separations

The separation techniques introduced in the foregoing section individually can separate a few tens of proteins up to at most a hundred, which falls far short of the 10,000–100,000 different protein species of a cell. The solution to this problem is to combine separations that are independent of one another, or orthogonal. The theoretical peak capacity is then the product of the two individual peak capacities.^[44] Combination of electrophoretic techniques (two-dimensional gel electrophoresis), and combination of chromatographic techniques are discussed.

Two-dimensional gel electrophoresis (2DGE), also two-dimensional electrophoresis (2DE), or two-dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (2D SDS-PAGE): For dealing with larger numbers of proteins, this technique of choice combines IEF with SDS PAGE.^[45] The sample is separated by isoelectric focusing in the first dimension. The resulting strip is treated with SDS and affixed to a slab of polyacrylamide gel. Thereafter, an electric field is applied orthogonal to the first separation, and proteins are separated by SDS-PAGE. The result is a two-dimensional pattern of protein spots, as illustrated in Figure 1.10. After the separation, proteins can be stained, the gel scanned and processed with image recognition analysis, and compared to other gels for relative quantification. Finally, spots of interest can be excised, purified, cleaved to smaller polypeptide fragments, and identified with mass spectrometry.

2D gel electrophoresis was developed over 30 years ago,^[29] and is still one of the most powerful protein separation tools in use. Either of the electrophoretic techniques of SDS gel electrophoresis or IEF can resolve a few tens of proteins up to at most a little over a hundred. The peak capacity of a 2D method is the product of the peak capacities of the component 1D methods.^[44] A theoretical peak capacity for an 18 × 18 cm gel is around 15,000 spots, and under the more realistic consideration of non-uniform distribution and wide abundance range the theoretical capacity is 2,000 spots.^[47] Depending on the type of stain (Coomassie, silver, fluorescent dyes) the sensitivity is in the range of a few nanograms to tens of nanograms protein mass, and the dynamic range is around 10–30

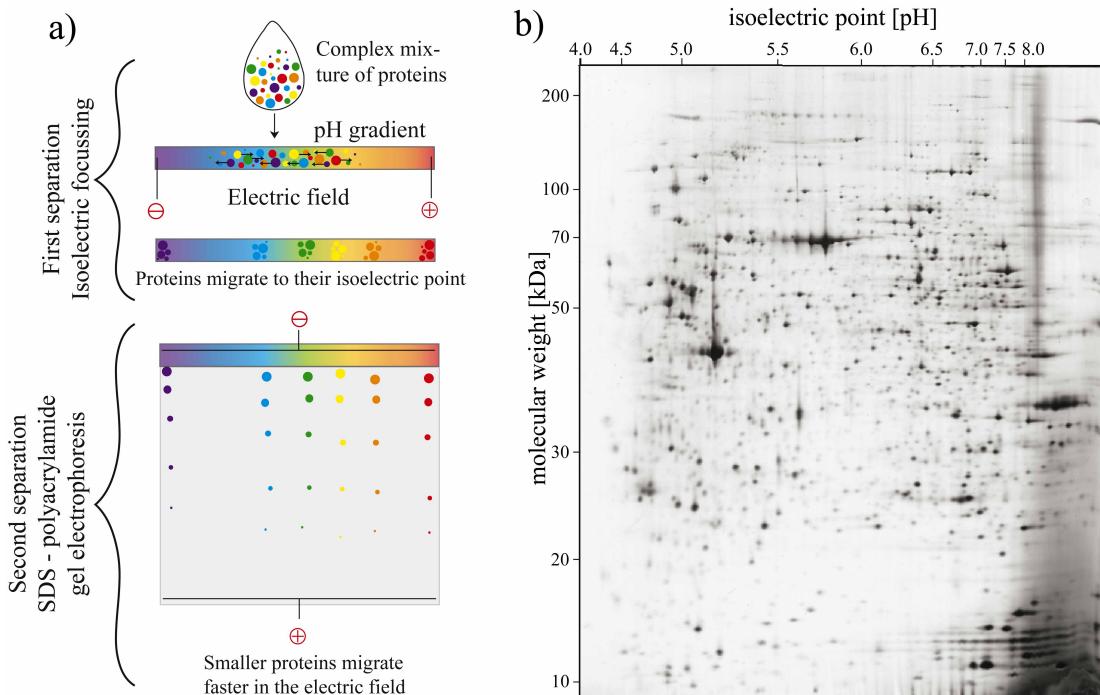


Figure 1.10: Two-dimensional gel electrophoresis. a) Schematic of a two-dimensional gel electrophoresis separation. After isoelectric focusing follows polyacrylamide gel electrophoresis. b) 2D gel image (human lymphoma, from ExPASy^[46]).

(higher for fluorescent dyes, up to 10^3). The dynamic range of proteins present in a cell is around 10^5 molecules.

In practice a 2D gel can resolve up to around 6,000 protein spots, making it the most powerful protein separation tool available, and one of the key techniques in proteomics. Despite its essential role in proteomics, 2D electrophoresis has a number of shortcomings, and it has been postulated repeatedly that it is time for a technology shift.^[48] The price for the outstanding resolving power is that 2D gel electrophoresis is very slow, labour-intensive and difficult to reproduce. It might easily take up to a week to obtain data from a sample, and there are many human interventions necessary, making it poorly suitable for automation.^[49] This low throughput is a significant bottleneck for protein expression screening. In many textbooks on the subject of proteomics, the reader encounters a statement to the effect of:

“Methodologies such as 2D PAGE are labour-intensive and do not lend themselves easily to the analysis and comparison of dozens of samples.”^[37]

Another obstacle is that 2D gels suffer moderate reproducibility and comparing different gels for quantification can be ambiguous. These drawbacks have inspired some innovations and much research, yet the need for a technique with better performance remains unsatisfied. One of the trends is the use of prefabricated gels for IEF and SDS

PAGE, which has improved reproducibility to some extent.^[50] With regards to protein quantification, differential gel electrophoresis (DIGE) allows better comparison and calibration. However, DIGE involves labelling proteins which can influence separation results.^[51,52] Although 2D gel electrophoresis is the gold standard of protein separation, low throughput and mediocre reproducibility remain significant problems.

Two-dimensional liquid chromatography (2DLC), also multidimensional liquid chromatography (as in multidimensional protein identification technology, MuDPIT), hyphenated liquid chromatography (LC-LC or (LC)ⁿ, or shotgun proteomics (coupled to mass spectrometry): Given the challenges of using 2D electrophoresis, separation by two coupled chromatographic techniques has been developed.^[16,36,53–55] Most commonly, reversed phase chromatography is coupled to ion exchange chromatography, either in one single column, in connected columns, or with fraction collection between two independent columns. The whole protein mixture is digested with trypsin prior to separation. The eluate can be coupled directly into a mass spectrometer with electrospray ionisation. Furthermore, comparison of two samples can be achieved by labelling one with a mass-coded tag, prior to performing 2D liquid chromatography on the combined samples.

The advantage of 2D liquid chromatography is that it can be fully automated, which implies fewer losses, less contamination, and greater ease of use. Although both 2D gel electrophoresis and 2D liquid chromatography struggle with particular proteins (2D gel electrophoresis with basic, large, hydrophobic proteins, 2D liquid chromatography with small, hydrophilic peptides), they are mutually complementary. Disadvantages of 2DLC are that the technique fails if the sample is too complex, and very high computational power becomes necessary to identify proteins with good confidence.

Table 1.2: Typical performance metrics for a variety of protein separation techniques.

	Liquid chromatography ¹	Protein chips ²	Gel electrophoresis ³	Capillary electrophoresis ⁴	Isoelectric focusing ⁵	2D gel electrophoresis ⁶	2D liquid chromatography ⁷
Peak capacity	50 ^[36]	100–200 (maximum 5800) ^[16]	65 ^[40]	65 ^[40]	130 ^[56]	2,000–10,000 ^[47]	100–500 ^[36]
Dynamic range (orders of magnitude)	3 ^[36]	semi-quantitative ^[57]	< 3 ^[45]	4 ^[58]	< 3 ^[59]	< 3 ^[36]	3 ^[36]
Sensitivity	< 100 pg ^[36]	1 pg ^[57]	< 0.01 pg ^[45]	400 pg ^[58]	100 pg ^[59]	100 pg ^[50]	20 pg ^[36]
Accuracy	n/a	n/a	1–10 % ^[38]	1–10 % ^[58]	0.15 pH unit ^[60]	1–10 % ^[38]	< 1 Da ^[38]
Sample load	μL ^[57]	10–60 μL ^[57]	< 50 mg ^[45]	nL ^[58]	50–100 μg ^[50]	50–100 μg ^[50]	μL ^[36]

¹ Reverse phase LC, nanoLC, UV detection² Fluorescent detection³ SDS-PAGE, silver stain⁴ Capillary gel electrophoresis, UV detection⁵ In immobilised pH gradient strip, silver stain⁶ Silver stain⁷ Ion exchange-reverse phase LC, nanoLC coupled to nanoESI, detection by MS

1.3 Protein electrophoresis

In this section different aspects of protein electrophoresis will be introduced in more detail. This includes fundamental laws governing electrophoresis as well as a more detailed discussion of different modes of electrophoresis. Particular attention will be given to capillary electrophoresis. Scaling laws will be discussed, and the technology of Label Free Intrinsic Imaging will be introduced.

1.3.1 Background

Under the influence of an external electric field, charged molecules migrate toward the electrodes. This is the principle of electrophoresis. Charged particles experience an electric force \vec{F}_E [N] in the direction of an electric field,

$$\vec{F}_E = q\vec{E} \quad (1.15)$$

where q [C] is the charge, and \vec{E} [V/m] is the electric field. In media, particles experience a friction force \vec{F}_D [N] in the opposing direction, as described by Stokes' law,

$$\vec{F}_D = 6\eta r\vec{v} \quad (1.16)$$

where r [m] is the Stokes' radius of the particle, η [sPa] is the viscosity of the medium, and \vec{v} [m/s] is the velocity. At constant migration, the electric force \vec{F}_E and the drag force \vec{F}_D are equal, and the electrophoretic mobility μ [$\text{m}^2/(\text{sV})$] can be described as:^[40]

$$\mu = \frac{|\vec{v}|}{|\vec{E}|} = \frac{q}{6\eta r} \quad (1.17)$$

The electrophoretic mobility is commonly used to describe the migration of charged species. In the case of proteins, which are not necessarily charged, the two widespread techniques of sodium dodecyl sulphate gel electrophoresis and isoelectric focusing employ different methods of fixing a charge to proteins, and this allows separation. In gel electrophoresis techniques, sodium dodecyl sulphate (SDS) confers a uniform charge per length, which affects the (q/r) term in Equation 1.17. All proteins have the same mobility, but they are made to migrate through a sieving matrix, which encumbers long proteins more than short proteins.

Isoelectric focusing does not employ additives, but relies on the intrinsic charge of proteins. The intrinsic charge of a protein in aqueous solution depends on the pH. Each protein has a characteristic isoelectric point (pI), which is the pH at which the protein

has zero net charge. In isoelectric focusing (IEF), proteins are introduced into a pH gradient and subjected to an electric field, which causes proteins to move toward their isoelectric point in the pH gradient. Individual proteins that diffuse away from their pI lose neutrality, and return to their pI. This effectively focuses and concentrates proteins at their isoelectric point.

The pH gradient in IEF can be generated by different means. The most common methods are use of carrier ampholytes (CA) and immobilised pH gradients (IPG). Carrier ampholytes are a large group of oligoamino-oligo-carboxylic acids that can be designed with varying pKa. They display a good buffering behaviour around their isoelectric point and therefore are good candidates for generating pH gradients. The pH gradients display drift, and can be uneven. CAs are rather expensive, can interact with protein samples, and are troublesome to eliminate. They reduce sensitivity of UV detection and interfere with mass spectrometry. Buffer mixtures other than ampholines have been suggested.^[61] However, they have largely been discontinued due to poorer pH gradient characteristics with exception of specialist applications.^[62–64] An immobilised pH gradient (IPG) is essentially a polyacrylamide gel that has copolymerised with carrier ampholytes.^[59] A gradient of ImmobilinesTM (acrylamido-derivatives with carboxyl- and amino-groups) is cast with a gradient mixer, and thereafter copolymerised with acrylamide.^[65]

1.3.2 Capillary electrophoresis

While the previously described electrophoretic techniques differ in how proteins are mobilised and separated, ‘capillary electrophoresis’ designates a particular format of separation, with associated characteristic behaviour. In capillary electrophoresis (CE), separations are performed in capillaries with an inner diameter of 20–200 µm with high field strengths of 500 V/cm and more. Conventional electrophoresis in gels such as polyacrylamide or agarose operates with slabs of tens of centimetres in length and field strengths of around 10 V/cm. By contrast capillary electrophoresis permits far faster separations with the same or better resolution. This is due to reduced Joule heating, improved heat dissipation, and less diffusion.

One phenomenon particular to capillary electrophoresis is electro-osmotic flow (EOF), in which surface charges on the capillary induce hydrodynamic flow in the capillary. This is described by the equation:

$$\mu_{EOF} = \frac{v_{EOF}}{E} = \frac{\epsilon\xi}{\eta} \quad (1.18)$$

where μ_{EOF} [m²/(sV)] is the electro-osmotic mobility, v_{EOF} [m/s] is the velocity, E [V/m] is the electric field, ϵ [Farad/m] is the dielectric constant of the liquid, ξ [V] is the zeta

potential of the liquid-solid interface, and η [sPa] is the viscosity of the liquid.^[40] Due to the surface charge on the silica walls, solute ions are not distributed homogeneously but form a gradient, as illustrated in Figure 1.11a.^[40,66,67] In particular the compact and diffuse layers (in the range of a few tens of nanometres from the interface) are of relevance, since they both have an overall charge and are mobile. Under the influence of an external electric field, these layers experience a force and migrate toward the electrodes. This causes a flow in the entire capillary. In the case of negatively charged silica walls it causes a flow toward the cathode with a mobility of around $10^{-8} \text{ m}^2/(\text{sV})$.

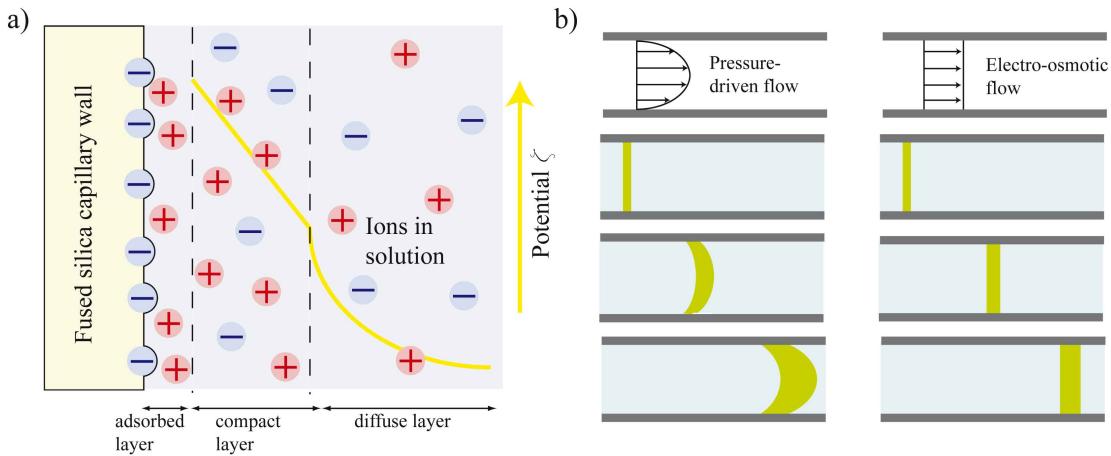


Figure 1.11: Electro-osmotic flow. a) Diagram of ion distribution near a charged surface. After Weinberger^[40]. b) Plug dispersion and flow profiles for pressure-driven and electro-osmotic flow.

While the flow profile across the capillary is parabolic for pressure-driven flow, the electro-osmotic flow profile is uniform across the channel, as illustrated in Figure 1.11b. Pressure driven flow is unwanted in a separation, as the parabolic flow shape adds extra band widening in addition to diffusion. Electro-osmotic flow on the other hand does not affect bands, and transports the undisturbed capillary filling along. In the most frequently encountered case of untreated fused silica capillary walls and negatively charged proteins (e.g. from SDS), the electro-osmotic flow direction is opposite to the protein migration direction. By convention the term ‘normal polarity’ refers to a separation configuration where injection occurs at the anodic end, and detection occurs at the cathodic end of the capillary. Here the net mobility of the analyte is toward the cathode. SDS-CGE separations are usually operated at reverse polarity, whereas EOF-based separations often proceed at normal polarity. Capillary surface charge and EOF can be tuned in direction and magnitude with a variety of surface treatments. Local adhesion of proteins to the surface causes discontinuity in the electro-osmotic flow, resulting in non-uniform flow profiles and band broadening. Surface charge and therefore electro-osmotic flow is influenced by buffer ionic strength and pH as well as a number of organic

modifiers and surface coatings. Controlling the surface properties to minimise variations in electro-osmotic flow is an important aspect of capillary electrophoresis.

Within the capillary format, different electrophoretic separation modes can be implemented. Capillary gel electrophoresis (CGE) and polyacrylamide gel electrophoresis (PAGE) share the same separation mechanism. In CGE, filling the capillaries with cross-linked polyacrylamide is less common, and preferentially aqueous solutions of entangled linear polymers are employed, for instance poly(ethylene oxide) (PEO), linear polyacrylamide (LPA), dextran, or polymers of the cellulose family.^[68,69] In capillary zone electrophoresis (CZE), proteins migrate depending on their intrinsic charge-to-size ratio in a suitable non-sieving buffer. In capillary isoelectric focusing carrier ampholytes are often used to generate the pH gradient, and the EOF can be used to move the focused proteins past a detector.

Several detection schemes are common in capillary electrophoresis. In absorption detection, the capillary is illuminated at a wavelength where the absorption of analyte and buffer differs. The absorption of light passing through a solution of concentration c [mol/m³], is described by the Beer-Lambert law:

$$I = I_0 \times 10^{-c\epsilon(\lambda)L_p} \quad (1.19)$$

where I_0 [W/m²] is the light intensity prior to entering the solution, I [W/m²] is the intensity after the light has passed through, L_p [m] is the path length the light travels through the solution, and $\epsilon(\lambda)$ [m²/mol] is the wavelength-dependent molar absorption coefficient.^[58,70]

From Equation 1.19 different approaches to improving sensitivity can be derived. The illumination must be optimised to avoid light passing beside the lumen of the capillary, which would cause an unwanted increase in the background signal.^[70,71] Increasing the path length the light travels through the solution also increases sensitivity. Configurations with Z-cells and bubble-cells have been suggested to this ends.^[40] Z cells include bends in the capillary such that the illumination occurs along the capillary axis for a stretch. Bubble cells feature a detection zone with a greater cross section area. Finally, as the molar absorption coefficients of both analyte and buffer depend on the illumination wavelength, this may be chosen to maximise the difference in absorption.

For proteins, the low ultraviolet wavelength is suitable (200–220 nm), because aqueous solutions are transparent and proteins absorb well. Diode array detector instruments can simultaneously detect at different wavelengths, although this is more common for detecting small molecules than proteins.^[72,73] Instead of using a wavelength where the analyte absorbs strongly, the background medium can be modified with an additive that

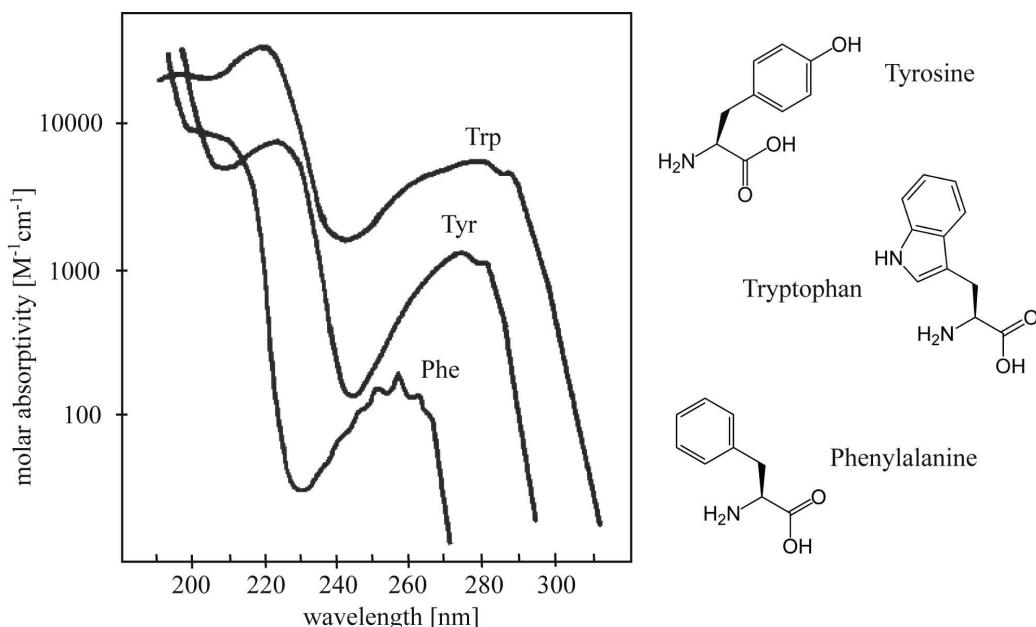


Figure 1.12: Absorption spectra of aromatic amino acids tyrosine, tryptophan, and phenylalanine. Absorption spectra from Neckers^[77].

absorbs strongly where the analyte is transparent (indirect absorbance detection).^[58] Another common detection scheme involves sample (or background) fluorescence. Although this can be very sensitive, it requires derivatisation of many analytes, including proteins.^[74] Other more unusual approaches have been reported, for instance radioactive detection, Raman scattering detection, refractive index detection, or electrochemical detection.^[40,58,75,76]

Proteins absorb at around 280 nm due to aromatic amino acid residue groups in tryptophan, tyrosine, and phenylalanine (Figure 1.12). Since different proteins have different amounts of aromatic residues, the absorption at 280 nm is not only dependent on the amount of protein in the detection volume, but also on composition. At lower wavelengths, below 220 nm and with a maximum at 190 nm, the peptide bond absorbs light. The peptide bond is equally present in all proteins, therefore the absorption below 220 nm is related to the quantity of amino acids in the detection volume. UV light at 214 nm is employed to avoid signal variation depending on protein species, and to remain within the limits of quartz transparency (which begins to drop below 200 nm).

The efficiency of capillary electrophoresis is often given in terms of the theoretical plate number. This is a figure that is related to the maximum number of peaks that could be observed (if the entire length of capillary were filled with consecutive peaks, each peak just distinguishable from the next). In the simplest of approximations, the only reason for peaks to have a width is molecular diffusion. The Einstein equation (1.20) describes

diffusion in liquids with:^[40]

$$\sigma_D^2 = 2Dt \quad (1.20)$$

where σ_D^2 [m²] is the zone variance, D [m²/s] is the diffusion coefficient, and t [s] the time. The theoretical plate number N is defined as:

$$N = \frac{L_s^2}{\sigma_D^2} \quad (1.21)$$

for a separation length of L_s [m]. This is the distance from the capillary inlet to the point of detection. A related figure of merit is the peak capacity n , with:

$$n = \frac{L_s}{4\sigma_D} \quad (1.22)$$

Customarily the time t in Equation 1.20 is replaced with:

$$t = \frac{L_s}{v} = \frac{L_s L_c}{\mu V} \quad (1.23)$$

using $v = \mu E$ (see also Equation 1.17), and $E = V/L_c$. Here v [m/s] is the migration speed, μ [m²/(sV)] the mobility, E [V/m] the electric field strength, V [V] the separation voltage, and L_c [m] the capillary length. Inserting Equations 1.20 and 1.23 into Equation 1.21 gives:^[58]

$$N = \frac{\mu V}{2D} \frac{L_s}{L_c} \quad (1.24)$$

Assuming $L_s \approx L_c$, the $\frac{L_s}{L_c}$ term becomes negligible, and the separation length does not affect the separation efficiency. The most important experimental factor is the applied voltage.

Beside molecular diffusion, numerous other factors can contribute to the width of a band. Common sources of band widening include injection width, capillary wall interactions, thermal effects, laminar flow, and geometric broadening. It is not always possible to accurately describe the overall band width analytically, but the equations above can provide a best-case estimate.^[40]

The techniques for sample injection are also particular to capillary electrophoresis. In gel electrophoresis, the sample is deposited in loading cavities, ‘wells’, in the gel. Since this is not possible in capillary electrophoresis, the techniques of electrokinetic injection and pressure injection are applied. For electrokinetic injection, the capillary end is immersed in the sample vial, and an electric field is applied. This causes the sample solution to migrate into the capillary. After a short time, the capillary end is removed from the sample and replaced in the reservoir vial. Now the separation of the sample plug in the capillary can begin. Electrokinetic injection has the advantage that it is fully compatible with the electrophoresis hardware that is already available. It can introduce

a bias in the sample, since not all species of a complex sample migrate into the capillary with the same speed. The alternative sample injection technique, pressure injection, forces a sample plug into the capillary under external pressure, either with pressurised gas, or by changing the liquid levels so that the sample end is higher than the capillary outlet for a short time. While this avoids sample biases, it requires more sophisticated instrumentation.

1.3.3 Dimensions and scaling laws

As previously indicated, electrophoretic separations benefit from size reduction for reasons of improved thermal characteristics. This can be seen from Equations 1.25 and 1.26, which describe the power input from Joule heating and power drain from convective heat transport. Joule heating:

$$P = \frac{\partial Q}{\partial t} = \dot{Q}_{in} = \frac{V^2}{R} = \frac{V^2 A}{\rho L} = \frac{V^2 d^2 \pi}{4\rho L} \quad (1.25)$$

with power P [W], energy Q [J], time t [s], voltage V [V], resistance R [Ω], cross section area A [m^2], resistivity ρ [Ωm], capillary length L [m], and capillary diameter d [m]. Newton's law of cooling:

$$\frac{\partial Q}{\partial t} = \dot{Q}_{out} = kS(T - T_\infty) = k\Delta T \pi d L \quad (1.26)$$

with heat transfer coefficient k [$\text{W}/(\text{m}^2\text{K})$], surface area across which heat is transferred S [m^2], temperature T [K], and environment temperature T_∞ [K]. The quotient between power input and drain, in dependence of capillary diameter, becomes:

$$\frac{\dot{Q}_{in}}{\dot{Q}_{out}} = \frac{V^2 d^2 \pi}{4\rho L k \Delta T \pi d L} = \frac{V^2 d}{4\rho k \Delta T L^2} = \frac{E^2 d}{4\rho k \Delta T} \quad (1.27)$$

with electric field strength E [V/m] ($E = V/L$). A large (input/drain) quotient is indicative of poor cooling, where a lot of the power remains in the capillary and causes the temperature to increase. From Equation 1.27, for the same electric field strength, the smaller the diameter is, the more efficient cooling becomes.

The advantages of miniaturisation are often demonstrated with scaling laws. Forces that are a function of the area decrease more slowly than forces that depend on the volume. For electrophoresis applications, heating is a critical factor, but other experimental parameters such as separation time or detection can be described in terms of scaling factors. In diffusion controlled systems, the time scales with (typical length)². For a 10-fold reduction of diameter, the typical separation time is 100-fold reduced.^[78]

The different detection schemes have different dimension dependencies. Absorption detection depends on a linear function of dimension. The absorbance (and detection signal) depends on the length of the light path, as seen in Beer-Lambert's law (Equation 1.19). In fluorescence detection, the signal depends on the volume of sample that fluoresces. Although fluorescent detection is more sensitive than absorbance detection, it suffers a greater decrease in signal with miniaturisation.^[79]

In addition to the benefits that scaling can bring regarding performance, small volumes mean low sample consumption, which can be of importance with scarce samples. Microfluidics also offers a platform for micro total analysis systems (μ TAS), where a variety of functionalities are combined and implemented on a single chip.^[80] This is more amenable toward automation, which in turn makes a technique easier to use and reduces operator variability. It also means that a multiplexed channel layout is easy to implement, which allows numerous separations in parallel.^[81,82]

1.3.4 Label Free Intrinsic Imaging

DeltaDOT Ltd. is a spin-out company from Imperial College London that has developed tools and techniques to enhance the sensitivity of UV absorbance detection in capillary electrophoresis. The core intellectual property is based around the concept of Label Free Intrinsic Imaging (LFII) and its associated analysis algorithms.^[83–88]

A significant limitation in CE is that when UV absorbance detection is used, only poor concentration detection limits are attainable. At the heart of the deltaDOT algorithms is the principle that accuracy can be improved by performing multiple measurements and subsequent signal averaging. Signal averaging may be achieved during a single run, by means of multiple point detection along the capillary. The signal-to-noise enhancement is equal to the square root of the number of diode detectors, if the detection is shot-noise- or white (random)-noise-limited.

Typically, linear photodiode arrays are employed with 512 consecutive pixels. The pixels are arranged linearly along the capillary. Each pixel produces an independent electropherogram, thus generating 512 electropherograms available for signal averaging. Generalised Separation Transform (GST) is the algorithm designed to combine all electropherograms, while retaining maximum information on shape and absolute peak size. This can bring about a signal-to-noise improvement of up to 22.6, theoretically (square root of 512). In fact, the improvement is typically less than this, as there are other factors (beside the diode white noise) that contribute to detection errors.

The Equiphase Vertexing Algorithm (EVA) is the core of the deltaDOT analysis package. In each individual electropherogram, a band passing that pixel coincides with a signal minimum. In an equiphase map, these minima are plotted in a space-time diagram, as seen in Figure 1.13. In this representation, analyte bands are seen as straight lines,

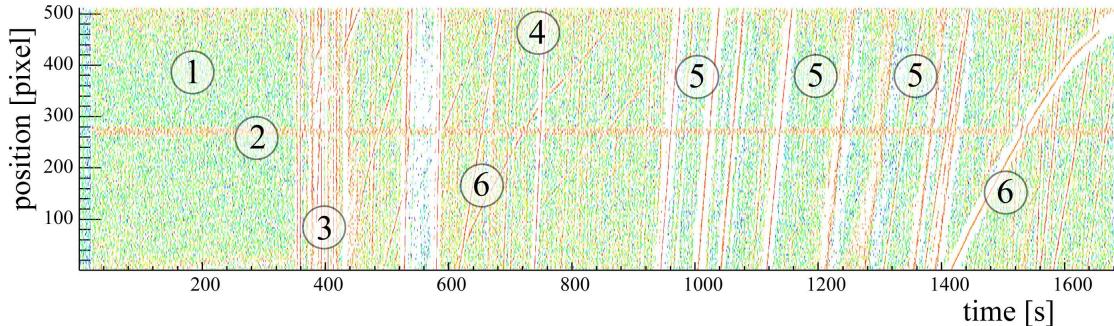


Figure 1.13: Equiphase map. 1) Background signal. 2) Marking on capillary for alignment. 3) Light saturation from opening and shutting instrument lid. 4) Bands from sample buffer chemicals. 5) Protein bands migrating past detection window. 6) Bubble or other non-analyte contaminant. Data from capillary gel electrophoresis separation (BioRad CE-SDS run buffer) of a commercial standard (8 proteins, BioRad CE-SDS Protein Size Standard 14-200 kDa). Experimental conditions: fused silica capillary, 75 μm ID / 375 μm OD, 20 cm separation length, total length 35 cm, separation voltage 12 kV, electrokinetic injection for 10 sec at 10 kV, detector pixels 25 μm wide.

because they progress past the detector in a linear fashion. Background signals appear as randomly distributed points. The slope of analyte lines in an equiphase map represents their velocity. All analytes that were injected into the capillary together share a common (time/position) coordinate in their trajectory. In a time-position plot of the migration trajectories, analytes that were injected together form a vertex. This is very helpful to distinguish analytes from other things that are detected, such as bubbles or particles in the background buffer. In Figure 1.13, number 6 labels two non-analytes that migrate past the detector on irregular trajectories. As all the bands from a common sample injection start at the same point in space-time, these lines will have a common vertex. A fit to this vertex can reduce the error on the fit of the velocities.^[89]

1.4 Microfluidics and electrophoresis

Electrophoresis in microstructures benefits from the reduced dimensions. As previously elaborated, smaller separation geometries inherently improve separation performance. Smaller channels mean greater surface area-to-volume ratios, which in turn improves heat dissipation. Better heat dissipation means higher voltages can be applied without a loss of resolution. As greater electric fields can be applied, faster separations can be achieved. With smaller dimensions, reduced sample volumes are required. Using a lab-on-a-chip approach, multiple functionalities (such as sample loading, injection,

and separation) can be implemented in the same microfluidic structure. This concept of micro total analysis system (μ TAS) is favourable with regards to automation, ease of use, and minimal user manipulation.^[90] Chromatographic techniques benefit from miniaturisation to a lesser degree. As a pressure driven separation, very narrow channels can cause high back pressures. Capillary electrochromatography (CEC) is a notable exception, where an essentially chromatographic separation is driven by electro-osmotic flow.

1.4.1 One-dimensional electrophoresis on-chip

Much work has been published on capillary electrophoresis in microfluidic chips.^[91] Many different substrate materials have been studied, such as silicon,^[92] glass,^[93,94] poly(methyl methacrylate) (PMMA),^[95] poly(dimethyl siloxane) (PDMS),^[74] SU-8 photoresist,^[96] cyclo-olefin copolymer (COC),^[97] and other polymers.^[98] To deal with problems such as protein adsorption and unfavourable electro-osmotic flow, a number of surface modifications have been reported.^[99,100] Chip-specific developments such as optimised sample injection,^[101] integrated detection,^[102] and sample stacking^[103] have also been studied. Caliper Life Sciences is marketing the LabChip^[104] for use in Agilent's 2100 Bioanalyzer for electrophoresis of proteins, DNA, and RNA.

Isoelectric focusing in microfluidic environments has been reported for a variety of approaches. One of the crucial elements is generating a pH gradient. With carrier ampholytes, a pH gradient forms under the influence of an electric field. Because this is compatible with many separation designs, isoelectric focusing with carrier ampholytes has been demonstrated in numerous microfluidic systems including in glass chips,^[105] in PMMA chips,^[106] PDMS chips,^[107] and COC chips.^[108,109] Optimised sample introduction has been studied,^[110] and multistage isoelectric focusing has been reported.^[111] A variation of isoelectric focusing, free-flow IEF, has been reported,^[112–114] along with other pI-based free flow separations.^[115,116]

In some instances carrier ampholytes are not the ideal choice to generate the pH gradient, in particular if the pH drift is unacceptable, or if carrier ampholytes interfere with the downstream analysis. Here immobilised pH gradients can be more useful.

Several techniques are reported for generating gradients in microfluidic structures, for example with ladder-like mixing regions,^[117,118] with multiple parallel channels,^[119] with interconnected reservoirs and electro-osmotic flow,^[120] or dynamically with buffer pulses.^[121] Though many of these approaches are more suitable to continuous flow gradients, they could be used to generate a pH gradient prior to immobilisation.

Among the reports dedicated to forming immobilised pH gradients on-chip, approaches using monolithic columns with carrier ampholytes grafted to the surface are found.^[122–124] Success has also been reported with diffusion-based gradients in acrylamide that are photopolymerised.^[56] As an alternative to smooth gradients, multiple chambers with discreet pH values have been used to demonstrate protein separation.^[125,126] Other methods to generate a pH gradient include thermal techniques,^[127–130] or electrolysis of water.^[131,132] Going beyond isoelectric focusing in a pH gradient, alternative electrofocusing techniques have been suggested, such as hydrodynamic flow gradient against electrophoretic migration, hydrodynamic flow against electrophoretic migration gradient, electrophoresis in a conductivity gradient (ion concentration gradient), or electrophoresis against a gradient of denaturant.^[133]

1.4.2 Two-dimensional electrophoresis on-chip

Academic research reflects the quest for improved protein separation techniques. Compared to conventional techniques, microfluidic structures have the advantages of convenience for automation, and ease of integration of downstream analysis and auxiliary functions. A number of publications report a variety of approaches for implementing two-dimensional electrophoresis on-chip. Previous work can be categorised as either 2D by sampling, or 2D by capillary arrays. An example of both can be seen in Figure 1.14.

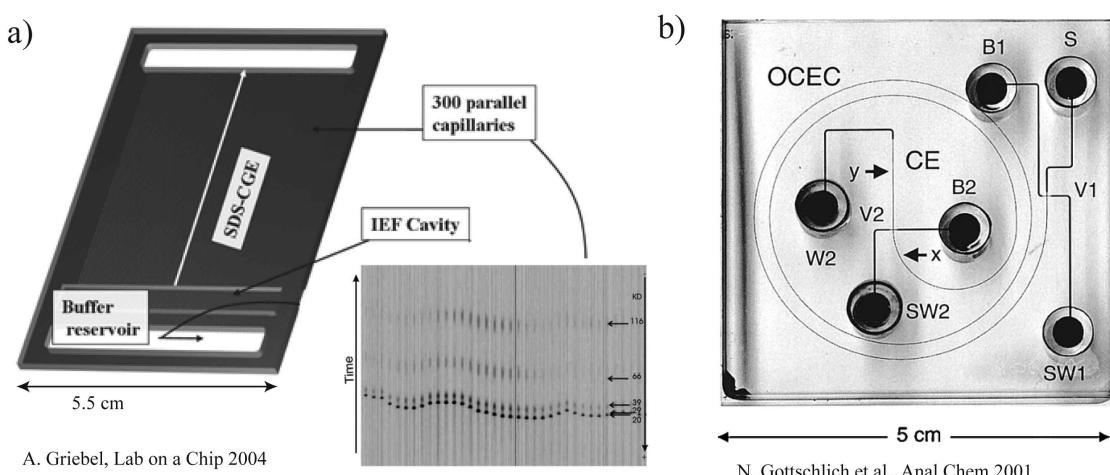


Figure 1.14: Examples of previous work on two-dimensional electrophoresis in microsystems. a) 2D by capillary array: the second separation takes place in an array of capillaries orthogonal to the first separation.^[134] b) 2D by sampling: the second separation takes place in a single cross-channel.^[135]

Microfluidic structures for 2D separations by capillary arrays consist of a channel for a first IEF separation and an array of orthogonal capillaries for a second SDS separation. This is in close resemblance to conventional 2D gel electrophoresis, with the slab gel replaced by a multitude of parallel channels. Capillary array systems have

been fabricated with poly(dimethylsiloxane),^[136] cyclic olefin polymer,^[137] quartz,^[138] poly(methyl methacrylate),^[139,140] and glass.^[141] The number of orthogonal capillaries in the array section has ranged from 10^[142] to 300.^[134] Variations have been reported, such as capillary zone electrophoresis in the first dimension,^[143] or native gel electrophoresis in the second dimension.^[144]

Uniformity over the capillary arrays is one challenge facing this type of design. Furthermore, the resolution is reduced to the capillary array ‘bin width’. If there are 10 orthogonal capillaries, the focused proteins are split into 10 groups for the second separation, so the resolution of the first separation is at best 1/10th the span of the first separation. Given the shortcomings of capillary arrays, there have been some attempts toward semi-microfluidic integration of isoelectric focusing and polyacrylamide slab gel electrophoresis.^[145–148] Although these systems generally perform faster than conventional 2D gel electrophoresis, they are still slower than microfluidic separations and do not achieve the resolution of conventional 2D electrophoresis.

The second category, 2D separations with effluent sampling, also termed ‘hyphenated’ separations, combine any two separation modes by collecting the effluent of the first separation with a given sampling frequency and feeding the batches into a second separation. The separations can be the classical combination of isoelectric focusing and capillary gel electrophoresis, for instance in poly(dimethylsiloxane) microfluidic chips^[149] or in poly(methyl methacrylate) microfluidic chips,^[150] or with a non-microfluidic setup.^[151] Other separations have also been combined, either in microfluidic channels or in capillaries. Such hyphenated separations include the combination of isoelectric focusing and transient isotachophoresis-zone electrophoresis;^[152] isoelectric focusing and capillary zone electrophoresis;^[150,153,154] micellar electrokinetic chromatography and capillary gel electrophoresis;^[155–159] micellar electrokinetic chromatography and capillary zone electrophoresis;^[160,161] micellar electrokinetic chromatography and capillary isoelectric focusing;^[162] electrochromatography and capillary electrophoresis;^[135] isoelectric focusing and hollow fibre flow field fractionation;^[163] and isoelectric focusing and liquid chromatography.^[164] There have even been reports of three-dimensional separations coupling capillary zone electrophoresis and isotachophoretic separations.^[165] Furthermore, capillary electrophoresis has been coupled with mass spectrometry^[166,167] which constitutes a two-dimensional separation. Other non-electrophoretic separations such as two liquid chromatographic separations have been coupled. Multi-dimensional protein identification technology (MuDPIT) is a well known representative of this approach.^[54]

One of the inherent weaknesses of sampling techniques is that the separation outcome after the first dimension is ‘binned’, reducing resolution. Although the resolution can be improved by increasing the sampling rate, this increases the number of individual

separations. In all two-dimensional hyphenated separations, there is an unfavourable trade-off between resolution and speed.

Finally, there have been reports of separations that measure two different properties of the analytes by unconventional means. In one instance, proteins were first subjected to isoelectric focusing, before being allowed to diffuse freely in the capillary. Diffusion is related to molecular weight, allowing an estimate of protein size.^[168] Other work used partial capillary filling^[169] and plugs^[170] to modify the chemical state or separation conditions mid-separation. Although these approaches make use of multiple properties of analytes, they only used them to improve the separation result, and not to record different analyte properties.

Of all the published work, none has achieved resolutions comparable to those of conventional 2D electrophoresis, due to the inherent limitations of temporal and spatial sampling approaches. There remains a need for fast, high-resolution multi-dimensional separations of complex protein samples.

1.5 Virtual two-dimensional separations

Instead of performing two separations orthogonally, virtual two-dimensional (V2D) electrophoresis performs two sequential separations collinearly, and employs time- and space-resolved measurement to track and reconstruct protein migration.^[171,172] This approach is derived from ‘vertexing’, which has its progenitor in recognition of short-lived decays in high energy physics. Unlike the techniques discussed above, the resolution is limited in the first instance by the resolution of the analytical reconstruction. The sequence of procedures is as follows: a first separation over isoelectric point (pI) is followed by a second separation over molecular weight (MW). This is illustrated in Figure 1.15.

A first separation is performed much the same as in conventional 2D electrophoresis. In a pH gradient, proteins migrate to their isoelectric point in an electric field. The pH gradient is immobilised in a gel-based separation matrix (immobilised pH gradient strip, IPG). Once the proteins are focused, the pH strip is subjected to a rinse with SDS, much as in conventional 2D electrophoresis, resulting in the electric charge necessary for the second (molecular weight) separation.

The second separation is performed in a collinear channel that extends beyond one end of the pH gradient. The second separation is exactly the same as a conventional capillary gel electrophoresis separation, with the difference that there is no common starting point, but the starting point of each protein is defined by its isoelectric point. In the course of the second separation small, fast proteins can overtake larger, slower proteins as they

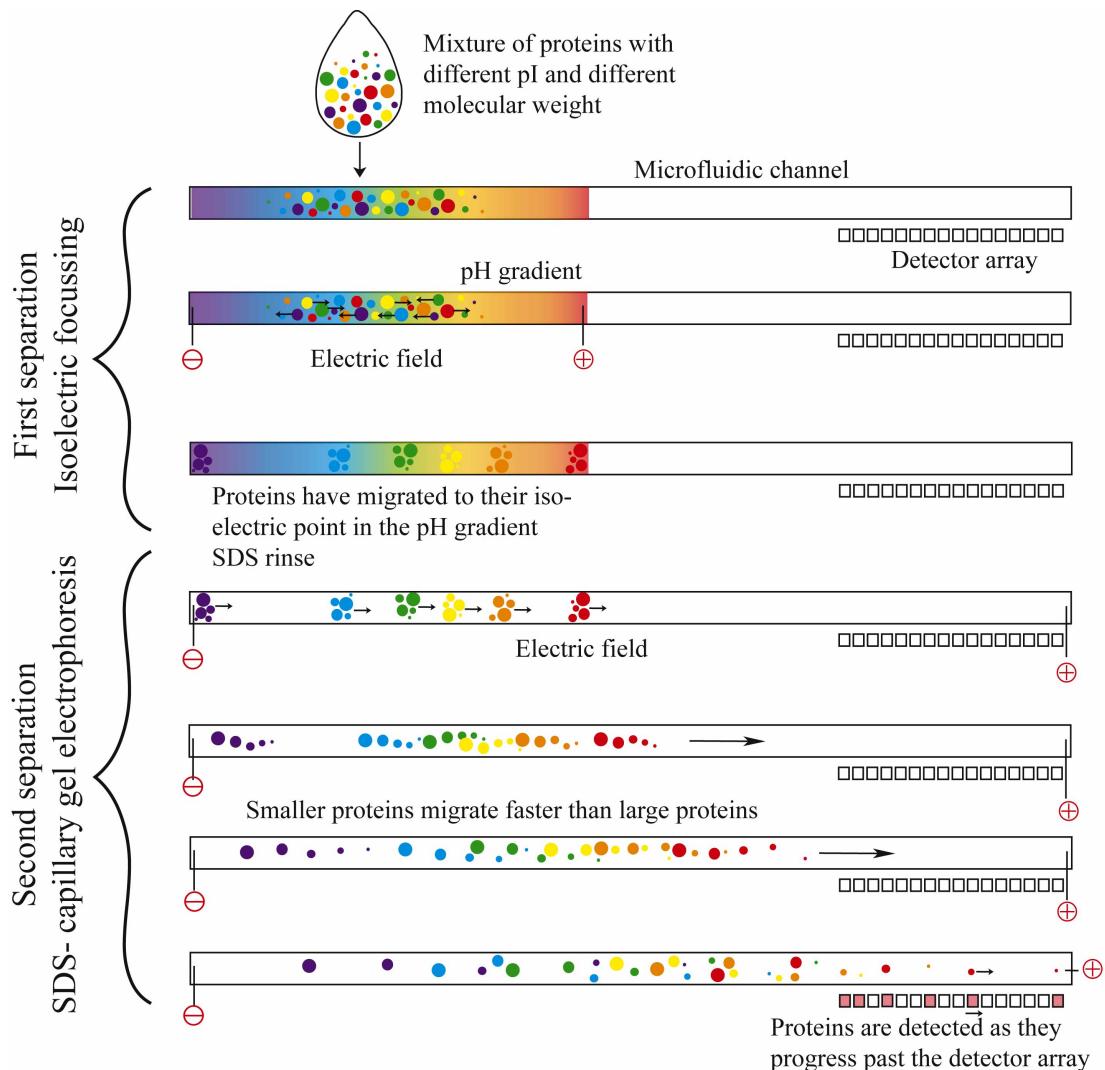


Figure 1.15: Schematic diagram of virtual 2D electrophoresis. A protein mixture is introduced into a channel with a pH gradient. Under an external electric field, the proteins migrate according to their charge. At the position in the pH gradient that corresponds to a protein's pI, the protein has net charge zero, and remains inert. Thus proteins are effectively focused at their pI position (IEF, isoelectric focus). Once this has been completed, an SDS-denaturing environment replaces the pH gradient. Upon application of an external electric field, the proteins begin their second separation by molecular weight. Proteins are detected as they progress along the channel; from these position-dependent measurements, the migration trajectory of a given protein can be reconstructed.

progress. At the end of the separation length, proteins are no longer simply ordered by either size or pI, but in a more complex manner.

As illustrated in Figure 1.15, protein bands are detected as they advance along the channel. Detection is by optical absorption of proteins at ultraviolet wavelength (214 nm). Passage of a band results in photonic intensity differences that are measured by an array of photo diodes that are placed along the channel. From the progression of a band, it is possible to reconstruct the starting point (which determines the isoelectric point), as well as the migration speed (which determines the molecular weight).

In Figure 1.16 the reconstruction of pI and MW values is illustrated. Reconstruction is done by recording the migration trajectory of each protein along the channel. From this, the travelling speed (or mobility) can be calculated, which correlates to the molecular weight. Given the migration trajectory of a protein in the detection window, its starting position can be evaluated, which is its isoelectric point in the pH gradient. From this simple process we can get two dimensions of information, in addition to information regarding quantity from the UV absorption data.

The V2D system will be implemented in a microfluidic chip-based format with integrated reservoirs, liquid handling elements, the separation geometry, and a detection window. Separation progression and data acquisition will all be software controlled. This minimises user intervention and eliminates a number of the problems surrounding 2D gel electrophoresis.

V2D separations have numerous advantages over other separations. Due to the capillary based format, very fast separations are possible. By bringing this into a microfluidic format, even faster separations are possible. The resolution is not limited by sampling rate or structural layout, but mainly by reconstruction accuracy, which can be improved by optimising the track fitting algorithms, and by increasing the number of time-space measurements (e.g. detection along a larger segment of the separation distance). Finally, the Beer-Lambert approach means that quantification is relatively bias-free—there is not the error associated with the use of a label and its perhaps incomplete uptake.

Procedures such as staining, destaining, and scanning, customary in conventional 2D gel electrophoresis, are unnecessary because the information for quantification is already included in the data collected for tracking. Since the separation is monitored online and the results are digitally reconstructed and stored in a numerical format, data analysis can be performed immediately. Graphical representation of a V2D separation in the traditional spot-image format is straightforward.

This combination of speed, resolution, and quantification is what sets V2D apart from other separation techniques. In Figure 1.17 V2D proteomics is set against conventional

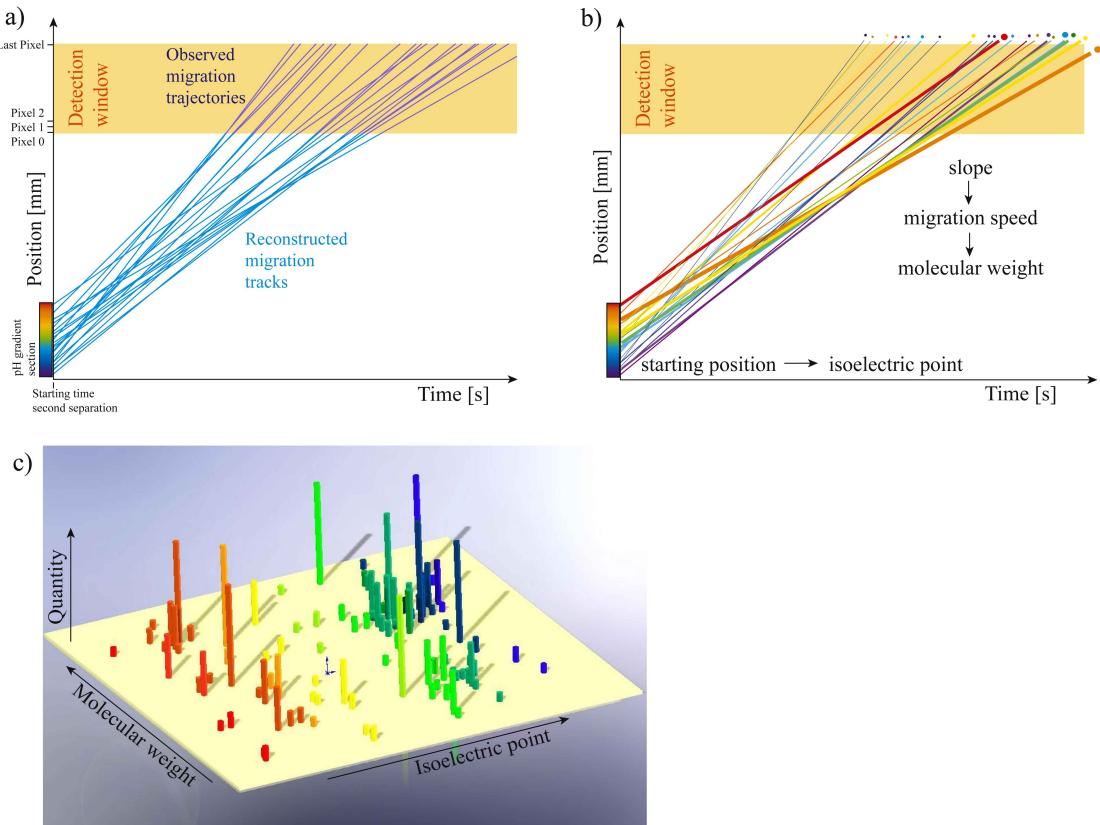


Figure 1.16: Reconstruction of isoelectric point and molecular weight with the migration trajectory of proteins. a) In the detection window, the migration trajectory of proteins is observed as they progress along the channel. From this a migration track is reconstructed. b) The reconstructed migration tracks provide information about the starting position in the pH gradient of any particular protein. This defines the isoelectric point of proteins. Additionally, the slope of the migration track provides the migration speed. From this a protein's mobility and molecular weight can be found. c) Using the reconstructed PI and MW of a track as well as the peak area, the separation data are presented in the same way as a conventional 2D gel.

electrophoresis-driven proteomics, and chromatography-driven proteomics. The processing steps for quantification in conventional gel-base proteomics (gel staining, destaining, imaging and image analysis) can be performed on-line in V2D proteomics, simplifying the workflow.

Embedding the complete separation on a single chip is an encouraging start for fully integrated proteomics. Thanks to the freedom of design of microfluidic chips, downstream functions can be added on, much in a pick-and-mix style. One option is selecting and switching out proteins for further scrutiny. After a protein has passed the detector, its further migration trajectory may be predicted from the observed trajectory. This allows determination of the exact time point a particular band reaches a downstream junction, or outlet. Flow switching technologies will enable selection and acquisition of specific bands for further analysis.^[173] To combat excessive peak broadening, the stream of effluent can be contained in droplets.^[174] Furthermore, the V2D approach lends itself to a

multiplexed approach, allowing many channels to be analysed simultaneously. This will lead to direct comparisons between samples not possible with SDS 2D PAGE systems. Combined, these elements enable downstream analysis after a V2D separation.

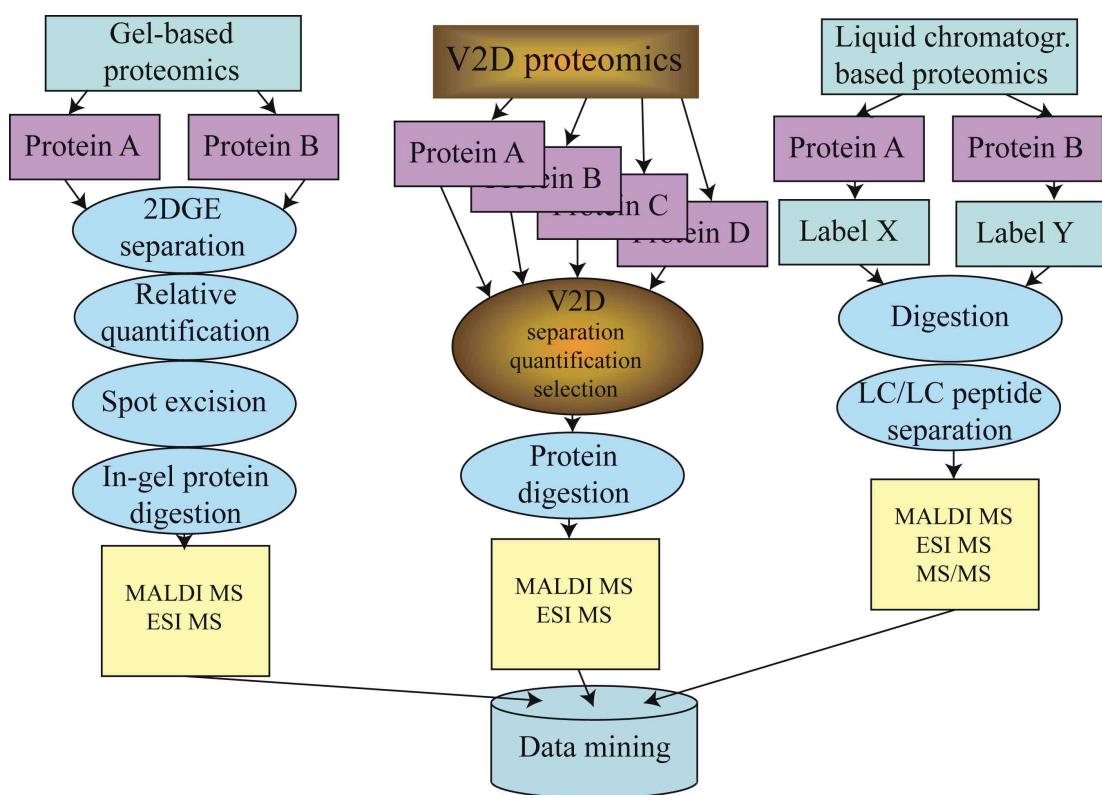


Figure 1.17: Comparison of different proteomics workflows.

Chapter 2

Error Analysis

Un physicien éminent me disait un jour à propos de la loi des erreurs: “Tout le monde y croit fermement parce que les mathématiciens s’imaginent que c’est un fait d’observation, et les observateurs que c’est un théorème de mathématiques”.

Henri Poincaré: Thermodynamique, Paris 1892

In this chapter, an error analysis for virtual two-dimensional separations is presented. A first section introduces the fundamental principles and statistical methods. The second section specifies the experimental methods including the model system, instrumentation, and data analysis algorithms. The third section of this chapter is dedicated to presenting the results of the error analysis. This includes a study of separations under standard conditions, followed by studies on systematic variations of the standard conditions. The last section of this chapter discusses the results. This includes developing and validating different models for error sources, and estimating the magnitude of errors from different sources. These findings are applied to produce mock-V2D data and assess the performance of separation reconstruction.

Virtual two-dimensional separations constitute an unprecedented approach to 2D separations, and as such conceal unknown challenges and weaknesses. Part of the purpose of this thesis is to identify and quantify error sources, circumvent them if possible, otherwise minimise their impact. To elucidate which parameters are critical for performance and give an idea of performance limits, an error analysis is necessary. V2D separations operate in a multivariate parameter space, with unknown interdependencies of parameters. To disentangle the effect of a palette of experimental settings from one another, accessible parameters were systematically varied and separation results were studied and compared. From these measurements, different sources of errors were identified

and characterised, leading to an optimisation strategy for V2D separations as well as a performance estimate.

2.1 Introduction

2.1.1 Errors

When measuring the natural world, it must always be kept in mind that any measurement is flawed and only an approximation. It is helpful to distinguish between systematic and random errors, both of which can contribute to the overall error in a measurement. Random errors create an unpredictable, but generally unbiased scattering of measurements around a central value. Standard procedure here is to average values from repeated measurements, so that the individual fluctuations cancel one another out. The mathematical basis for this is the central limit theorem, which states that the sum of independent variables approaches a Gaussian distribution with an increasing number of variables. This means that random errors, which may be considered the sum of many minute effects, follow a Gaussian distribution around the true value. For a Gaussian distribution, the mean of repeat measurements provides a good estimator for the true value, hence averaging can minimise the impact of random errors.^[175] In Figure 2.1 a schematic representation is shown of 20 measurements of a value with random errors. The random measurements follow a Gaussian distribution around the true value, for which the mean gives a good estimator.

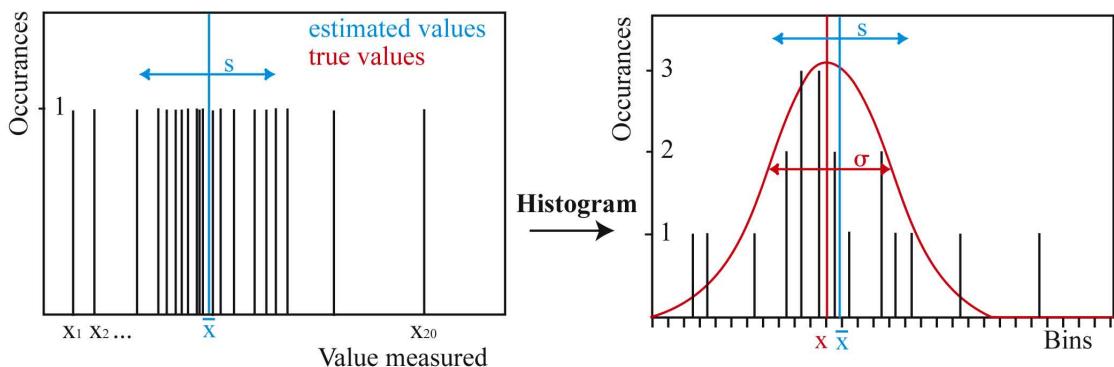


Figure 2.1: Schematic of 20 measurements x_1, x_2, \dots, x_{20} with random errors. In blue the estimator \bar{x} for the true (unknown) value of x , and the estimator s for the standard deviation σ . In the schematic on the right, the measured values are binned into a histogram, illustrating that they follow a Gaussian distribution around the true value x with true standard deviation σ (red).

The mean, or average, \bar{x} of n measurements is:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.1)$$

with x_i being the i-th measurement. The term ‘error’ refers to the Gaussian standard deviation σ , which is the square root of the variance V :

$$V(x) = \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.2)$$

For repeat measurements with random error, the unbiased estimate of the variance, \hat{V} , is:

$$\hat{V}(x) = s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.3)$$

The variance of this estimate is:

$$V(\hat{V}(x)) = \sigma_s^2 \approx \frac{s^2}{2(n-1)} \quad (2.4)$$

for Gaussian distribution and large n .^[175]

Systematic errors behave differently from random errors. Systematic errors arise when there are discrepancies between model and experiment, and not all factors are accounted for. Repeat measurements will not help avoid systematic errors, but they will allow the experimenter to recognise systematic effects more clearly. In an error analysis, systematic errors are indicative of significant influences on the experiment that need consideration. Sources of systematic errors can also be related and so can generate highly correlated errors.

2.1.2 Estimators

The aim of quantitative science is to describe a set of measurements with a mathematical model. The model is a set of equations with parameters of unknown value. These parameters need to be estimated so that the data and the model fit with minimal overall error. The method of least squares is a generally useful—certainly the most widely used and popular—approach to the task of linear regression. Alternative regressions include least-absolute-deviations regression, M-regression, and rank-based estimates.^[176] As the name suggests, the sum of the squared difference between data and model is minimised while varying the model parameters, giving an optimal estimate for those parameters. Although different curves can be fitted with this approach, a straight line fit is relevant to the studies herein, since the space-time trajectory of the observed

protein migration is expected to be linear.^[40] A straight line that is fitted to a given set of data requires determination of two parameters: intercept **a** and slope **b**. This is a well-characterised problem, and an extensive analysis of linear fit approaches for different statistical categories of data is available.^[177] The most common model is that where of the n paired values x_i and y_i , one is assumed to be without error and the other has constant variance, as illustrated in Figure 2.2.

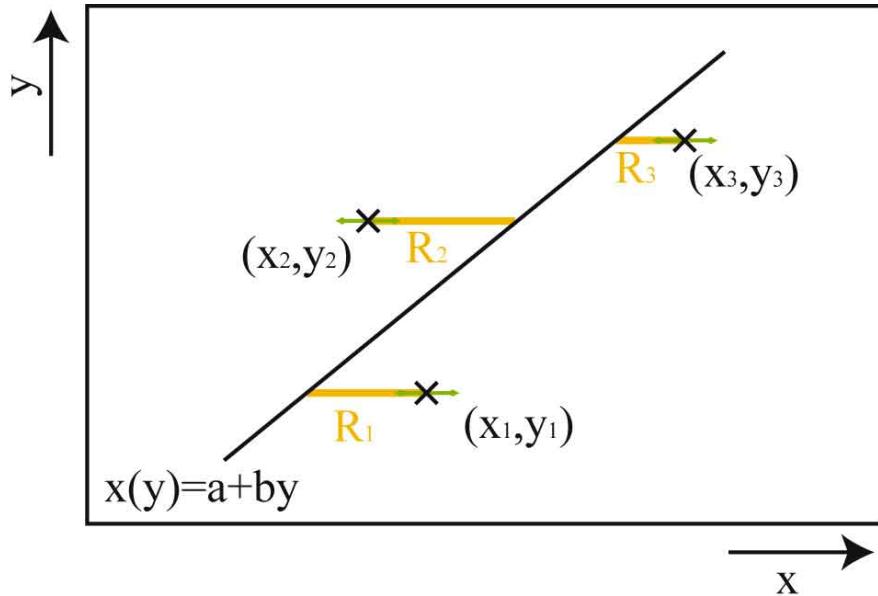


Figure 2.2: Schematic of a least squares estimate for a straight line fit to three measurement values. The measured value pairs (x_n, y_n) have x-values with a constant error (indicated in green) and precisely known y-values. The sum of the horizontal distances R_n (orange) is minimised to obtain estimates for the intercept **a** and slope **b**.

To fit a straight line $x(y) = \mathbf{a} + \mathbf{b}y$ to n measurement points $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ the sum of the horizontal distances is minimised. The sum R^2 of horizontal squares is:

$$R^2(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^n (x_i - (\mathbf{a} + \mathbf{b}y_i))^2 \quad (2.5)$$

R^2 is minimised with

$$\frac{\partial R^2}{\partial \mathbf{a}} = 0 \quad \text{and} \quad \frac{\partial R^2}{\partial \mathbf{b}} = 0 \quad (2.6)$$

By solving Equations 2.5 and 2.6 for **a** and **b**, the following estimates for the fit parameters are obtained:

$$\mathbf{b} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2} \quad (2.7)$$

$$\mathbf{a} = \bar{x} - b\bar{y} \quad (2.8)$$

The estimates for the errors associated with the estimates for \mathbf{a} and \mathbf{b} are:^[175]

$$\sigma_{\mathbf{a}}^2 = s^2 \left(\frac{1}{n} + \frac{\bar{y}^2}{\bar{y}^2 - n\bar{y}^2} \right) \quad (2.9)$$

$$\sigma_{\mathbf{b}}^2 = \frac{s^2}{\bar{y}^2 - n\bar{y}^2} \quad (2.10)$$

$$\text{with } s^2 = \frac{(\bar{x}^2 - n\bar{x}^2) - \mathbf{b}(\bar{y}^2 - n\bar{y}^2)}{n - 2}$$

The least squares fit model is based on two assumption regarding the paired measurement values x_i and y_i . First, it is implicitly assumed that y_i values are known without error, and second, it is assumed that the variance of the x_i values is constant. If this is not the case, a weighted (by variance) least squares fit is more appropriate. Use of a least squares fit requires validation of these assumptions.

For the purposes of the V2D error analysis, the paired measurement values are in the form of space/time coordinates of protein detection, as will be discussed in more depth in Section 2.2. The space component of these coordinates, the pixel number, is known with negligible error, since pixel spacing is precise, accurate, and does not change (pixels do however have a finite imaging width, which must be taken into account, as well as a systematic error insofar as the detector might be positioned at an offset).

The other assumption, constant variance of time coordinate measurement, requires closer consideration. As peaks move from one pixel to the next, they get progressively flatter and wider, due to molecular diffusion and other effects.^[40] For flatter, wider peaks, the coordinates of the minima vary within a greater range. This implies that the variance of the time coordinates is not constant, but increases. To ensure that a least squares fit is still applicable, the increase of the time coordinate variance must be negligible within the observation space. This is confirmed in Figure 2.3, where the observation coordinates (relative to a fitted line) are plotted for five protein peaks migrating past the detector. Here it is seen that the time coordinates fall within a band of constant width (indicated with green dashed lines in Figure 2.3). We conclude that any increase in the time-dimension variance along the detection window is indeed negligible. Hence the proposed least squares fitting model is applied, despite aforementioned caveat regarding this fit's implicit assumptions.

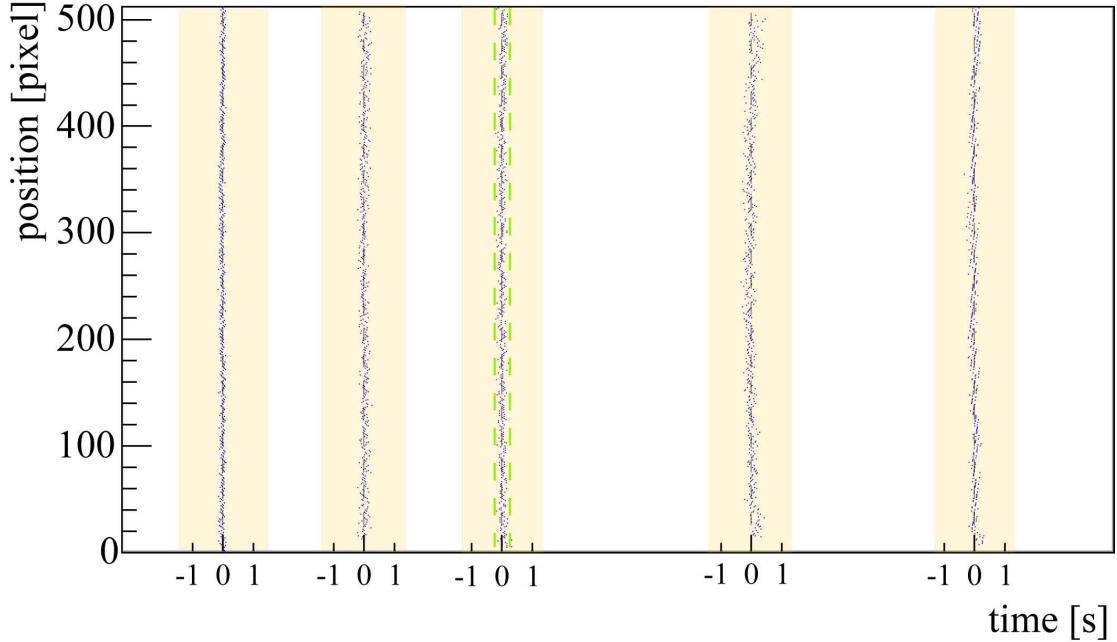


Figure 2.3: Coordinates of measurement pairs for five protein bands. To exaggerate the distribution of coordinates, not actual detection coordinates are plotted, but the time values (x-axis) are plotted relative to a straight line (residual plot, illustrated in Figure 2.13a). The distribution does not grow wider with increasing pixel number, as illustrated with dashed green lines for the third band. This supports the assumption that increases in variance are negligible.

To ensure that the linear least squares fit describes the measurement data well, goodness-of-fit estimates are consulted. The correlation coefficient ρ gives the proportion of data that is accounted for in the fit. For the tracks plotted in Figure 2.3 above, values for ρ are in the range of 0.9991 to 0.9999, indicating an excellent quality of fit.

Another quantity to describe the goodness of a fit is the χ^2 ('chi square') value. This is the squared difference between observed values and the corresponding value of the fitted line, weighted by the errors of the measurements. Given n paired measured values x_i and y_i , the fitted function $f(x)$, and a measurement error of σ_i , then χ^2 of the fit is:

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - f(x_i))^2}{\sigma_i^2} \quad (2.11)$$

For large n , the probability distribution of χ^2 has a mean that is equal to the number of degrees of freedom. The degree of freedom is defined as the number of measurements (n) minus the number of variables that have been optimised for the fit (for a linear fit: two variables, intercept **a** and slope **b**). Applied to a track with 512 data points, the degrees of freedom are 510. In this case, the most likely value for χ^2 is 510. If χ^2 is of similar magnitude as the degree of freedom, the fit function agrees well with the measured values.

For a given degree of freedom, the probability of observing a certain χ^2 value follows a defined distribution, the χ^2 distribution. For 510 degrees of freedom the χ^2 distribution is shown in Figure 2.4. This can be used to assign a confidence level to the goodness of a fit.

The confidence level for a set χ^2 value gives the probability that an observed χ^2 value exceeds the set χ^2 value by chance, even for a correct model. For instance, for 510 degrees of freedom, there is a 5 % probability that if the fit is correct, then a χ^2 value greater than 563 is observed; there is a 0.01 % probability that if the fit is correct, then a χ^2 value greater than 683 is observed.

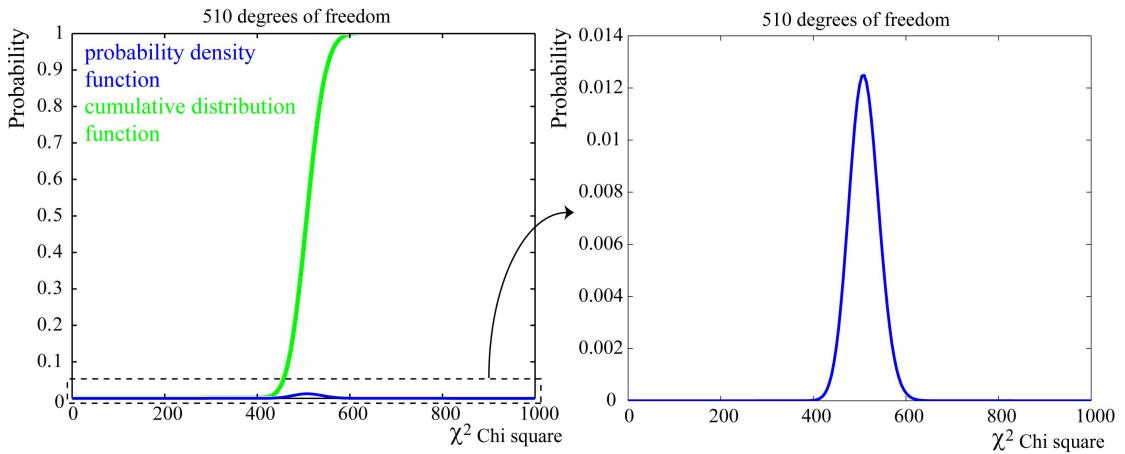


Figure 2.4: χ^2 distribution for 510 degrees of freedom. Blue probability density function, and green cumulative distribution function.

2.1.3 Combination of errors

The outcome from fitting lines to sets of data is an estimate for values of slope and intercept of these lines. Since it is not necessarily the values of slope and intercept that are of interest, further mathematical operations are performed. Depending on what operations are performed, the errors on the estimated values can combine and propagate according to the law of combination of errors.^[175] For the case of a single function $f(l, m)$ with two independent variables l and m , the law of combination of errors is:

$$\sigma_f^2 = \left(\frac{\partial f(l, m)}{\partial l} \right)^2 \sigma_l^2 + \left(\frac{\partial f(l, m)}{\partial m} \right)^2 \sigma_m^2 \quad (2.12)$$

Figure 2.5 illustrates how the combination of errors pertains to the problem of fitting a migration trajectory through a set of detection data. From the least squares fit estimators are obtained for the slope **b** (Equation 2.7) and intercept **a** (Equation 2.8) of

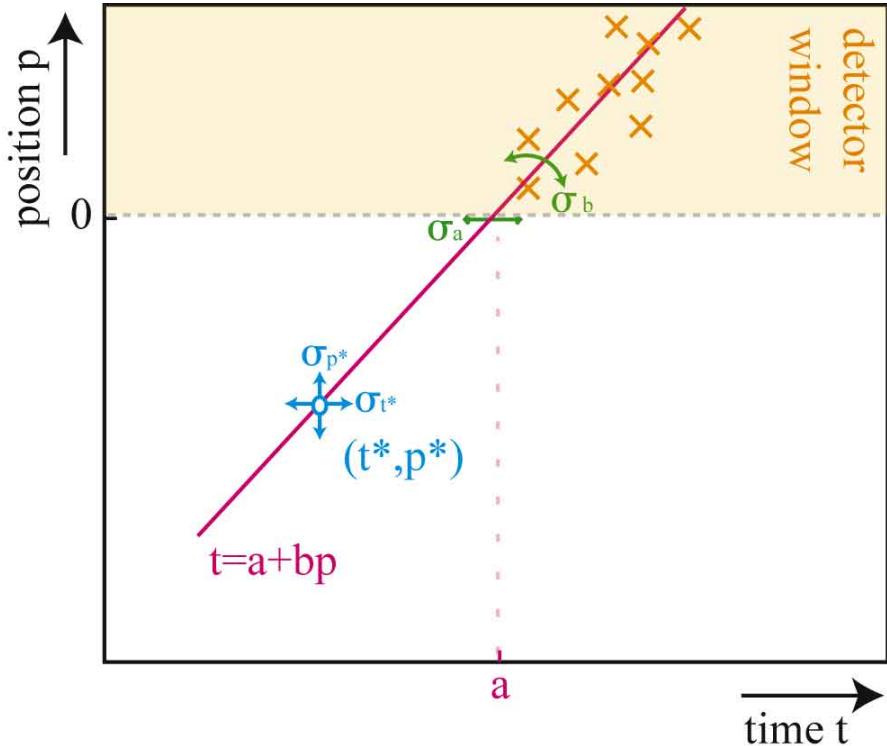


Figure 2.5: Schematic of protein detection coordinates (orange), fitted straight line (pink), and different errors σ . The fitted parameters for the straight line (intercept \mathbf{a} and slope \mathbf{b}) have associated errors $\sigma_{\mathbf{a}}$ and $\sigma_{\mathbf{b}}$ (green). The errors of the fitted line at position (t^*, p^*) are σ_{t^*} and σ_{p^*} (blue).

the fitted lines, along with estimators for the errors associated with the slope $\sigma_{\mathbf{b}}$ (Equation 2.10) and intercept $\sigma_{\mathbf{a}}$ (Equation 2.9). With these values, Equation 2.12 allows propagation of the error on the fitted line to any point of interest. In the case of protein migration in the space-time coordinate system, this means that the errors associated with an arbitrary point on the line (t^*, p^*) can be found, see Figure 2.5.

The linear equations describing the migration of proteins in the time-position coordinate system are:

$$t = \mathbf{a} + \mathbf{b}p \quad (2.13)$$

$$p = \frac{t - \mathbf{a}}{\mathbf{b}} \quad (2.14)$$

where \mathbf{a} is the intercept and \mathbf{b} is the slope estimated with a least squares fit to detector data, and t and p are time and position coordinates, respectively. Note that in this coordinate system \mathbf{a} [s] is the intercept with the 0-pixel axis (i.e. the time a band enters the detection window) and \mathbf{b} [s/mm] is the inverse of the migration speed. From the least squares fit, we also obtain estimates for the errors of intercept and slope, $\sigma_{\mathbf{a}}$ [s] and $\sigma_{\mathbf{b}}$ [s/mm]. Starting with the errors on intercept and slope, the law of error propagation

is applied to yield the error on time and position:

$$\sigma_t^2 = \sigma_a^2 + p\sigma_b^2 \quad (2.15)$$

$$\sigma_p^2 = \frac{\sigma_a^2}{b^2} + \left(\frac{a - t}{b^2} \right)^2 \sigma_b^2 \quad (2.16)$$

The error on position σ_p^2 [mm] is particularly useful, since the position is the parameter from which the isoelectric point is determined in V2D data.

2.1.4 Error sources in capillary gel electrophoresis

For conventional capillary gel electrophoresis (CGE), a body of work examining the impact of various factors on resolution is available.^[40,178–181] In conventional CGE, error analysis is somewhat simplified compared to multiple pixel detection systems. A single detector records changes in a given signal (e.g. intensity of transmitted light), and from this electropherogram peaks are distinguished from the background and assigned a migration time. Important factors influencing detection are peak width and peak height, i.e. theoretical plate number and signal-to-noise value.

Peak width is affected by sample injection, molecular diffusion, geometric broadening (for microfluidic chips with turns), band dispersion from non-uniform flow, temperature (Joule heating), surface interactions, gravity, as well as detector width and data acquisition frequency. Peak height is affected by detection mode (i.e. fluorescent label detection or intrinsic absorption detection, as introduced in Section 1.3) and detector sensitivity, detection length, as well as buffer and sample properties.

The errors from aforementioned sources can be reduced by choice of buffer formulation (transparency, conductivity, surface control and electro-osmotic flow); by experimental settings (voltage, separation length,^[182] or injection mode;^[183]) and by hardware setup (detector specifications,^[184,185] capillary geometry,^[186] or temperature control.^[187]) Furthermore, advanced data analysis has been proposed to achieve better data quality. This includes deconvolution of overlapping peaks with calibration data and partial least-squares regression,^[188] convolution of detection data with a Shah function or a sine function and then filtering with Fourier transform,^[189–193] and signal enhancement by pseudorandom multiple sample injection and data decoding with a Hadamard transform.^[194]

In multipixel detection, the same error sources affect peak detection on a single pixel level. By combining data from multiple pixels, random errors (noise) of detection can be reduced and the signal-to-noise ratio improved. Taking advantage of the fact that

bands migrate in a linear fashion (as defined in Section 1.3, Equation 1.17), additional constraints may be applied to band detection.

In single pixel detection, the average migration speed is calculated with the migration time and the migration distance to the detector. On the contrary, multipixel detection allows direct measurement of the actual migration speed at the detector. From studying the migration trajectory across the detection window, non-linear migration behaviour can be studied. This allows new insights into fundamental aspects of protein separations.

For an error analysis of V2D separations, the error sources may be categorised as random errors, detection errors, and non-linear migration errors. While random and detection errors are expected to behave in a manner similar to conventional CE, non-linear migration presents a new and uncharacterised error source.

2.2 Experimental methods

2.2.1 Multiple injections as a model system for V2D error analysis

In a V2D separation proteins first undergo isoelectric focusing, and thereafter separation by CGE. At the end of the CGE separation proteins migrate past a photodiode array, allowing direct observation of migration. From the migration trajectory, the migration speed and the starting position can be reconstructed. These parameters are directly related to molecular weight (migration speed), and isoelectric point (starting position). The correlation of the molecular weight MW and electrophoretic mobility μ is described by:

$$\log(\text{MW}) = p + q \left(\frac{1}{\mu} \right) \quad (2.17)$$

where p and q are unknown parameters that are determined on the basis of a calibration run.^[195]

For the purpose of evaluating the errors on migration speed and starting position, conventional CGE separations present a suitable model system. Studying CGE reduces the complexity of the error analysis and enables use of very simple equipment (without the need for a pH gradient, capillaries instead of microfluidic chips suffice). One aspect must be carefully considered, namely that in CGE separations protein migration trajectories do not intersect, since they originate from the same point of injection. In V2D separations individual protein bands can overtake one another. It is unknown whether these band crossings affect protein migration.

To ensure effects from band crossings are not neglected, separations with multiple injections are studied. In multiple injections, sample is injected more than once, from the same starting point. Following an initial injection of sample, the separation progresses the same as standard runs. However, after a set time, a second injection of sample is introduced into the same capillary, with minimal interruption of the already ongoing separation of the first injection. Depending on the time interval between the injections, slow proteins from the first injection can be overtaken by fast proteins from the second injection, as illustrated in Figure 2.6. The outcome is the same if different proteins start at different positions (as with V2D), or at different times (as with multiple injections). Therefore separations with multiple injections of sample are a valid model for error analysis of V2D separations, with the advantage of experimental simplicity.

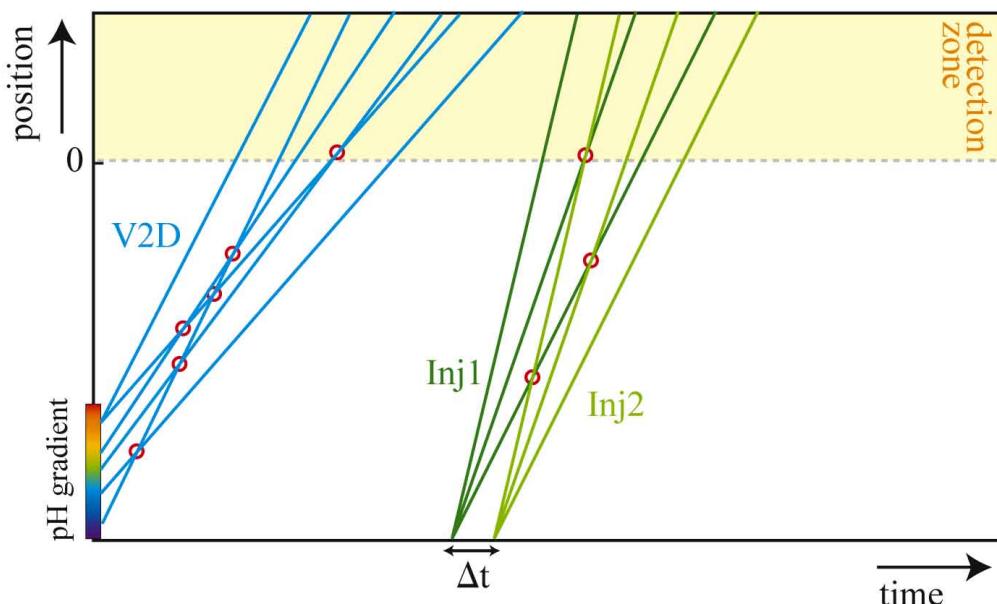


Figure 2.6: Schematic of protein bands crossing (marked with red circles). In blue on the left are the migration trajectories of six bands in a V2D separation. In green on the right three protein bands from a first injection (Inj1) intersect with bands from a second injection of the same sample (Inj2), injected after an interval of Δt .

2.2.2 Instrumentation

A custom-built capillary electrophoresis system was used for standard measurements. A translation/tilt adjustable stage for the capillary was fixed on an optical rail. The capillary ends were immersed in vials with buffer and platinum electrodes (Alpha Aesar, 99.9 % Pt) to a high voltage power supply. The optical train can be seen in Figure 2.7; in essence, a light source (Heraeus-Nobellight DS225/05 J deuterium lamp) was filtered (bandpass filter at 214 nm, 22 nm bandwidth) and focused onto the capillary in the stage.

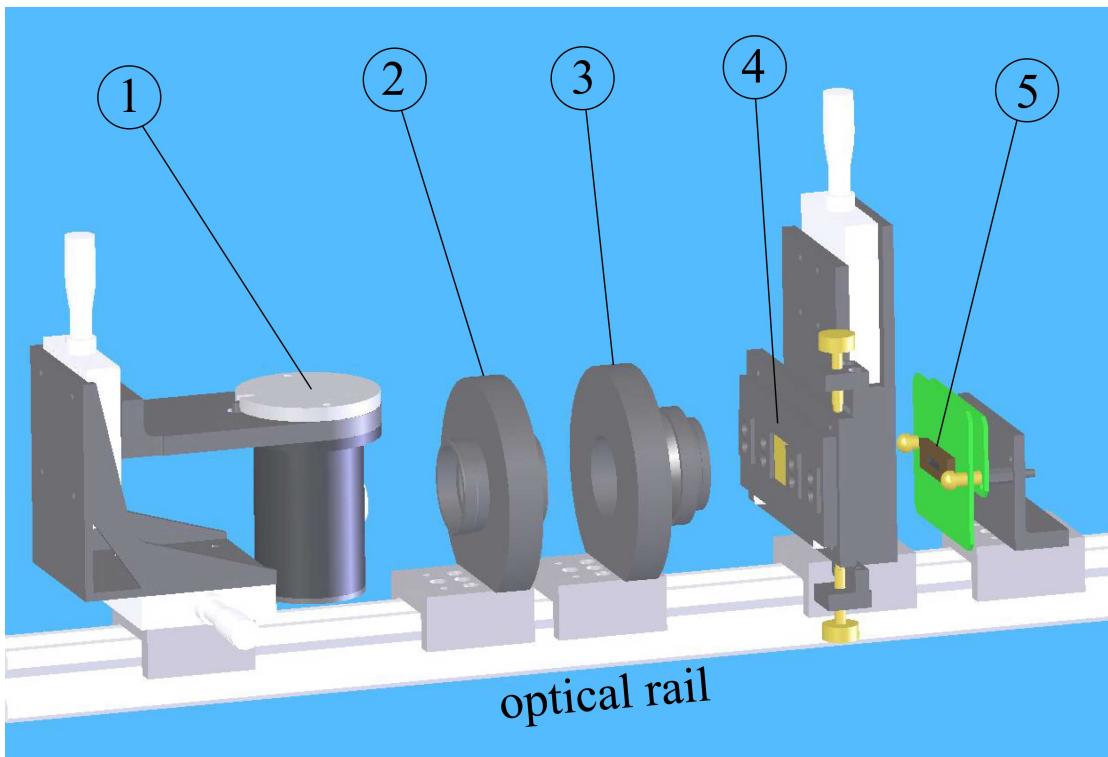


Figure 2.7: Optical train of capillary electrophoresis instrument. 1) Deuterium lamp in lamp holder. 2) Collimating lens and filter. 3) Cylindrical lens. 4) Capillary stage. 5) Photodiode array detector.

A photodiode array (PDA) was positioned immediately behind the capillary, without imaging optics. This was chosen to ensure that optical calibration (and possibly distortion) does not affect the detected migration trajectory.

Optical alignment was performed for each new capillary to reduce excessive background (i.e. from light passing beside the capillary lumen). The far UV light that is focused onto the capillary lumen is absorbed by passing proteins. UV light at 214 nm was employed, as discussed in Chapter 1.3.2. The capillary was projected onto a PDA to record the UV light intensity. The light ray diagram for illumination and detection is shown in Figure 2.8.

The PDA was a self-scanning linear array of 2048 pixels, each 2.5 mm by 25 μm (active area 20 μm), from Hamamatsu (S3904-2048Q). Photodiodes were N-channel MOS transistors, current output type, with low dark current (0.1 pA) and high saturation charge (25 pC) for long integration times and a wide dynamic range. The PDA was mounted on a driver board/pulse generator, also from Hamamatsu (C7884G-01), that supplied start and clock pulses and charge-to-voltage conversion. The PDA frequency and integration time can be adjusted (set to 62.5 kHz, 100 ms).

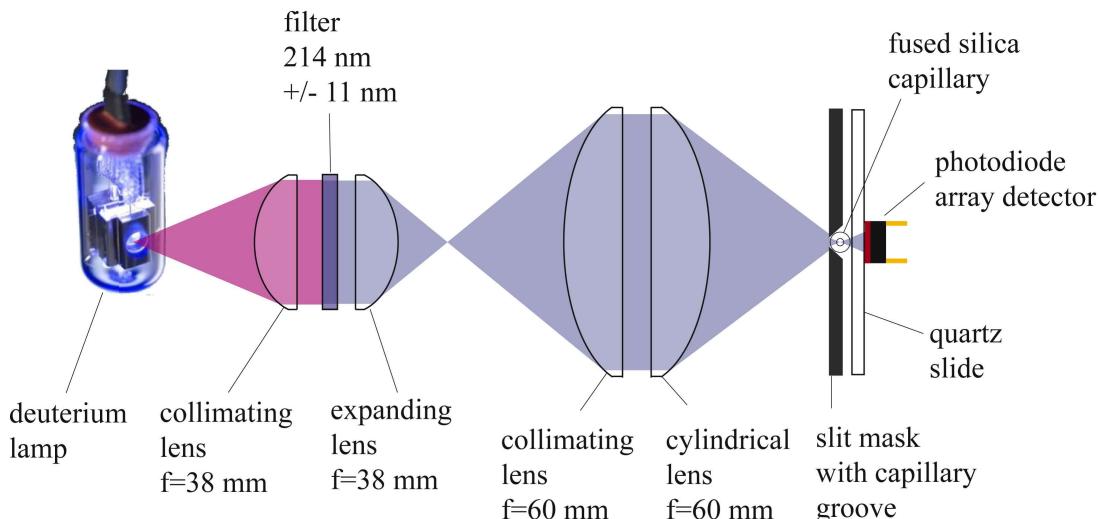


Figure 2.8: Light ray diagram for illumination and detection.

Data acquisition readout was set to 10 Hz. The analog PDA signal was read out to a USB data acquisition function module (Data Translation DT9800) which also recorded trigger and clock pulses from the PDA board. The data acquisition module performed 16 bit analogue-to-digital conversion and was linked to a computer via a USB interface, from which C++ code was used to control data acquisition and storage. For the 2048 pixel PDA, data from four adjacent pixels was averaged to curb the data volume. The overall effect of this averaging is the same as using a 512 pixel PDA with 100 μm pixel width.

One consideration with regards to error analysis was that the pixels need to be small enough, and data acquisition fast enough, to resolve the expected signal. The instrumentation used employed 0.1 mm pixels at 0.1 seconds per scan, and produced data volumes in the range of 40 Mbytes per run. In single-detector CE equipment, detector dimensions are in the range of 1–5 mm (which might differ from detected dimensions), and readout times of 0.5–2.5 seconds are usual.^[40] The typical extent of a protein band in a capillary was in the range of 4–20 mm, and under routine migration conditions a band took from 4–120 seconds to pass the detector. Compared to even very quick and narrow protein bands, the detector sampling in space and time dimension was at least an order of magnitude smaller.

2.2.3 Data analysis: hit-finding and track-fitting

A set of custom offline analysis packages (ENIUS, deltaDOT, developed by Phil Lewis) were used and adapted to the requirements. Code was written in C++,^[196] extended by the ROOT framework (version 5.12/00).^[197] To outline the data analysis process as illustrated in the flow chart in Figure 2.9, the program begins with loading detector data from a file, then identifies hits where something has passed the detector, and finally fits tracks through the found hits. Once this is done, various operations can be performed, for instance plotting the residuals for tracks.

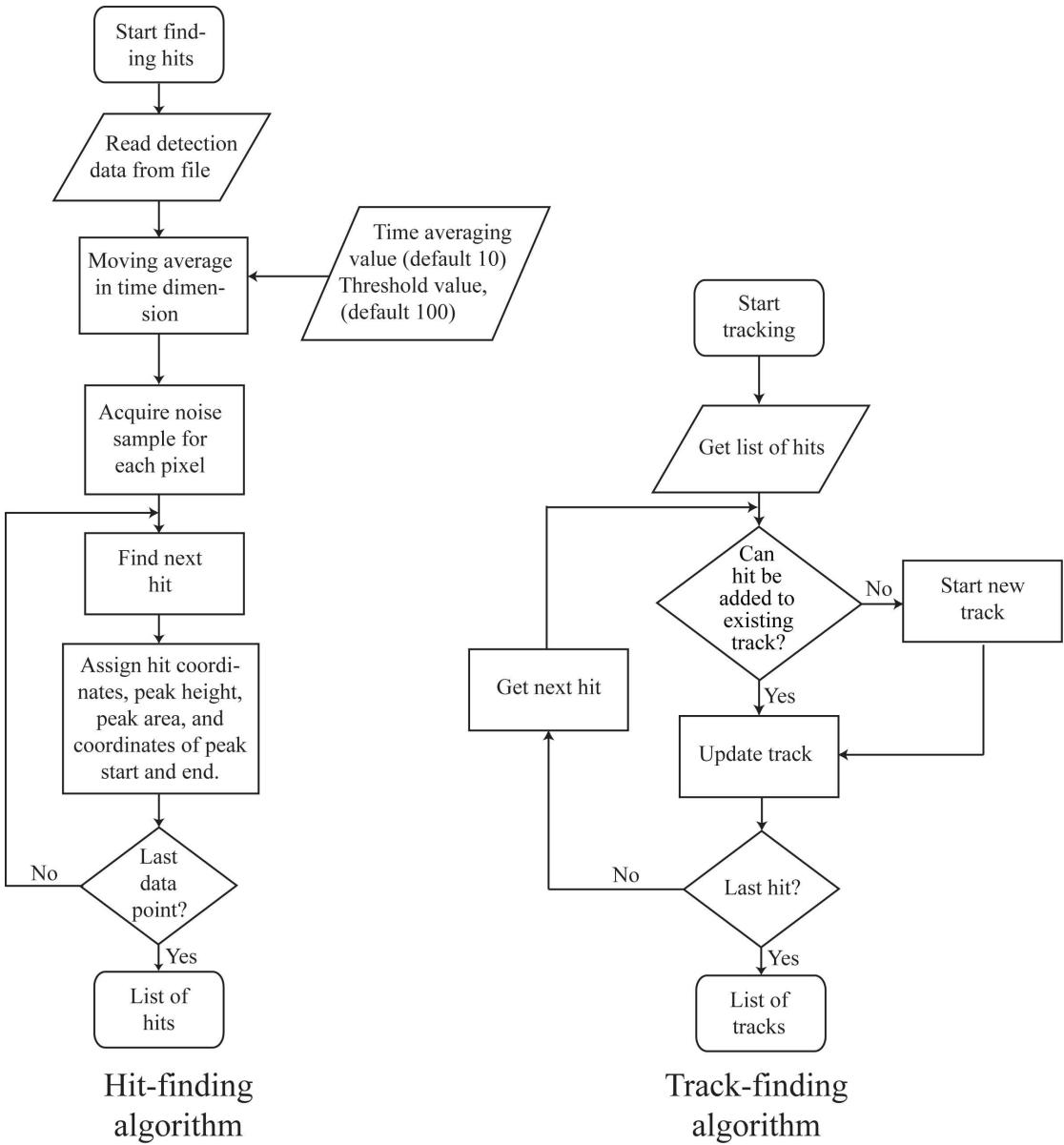


Figure 2.9: Flow chart for hit-finding (left) and subsequent track-finding (right) algorithms.

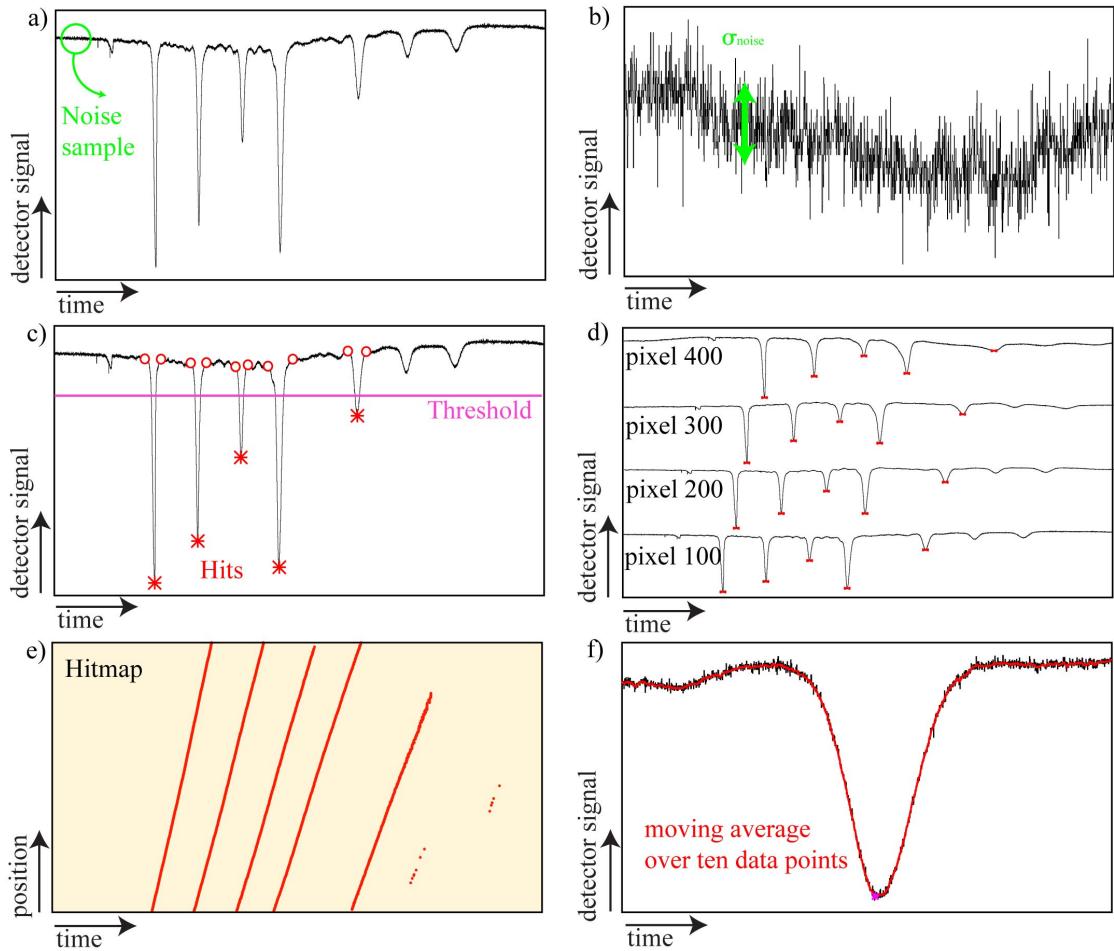


Figure 2.10: Illustration of hit-finding algorithm on measurement data (a-e), and time averaging (f). a) Electropherogram of a pixel. b) Noise sample at beginning of electropherogram. c) The noise value is multiplied by a threshold value to give a hit threshold (pink line). Hits are minima below the threshold (red asterisk symbols). For each hit the nearest maxima are also determined (red circle symbols). d) The hits of all pixels are determined, resulting in e) a heatmap. f) Time averaging: moving average of 10 time values along an electropherogram (raw data black, time averaged red).

In Figure 2.10a-e the hit-finding algorithm is illustrated. Prior to finding hits, a five second noise sample is collected for each pixel (Figure 2.10a). The photocurrent data values for each pixel are normalised to this initial intensity. This is necessary because the absolute photocurrent level varies from pixel to pixel, as does the noise level. Next, one pixel after the other is processed for hits (Figure 2.10c). A hit is a minimum in photocurrent (where a protein band absorbs some of the UV illumination), that exceeds the noise value multiplied by a threshold value. The default threshold value is set to 100 (representing $100 \times$ the noise floor of the data), but can be modified with a graphic user interface. A threshold value of 100 is very high compared to the limit of detection threshold slightly above 3.^[198] This can be useful to distinguish true proteins bands from other signals, as demonstrated in Figure 2.12. Hit-finding can be assisted by applying a moving average filter, again adjustable with a graphic user interface (with the default

setting, 10 adjacent scans are averaged) (Figure 2.10h). For each hit, along with pixel and scan and signal value, the preceding and following maxima are recorded (giving front and back widths of the peak in the time-dimension), as well as the area of the peak.

Once all the hits have been determined, hits are assigned to tracks. One by one, hits are tested as to whether they can be added to an existing track, and if not, they are used to start a new track. Hits are added to a track if they lie within an acceptance window, as illustrated in Figure 2.11. For the first few hits of a track, the acceptance window is related to the width of the peak of the previous hit. Once there are enough hits to provisionally fit a track, the acceptance window is centred on the provisional track. Appropriate settings for hit testing and acceptance widths are crucial and can quite easily cause failure if not well chosen. These settings were adjusted for a standard run and not modified thereafter. Once all hits have been tested, tracks with too few

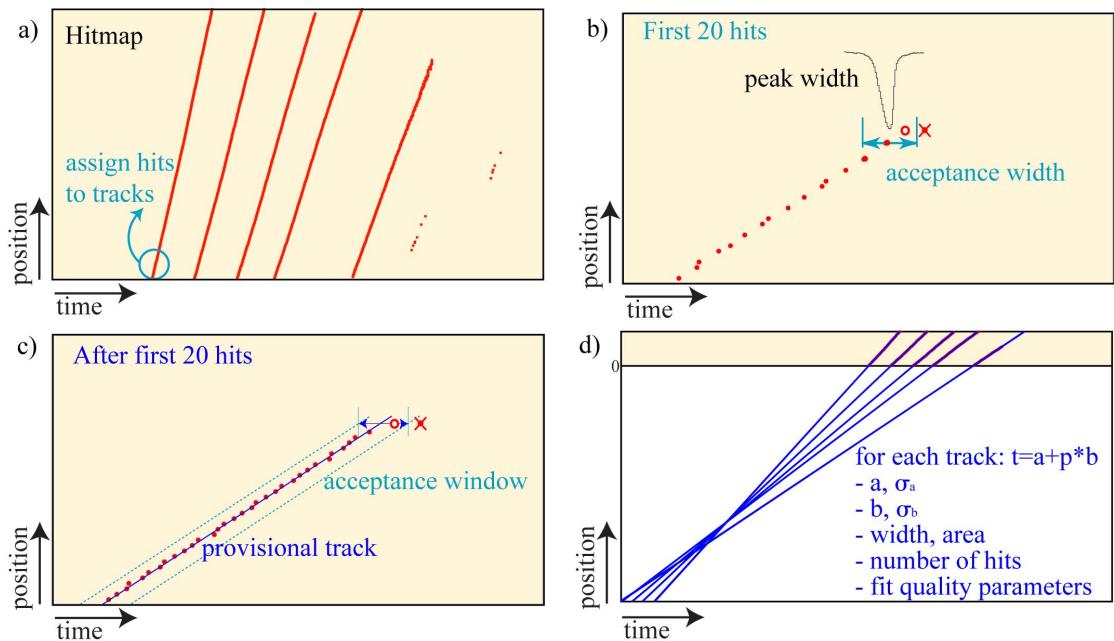


Figure 2.11: Illustration of track-finding algorithm. a) Starting with all hits found in a hitmap, hits are analysed in ascending time. A hit is either added to an existing track, or it is used to start a new track. b) If a track has fewer than 20 hits, then the acceptance criterion is related to the width of the last hit added to that track. Blue arrows indicate the acceptance window for the illustrated peak; the orange circled hit lies within the acceptance window, and is added to the track. The crossed out hit is outside the acceptance window and is used to start a new track. c) Once a track has more than 20 hits, these hits are used to lay a provisional track (with least squares fitting). The projected time coordinate of a pixel on the provisional track is given an acceptance window (blue arrows) for testing hits. The acceptance width is related to the spread of hits around the provisional track, but is capped at a set value. d) Once all hits have been assigned to a track, the tracks with only few hits (< 150) are deleted. Fitted tracks (blue lines) are plotted and track parameters are available for further analysis.

hits are discarded to avoid inclusion of unmeaningful tracks. After every addition of a hit to a track, the fit values are updated. This includes the estimators for slope and intercept, along with error estimates, as well as the sum of squared deviations (for all hits belonging to a track, in time dimension). This last value gives an estimate of track quality and can be used to discard poor quality tracks.

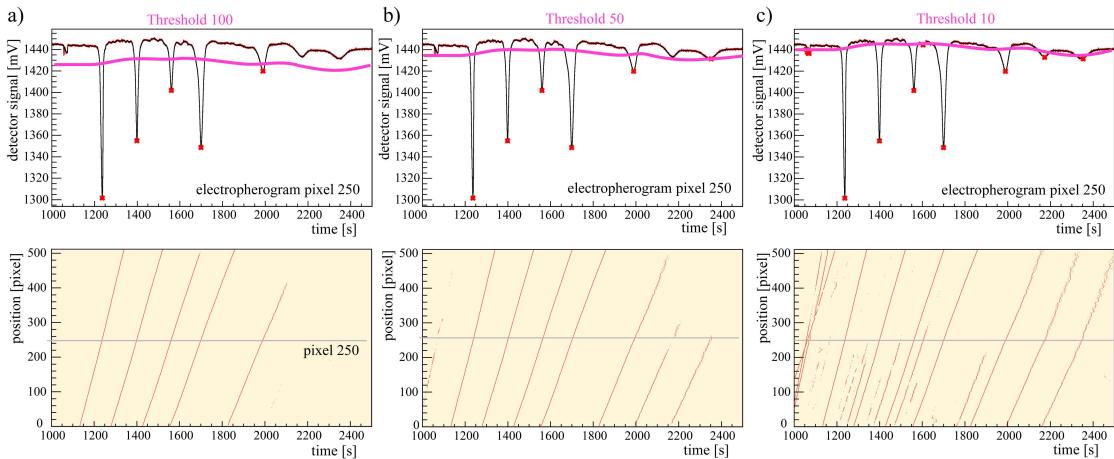


Figure 2.12: Electropherogram and heatmaps for different threshold values. a) Default threshold 100. Note only five of the expected seven bands are found at this threshold. Using the graphic user interface the threshold could be reduced until the missing two bands are found. b) At a threshold value of 50 all seven bands are found. c) Threshold value is 10, which is near the limit of detection. Here excessive secondary bands are found. The data is from a standard separation (200 mm separation length, 350 V/cm electric field, 75 μ m inner diameter capillary, Beckman SDS gel buffer), with the standard sample (7 proteins, 10-225 kDa, 3.2 mg/mL).

Once track data have been determined, a variety of further analyses can be performed, as illustrated in Figure 2.13. Tracks can be plotted and a vertex verified. The residuals (distance hit-track in time dimension) can be plotted to assess migration trajectory shape. The peak half-width (i.e. toward front and back) of any given peak can be used to describe and define peak shape and asymmetry.

Given the start time of migration, the track error estimators can be propagated back to the beginning of migration, to give a start position estimated error. Finally, since the actual starting coordinates are known, the distance between real and reconstructed starting position can be used for comparison. These two error values will be distinguished as ‘track error’ and ‘starting position error’. Note in a V2D separation, only the track error is available, but the starting position error value is necessary for correction.

In plots of migration trajectories, the x-axis is the time axis (in units of seconds) and the y-axis is the position axis (in units of millimetres). For track fit plots, **a** is the intercept with the 0-pixel axis (i.e. the time a band entered the detection window) and **b** is the inverse migration speed. For error plots, the units are in millimetres and seconds, and position error values are propagated to the start time (illustrated in Figure 2.13c).

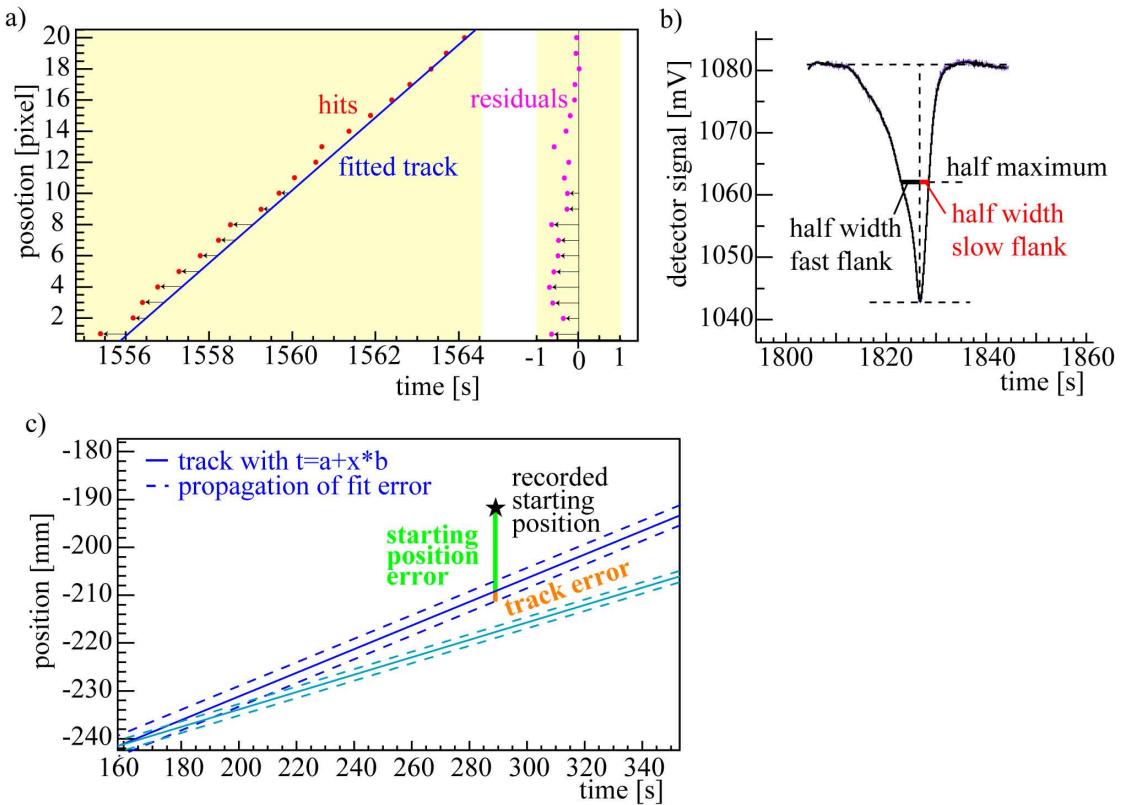


Figure 2.13: Further analysis to evaluate track data. a) Residual plot: distance in time dimension between hits and the fitted track. Only the first 20 pixels of the 512 in total are shown. b) Half widths at half maximum of a hit: toward front (slow flank) and back (fast flank). c) Track error and starting position error: propagated fitting error at starting time, and distance between recorded starting position and track.

The precise (to 0.1 seconds) start time is read out of the PDA data, which show a characteristic drop when the external high voltage is switched on.

2.2.4 Experimental parameters

A set of standard experimental conditions is defined along with a sample preparation protocol. The same preparation protocol, buffer, and equipment are used unless otherwise indicated. A selection of experimental and analytical parameters that are expected to have an impact on errors is chosen and systematically varied. Many of these parameters affect multiple error sources which compete or are compounded, giving rise to complex error behaviour patterns. These data have been studied to understand how proteins migrate and, more specifically, how this migration behaviour affects V2D separations. Varied experimental parameters include separation voltage, capillary diameter, separation length, separation buffer formulation, sample dilution, hit threshold, and pixel

inclusion. Since it is desirable to decouple error sources as much as possible, errors arising from band crossings (after multiple injection of sample) are studied independently from other effects.

Experimental procedures and reagents Unless otherwise noted, fused silica capillaries (PolyMicro Technologies) with 75 μm inner diameter and 375 μm outer diameter are used. Standard separation length (to first pixel) is 20 cm, with a 5.5 cm imaging window and 35 cm total length. Capillaries are subjected to a cleaning routine immediately prior to every separation (flush 3 minutes 0.1 M NaOH, 3 minutes 0.1 M HCl, 3 minutes ultrapure water, 7 minutes separation buffer, and 3 minutes separation buffer into anode vial) as well as a special cleaning routine for new capillaries (5 min methanol, 3 min ultrapure water, 5 min 1 M NaOH, 3 min ultrapure water, 5 min 0.1 M HCl, 5 min ultrapure water).

Standard separation buffer is a SDS-MW gel buffer for protein CGE (Beckman Coulter Inc., proprietary polymer buffer formulation, pH 8.0, 0.2 % SDS). The sample used is a 7-protein SDS-MW sizing standard 10-225 kDa (Beckman Coulter Inc., recombinant proteins, 16 mg/mL total protein concentration, MW 10, 20, 35, 50, 100, 150, 225 kDa). The sample is diluted to 3.2 mg/mL. Sample buffer (Beckman Coulter Inc., 100 mM Tris-HCl, pH 9.0, with 1 % SDS) is diluted with ultrapure water so that in the final sample, sample buffer is diluted 1:1. To reduce and denature proteins, 5 % v/v 2-mercaptoethanol (Sigma-Aldrich Co.) is added to the sample and heated above 90 °C for 5 minutes. Sample and separation buffer are degassed under vacuum prior to use.

At the beginning of each run at least 1 minute is allowed for noise data collection. Sample is injected electrokinetically (as described in the Section 1.3.2), with 200 V/cm field strength for 15 seconds. Standard separations proceed at 350 V/cm field strength. The UV light source is allowed at least 15 minutes to stabilise, and not left on over 3 hours to avoid intensity fluctuations.

2.3 Results

In the first part of the results from the error analysis, standard separations—and their associated errors—are characterised. This includes studying run-to-run variations, as well as comparing estimated position errors with actual errors of reconstructed starting position. The shape of the migration trajectories from detection data is examined to validate the track-fitting model. The errors for different molecular weights are compared. Next, the effects of hit-finding and tracking parameters on errors are explored. In particular the significance and dependence on hit threshold and detection length are considered. Thereafter variations of separation parameters are studied. A first study will characterise the effect of four different separation conditions, namely sample dilution, separation voltage, separation length, and capillary diameter. Furthermore, different separation buffers are compared. Finally, the effect of overlapping analytes—the so-called band crossings—are included in the analysis by studying errors from multiple injection separations.

In this thesis, the verb ‘to vertex’ will be used to refer to the situation where a cluster of space-time correlated tracks apparently emanate from a single space-time coordinate. If proteins share a common injection coordinate, then the corresponding tracks vertex. The term ‘measured error’ will be used to refer to the position error from the least squares fit, propagated back to the starting time of the separation. The term ‘starting position error’ will be used to refer to the difference in position between the recorded starting position (at the start of the capillary) and the reconstructed starting position (from the observed migration trajectory).

2.3.1 Characterisation of errors for standard separations

2.3.1.1 Vertex analysis

Before probing the various separation parameters, the errors of a standard separation are scrutinised. In Figure 2.14 the fitted tracks for a standard separation can be seen. The recorded vertex is marked with a black star. As can be seen, although some tracks do appear to converge at a vertex (indicating a common starting position and time), the vertex does not coincide with the recorded starting coordinates. The track pattern displayed is characteristic, with the first four bands (10, 20, 35, 50 kDa) vertexing better than the last three (100, 150, 225 kDa). If the vertex region is inspected in more detail as shown in Figure 2.14b, it becomes clear that the first tracks do not vertex either, but intersect at staggered coordinates. This indicates that migration may not be linear, or that other unknown systematic errors could be affecting the results.

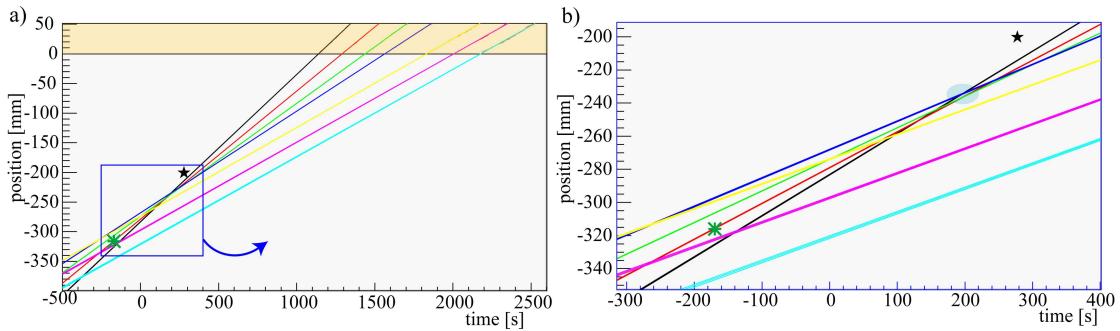


Figure 2.14: Track plot: reconstructed protein migration trajectories. a) Entire separation space (with detector zone shaded orange). b) Enlarged vertex region of tracks. The recorded injection coordinates are indicated with a solid black star symbol. The ‘best fit’ vertex with minimal summed distances of tracks is indicated with a green asterisk symbol. The black track indicates the lowest molecular weight protein (10 kDa), and the cyan track labels the highest molecular weight protein (225 kDa). The separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

To ensure that the discrepancy between the recorded starting coordinates and the observed vertex is not simply a problem of positioning the imaging window incorrectly or other instrumentation bias, a ‘minimum-error’ vertex was evaluated. The minimum-error vertex is at the coordinates where the summed track distances is minimal. In the track plots in Figure 2.14, the minimum-error vertex is displayed as a green eight-point asterisk. Note the coordinates of the minimum-error vertex: sub-zero time coordinate, indicating sample injection prior to starting data acquisition, and position coordinate in the neighbourhood of 31 cm, 11 cm away from the capillary inlet at 20 cm. Both coordinates are beyond reasonable errors for the recorded vertex. This minimum-error vertex is not very credible, indicating that it is not a simple problem of detector positioning for instance.

In Figure 2.14 the last three tracks do not converge at all, but the first four tracks might give a more meaningful vertex. If only the first four tracks are considered, the minimum-error vertex (in the blue shaded area in Figure 2.14b) is closer to the recorded starting coordinates. The position coordinate of this minimum-error vertex is then around three centimetres beyond the recorded starting position (capillary inlet). The time coordinate of the minimum-error vertex is around 75 seconds too early, compared to the recorded starting time. This minimum-error vertex is farther from the recorded starting coordinates than what could be explained by a simple positioning mistake or inaccuracy in the starting time from capacitive effects.

Obviously the choice of vertex has an effect on the errors of the reconstructed starting positions. This is displayed in Figure 2.15. While the track errors change only very little, the behaviour of starting position errors depends on the choice of vertex. The four

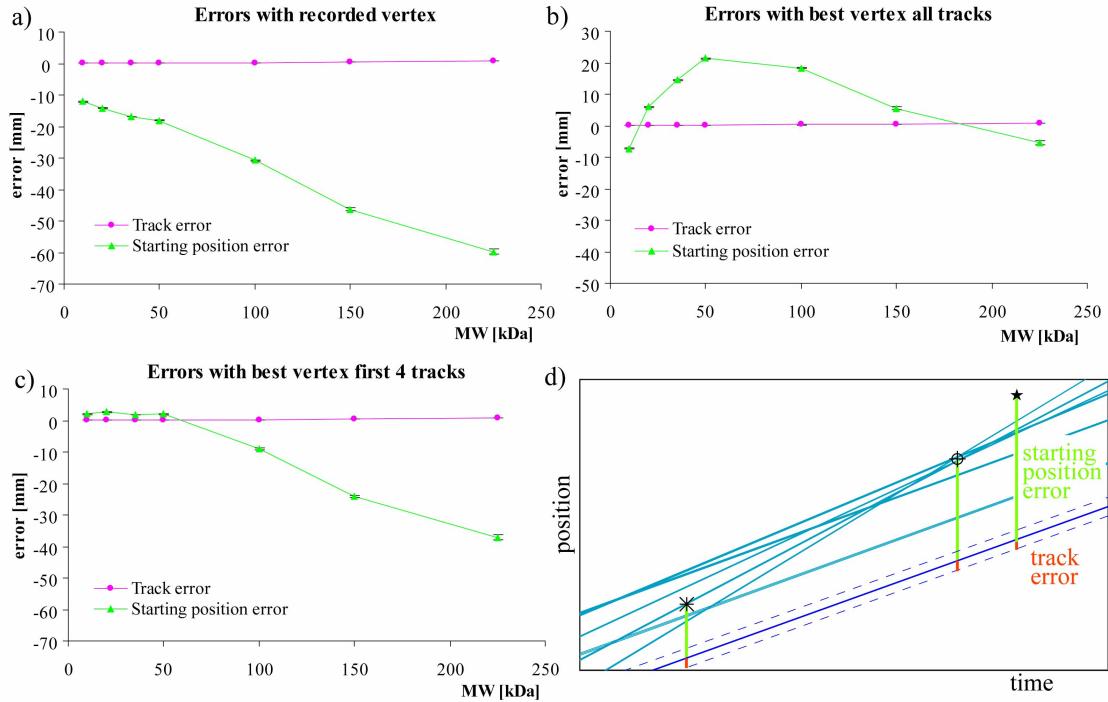


Figure 2.15: Track errors and starting position errors for different vertices. a) Errors with recorded vertex (solid star). b) Errors with ‘best’ vertex (minimum-error) of all 7 tracks (asterisk). c) Errors with minimum-error vertex for the first four tracks (crosshairs). d) Tracks around vertex region with different vertices (full star—recorded starting coordinates, asterisk—‘best’ vertex, crosshairs—minimum error vertex of first four tracks) and errors (green—starting position error, orange—track error). The separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

graphs in Figure 2.15 reflect nothing other than the linear nature of the tracks. Furthermore, although the graphs illustrate that different vertices can be found to minimise the starting position errors for the first four tracks (Figure 2.15c), or minimise the starting position errors for all tracks (Figure 2.15b), they also display that for any choice of vertex there remain considerable errors. This is also obvious from the track plot, where there is no single clear vertex. This indicates that the difference between recorded and observed vertex is not a simple mistake in capillary length, and may not be dismissed as irrelevant. Another important consideration is that in a V2D separation, the tracks will not vertex, and no other information besides recorded starting time is given. We are considering vertexing merely to judge the likely errors in the reconstructed V2D starting coordinate. For this reason, the recorded starting coordinates will always be used in all analyses.

2.3.1.2 Track errors and starting position errors

One of the first questions that springs to mind on observing track plots pertains to the large discrepancy between track error and starting position error. Because these errors are of a different order of magnitude, separate graphs are necessary to display their behaviour. Comparing track error and starting position error for different tracks, as seen in Figure 2.15a, it is seen that the starting position errors are roughly a factor of ten greater than the track errors. This indicates that while the migration across the detection window is approximately linear, there remains unaccounted non-linear migration in the course of the separation that needs to be investigated.

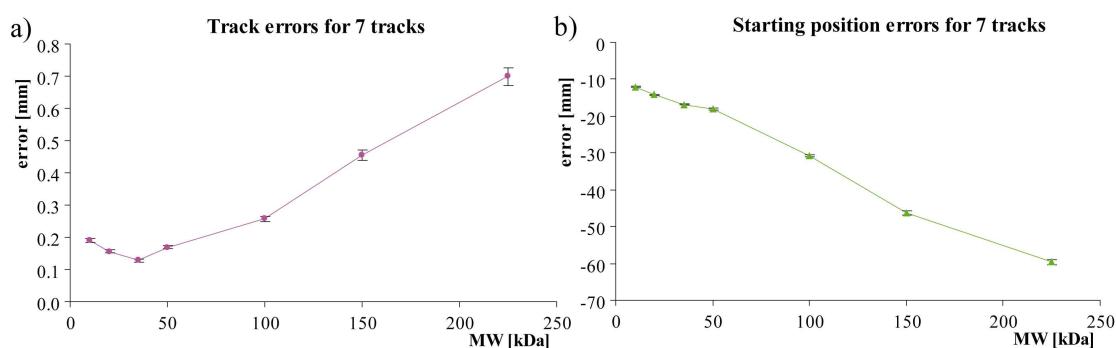


Figure 2.16: Track error and starting position error of seven proteins, plotted against molecular weight. a) Track errors. b) Starting position errors. The separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

On comparing track errors for increasing molecular weights in Figure 2.16a, one observes a slight decrease in tracking error for the lower 10–50 kDa range, and thereafter the track error increases again, from 0.1 mm to 0.7 mm. The track error indicates how well the migration trajectory fits a straight line. It is expected that heavier, slower proteins experience greater band broadening and display flatter peaks,^[181] and therefore the errors for their tracks are expected to be greater than for smaller, faster proteins.

While peak diffusion explains the track error increase above 50 kDa, track errors in the lower molecular weight region remain unexplained. Although it is a possibility that the sample is not well balanced and the greater errors for the first bands are caused by relatively smaller peaks, a quick check of a single pixel electropherogram (Figure 2.17) shows that this is not the case. These data indicate that there might be some unexpected reasons why migration trajectories differ from a straight line, and that the migration trajectory needs closer examination. Overall, a track error in the range of 0.1–0.7 mm is not very large and indicates a good track fit.

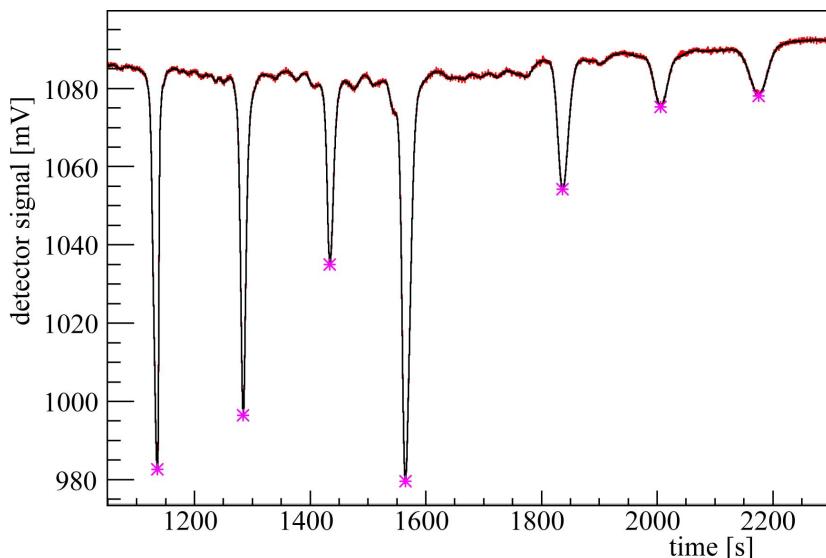


Figure 2.17: Electropherogram of a standard separation with seven proteins, 10–225 kDa. Pink stars indicate ‘hits’ as determined by the tracking algorithm. The separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer).

As mentioned above, the starting position errors are around a factor 10 greater than the track errors. For increasing molecular weight, the starting position error increases from around 10 mm up to 50 mm. The negative values of the starting position errors in Figure 2.16b indicate orientation relative to the recorded starting coordinates. The error increase is not simply linear, but shows an exaggerated change above 50 kDa.

To ensure that the observations made are representative and reflect day-to-day variations, data from repeated runs were collected. These runs were not performed consecutively, but interspersed with other experiments over months. In Figure 2.18 the fit parameters for all the tracks are plotted, along with error estimates.

In the plot of fit parameters, the values lie on a straight line when the tracks form a vertex, and the values that deviate from the straight-line trend are from tracks that do not have a common vertex with the others. Although the separations show some variation, values from different runs show a similar flattening off, away from a straight line.

The separations where the values lie outside the general trend coincide with the separations that are of poor quality, where bubbles intersect migration tracks or the sample is slightly depleted. Figure 2.19 shows how a bubble intersects with migration trajectories, adding false hits to a track. Least squares fits are known to perform poorly when the data include outliers. Inclusion of some outliers is a trade-off for an acceptance width for hits around a track that is wide enough to include most of the correct hits.

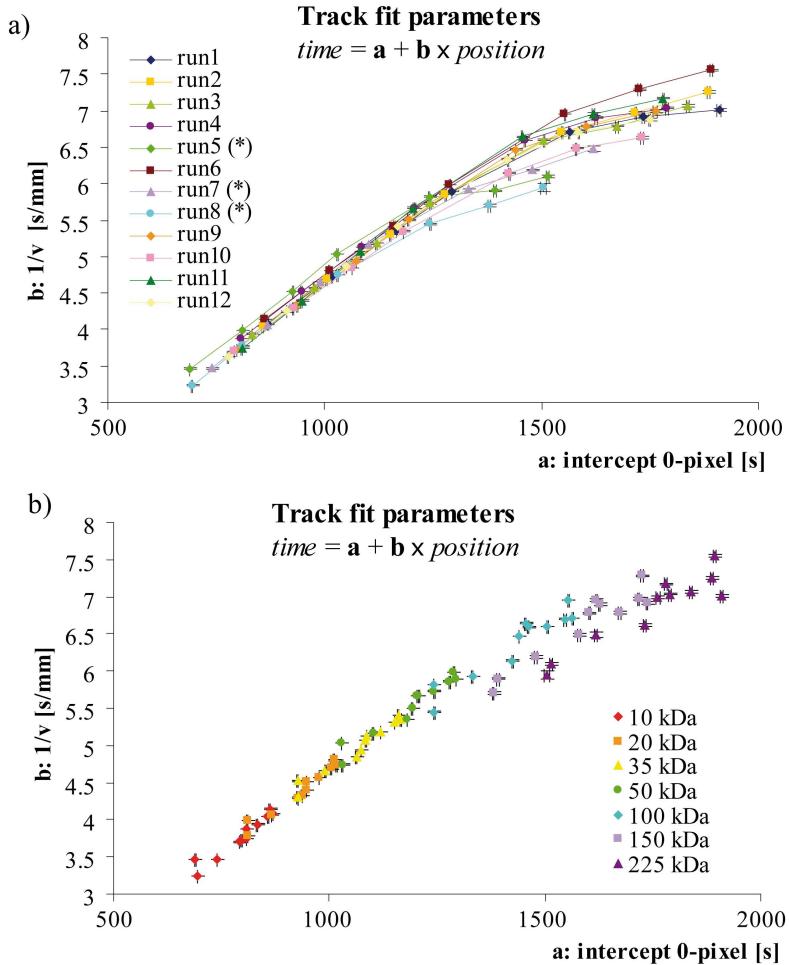


Figure 2.18: Track slope and intercept for 12 repeat separations. Track equation ($time = \mathbf{a} + \mathbf{b} \times position$) with intercept \mathbf{a} and slope \mathbf{b} . a) Grouped by run. b) Grouped by band molecular weight. Runs with poor data quality are marked with an asterisk in the legend. The separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

This illustrates the problem of track interference, and the impact of small experimental irregularities.

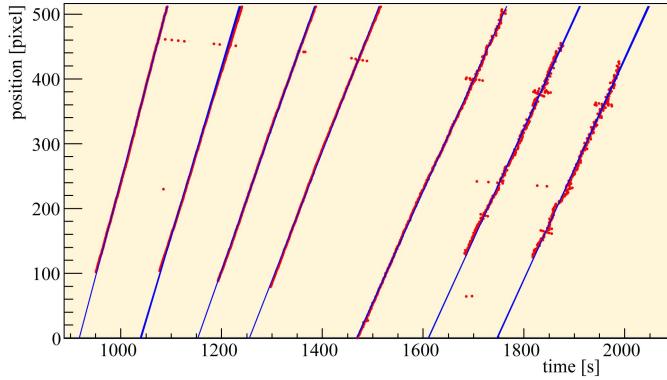


Figure 2.19: Hitmap illustrating a bubble interfering with protein tracks. Hits (red) and fitted tracks (blue) are shown. In the upper half a string of ‘false’ hits from an unknown disturbance such as a bubble are included in tracks. The data shown are from run 5 of the standard separation series in Figure 2.18.

2.3.1.3 Error correlation

With the collected data from 12 repeat runs, it is possible to study the dependence of errors on various properties. This is important in view of V2D separations, where the error of the reconstructed starting coordinates is not known. If these errors are correlated with some other measure, such as mobility or migration time, this can be used to estimate errors, and correct the reconstructed starting position.

From the track-fitting, two track properties are readily available: intercept and slope, or migration time and mobility. An initial hypothesis is that errors (starting position error as well as track error) are correlated to migration time. This is explained by peak broadening with time-dependent molecular diffusion. For molecular diffusion, the diffusion length σ_D is dependent on the diffusion coefficient D and the time t as described by the Einstein equation $\sigma_D^2 = 2Dt$ (Eq. 1.20 in Section 1.3.2).

Figure 2.20 shows that although errors do increase with migration time, and a quadratic function can be found to fit values of the same run, the differences between runs are considerable and a global correlation between starting position error and migration time is difficult. Since migration time and migration speed are correlated for this experiment, migration speed (i.e. mobility) behaves similarly.

Although migration time or mobility can be replaced with relative migration time, or relative mobility, respectively, to achieve better alignment, the run-to-run variance is still too large for reliable calibration, as shown in Figure 2.21.

Neither migration time nor mobility is a convincing candidate for calibration to correct the reconstructed start position error for an unknown track. Another readily available

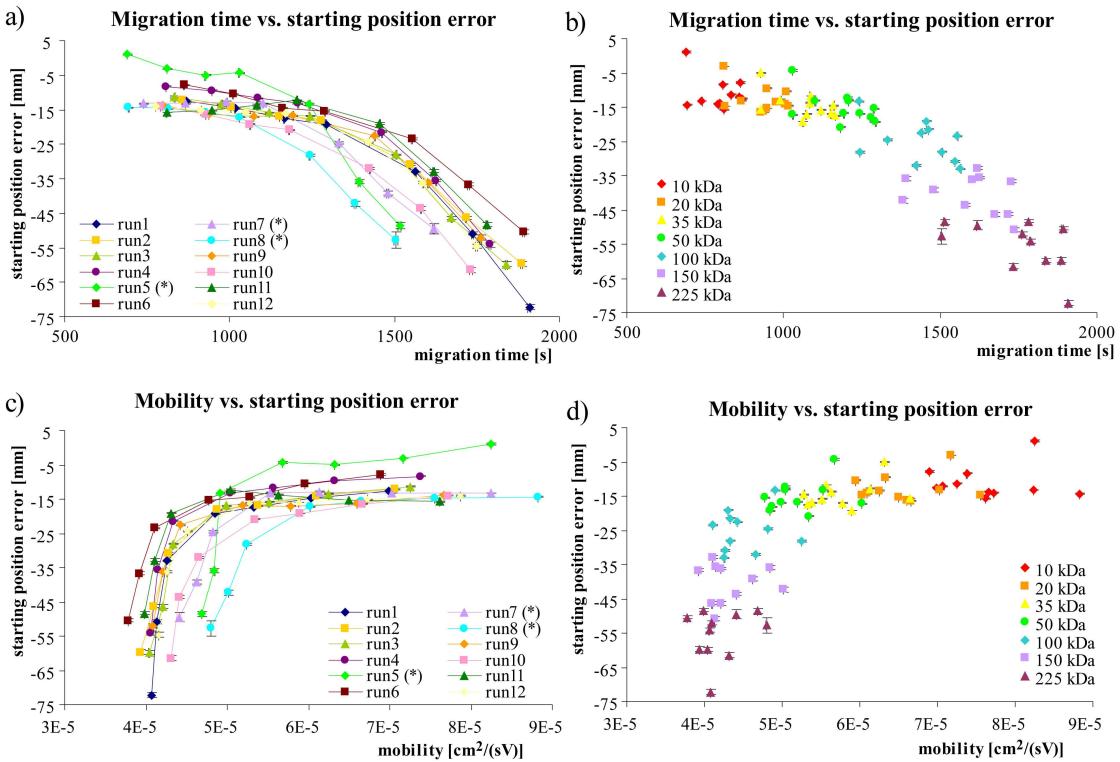


Figure 2.20: Starting position errors plotted against migration time and mobility. a) Starting position error vs. migration time, grouped by run. b) Starting position error vs. migration time, grouped by band molecular weight. c) Starting position error vs. mobility, grouped by run. d) Starting position error vs. mobility, grouped by band molecular weight. Runs with poor data quality are marked with an asterisk in the legend. The separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

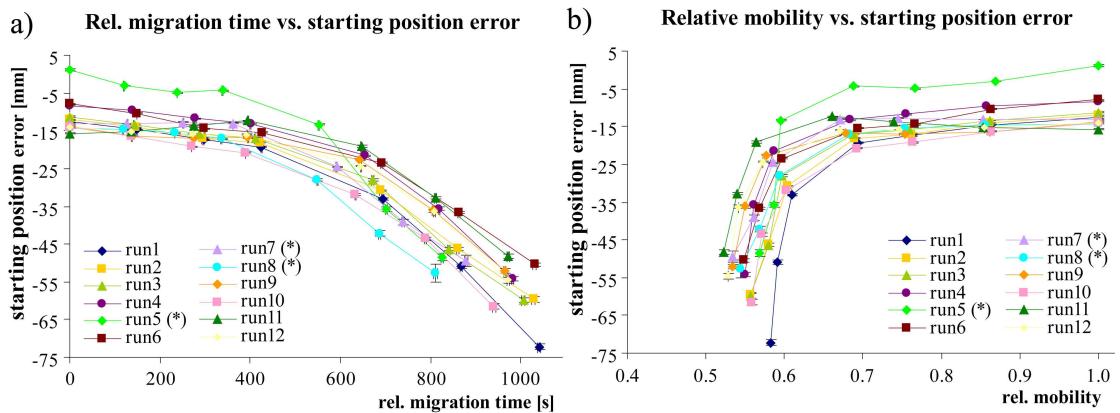


Figure 2.21: Starting position errors plotted against relative migration time and mobility. a) Starting position error vs. relative migration time, grouped by run. b) Starting position error vs. relative mobility, grouped by run. For calibration the first peak of a series is used as reference. Runs with poor data quality are marked with an asterisk in the legend. The separation conditions are standard, with standard sample.

quantity is the track error. If the starting position error were correlated to the track error, this would enable estimation and correction of unknown position errors. Figure 2.22 shows starting position errors plotted against track errors for 12 runs with seven tracks each, grouped either by run (Figure 2.22a) or by molecular weight (Figure 2.22b).

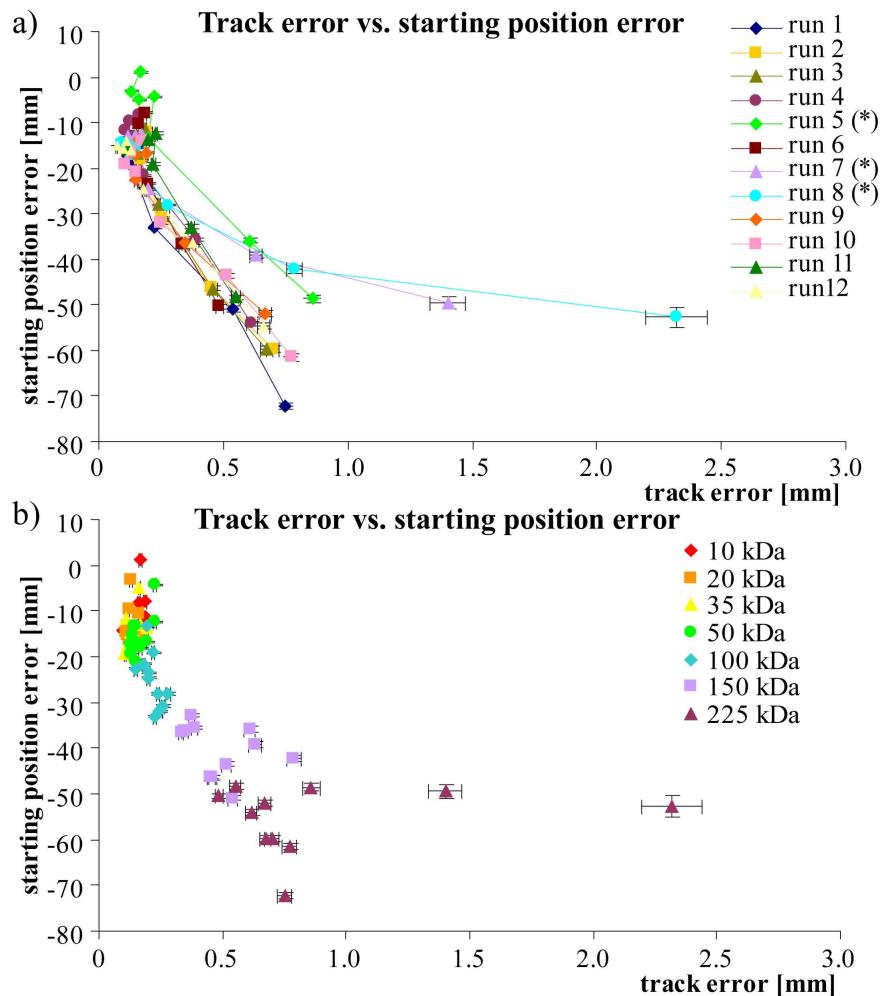


Figure 2.22: Starting position errors plotted against track errors. a) Starting position error vs. track error, grouped by run. b) Starting position error vs. track error, grouped by molecular weight. Runs with poor data quality are marked with an asterisk in the legend. The separation conditions are standard, with standard sample (7 proteins, 10–225 kDa).

As observed earlier, the track error is sensitive to outliers, which can occur in the form of false hits from bubbles or other background signals that are added to tracks. This affects the graphs in Figure 2.22, where runs 5 and 8 shear out from the rest. As for the majority of separations, the starting position error does appear to correlate with the track error, and this might possibly be used for error correction.

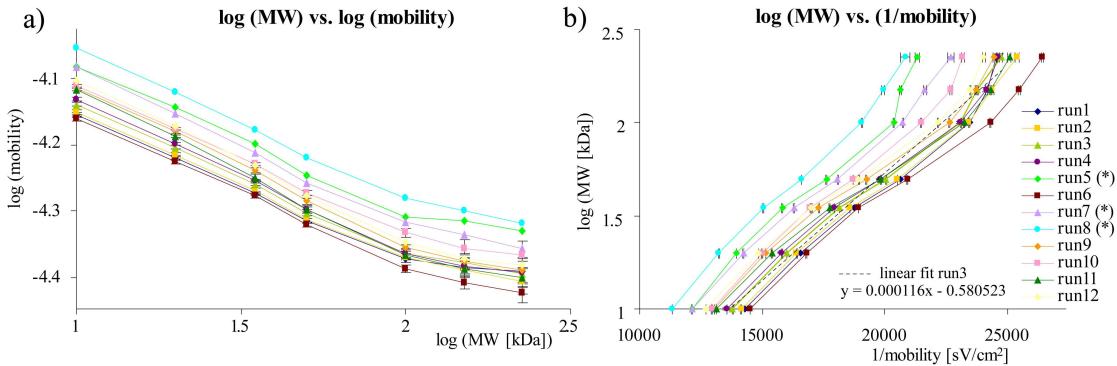


Figure 2.23: Molecular weight versus mobility. a) log (molecular weight) vs. log (mobility). b) log (molecular weight) vs. inverse mobility. The dashed black line indicates a straight line fit for data from run 3. Data were taken from 12 separations (runs with poor data quality are marked with an asterisk in the legend) performed under standard conditions (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa).

The collected data are used to verify that the buffer is not approaching loss of separation for the 225 kDa proteins,^[69] and that the molecular weight may be calibrated to the mobility. Figure 2.23 shows two versions of molecular weight-mobility plots.

In the log (mobility) vs. log (molecular weight) plot (the conventional format for analysing separation performance), it can be seen that even for the highest molecular weight protein, the curve has not flattened out entirely, and sieving is still effective. It does however appear that 225 kDa is near the upper limit of sample molecular weight for which this buffer is effective.

In the log (molecular weight) vs. inverse mobility plot, a straight line can be fitted as a calibration function in the form of Equation 2.17, $\log(\text{MW}) = p + q \left(\frac{1}{\mu} \right)$ with molecular weight MW [kDa], mobility μ [$\text{cm}^2/(\text{sV})$], and fitted intercept p and slope q .^[199] The calibration curves for different runs demonstrate the same shape, but with an offset. For a purposeful calibration, samples would need to include a calibration marker, so that the relative offset could be corrected. This is common practice in many protein separation techniques.

2.3.1.4 Goodness of fit analysis

To study the quality of the fit obtained, different quantities are compared. In Figure 2.24, the correlation coefficient ρ (proportion of data that is accounted for in the fit), the χ^2 (chi square) value, and standard deviation σ_t of hit coordinates relative to the fitted lines (in the time-dimension) are shown for five sets of tracks.

In Figure 2.24a, the correlation coefficient of different tracks is shown for five separations under standard conditions, with tracks plotted by molecular weight. The correlation

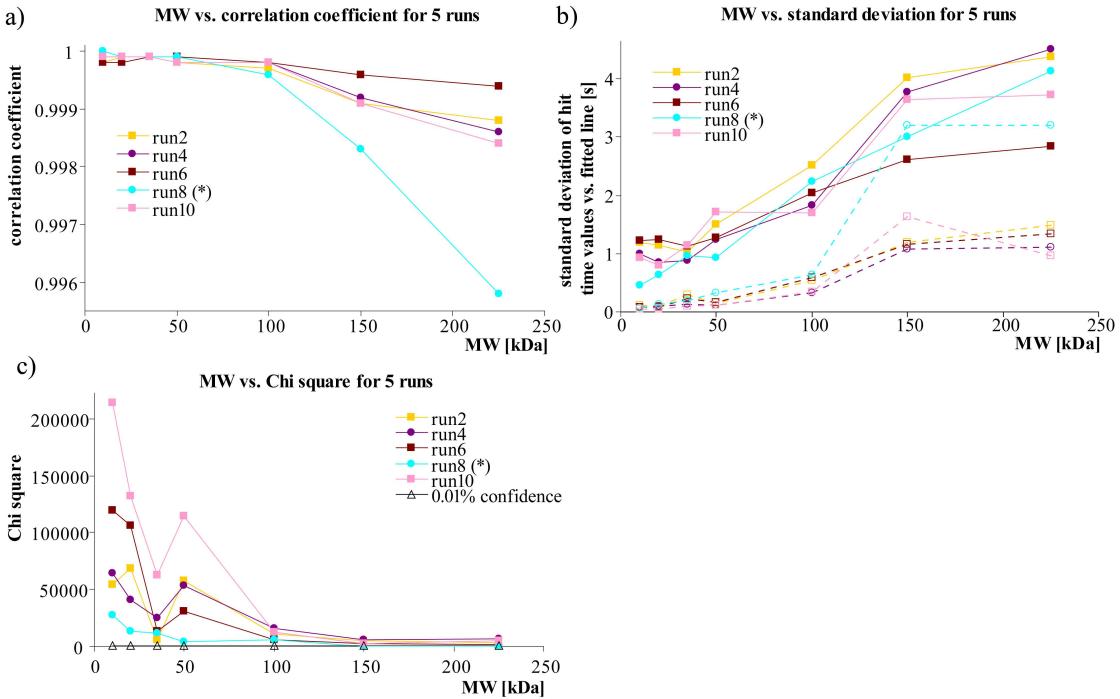


Figure 2.24: Goodness of fit for data from five standard separations. a) Correlation coefficient for protein tracks, plotted by molecular weight. b) Standard deviation of hit time coordinate relative to fitted line. Solid lines represent all hits of a track, dashed lines indicate where only the first 20 hits of a track are considered. c) χ^2 (chi square) values of fits, plotted by molecular weight. In black the χ^2 value that is associated with a 0.01 % confidence level of the fit being correct. Runs with poor data quality are marked with an asterisk in the legend. All separations are performed under standard separation conditions (200 mm separation length, 350 V/cm electric field strength, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa).

coefficient indicates how strongly the two coordinates (time, position) of each hit are correlated. Correlation coefficient values fall between 0.996 and 0.999, which may be interpreted as a high degree of linear dependence between time and position coordinate of hits. It is noteworthy that the correlation decreases slightly with higher molecular weight bands, which is plausible because these bands have flatter, wider peaks and the exact hit coordinate varies slightly more. The run with especially poor correlation is one of the runs that was already earlier identified as a particularly poor quality run with bubbles interfering with protein tracks.

In Figure 2.24b, the standard deviation of hit time-coordinates relative to the fitted line is shown for different tracks. This standard deviation is plotted for inclusion of all the hits belonging to a track (solid lines), and for inclusion of only the first 20 hits of a track (dashed lines). With increasing molecular weight, peaks get wider and flatter, and hits are spread further around the fitted track, hence the standard deviation increases with molecular weight. The standard deviation of the first 20 hits is smaller than for the complete set of hits belonging to a track. This is explained by the non-linear migration

discussed above. The standard deviation of the first 20 hits indicates the (random) error on the time coordinate of hits. The standard deviation of the complete set of hits indicates the (systematic) error of hits relative to a track. The χ^2 values express the discrepancy between these two standard deviations and indicate how much of the track lies within a reasonable distance (the expected standard deviation) of the measured values.

In Figure 2.24c, χ^2 values are plotted for increasing molecular weight. The number of degrees of freedom for most tracks is 510, with certain poor quality tracks below that (only in three instances below 300). For a good fit, the χ^2 values are slightly higher than the value for degree of freedom. The black line (open triangle symbols) indicates critical values for χ^2 at a 0.01 % confidence level that the track fits the measurements (0.01% probability that an observed χ^2 exceeds the critical χ^2 values by chance, even for a correct model). These critical χ^2 values fall between 280 and 640 for the plotted tracks. In contrast, the obtained χ^2 values are orders of magnitude greater. This means that only very few of the measured values actually coincide with the track.

To summarise, although time and position coordinates are highly correlated and indicate a good linear fit, the measurement values do not actually follow a linear migration track, but more complex non-linear trajectories. Although the fitted lines represent the data well, they are only an approximation of the non-linear migration trajectory of proteins.

2.3.1.5 Shape of migration trajectories

From analysis of errors so far it has become clear that the trajectory shapes need close scrutiny. For this purpose, the residuals, i.e. the distance in the time-dimension between each hit and the fitted track (as illustrated in Figure 2.13), are plotted for all seven tracks of a separation. In Figure 2.25a the hits and fitted track for a 50 kDa band are shown, and in Figure 2.25b the residual plots for all seven tracks are shown.

In conventional capillary gel electrophoresis, the relationship between position and time is generally assumed to be linear.^[40] With single-point detection there is no evidence to the contrary. With detection at multiple positions along the capillary, it appears that the trajectories are not straight. Contrary to the expected straight lines, the residuals plots are curved and bend at certain points. This is confirmed for all separations. The residuals of different tracks are positioned correctly relative to one another, so the observation that the bends in migration trajectories lie in a straight line deserves further investigation.

Figure 2.26 displays the heatmap with migration trajectories of the 100, 150, and 225 kDa bands. On closer inspection of this heatmap, a bend in the migration trajectories

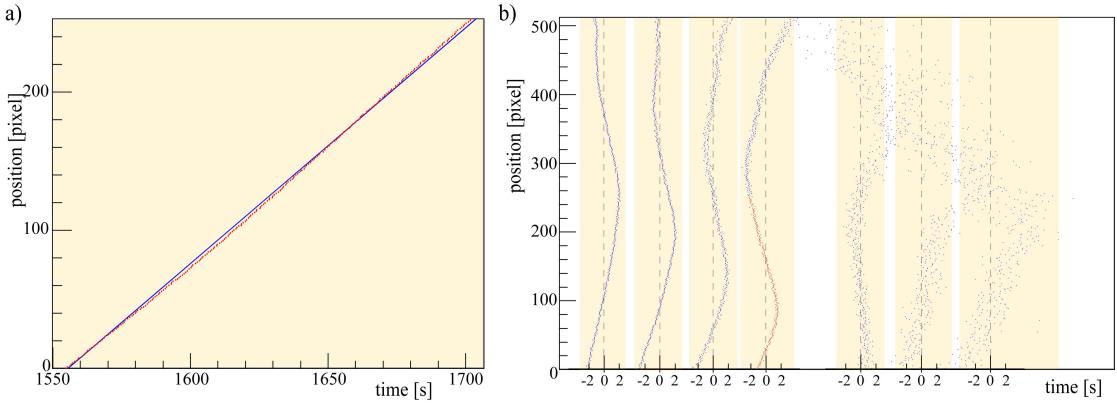


Figure 2.25: Shape of migration trajectories. a) Hits (red) and fitted track (blue) for close-up of 50 kDa band trajectory. b) Residuals of seven tracks are plotted. Residuals are positioned correctly relative to one another. The section shown in Figure a) is coloured red in Figure b). Standard separation conditions and sample apply.

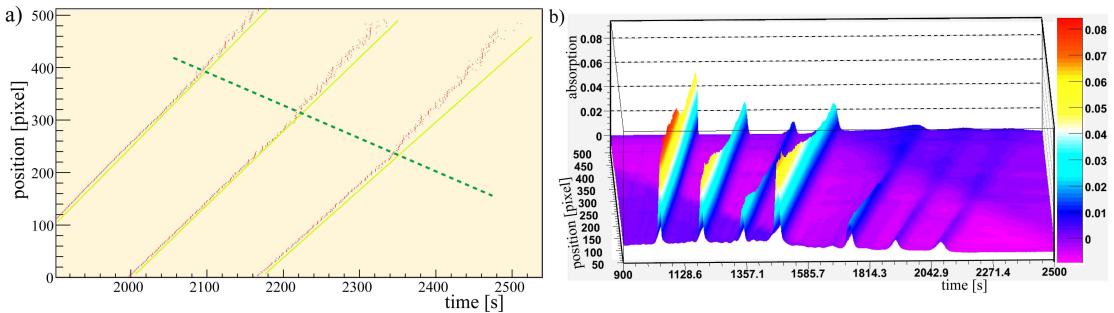


Figure 2.26: Bend in migration trajectories. a) Section of heatmap showing last three bands (100, 150, and 225 kDa). Hits are shown in red; light green indicates a straight line for reference; the kinks lie on a line, indicated with a dashed dark green line. b) Three-dimensional electropherogram. A lighter background band can be seen from front right to back left. Standard separation conditions and sample apply.

is visible, especially when compared to a straight line (light green line). In this plot, it can be seen that the bending point of the three trajectories lie on the same line (dashed dark green line in Figure 2.26a).

Further evidence is found in the three-dimensional electropherogram plot (Figure 2.26b). Here the traces from all pixels are included to give a comprehensive image of the separation, with the colourscale additionally illustrating height or z-coordinate (absorbance). The seven protein bands are seen as ridges from the front toward the back right. Note the lighter stripe along the baseline that moves from front right to back left, against the protein bands, at a shallower angle than the proteins.

This counterband is not an image artefact, but can be seen in the single pixel electropherograms as a step. The point of intersection of this counterband with the protein peaks coincides with the trajectory kink. The counterband can be seen in a blank run, where no protein sample is injected, as displayed in Figure 2.27b.

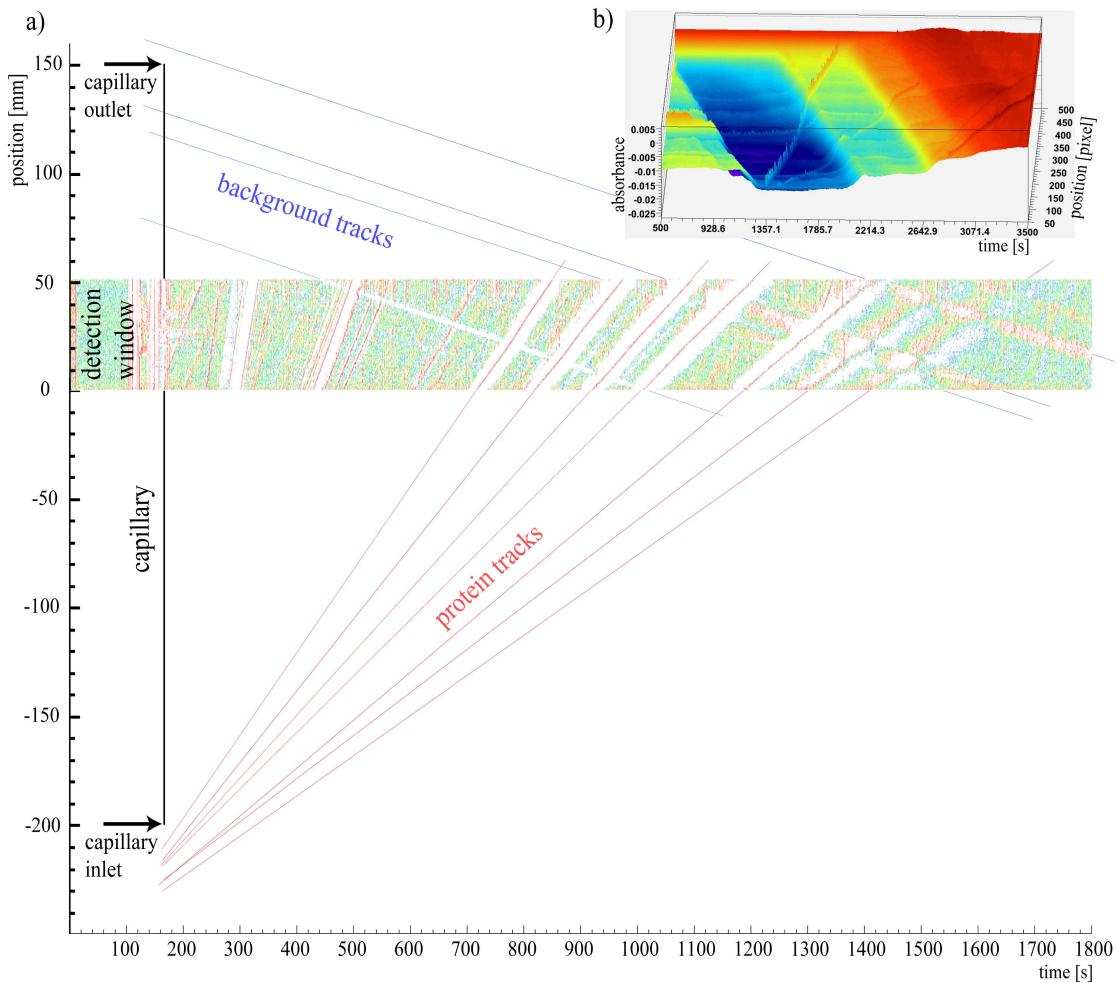


Figure 2.27: Background disturbance. a) 2D projection of 3D electropherogram. Time-position representation of detection data, with colour representing peak height. Seven proteins migrate from left to right (red line in blank zone). Four counterbands migrate from right to left (blue lines in blank zones). In black on the left the length of the capillary is represented, positioned at the coordinate of the injection time. Tracks are laid through the protein bands (red) and counterbands (blue). The point of intersection of each blue track with the black capillary indicates its position at the start of the separation. b) (Inset) Three-dimensional electropherogram of a blank run with no sample injected. Standard separation conditions and sample apply.

The counterband is a background disturbance that migrates from the anode to cathode, in the opposite direction as the proteins. Counterband migration is relatively slow compared to protein migration. The counterband causes non-uniform migration conditions and interferes with protein migration. Due to the step shape of the counterband in electropherograms, it is not detected by peak-finding algorithms.

In a two-dimensional time-position image of detection data, with colour representing peak height, the counterband steps are seen as blank zones (in a slope the peak-finding algorithm fails and no peak heights can be assigned) in Figure 2.27a.

A graphic analysis of the counterbands indicates that they migrate at the same speed, and that the starting positions are in the capillary for the first bands, and in the anodic vial for the last band. The background migration along with the non-linear protein migration it appears to cause will be discussed in further detail in Section 2.4.2.

2.3.1.6 Peak shape analysis

In Figure 2.28 an enlarged section of the residual plot of the second (20 kDa) track is shown. The section displayed is only from the first 150 pixels, before the counterband

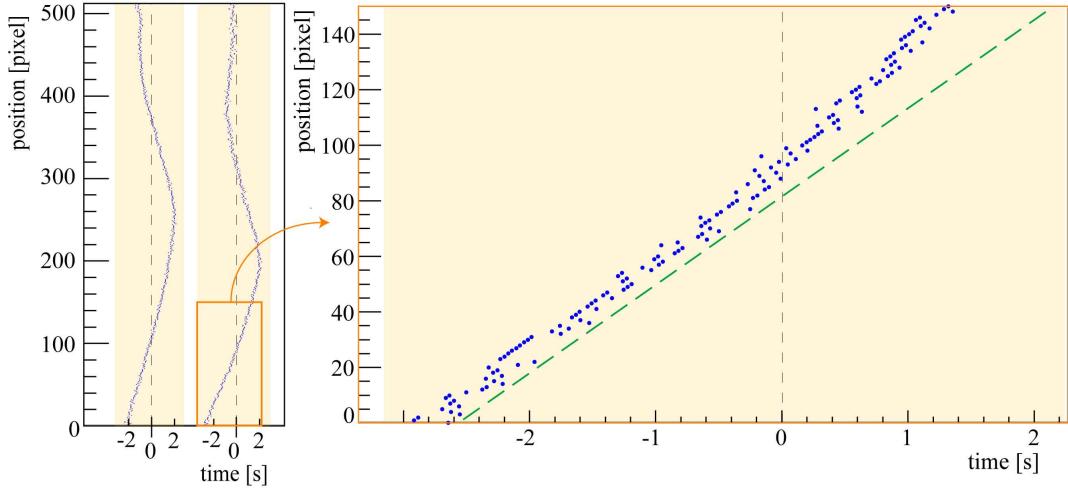


Figure 2.28: Non-linear migration trajectory in an enlarged section of the residual plot. The first 150 hits (blue) of the second (20 kDa) track of a standard separation are shown. The straight reference line (dashed green) illustrates that this band is not undergoing linear migration, but accelerating.

intersects with this band. If this band migrates in strictly linear fashion, as expected, the shape of the residual is a straight line. Comparing the residual shape in Figure 2.28 with a straight line (dashed green) shows that this band does not migrate in a linear motion, but accelerates as time progresses. At the same time, the shape of peaks in the single pixel electropherograms changes, from symmetrical to asymmetrical, as seen in Figure 2.29.¹

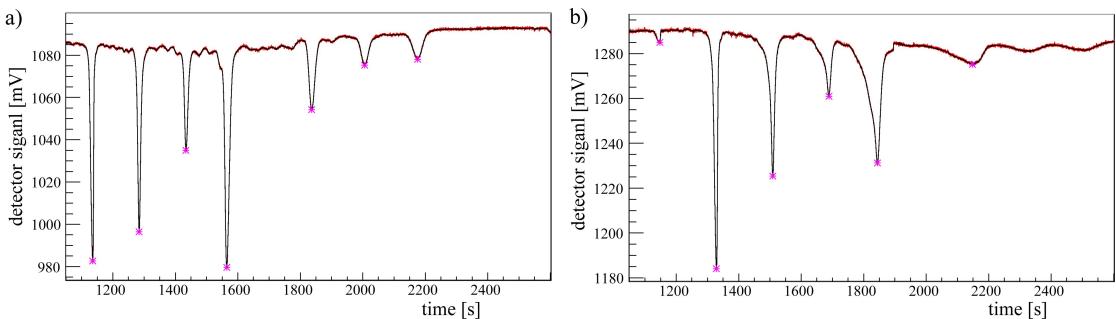


Figure 2.29: Electropherograms of pixels at different positions along the detector. a) Pixel near beginning of detection window. b) Pixel near end of detection window. Approximately 50 mm lie between the two pixels. Standard separation conditions and sample apply.

¹This effect was first observed in 2003 by deltaDOT Ltd. We are indebted to Dr Nathan Lacher of Pfizer Global R+D for bringing it back to attention in 2006.

To quantify the peak shape, the values for half width at half maximum (HWHM) in time-dimension toward the front ('slow' trailing edge) and back ('fast' leading flank) of the maximum are compared (see Figure 2.30a). If the peak is symmetrical, HWHM toward front and back are same. If the peak shape remains the same for different electropherograms from different pixels, HWHM develop in the same way. In Figure 2.30b the HWHM are displayed for seven bands. For each hit, the HWHM toward fast (black) and slow (red) edge of a peak is plotted against the position coordinate of the hit (the pattern remains the same if plotted against the time coordinate). In Figure 2.30 it can be seen that the front-end width and back-end width diverge as peaks migrate across the imaging window.

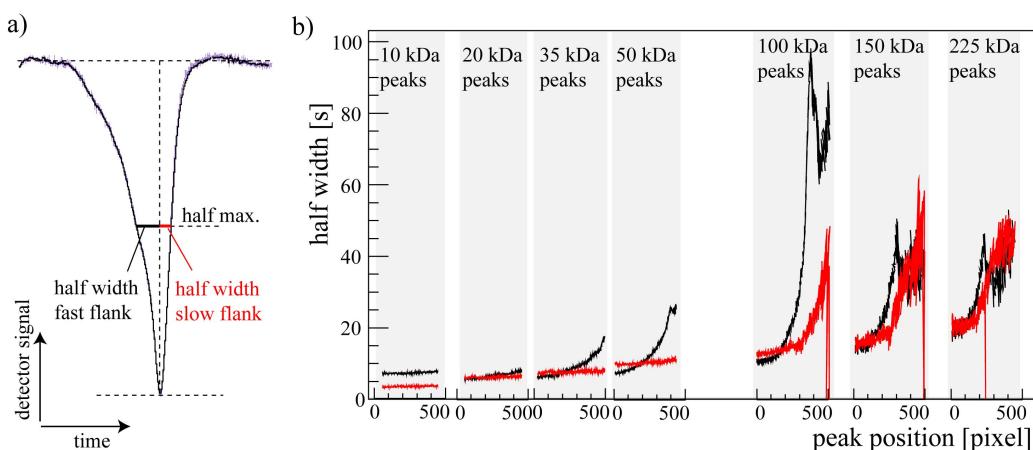


Figure 2.30: Peak shape as migration across detection window progresses.
a) Schematic of half maximum, and half width at half maximum (HWHM) toward front (slow) and back (fast) flank of a peak. b) HWHM plot for a standard separation. Each line pair represents a band, with red for the 'slow' HWHM and black for the 'fast' HWHM.

To find the cause for the development of peak asymmetry, a comparison is made for the same separation, but with the midsection of the imaging window shielded from illumination. In Figure 2.31 the HWHM plot of the first five bands is shown for this separation, along with its control (PDA blocked out but capillary illuminated).

The effect of blocking UV illumination in the midsection of the imaging window is that the front-end width and back-end width do not diverge. This indicates that non-linear migration which goes together with peak asymmetry is a UV-mediated effect. UV effects will be discussed in further detail in Section 2.4.3.

The aspect of non-linear migration in capillary gel electrophoresis of proteins has not been explored prior to this work. As will be obvious, such non-linearities represent at the very least a major challenge to the V2D methodology, and ultimately may provide the limits in performance for the approach.

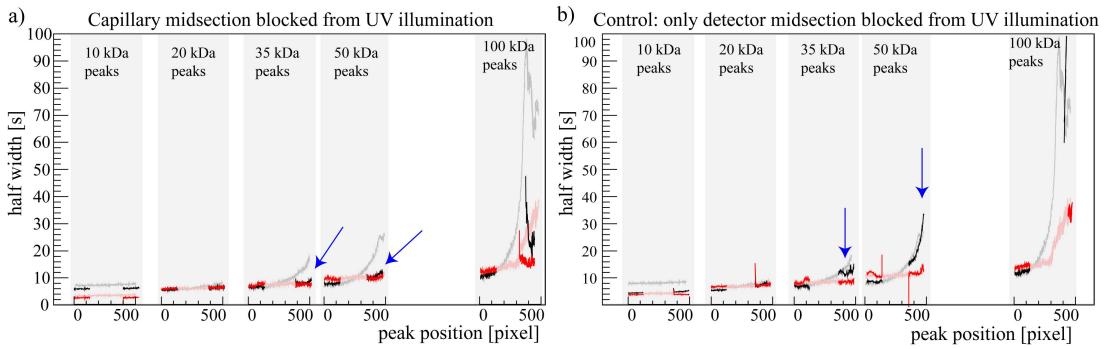


Figure 2.31: Peak shape for five bands with and without capillary midsection covered. Peak shape data from a standard separation is underlaid (pale traces). a) Capillary and detector midsection blocked from UV illumination. The traces deviate from the standard separation, as indicated with arrows. b) Control: only detector midsection blocked from UV illumination, the capillary remains illuminated. Traces behave same as for the standard separation (pale traces). Two line pairs represent a band, with a gap in the middle from the blocked detector section. Red lines indicate the ‘slow’ HWHM and black lines the ‘fast’ HWHM.

2.3.2 Effect of hit-finding and tracking parameters

To assess how the track-finding parameters affect results, a set of data from one separation is evaluated with different settings. A first study focuses on the threshold value for finding hits. Thereafter, the number of pixels included in the tracking procedure is varied, and finally the tracking results of pixel subsets at different positions are compared.

2.3.2.1 Hit-finding threshold

As described in Section 2.2.3, a hit is a minimum in the electropherogram that exceeds the noise value multiplied by a threshold value. The choice of threshold influences which hits are included in a track, as illustrated in Figure 2.32, where two different settings of threshold are compared.

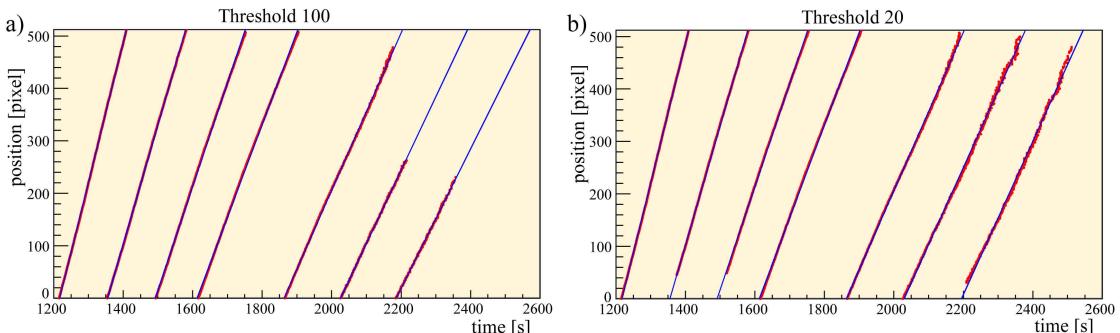


Figure 2.32: Hits (red) and tracks (blue) for different hit thresholds. a) Threshold 100. b) Threshold 20. Hit-finding and tracking are performed on the same set of data from a standard separation (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa).

For the lower threshold (Figure 2.32b), the hits extend farther along the track. This has to do with the relative signal strength of the different proteins in the sample. The first four tracks belong to low molecular weight proteins (10–50 kDa), which have far higher signals than the high molecular weight proteins (see electropherogram in Figure 2.29). Peak-finding for low molecular weight proteins is only affected when the threshold is very high, over 500. By contrast, hit-finding of the high molecular weight proteins (100–225 kDa) responds more strongly to the threshold value.

At a constant threshold, fewer and fewer hits are incorporated as the molecular weight increases. This is not unexpected, however the impact on track errors is instructive. In Figure 2.33, the errors for different molecular weights with varying threshold are displayed. As expected, for 10–50 kDa proteins no great variations are seen with increasing threshold. These peaks are well above even high thresholds. The track errors of 100–225 kDa proteins first decrease with decreasing threshold, then after a minimum around 100 the track errors increase with yet lower thresholds.

With lower thresholds, more hits are included in a track, and due to non-linear migration effects from counterbands and UV effects, these hits deviate from the track more and so increase the track error. The same behaviour is seen in the starting position errors. For lower thresholds, more hits are included in a track, and the non-linear parts of the migration shift the track.

Recalling the sinusoidal shape of residual plots (Figure 2.25), and observing the hits in Figure 2.32, if a certain number of hits is included, the non-linear errors can balance out to reduce the track error again. This is seen in the 150 kDa track, where with lower threshold the error generally increases, but has some regional minima. For thresholds over 200 the relatively small peaks of high molecular weight proteins are no longer found by the hit-finding algorithm, and no tracks are identified.

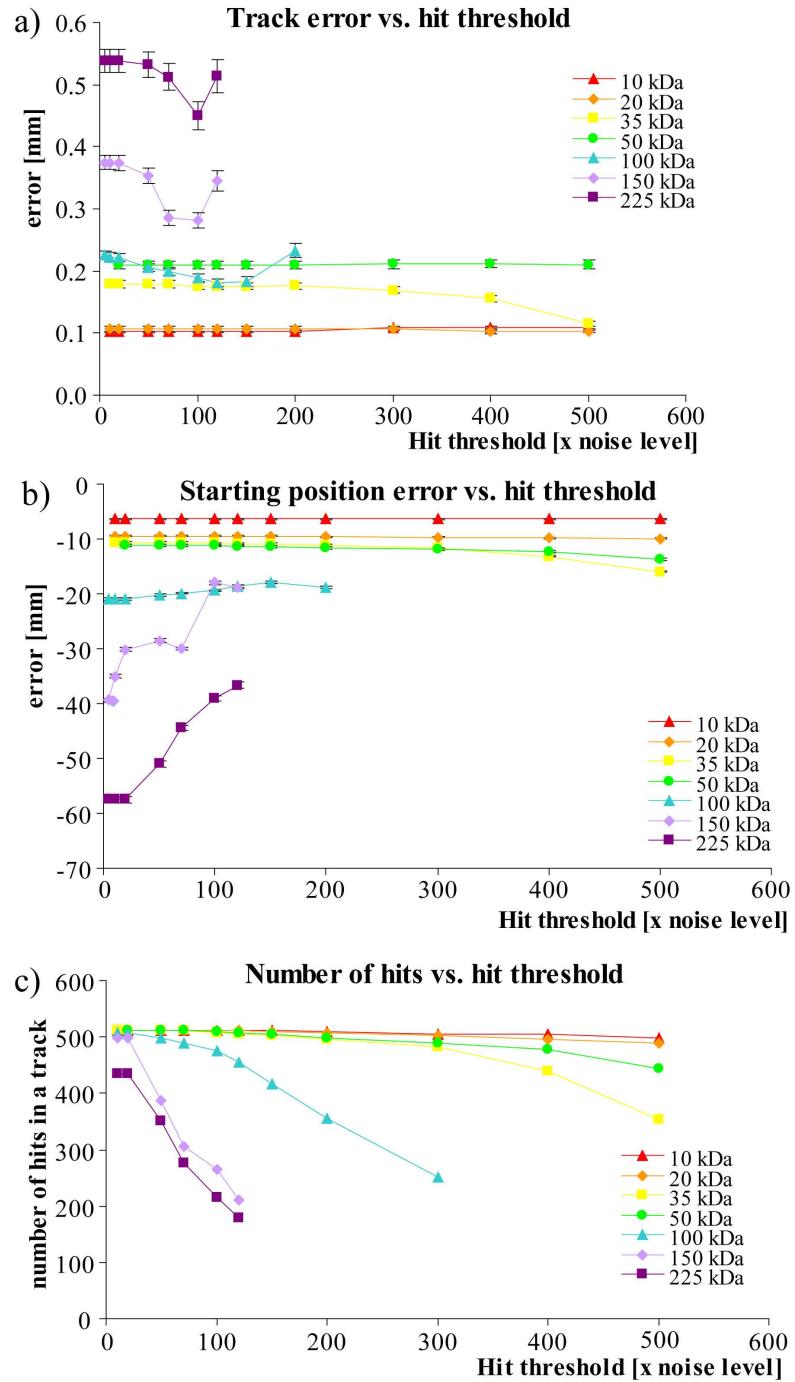


Figure 2.33: Track error, starting position error, and number of hits per track for different hit threshold levels. a) Track error vs. hit threshold. b) Starting position error vs. hit threshold. c) Number of hits per track vs. hit threshold. Hit-finding and tracking performed on the same set of data from a standard separation (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

2.3.2.2 Pixel subsets: length of subset

It is useful to study errors depending on how many pixels are included in the analysis data. Again the same data set is evaluated, and thresholds are adjusted to ensure hits are present along the entire pixel range of a track. This is illustrated in Figure 2.34.

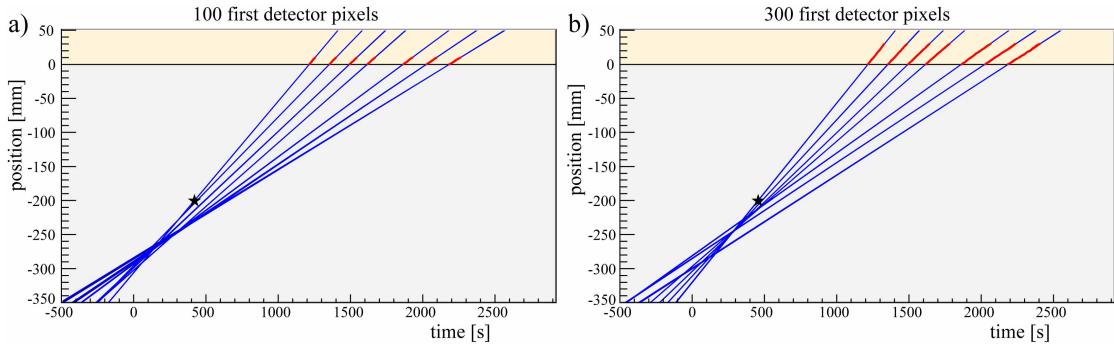


Figure 2.34: Hits (red) and tracks (blue) for different lengths of pixel subset. a) 100 pixel subset (pixel 0–99). b) 300 pixel subset (pixel 0–299). The recorded vertex is indicated with a black star. Hit-finding and tracking are performed on the same set of data from a standard separation (200 mm separation length, 350 V/cm electric field strength, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

In Figure 2.35 the track error and starting position error for different molecular weights with varying detection length are compared. For a very short tracking distance, the track errors are large, but beyond 20 mm tracking distance the track errors do not change much and are all below 1 mm. The decrease in track error can be taken as an indication that from a track-fitting point of view, there are considerable gains for increasing the tracking distance above 20 mm. Improvements come from the increased number of measurement values included in a fit (the error of the fit decreases with $1/\sqrt{N}$), but also with geometrical factors (extrapolation distance relative to tracking distance, smaller slope error for greater tracking distance).

The starting position errors seen in Figure 2.35b show a more varied behaviour than the track errors. Small proteins, 10–35 kDa, show increasing error with increasing detection length. Midrange proteins, 50–100 kDa, have decreasing starting position errors with increasing detection length. Large proteins over 150 kDa behave similarly to small proteins and show increasing errors.

The best explanation for this unexpected pattern is that different causes for non-linear migration and their effects on migration trajectory counteract each other. In the first few pixels, the effects from counterbands and UV degradation have not yet had much impact, and only thermal effects should cause much of a starting position error. In the residual plot (Figure 2.25) it can be seen that the first counterband intersects with the

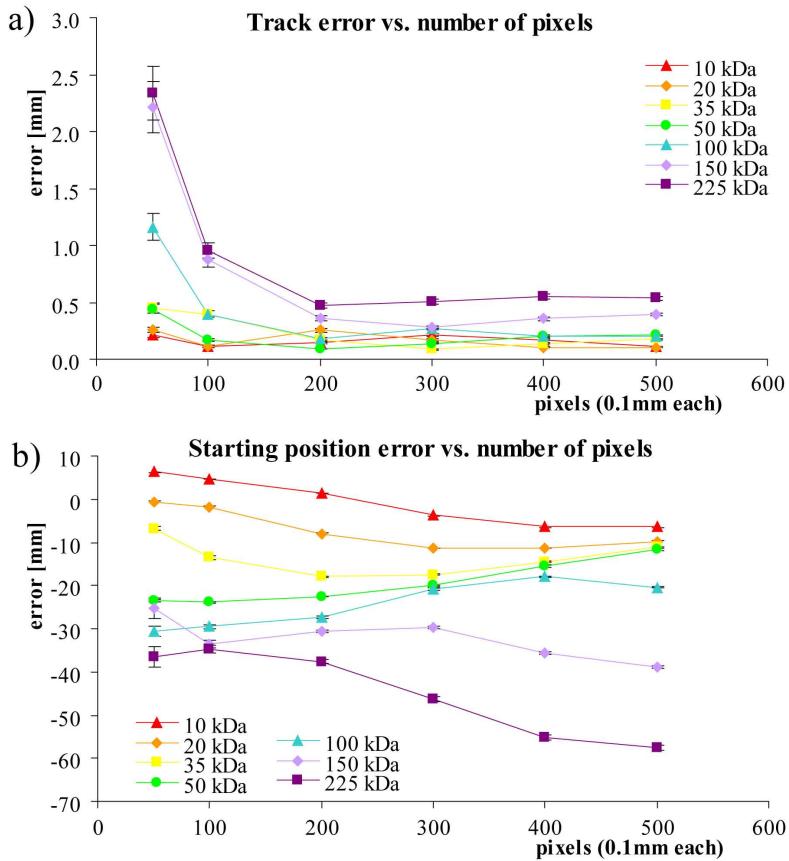


Figure 2.35: Track errors and starting position errors for different lengths of pixel subsets. a) Track errors vs. number of pixels. b) Starting position errors vs. number of pixels. Hit-finding and tracking performed on the same set of data from a standard separation (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with the standard sample (7 proteins, 10–225 kDa).

10 kDa track around the 250-th pixel, the 20 kDa track around the 200-th pixel, and the 35 kDa track around the 125-th pixel. In the starting position errors it is observed that with progressively longer detection length the 35 kDa error takes the first drop, followed by the 20 kDa error, and finally the 10 kDa errors. It is noteworthy that non-linear effects can combine to reduce the overall error, as the 50 and 100 kDa tracks demonstrate.

Another unexpected observation is that at full detection length the lower molecular weight bands have roughly the same starting position error, i.e. they form a vertex, if not at the recorded starting coordinates. From these observations it becomes clear that the migration departs from a linear trajectory in a complex manner.

In studying the goodness of fit for standard runs (Section 2.3.1.4), it was determined that although linear tracks represent the migration trajectory of proteins well, they are only an approximation of more complex non-linear migration. It is likely that if only a very small sub-section of the tracks is considered, then linear tracks become a very good

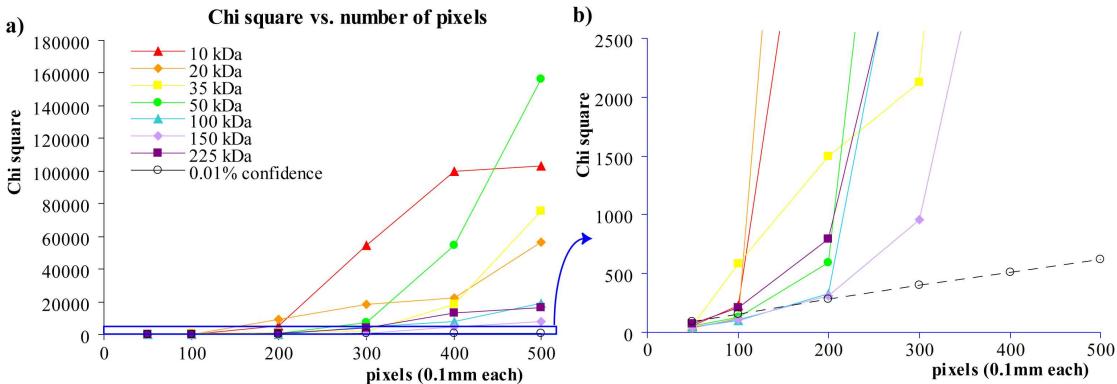


Figure 2.36: Goodness of fits as quantified by the χ^2 values, plotted by number of pixels considered for the track-fitting. a) Full range of χ^2 values. b) Enlarged view of the range up to 2500. In black the χ^2 value that is associated with a 0.01 % confidence level of the fit being correct. Data are from a standard separation (200 mm separation length, 350 V/cm electric field strength, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

fit for that sub-section of migration trajectory. To assess this, the track χ^2 values are compared for data of the same separation, but increasing numbers of pixels are included for the fitting algorithm. Figure 2.36 shows χ^2 values grouped by molecular weight of the protein band. For all tracks, non-linear migration becomes dominant beyond a pixel length of 250. Already at very short pixel lengths the χ^2 value is very high. At 100 pixels half the tracks are outside the 0.01 % confidence limit. Furthermore, it appears that high molecular weight proteins retain a good fit for longer. This is due to greater initial spread of hit coordinates, so the band within which a trajectory is accepted as adhering to the linear fit is wider.

2.3.2.3 Pixel subsets: position of subset

In the preceding study the number of pixels included in the analysis was varied, with all such artificial detection windows starting at the same position, the 0-pixel. A slightly different but similar approach takes always the same number of pixels, but shifts the starting position along the capillary, to capture different sections of the full detection window.

A subset width of 150 pixel was chosen, because the data from the previous experiments indicates that above this the track errors show little further decrease. The starting pixel of the 150-pixel interval was increased in 50-pixel intervals, considering pixels 0–150, then 50–200 and so on. Thresholds were chosen such that all tracks incorporated hits along the full 150-pixel window. Figure 2.37 shows the effect on track errors and starting position errors for different positions of such 150-pixel windows.

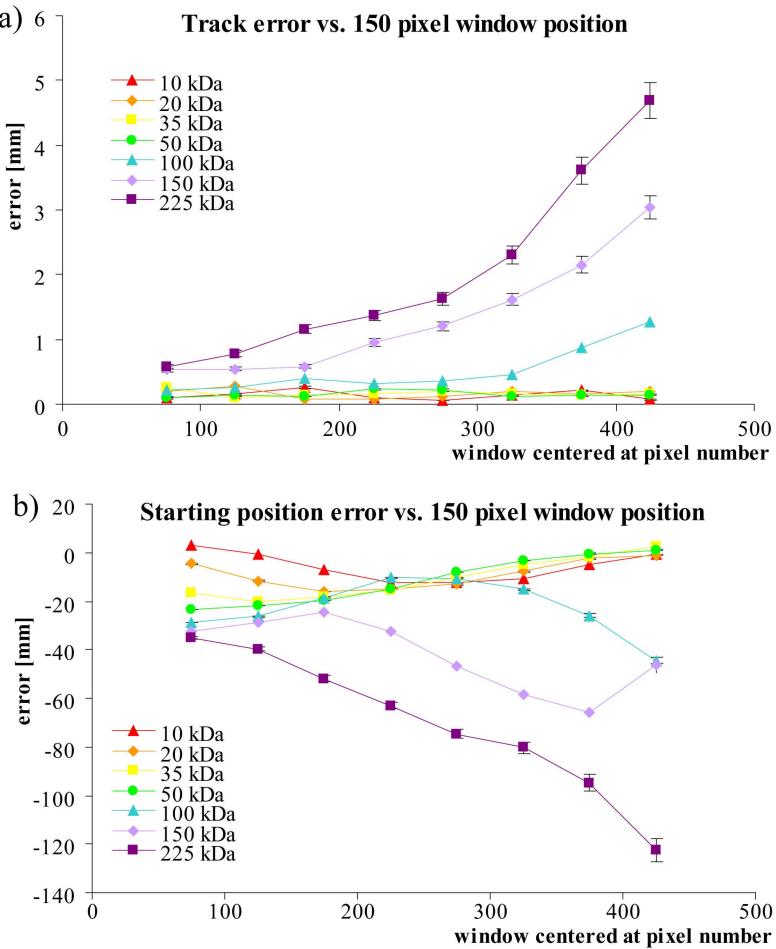


Figure 2.37: Track errors and starting position errors for different 150-pixel data subsets starting at different positions along the capillary. a) Track error vs. position. b) Starting position error vs. detector position. Data were taken from separations performed under standard conditions (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with the standard sample (7 proteins, 10–225 kDa).

Track errors increase as the 150-pixel subgroup moves in the direction of migration (i.e. increasing start pixel number), particularly for high molecular weight proteins. This is presumably due to molecular diffusion and UV effects. As described in Section 2.3.1, electropherogram peaks get wider and flatter across the detection window, especially for high molecular weight bands. This results in a wider spread of hits along a track and consequently greater track errors.

Starting position errors display more complex variations, because of the non-linear migration of proteins across the capillary window. Recalling the sinusoidal residual plot observed for standard separations (Figure 2.25), if only a subset of data points is taken, the fitted track varies depending on data section. This is again reflected if the χ^2 values of the fitted tracks are compared. Figure 2.38 shows χ^2 values of the fitted tracks for 150-pixel data subsets starting at different positions along the capillary.

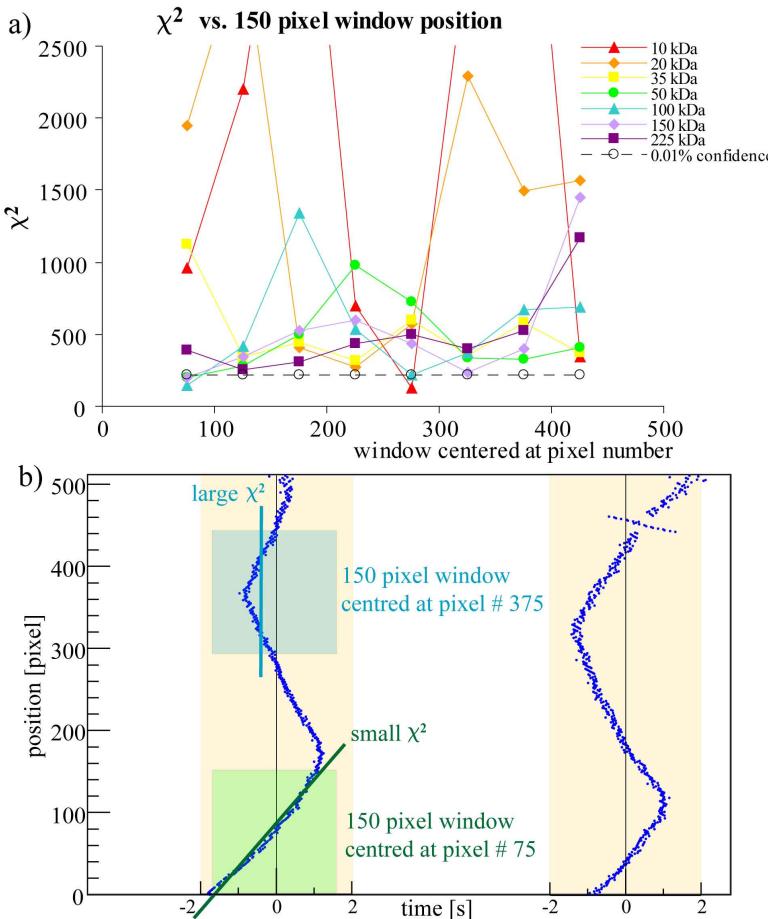


Figure 2.38: Goodness of fit for short data subsets along detection window. a) χ^2 values of fits, plotted for 150-pixel data subsets starting at different positions along the capillary. In black the χ^2 value that is associated with a 0.01 % confidence level of the fit being correct. Only the range of up to 2500 for χ^2 values is shown. b) Residual plot of two migration trajectories (10 kDa and 20 kDa), with tracks for different pixel subsets. Data are from a standard separation (200 mm separation length, 350 V/cm electric field strength, 75 μm ID capillary, Beckman SDS gel buffer), with the standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

In Figure 2.38a the χ^2 values buckle up, depending on the position of the 150-pixel subset. This can be explained in light of the sinusoidal shape seen in residual plots, as illustrated in Figure 2.38b. If the pixel subset happens to include the apex section of the non-linear migration trajectory, then the χ^2 value will be particularly high, since the migration trajectory is in particularly poor agreement with the linear track. Conversely, if the 150-pixel subset falls on a section of the trajectory that happens to be reasonably linear, then the χ^2 value is relatively low. High χ^2 values also coincide with zones with large variation in starting position error values. If the apex section of the non-linear migration trajectory falls in the 150-pixel subset, then the neighbouring subsets will have a rather different section of the trajectory and a different track will be fitted, resulting in a change in starting position error.

Nearly all χ^2 values are above the 0.01 % confidence level, indicating that non-linear migration effects dominate the track error. This is in agreement with the results plotted in Figure 2.36, where it was seen that for short pixel subset lengths (100 pixel), χ^2 values are high. Finally, the χ^2 baseline seems to be increasing slightly as the 150-pixel subset position is shifted along the detection window. This could be explained by a general increase in band width (and hit scatter), either by molecular diffusion, or by UV effects (as discussed in Section 2.4.3).

2.3.3 Effect of different separation conditions

It is well established how various separation conditions such as capillary length and separation voltage influence separation.^[40,181] On the single pixel level, detection is related to peak width and peak height (plate number and signal-to-noise ratio).

Given instrumentation and separation chemistry, the choice of separation conditions balances the following factors: high separation voltages give minimal duration and low molecular diffusion, but also high Joule heating. Long separation distances give high plate numbers, but also high molecular diffusion. Large capillary diameters give high signal-to-noise ratio, but also high Joule heating.

With the help of multipixel detection, it becomes possible to study the effect of these parameters on migration trajectories and migration track reconstruction. For this purpose the sample dilution, separation length, separation voltage, and capillary diameter are systematically varied, and the effect on track reconstruction errors is studied.

2.3.3.1 Sample dilution

Separations with varying sample dilution help to establish how the signal-to-noise ratio affects errors. The stock sample of 16 mg/mL protein was diluted to between 0.2 % and 20 % stock concentration. The 20 % dilution (3.2 mg/mL) corresponds to the standard conditions as defined in Section 2.2.4.

In Figure 2.39 the starting position errors and track errors are shown. The track errors initially remain constant with decreasing concentration. Beyond a certain threshold concentration, track errors begin to increase. At very low concentrations the track ceases to be detectable. Large proteins are already near their detection threshold at 20 % sample concentration.

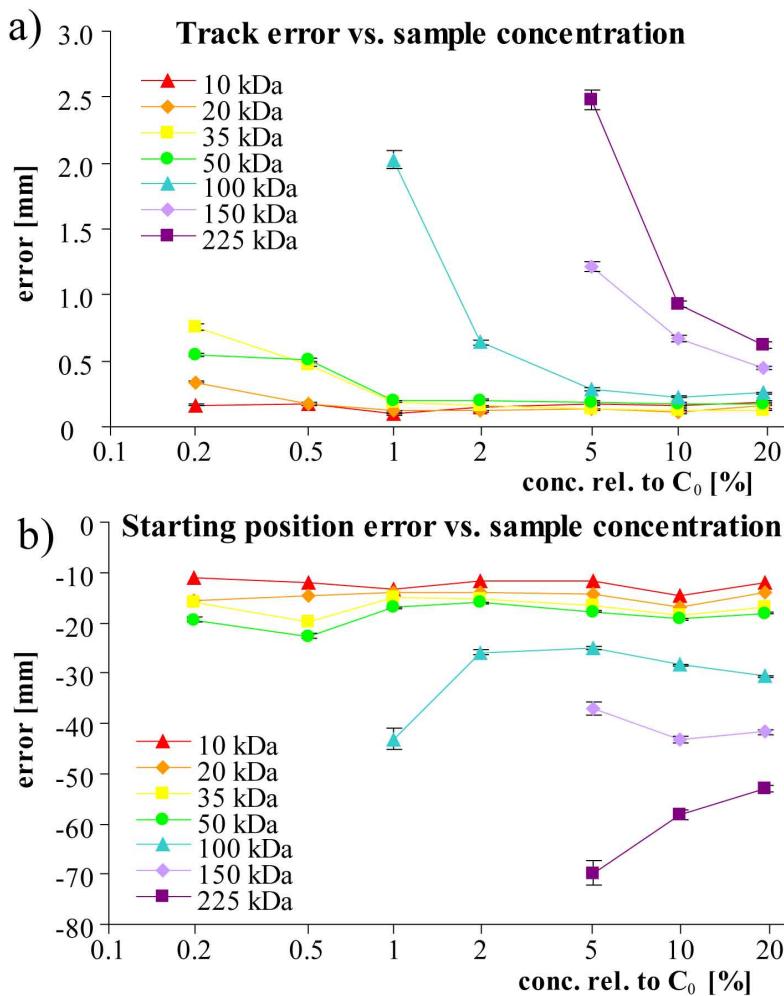


Figure 2.39: Track errors and starting position errors at different concentrations of sample. a) Track error vs. sample concentration, grouped by track molecular weight. b) Starting position error vs. sample concentration, grouped by track molecular weight. Sample concentration is given relative to stock concentration (16 mg/mL). All other separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with the standard sample (7 proteins, 10–225 kDa).

The starting position error follows largely the same pattern. If the track error is much greater than 1 mm, then detection errors appear to bear an influence on the starting position errors.

2.3.3.2 Separation length

A next set of experiments is dedicated to varying the separation length of capillaries. This is the distance between capillary inlet and the first pixel of the detector. The total capillary length is 15 cm longer than the separation length, 5.12 cm for the detection window and slightly less than 10 cm between the last pixel and the capillary outlet.

The separation length was varied from the standard 20 cm to lengths between 10 and 35 cm, with all other separation conditions standard. Capillaries shorter than 10 cm are difficult to handle, whereas separation lengths over 35 cm need voltages above 20 kV to maintain the standard separation field, and are prone to arcing. In Figure 2.40 the track errors and starting position errors can be seen.

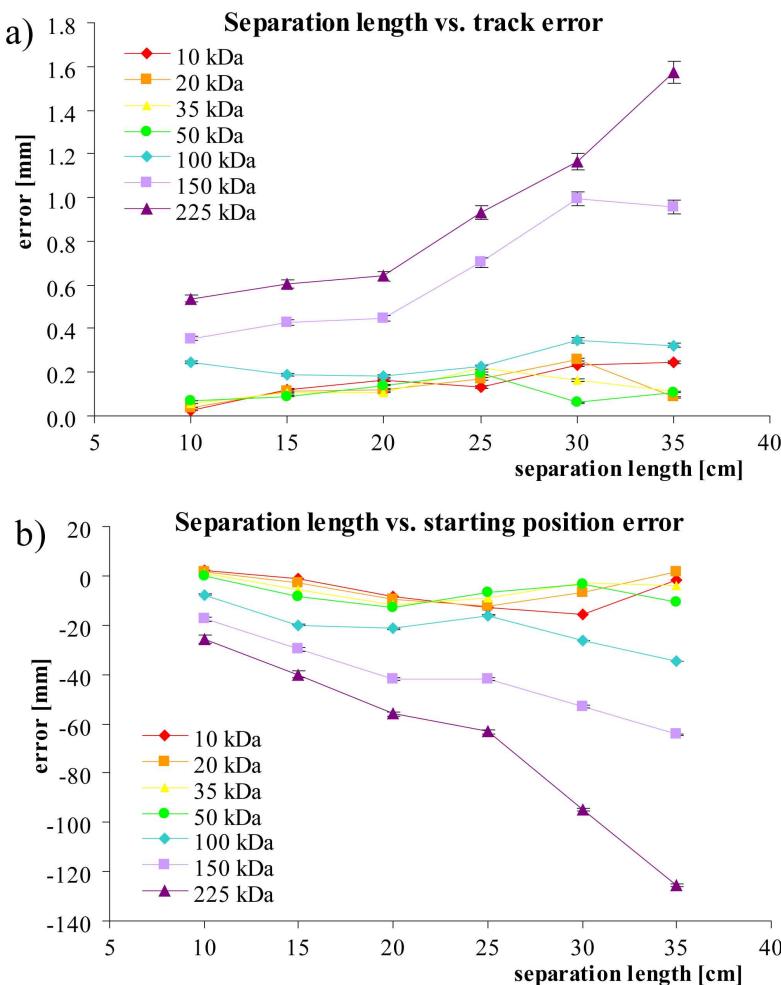


Figure 2.40: Track errors and starting position errors at different separation lengths.
a) Track error vs. separation length, grouped by track molecular weight. b) Starting position error vs. separation length, grouped by track molecular weight. All other separation conditions are standard (350 V/cm electric field strength, 75 μ m ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

The track errors increase steadily with longer separation lengths. It must be noted that the error values σ_a and σ_b that belong to the track fit parameters **a** and **b** do not change very much. Here, the increasing track error is explained by the error propagation distance. Equation 2.16 from Section 2.1.3 ($\sigma_p^2 = \frac{\sigma_a^2}{b^2} + (\frac{a-t}{b^2})^2 \sigma_b^2$) describes how the slope and intercept errors combine for an arbitrary migration distance. If σ_a and σ_b vary only little with increasing separation length, but the $(a - t)$ term (which is the migration time) increases, then the overall effect is increasing track errors.

The higher molecular weight proteins are expected to show an extra increase in track error, because those peaks are in the range where detection errors become relevant, as illustrated in the dilution series above. Molecular diffusion scales with migration time, so longer separation distances will cause an increase in track error especially for 150–225 kDa proteins.

The starting position errors follow the same trend and increase with increasing separation distance. Some variations are seen, in particular the lower molecular weight errors display a staggered local error minimum for certain lengths. The cause for this may be found in the way counterband interactions affect band migration. In Figure 2.41 the residual plots are shown for different separation lengths.

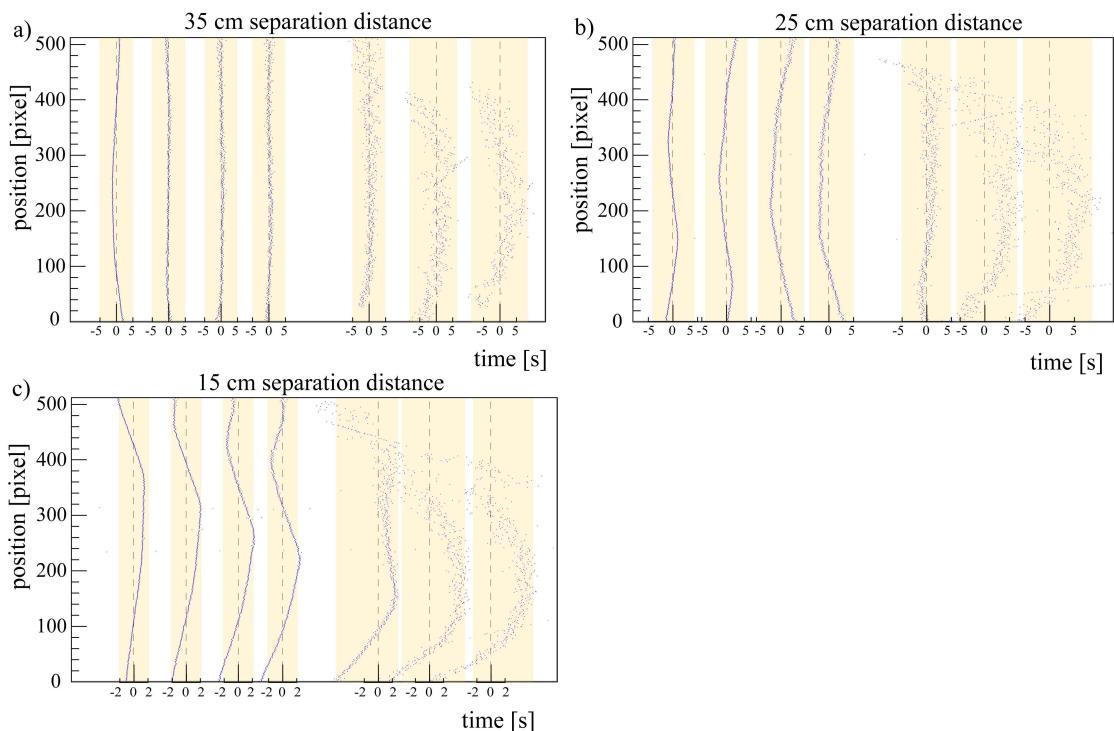


Figure 2.41: Residual plots for different separation lengths. a) Residual plot of seven tracks from separation over 35 cm capillary. b) Residual plot of seven tracks from separation over 25 cm capillary. c) Residual plot of seven tracks from separation over 15 cm capillary. All other separation conditions are standard (350 V/cm electric field strength, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

It seems that the first counterband intersects with tracks at different points in their trajectory, depending on separation length (35 cm, 25 cm, 15 cm). This is to be expected. Because the capillary length from last pixel to outlet remains unchanged, the counterbands should always pass the detector at the same times, whereas the protein bands pass the detector earlier for shorter separation length and later for a longer separation length.

As observed from studying detection length, non-linear migration from counterband interactions can counteract and correct for other non-linear effects. Depending on how much of a migration kink is included in the track-fitting data, the errors can unexpectedly be reduced (such as for the 50 kDa band for 30 cm separation length). Considering the entire molecular weight spectrum, it is safe to conclude that short separation lengths help reduce errors (at the cost of resolution).

2.3.3.3 Separation voltage

The separation voltage affects migration in two ways. Higher field strengths mean faster separations with less molecular diffusion, as will be demonstrated. This competes with increasing Joule heating with higher separation voltage, which can affect migration.^[181,200,201]

Equations 1.20 and 1.23 from Section 1.3.2 are used to express the migration time relationship to the electric field strength:

$$t = \frac{L_s}{\mu E} \quad (2.18)$$

where t [s] is the migration time, μ [$\text{m}^2/(\text{sV})$] the mobility, E [V/m] the electric field strength, and L_s [m] the separation length. The diffusion width of bands is given by the Einstein equation (1.20, Section 1.3.2):

$$\sigma_D^2 = 2Dt \quad (2.19)$$

where σ_D^2 [m^2] is the zone variance, σ_D [m] the diffusion width, and D [m^2/s] is the diffusion coefficient. With Equations 2.18 and 2.19 the diffusion width as a function of the electric field strength becomes:

$$\sigma_D = \sqrt{2D \frac{L_s}{\mu E}} \quad (2.20)$$

Equation 2.20 shows that with increasing separation field strength, the diffusion width decreases.

Increasing the separation voltage also increases the Joule heating. Considering only the thermal power input from Joule heating in the capillary:

$$P = \dot{Q} = IV \quad (2.21)$$

with Equation 2.18, $I = V/R$, and the resistance $R = \frac{\rho L_c}{A}$ we obtain:

$$dQ = IV dt = \frac{\mu IV^2}{L_s L_c} = \frac{\mu V^3 A}{\rho L_s L_c^2} \quad (2.22)$$

where Q [J] is the energy contained within the volume, dt [s] the migration time, I [A] is the current, V [V] the voltage, μ [$\text{m}^2/(\text{sV})$] the mobility, L_s [m] the separation length, L_c [m] the capillary length, ρ [Ωm] the specific resistance of the buffer, and A [m^2] the cross sectional area. The fact that the separation voltage occurs to the third power in Equation 2.22 indicates how influential a parameter it is.

In Figure 2.42 the track errors and starting position errors for separation voltages between 300 and 500 V/cm are shown. To improve data quality, each run was performed

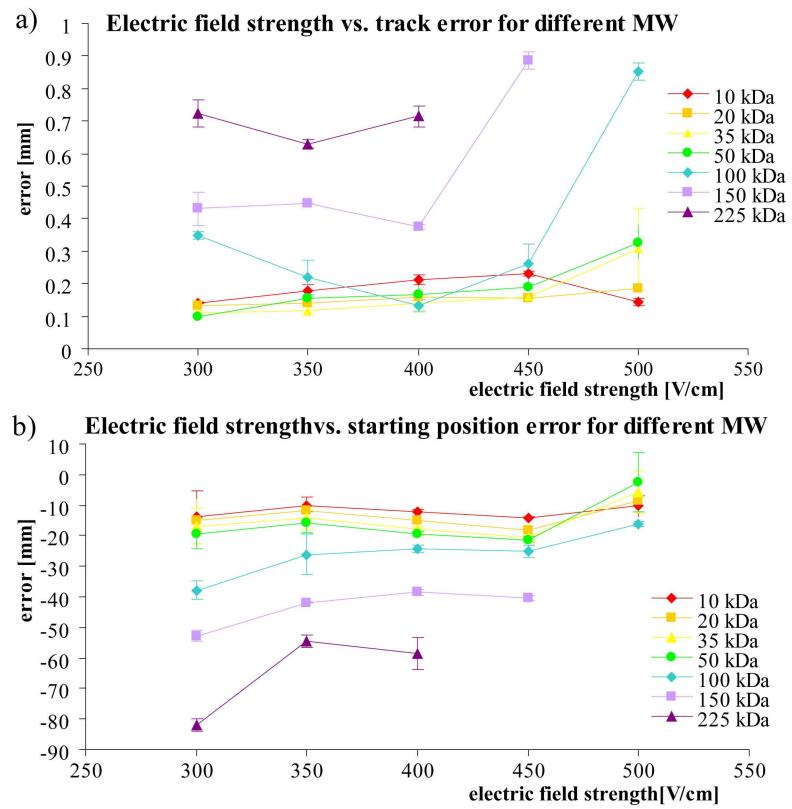


Figure 2.42: Track errors and starting position errors at different electric field strengths. a) Track error vs. electric field strength, grouped by track molecular weight. b) Starting position error vs. electric field strength, grouped by track molecular weight. All other separation conditions are standard (200 mm separation length, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa).

twice, and error values were averaged. With increasing separation voltage, improved detection due to faster migration and reduced diffusion (Equation 2.18) competes with greater non-linearities in migration from Joule heating (Equation 2.22).

Regarding the track errors, high error values are observed for very low voltages and very high voltages, with a minimum in between. An explanation for this is that diffusion and detection errors dominate for low voltages, while heating effects dominate for higher voltages, particularly for proteins over 100 kDa.

Regarding starting position errors, non-linear migration errors from thermal effects are expected to increase with increasing voltage. In Figure 2.42b it can be seen that this does not describe the behaviour very well. On the contrary, starting position errors tend to decrease slightly with increasing voltage. It appears that errors from non-linear migration due to Joule heating play a subordinate role. Instead, the starting position errors reflect the behaviour of track errors.

In Figure 2.43 the track errors plotted against the starting position errors demonstrates that there is a voltage dependence for errors. This means that for low voltages, the starting position error is well correlated to the track error. Conversely, for high voltages starting position errors do not depend strongly on track errors.

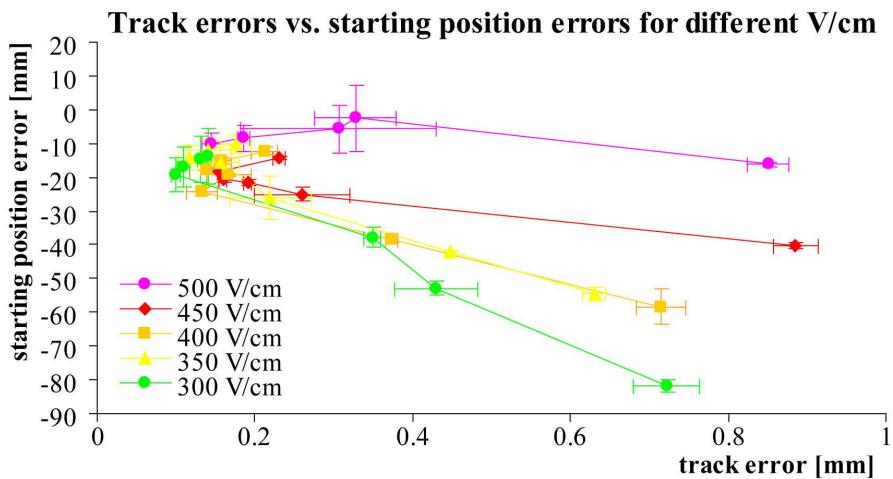


Figure 2.43: Track errors against starting position errors at different electric field strengths. All other separation conditions are standard (200 mm separation length, 75 μ m ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL)

A comparison of residual plots shows that counterbands do not intersect protein bands at the same position coordinate for different voltages. Figure 2.44 shows residual plots for separations with 300, 400, and 500 V/cm electric field strength. In the 500 V/cm separation, counterbands do not make their typical appearance. This is possibly because

they have already passed, which is further supported by the wide distribution of hits compared to the usual tight trajectories.

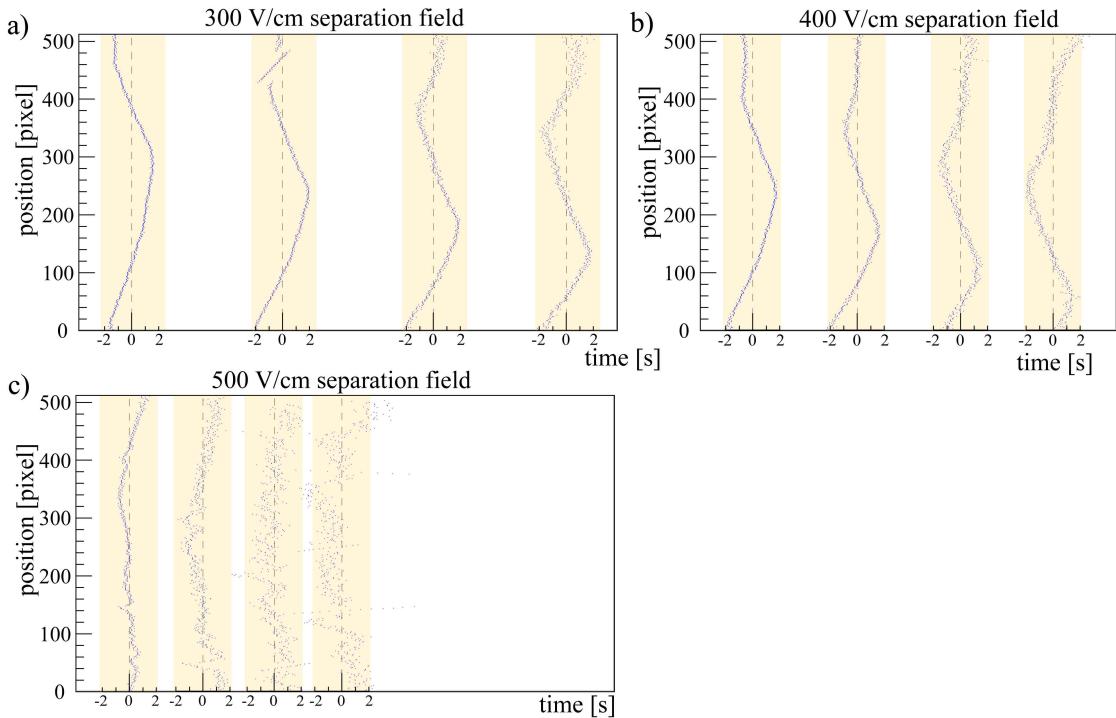


Figure 2.44: Residual plots for different electric field strengths. a) Residual plot of first four tracks from separation with 300 V/cm electric field strength. b) Residual plot of first four tracks from separation with 400 V/cm electric field strength. c) Residual plot of first four tracks from separation with 500 V/cm electric field strength. All other separation conditions are standard (200 mm separation length, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL)

2.3.3.4 Capillary diameter

Another experimental parameter that is closely connected to thermal effects is capillary diameter. As seen from Equation 2.22, the energy deposited in the capillary during a separation increases with current, which for the same buffer scales with area, or (diameter)². Hence large bore capillaries are unfavourable due to significant Joule heating. On the other hand, detection depth scales with diameter, so signal strength improves with diameter. This is counterbalanced to some degree by the lensing effects in smaller capillaries, i.e. with the greater curvature light is bundled into the capillary, which improves signal strength again.^[70,71]

Capillaries of the same outer diameter (all 375 μm) with an inner diameter (ID) between 25 and 250 μm were used for otherwise standard separations. Figure 2.45 shows track errors and starting position errors.

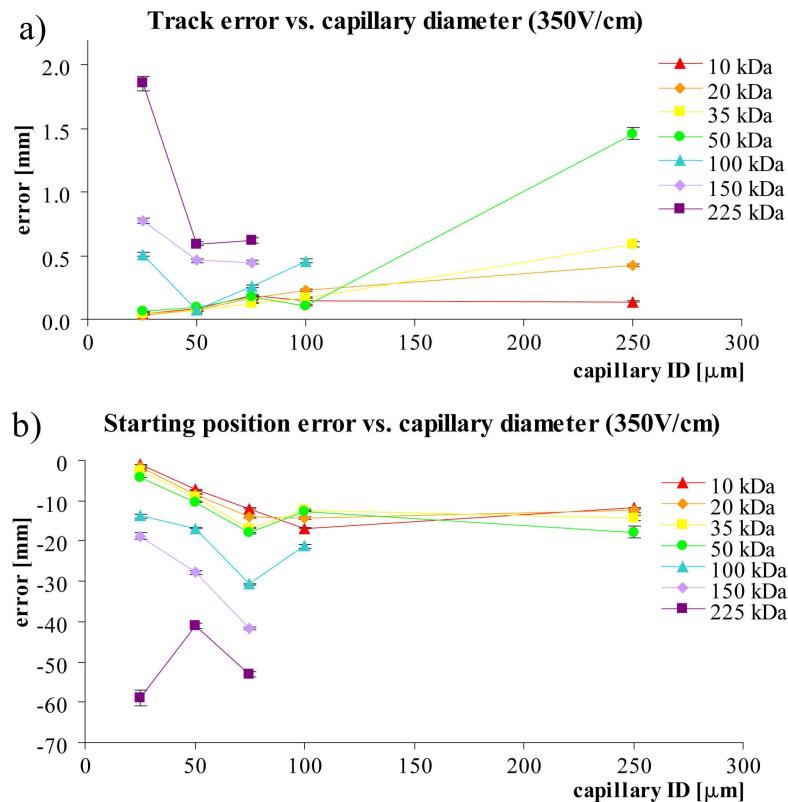


Figure 2.45: Track errors and starting position errors for different capillary diameters. a) Track errors for capillaries from 25–250 μm inner diameter, grouped by molecular weight. b) Starting position errors for capillaries from 25–250 μm inner diameter, grouped by molecular weight. All separations are performed under otherwise standard separation conditions (200 mm separation length, 350 V/cm electric field strength, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

Track errors decrease from 25 to 50 μm ID, and increase again over 100 μm ID. Especially large proteins, which already in a standard run are approaching the detection threshold, are affected by poorer detection, explaining the high track errors for 25 μm ID. For an ID of 100 μm and over, proteins that are larger than 150 kDa are again difficult to detect, this time presumably due to thermal dispersion.

The increase in current from 25 to 50 μm ID is 4-fold, and again 4-fold from 50 to 100 μm ID. At 250 μm ID the current (for a given voltage and therefore field) would be 100-fold what it was at 25 μm ID, and some compromises regarding voltage were necessary because the power supply limits current to 380 μA. Therefore the error values are on the conservative side—with even higher power the already considerable heating would be far more severe.

The track errors for 250 μm ID become very great, influencing slower proteins most. This appears to be a molecular diffusion effect, as the starting position errors do not increase nearly as much—indicating that here non-linear migration is less of an important factor than molecular band diffusion.

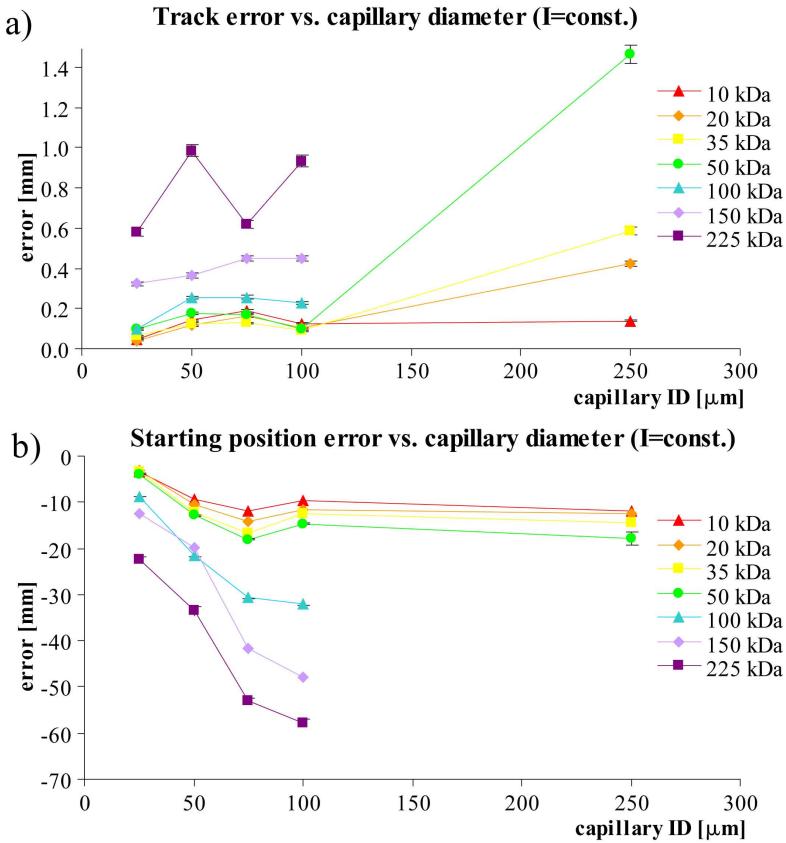


Figure 2.46: Track errors and starting position errors for different capillary diameters under constant current conditions. a) Track errors for capillaries from 25–250 μm inner diameter, grouped by molecular weight. b) Starting position errors for capillaries from 25–250 μm inner diameter, grouped by molecular weight. Separation voltages are adjusted to obtain the same electrophoretic current (90 μA).

Studying the starting position errors between 25 and 100 μm ID, it seems that the general trend is an increase in error (i.e. more non-linear migration) with increasing diameter. It is noteworthy that for the smallest capillary, 25 μm ID, the lower molecular weight bands (≤ 50 kDa) have only very small errors of 1–4 mm and very nearly form a vertex at the recorded injection coordinates. This indicates that non-linear migration due to thermal effects is the primary cause for the starting position errors observed in low molecular weight proteins.

For comparison Figure 2.46 shows the same series of experiments, this time with constant current and varying electric field strength. Here track errors remain in the same range, while starting position errors steadily increase. This indicates that errors due to thermal non-linear migration are quickly dominated by other error sources, including peak detection and UV effects.

2.3.4 Effect of gel formulation

As specified in the experimental details, the standard separation buffer is a SDS-MW gel buffer for protein CGE (Beckman Coulter Inc., ProteomeLab, proprietary polymer buffer formulation, pH 8.0, 0.2 % SDS). Since BioRad have discontinued their CGE buffer product, Beckman ProteomeLab CGE buffer is the most widespread commercial product. Despite this, alternative protein CGE buffers have been reported.^[202–208] The only alternative commercial buffer found was ‘Optima Protein Run Buffer for Molecular Weight Standard’ from Groton Biosystems.^[209] Buffer development for capillary gel electrophoresis is a discipline in its own, and buffer behaviour is a critical factor in separation performance.

To get an idea of the impact different buffers have on the track reconstruction of a separation, separations under same conditions with different buffers are compared. Buffers compared are Beckman ProteomeLab buffer, Groton Optima buffer, and two custom prepared dextran-based buffers. Dextran based buffers contain 10 % dextran 2,000 kDa, 1.2 M tris-borate, 0.2 % SDS and optionally 10 % glycerol (all from Sigma-Aldrich Co.). Because the composition of both commercial products is proprietary, very little is known about their formulation. Beckman buffer contains 0.2 % SDS, presumably in tris-HCl (the same as in the sample buffer), while Groton buffer is assumed to be tris-borate.

Groton buffer is the least viscous of the four buffers and has the lowest conductivity. The other three buffers have comparable viscosity and conductivity.

A first observation is that not all buffers have similar electropherograms, but in particular custom prepared buffers show poor peak shape, especially for high molecular weight proteins, see Figure 2.47. Plausible causes of this behaviour are absorption of proteins to capillary walls and poor control of electro-osmotic flow.

Next it is observed that the sieving strength (quantified with the slope in molecular weight–mobility plot) of the Groton buffer is significantly lower than the other buffers. Figure 2.48 shows the logarithm of the mobility plotted against the logarithm of the molecular weight, and is useful for distinguishing different sieving mechanisms. Figure 2.48b shows the logarithm of the molecular weight plotted against the inverse mobility and is useful for calibration. Both show that the separation of the dextran and dextran-glycerol buffers is similar to that of Beckman buffer. Groton buffer on the other hand shows weaker sieving. Under the same conditions protein resolution is different for each buffer. This was already seen in the electropherograms in Figure 2.47b.

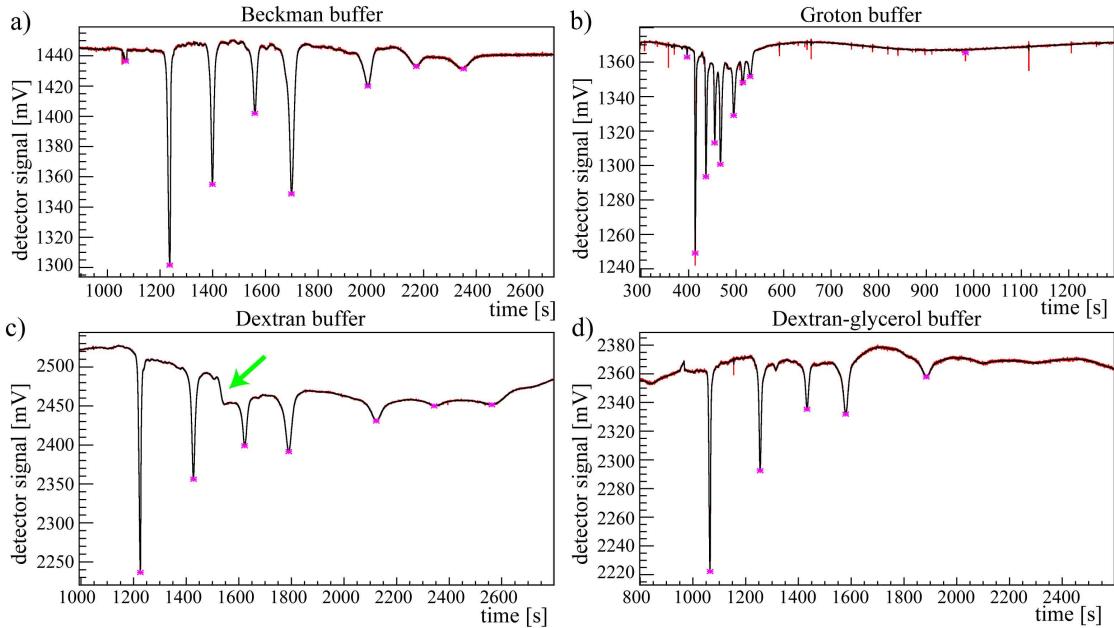


Figure 2.47: Electropherograms for separations with different buffers. a) Beckman buffer, b) Groton buffer, c) dextran buffer, d) dextran-glycerol buffer. The green arrow indicates a particularly pronounced step from background interference. For all separations standard conditions apply (200 mm separation length, $E = 350 \text{ V/cm}$, 75 μm ID capillary), with standard sample (7 proteins, 10–225 kDa). Electropherograms are from pixel number 250.

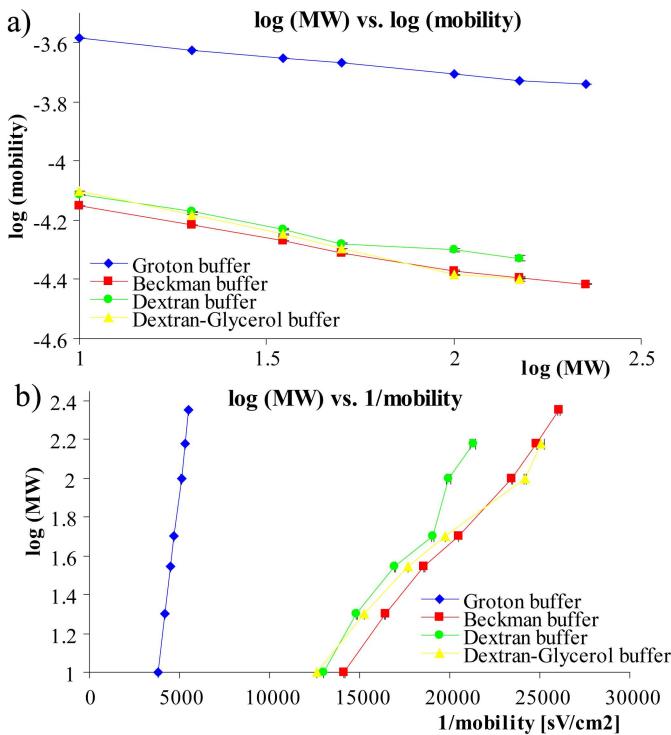


Figure 2.48: Mobility-molecular weight calibration plots for different buffers. a) $\log(\text{mobility})$ plotted against $\log(\text{molecular weight})$ for Beckman, Groton, dextran, and dextran-glycerol buffer. b) $\log(\text{molecular weight})$ plotted against $(1/\text{mobility})$ for Beckman, Groton, dextran, and dextran-glycerol buffer. All separations were performed under standard conditions (200 mm separation length, 350 V/cm electric field strength, 75 μm ID capillary), with standard sample (7 proteins, 10–225 kDa).

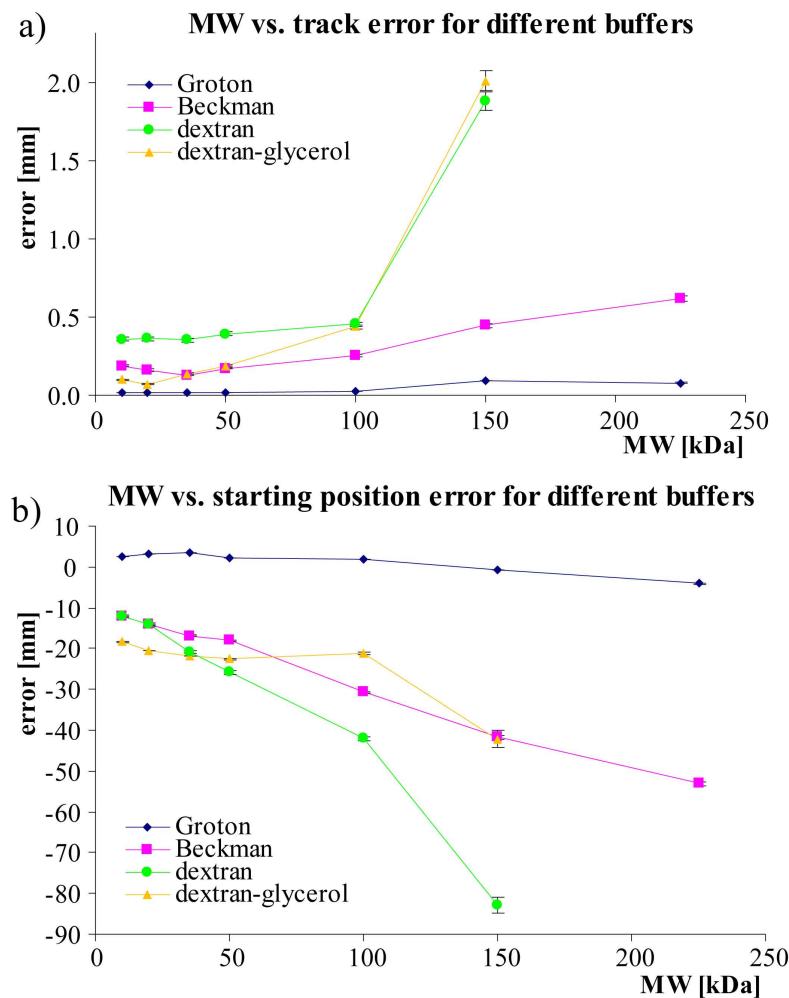


Figure 2.49: Track errors and starting position errors for different buffers. a) Track error plotted against molecular weight for Beckman, Groton, dextran, and dextran-glycerol buffers. b) Starting position error plotted against molecular weight for Beckman, Groton, dextran, and dextran-glycerol buffers. All separations are performed under standard separation conditions (200 mm separation length, 350 V/cm electric field strength, 75 μ m ID capillary), with standard sample (7 proteins, 10–225 kDa).

In Figure 2.49 the track errors and starting position errors for the different buffers, again under standard separation conditions, are seen.

In Figure 2.49a the track errors for both dextran buffers are large, indicating peak deterioration, especially for proteins over 100 kDa. This is not due to higher electrical current, as conductivity for dextran buffers is similar to Beckman buffer. Addition of glycerol achieves a slight improvement, especially in the lower molecular weight range. Glycerol and other alcohols have been reported to have beneficial effects on surface control in capillary electrophoresis.^[40,210] Notably the Groton buffer has very small track errors (all remain below 0.1 mm). This is in part because migration time and molecular diffusion are minimal.

Starting position errors reflect the same trends as track errors. Again Groton buffer displays very small starting position errors. Tracks lie within 5 mm from the recorded start coordinate.

In plain dextran buffer, background bands in both directions interfere with migration. This effect is reduced with the addition of 10 % glycerol. This difference of background bands explains the difference in track errors between the two dextran-based buffers. The effect of background bands is seen in the residual plot in Figure 2.50, taken for separations under standard conditions. For the custom prepared dextran buffer the background step can easily be seen in the electropherogram, indicated with an arrow in Figure 2.47.

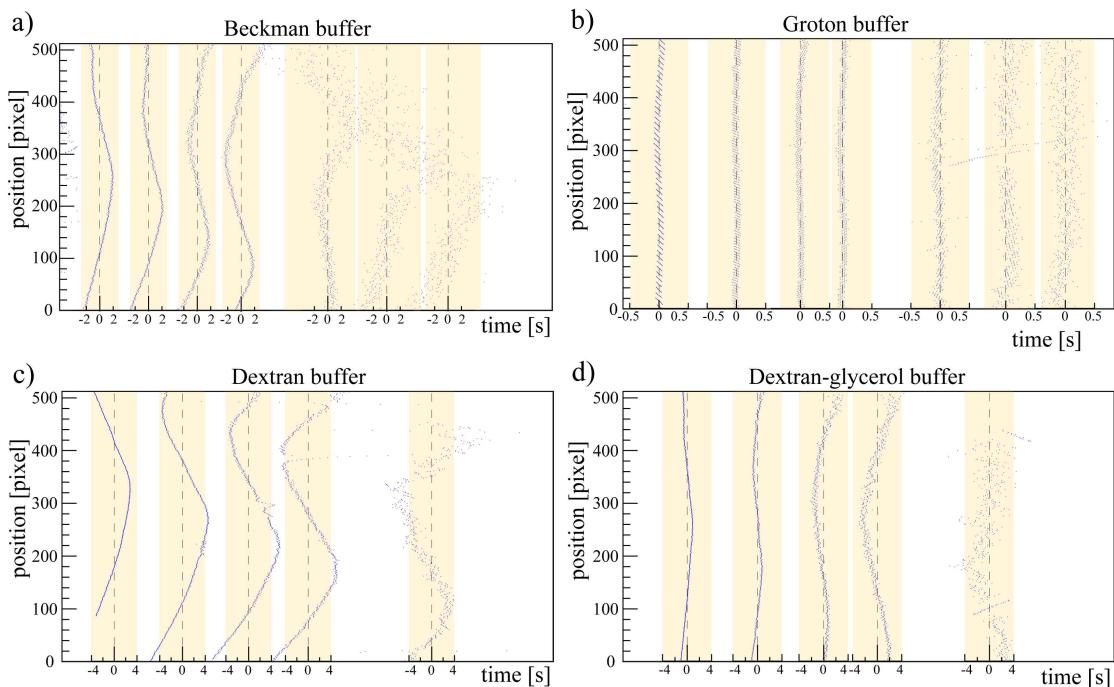


Figure 2.50: Residual plots of separations with different buffers. a) Beckman buffer, b) Groton buffer, c) dextran buffer, d) dextran-glycerol buffer. The standard separation conditions apply.

To compare peak shape development as peaks migrate from the first to the last pixel, the separation voltages are adjusted so that the migration speeds are same. In Figure 2.51 the peak asymmetry (half width at half maximum toward front and back of a peak) in different buffers is compared.

In dextran buffers, in particular without addition of glycerol, the peak shape is irregular and HWHM plots appear jagged. The large spikes in the backend (black) occur when a background band has passed through the peak, and the baseline changes. This changes the maximum and the width of the peak, and the way the algorithm works the backend is more strongly affected. Observing the magnitude of HWHM, it appears that proteins

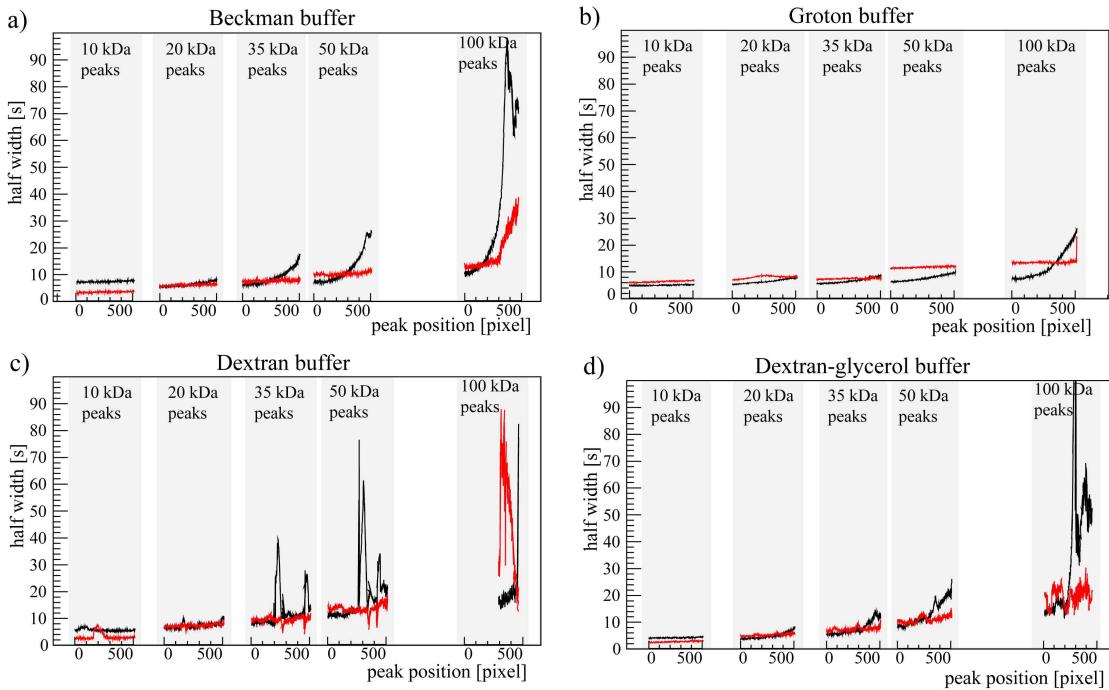


Figure 2.51: Electropherogram peak shapes along detector (HWHM plot) for different buffers. Black indicates the faster, leading flank of a peak, and red the slower, following flank. a) Beckman buffer, b) Groton buffer, c) dextran buffer, d) dextran-glycerol buffer. Only the first five tracks are included, because in dextran buffers the last two peaks are not detected. Separation voltages are adjusted to ensure comparable migration speed and UV exposure.

in Beckman buffer suffer the severest UV degradation, and proteins in Groton buffer are least affected.

Overall, custom prepared buffers have similar performance as Beckman buffer in the 10–50 kDa range, but performance falls off above 50 kDa. Groton buffer displays favourable behaviour for track reconstruction, i.e. little effect of non-linear migration, but has the drawback that sieving strength is comparatively low, and small molecular weight differences might not be resolved. For molecular weights over 100 kDa, where non-linear migration burdens tracking for the Beckman buffer, the Groton buffer seems to be a useful alternative.

2.3.5 Effect of band crossings and multiple injections

All experiments discussed so far are based on migration tracks that intersect at the common starting coordinates, but otherwise do not interact. In a V2D separation, protein bands migrate through one another in the course of a separation, as faster proteins overtake slow proteins that started at a position farther along the pH gradient. Therefore studying tracks that do not cross might produce results that are not representative of real V2D separations.

To introduce a certain degree of band interaction and crossing, a series of experiments was performed where sample is injected twice at different times, allowing fast proteins from the second injection to migrate through slow proteins bands of the first injection. This gives an approximation of the impact of band interactions that occur in a V2D separation. Tracks can be assigned to an injection, and the number of track crossings recorded.

If band interactions occur and affect protein migration, then the migration trajectory will be influenced by hitherto unaccounted for non-linear migration. In this case the track fitted from the observed migration section will be particularly ill-fitted and the discrepancy between recorded injection coordinate and reconstructed starting coordinate will reflect this. Another consideration is that if bands are very close together in the detection section of the capillary, then the fitted track can be influenced by the proximity of other peaks, similar to the interference of bubbles. Both of these effects, band intersection and band distinction, have an impact on performance of a V2D separation.

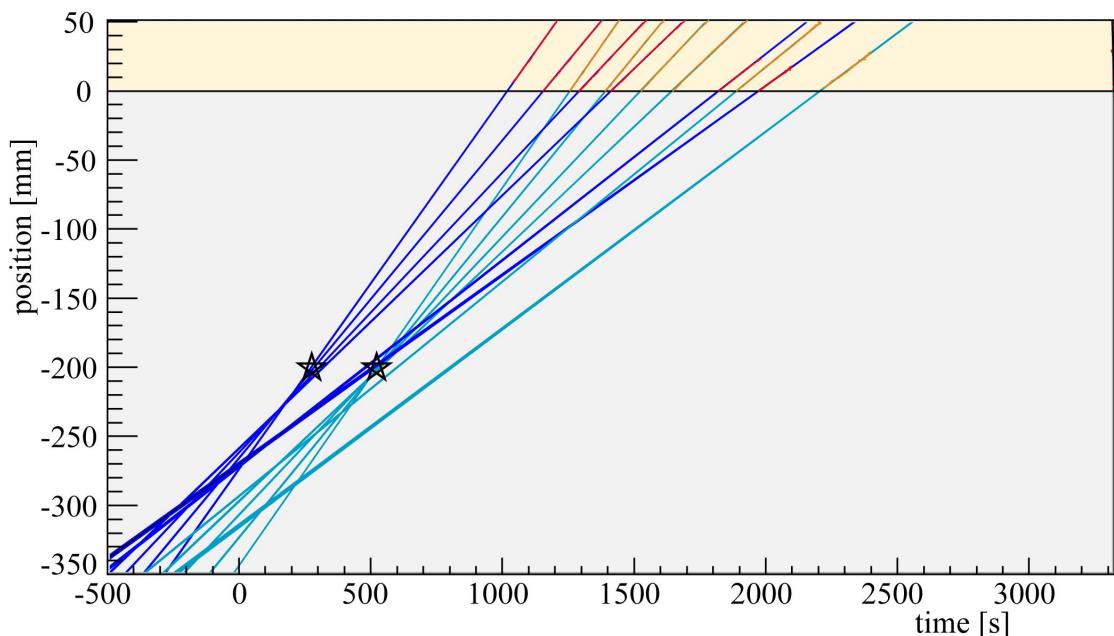


Figure 2.52: Plot of tracks for a double injection separation. Sample (7 proteins, 10–225 kDa, 3.2 mg/mL) was injected twice with an interval of 4 minutes. For illustration purposes tracks of the second injection are plotted in a lighter shade. The star symbol indicates the recorded injection coordinates.

In Figure 2.52 the tracks are shown for a standard double injection. Only 12 of the expected 14 tracks are displayed. The remaining two tracks, from the last band of each separation, are below the hit-finding threshold for the chosen sensitivity setting. In the region of the injection coordinates two different vertices can be distinguished, and tracks can be assigned to either the first or second injection. Also, the number of

band crossings can be deduced, whereby only protein bands are considered (including undetected bands, but not for instance the sample ion fronts).

An initial series of investigations is dedicated to establish whether there are any systematic effects that change the separation from the first to the second injection. The data from seven double injection separations are compared in Figure 2.53. In Figure 2.53a, the track parameters for all tracks are overlaid. No tendency can be discerned for the two different injections, and the variations appear to fall within the ordinary run-to-run variations. To examine in closer detail how tracks change from the first to the second injections, Figure 2.53b and 2.53c show the relative changes in track parameters for

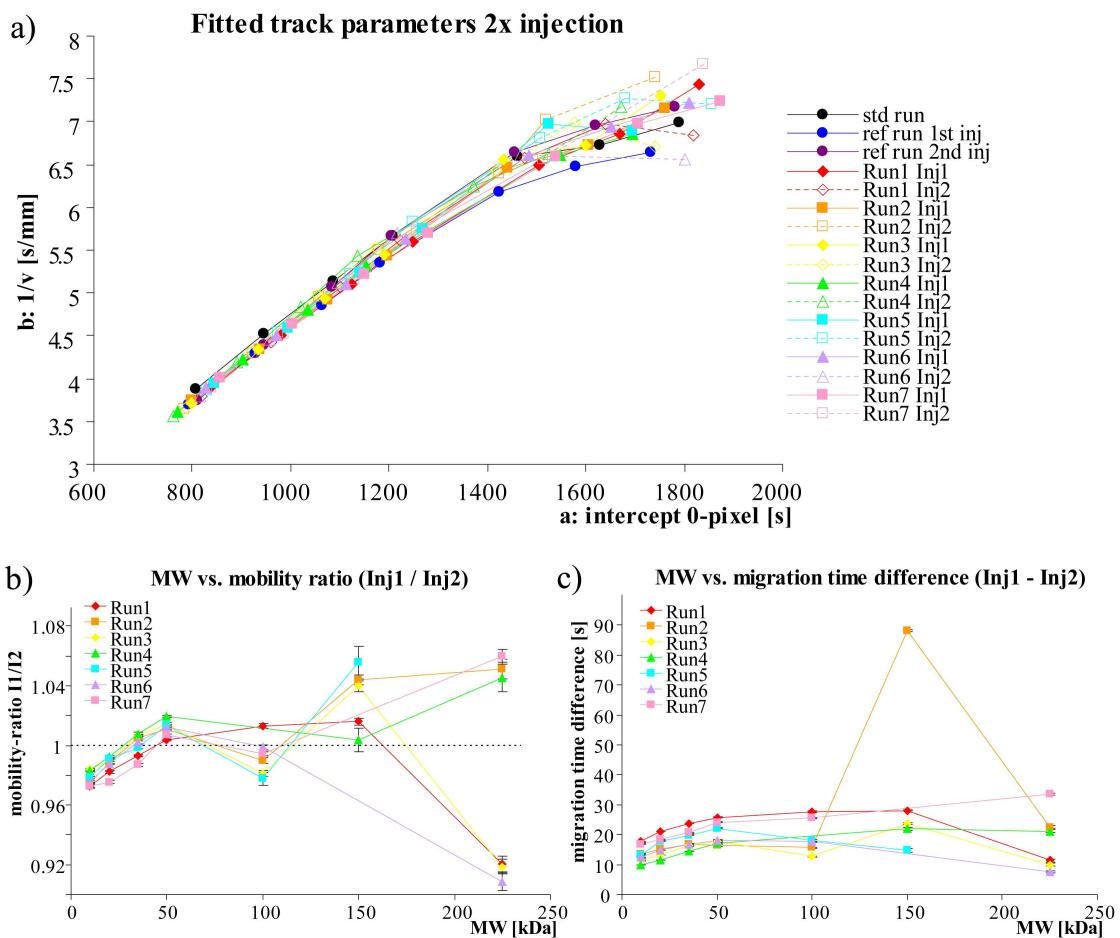


Figure 2.53: Track fit parameters for seven repeat double injection separations.

a) Track intercept **a** and slope **b** (from track equation $\text{time} = \mathbf{a} + \mathbf{b} \times \text{position}$ with **a**—track intercept, or migration time [s], and **b**—track slope, or inverse migration speed [s/mm]), grouped by run. b) Quotient of mobilities. Within the same separation, the mobility of a protein from the first injection is divided by the mobility of the corresponding protein from the second injection. c) Difference of migration times. Within the same separation, the migration time of a protein from the second injection is subtracted from the migration time of the corresponding protein from the first injection. Sample (7 proteins, 10–225 kDa, 3.2 mg/mL) was injected twice with an interval of 4 minutes under otherwise standard conditions.

tracks of the same run, always comparing two corresponding tracks belonging to the same proteins.

In Figure 2.53b the quotient of mobilities (track slope) is shown, with values smaller than one if the mobility of the second band is faster than the first band (proteins from second injection appear accelerated). For the first four proteins, 10–50 kDa, this quotient steadily increases from below unity to slightly above. For proteins above 50 kDa, no trend can be discerned. This indicates that the effects dictating the observed changes between first and second injection are more complex than for instance simple heating.

Figure 2.53c shows how the migration times change from the first to the second separation. Bands from the second injection arrive at the detector always slightly faster than the same band from the first separation. An explanation for this is that the instrument gradually heats up, causing systematically quicker migration in proteins of the second injection. The smallest four proteins show a trend toward gradually increasing differences, but the proteins larger than 50 kDa behave more irregularly. Possibly the high molecular weight proteins, which have more poorly-defined peaks, suffer the compound effects of more than one error source (background migration effects, UV effects, capillary surface modification, or other uncharacterised effects). Superimposition of different effects could cause the observed randomisation of behaviour. The low molecular weight peaks have well-defined shapes, and would only suffer strong effects.

2.3.5.1 Track errors

In Figure 2.54 four graphs with track errors for bands with different molecular weight are shown. Figure 2.54a displays the track errors for seven double injections, grouped by run, with full symbols for tracks of the first injection and open symbols/dashed lines for the second injection. Figure 2.54b shows the same data, but this time grouped by the number of crossings, with open symbols/dashed lines where tracks intersect in the detection window. Figure 2.54c shows the track errors of 12 standard runs, for comparison.

Compared with standard separations, the track errors for double injection separations appear to vary more. Track errors of the same run do not seem to correlate very well, nor do tracks of the second injection show consistent behaviour relative to tracks from the first injection. On closer inspection one might discern a slight trend toward higher track errors for the second injection.

Within a 95.8 % confidence interval the track errors for the second injection are greater than the track errors for the first injection (paired one-sided Students t-test). This is not

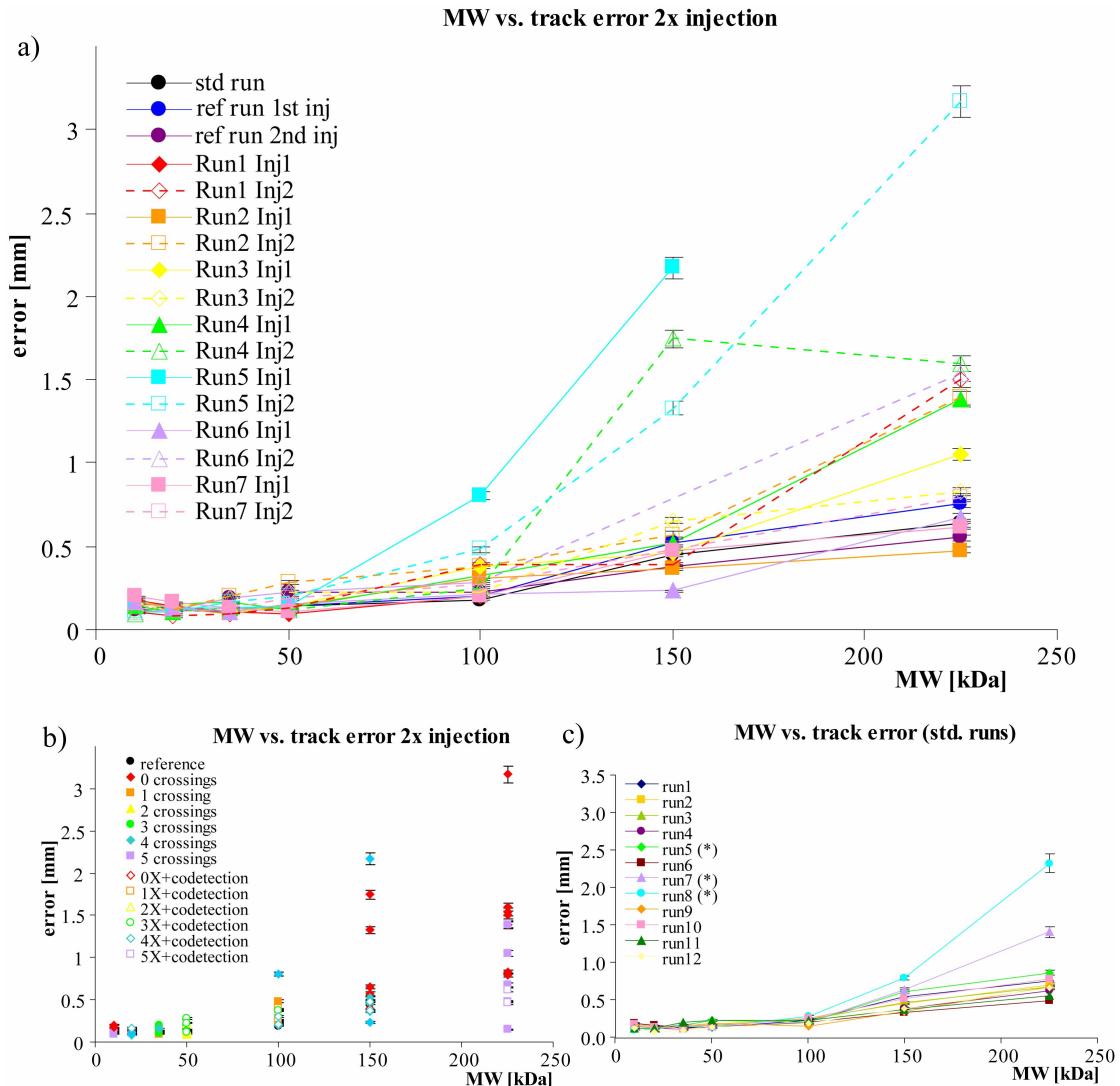


Figure 2.54: Track errors plotted against molecular weight for 7 double injection runs. a) Grouped by run with full symbols for first injection and open symbols/dashed lines for second injection. b) Grouped by crossing number with open symbols for band co-detection. c) Enlarged section of graph a) with track errors plotted against molecular weight for double injection runs. d) Data from standard single injection runs for comparison. Sample (7 proteins, 10–225 kDa, 3.2 mg/mL) was injected twice with an interval of 4 minutes under otherwise standard conditions.

due to longer migration times and greater molecular diffusion for tracks of the second injection. On the contrary, the migration times of the second injection are systematically shorter than for the first injection (as illustrated in Figure 2.53c). Tentative explanations might be that as time progresses the capillary suffers protein surface adsorption, and EOF and electric field strength uniformity are poorer, resulting in increased band broadening; or that the separation environment has heated up, from Joule heating or UV absorption, causing faster molecular diffusion and wider bands.

With regards to the track errors for different numbers of crossings in Figure 2.54b, it is more difficult to discern a pattern. There is a natural correlation between molecular

weight and crossing number, i.e. the last protein band of the first injection always crosses with the five fastest proteins in the second injection, whereas the last protein band of the second injection does not intersect any other bands (as can be seen from Figure 2.52). If anything, it appears that the track error is greater for tracks with the fewest crossings. This could be simply the same effect as the observation that tracks of the second injection have greater errors.

To study the effect of crossing number, proteins are grouped by molecular weight. The confidence level that the population of (paired) error values differs between the first injection compared to the second injection is evaluated with a Student's t-test (one sided, paired).

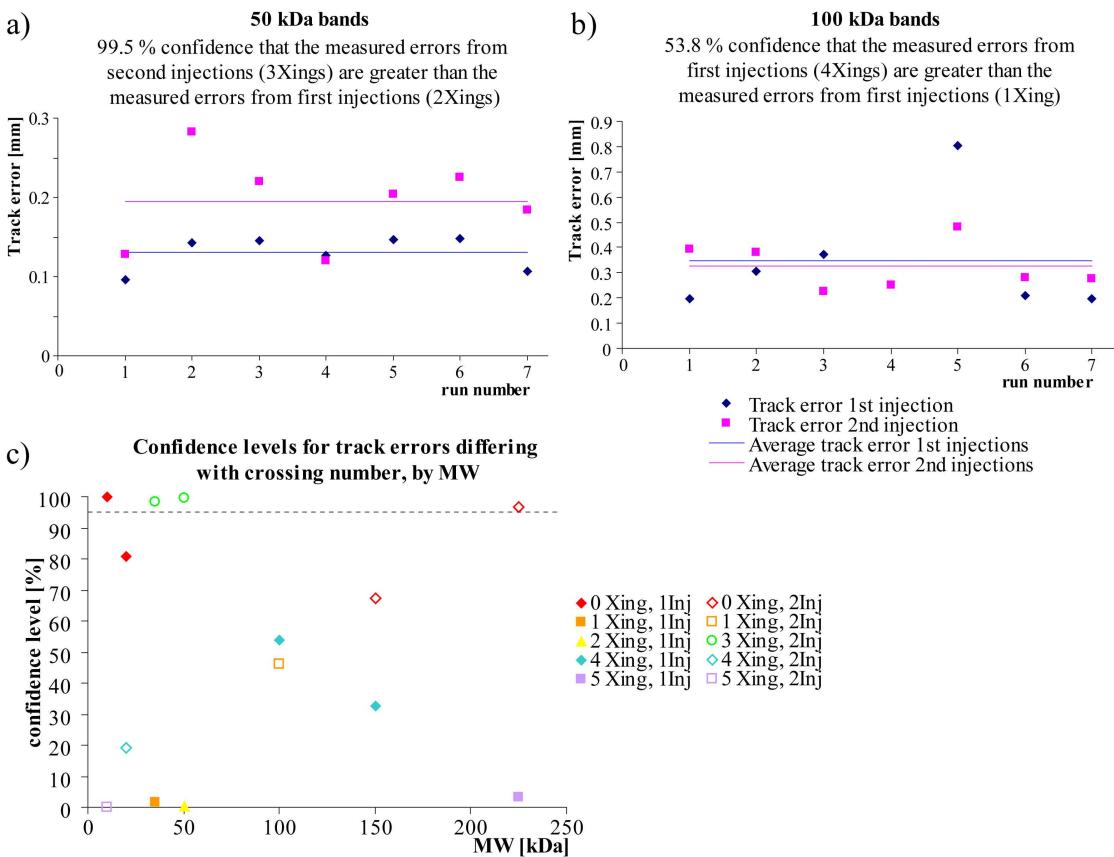


Figure 2.55: Confidence levels for track errors differing with crossing number, grouped by molecular weight. a) Track errors of the 50 kDa proteins for the first and second injection in 7 independent runs. b) Track errors of the 100 kDa proteins for the first and second injection in 7 independent runs. c) Summary of confidence levels for all the different molecular weights. The legend indicates which track error group is associated with the confidence value of greater track errors. Full symbols indicate proteins from the first injection, and open symbols indicate proteins from the second injection; the colour indicates number of crossings (red—zero crossings, violet—five crossings).

To illustrate the test procedure Figures 2.55a and b show the track errors of the 50 kDa proteins and the 100 kDa proteins for the first and second injection, in seven independent

runs. The number of crossings each set of proteins undergoes is also noted. For the 50 kDa proteins, the track errors from the second injection (with three crossings) are greater than the track errors from the first injection (with two crossings) with good confidence. For the 100 kDa proteins, the track errors from the second injection (with four crossings) are not distinguishable from the track errors from the first injection (with one crossing).

In Figure 2.55c the results from all seven sample proteins are summarised. The legend indicates which group (first or second injection), and the number of crossings associated with that group. The tracks associated with the first injections are shown with a full symbol, from the second injection with an open symbol. Low crossing numbers are shown in red/orange, high crossings in blue/green symbols.

The aim of this analysis is to identify whether errors from the second injection are consistently greater, or whether errors with high crossing numbers are consistently greater. Values below 95 % confidence (dashed line in Figure 2.55c) are disregarded because the difference between the two sets of errors is not significant enough. Figure 2.55c shows how ambivalent the track errors are with regards to differences depending on either injection number or number of crossings.

It is equally convincing that tracks with fewest crossings tend to have greatest errors (tendency toward red/orange symbols in high-confidence zone), and that tracks from the second injection have greatest errors (tendency toward open symbols in high-confidence zone). Neither statement is consistently observed, and what is more, the statements are in conflict with one another. We therefore conclude that any correlation between crossing number and track error is very weak, and band crossings do not significantly influence track errors.

One final track attribute is drawn into the considerations in the search for an explanation for greater track error variance in double injections. In nearly every run there are peaks that run together in a part of the detection window. These are always peaks of different injections, and in most cases the two peaks separate from one another in the detection window and allow two distinct tracks to be laid. In the band merging region, it is to be expected that the hit allocation to the tracks is influenced by the peaks merging. To estimate the impact this has on track reconstruction, track errors are studied for bands that co-migrate in the detection window, depicted with open symbols in Figure 2.54b. Here merged bands do not seem to have any characteristic behaviour, and if there is a difference between merged tracks and free-standing tracks, there are not enough data to indicate this.

To summarise, although it is observed that track errors show greater variance with double injections than for standard separations, this is not well correlated with number of crossings, nor injection number, nor track co-detection.

2.3.5.2 Starting position errors

Because the starting position errors are indicative of non-linear migration, studying starting position errors is expected to give some answers to the question whether band crossing affects protein migration. In Figure 2.56 the starting position errors for the same seven double injection separations as above are shown. Again, sample was injected twice with an interval of 4 minutes, and the errors that are associated with a band are either grouped by run (Figure 2.56a, full symbols for the first injection, open symbols for the second injection) or by crossing number and co-detection (Figure 2.56b). For

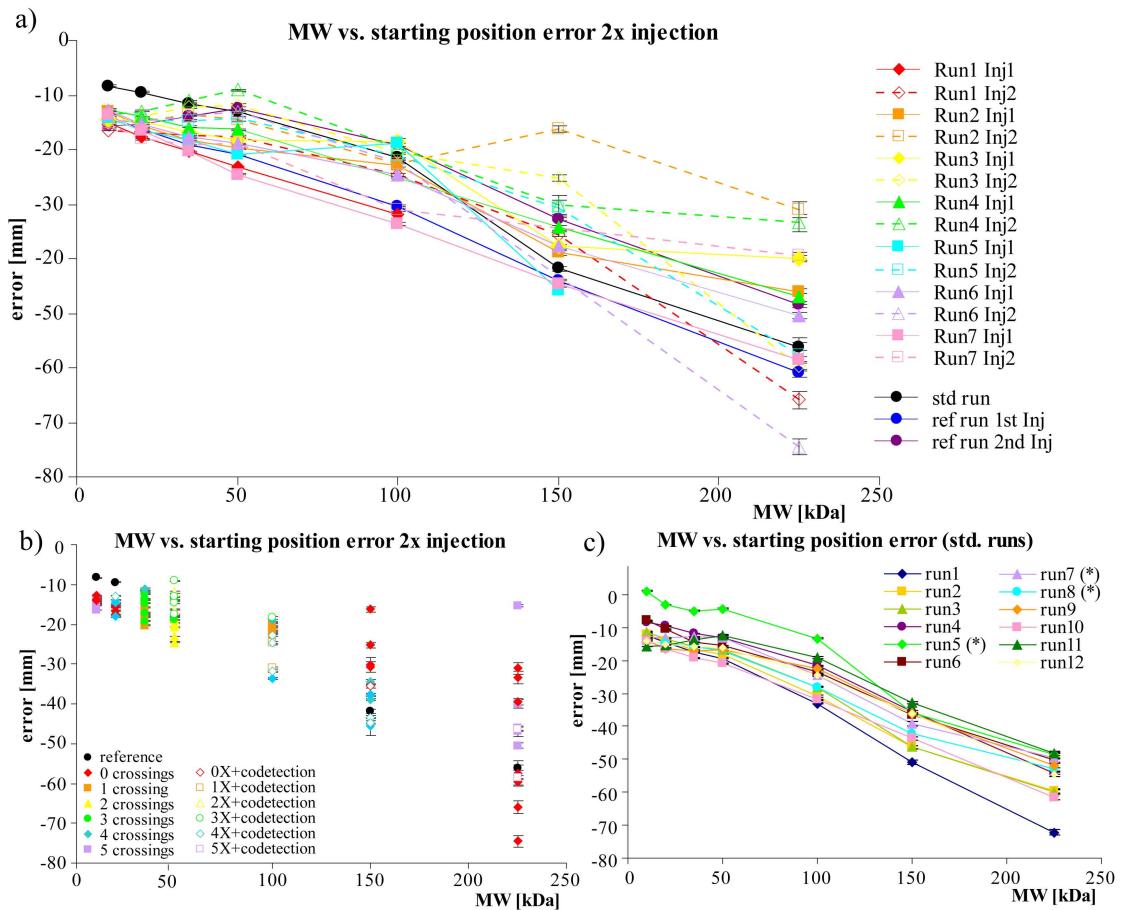


Figure 2.56: Starting position errors plotted against molecular weight for seven double injection runs. a) Grouped by run with full symbols/solid lines for the first injection and open symbols/dashed lines for the second injection. b) Grouped by crossing number with open symbols for band co-detection. c) Data from standard single injection runs for comparison.

comparison the starting position errors of 12 standard (single injection) runs are shown in Figure 2.56c.

Similar to the track errors discussed above, a first observation is that the variance in starting position errors of double injection separations is greater than for standard separations, especially for high molecular weight, slow migration proteins. On close inspection there appears to be a trend toward greater starting position errors associated with the first injection. This is confirmed with a paired Students t-test; the confidence level is 97.6 %. This is different from the situation with the track errors, where the errors associated with the second injection are greater.

As with the track errors discussed above, the starting position error behaviour with regards to crossing number (Figure 2.56b) is not easily described. Again the starting position error groups are tested for differences depending on crossings with a t-test. In Figure 2.57 the confidence levels for differences between data groups are shown, with the symbols indicating the set with the greater starting position error. Disregarding

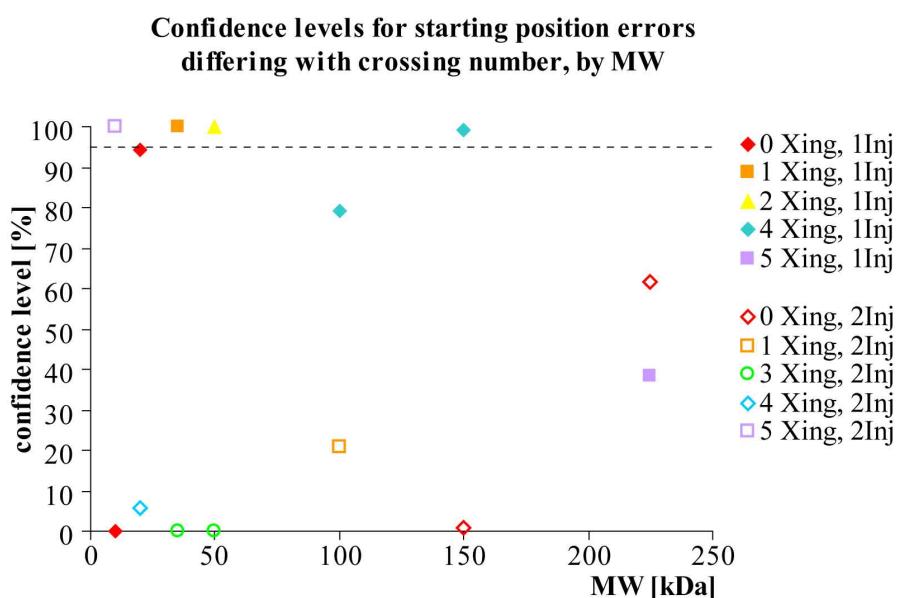


Figure 2.57: Confidence levels for starting position errors differing by crossing number, grouped by molecular weight. This shows the confidence level that proteins with a particular molecular weight from the first injections (with a given number of crossings) have greater starting position errors than the same proteins of the second injection (with a different number of crossings). The legend indicates which starting position error group is associated with the confidence value of greater starting position errors. Full symbols indicate proteins from the first injection, and open symbols indicate proteins from the second injection; the colour indicates number of crossings (red—zero crossings, violet—five crossings).

confidence levels below 95 %, it is observed that there might be a slight trend toward greater errors for higher crossing numbers. For extreme molecular weights where significant differences are observed (10, 150 kDa), the set with most crossings has the greater

error (five vs. zero crossings, respectively four vs. zero crossings). In the mid-region of molecular weight (35, 50 kDa), the sets with fewer crossings show greatest starting position errors (one vs. three, respectively one vs. two crossings), but here the differences between crossing numbers of the two sets are fairly small. In these sets although there is a significant difference between data groups, the group average values are very near to one another (-18 mm vs. -15 mm, respectively -20 mm vs. -14 mm average starting position error). Our interpretation of these findings is that there is indeed a tendency toward greater errors with increasing number of band crossings, but the differences are small and this is not consistently true for all molecular weights.

Band overlapping and co-detection effects (open symbols in Figure 2.56b) do not show any characteristic pattern. Only in one instance, where bands co-migrate along the majority of the detection window and the track finding algorithm returns only one track instead of two, does this show any effect. This is the outlier value in Figure 2.56b with a very low starting position error of the 50 kDa protein. Overall, it is observed that although starting position errors vary more for double injection separations, the variation does not depend simply on crossing effects or detection effects, but behaves in a more complex manner.

Though the difficulties of correlating errors with a measurable property are unfavourable, a calibration is conceivable if starting position errors correlate with track errors. In Figure 2.58 the starting position errors are plotted against track errors, grouped either by run (Figure 2.58a), or by crossing number (Figure 2.58b), with standard run data for comparison (Figure 2.58c). Here the same trends as already discussed are seen again. Values from first and second injections form distinct clusters, but no obvious dependence on crossing can be discerned. Without prior knowledge of a protein band's starting coordinates, it is not possible to deduce an estimate for starting position errors for a given track error, and correction of a reconstructed starting coordinate will fail.

Data from 94 tracks were analysed, with 18 tracks displaying an overlap in the detection window, under conditions such that tracks intersect between zero and five times. Of the maximum 98 tracks 4 were undetectable, in two cases due to poor signal, in one case due to prevalent co-migration with another band and in the last case due to bubble interference.

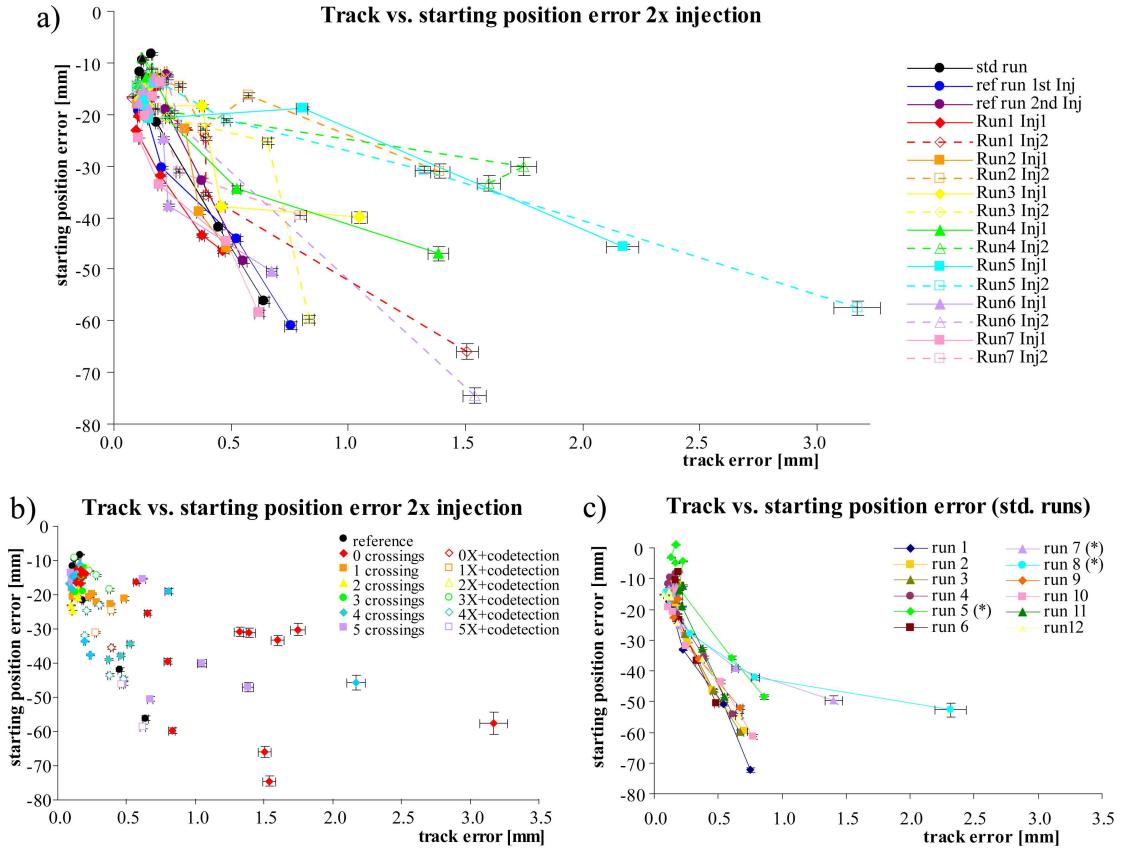


Figure 2.58: Starting position errors are plotted against track errors for seven double injection runs. a) Grouped by run with full symbols for first injection and open symbols/dashed lines for second injection. b) Grouped by crossing number with open symbols for band co-detection. c) data from standard single injection runs for comparison.

From the data presented it appears that greater migration variations are introduced with band crossings. It is observed that tracks from the second injection show a slight tendency toward greater track errors (i.e. poorer hit-track alignment), perhaps from gradual degradation of capillary conditions. It is also observed that increasing band crossings show a slight tendency toward greater starting position errors (i.e. more non-linear migration effects).

Despite these observations the variations in migration remain difficult to predict. This might be a problem of experimental setup, in so much as the second injection interrupts the first injection, and although the recorded starting coordinates are corrected for this, it might introduce some unwanted effects such as residual electric field from capacitive behaviour. Another consideration is that only protein bands were given attention, while for instance the ion front from the sample buffer of the second injection might affect migration.

2.4 Discussion

The experiments and data presented above allow the conclusion that the greatest challenge in tracking protein migration is that proteins do not move on a straight-line trajectory in the time-space coordinate system, but are subject to non-linear migration. This is illustrated in the schematic in Figure 2.59. In multipixel detection, the momentary migration speed in the detection window is evaluated. In the case of non-linear migration, the momentary speed varies from the average speed (as evaluated in single point detector setups). The migration speed at the beginning of the separation, v_0 , is not observed, nor is the majority of the migration trajectory.

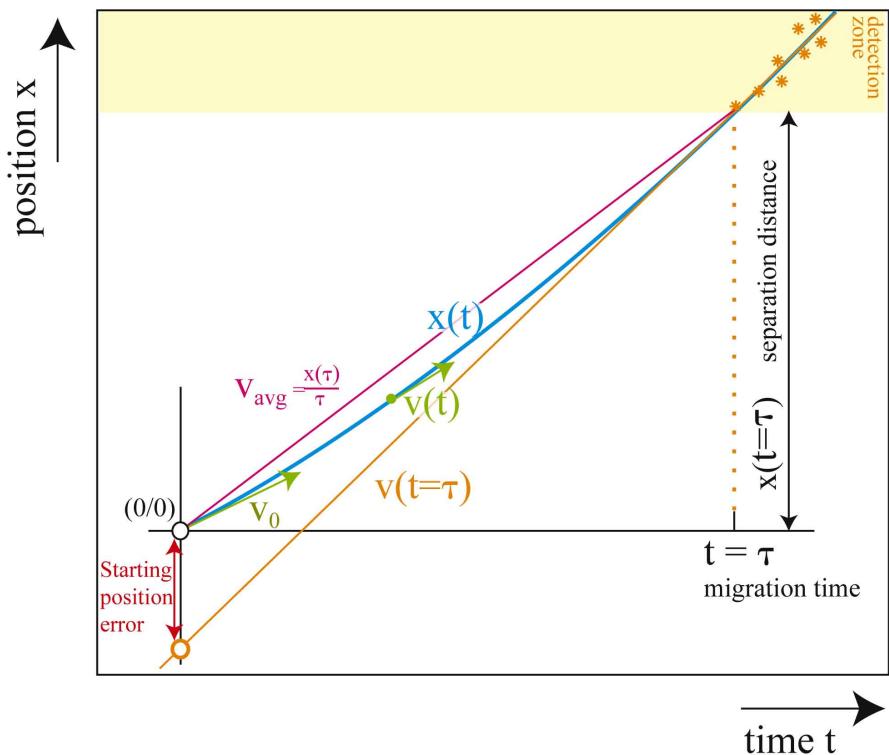


Figure 2.59: Schematic of a migration trajectory $x(t)$ with initial (v_0), average (v_{avg}), and observed ($v_{t=\tau}$) migration speeds, as well as the reconstructed starting point from the fitted migration track (orange).

Various causes for non-linear migration, with different characteristic behaviour, have been distinguished. Thermal effects make proteins accelerate as the separation progresses, and affects protein migration depending on migration time. Counterbands interfere with the migration trajectory, causing deviations from a straight line. UV related effects occur in the detection window, making bands migrate quicker as they progress. Finally, protein band crossing is thought to influence migration trajectories.

In the following discussion the findings of the experiments will be used to develop a model for the different categories of non-linear migration. The causes for non-linear migration will be characterised and quantified. The findings will be summarised to find optimal conditions for a V2D separation for maximum accuracy, and estimate performance limitations. The data from double injection separations will be used to construct a mock-up of a V2D separation and use various calibrations to correct for known error sources.

2.4.1 Errors due to thermal effects

In the course of separations a gradual increase in electric current is observed. This is attributed to an increasing temperature in the capillary (as electric conductivity changes with temperature). Temperature affects CGE separations in different ways: molecular diffusion is temperature dependent; medium viscosity is temperature dependent; and sieving in a polymer network is affected by temperature.^[187] Manifold publications study the axial and radial temperature distribution in a capillary during electrophoresis, with particular focus on peak broadening effects.^[200,201,211–217] The time-dependent temperature development has received little attention. Because conventional detection does not have the means to measure and compare the migration speed at the point of detection with the average migration speed (as illustrated in Figure 2.59), there were no grounds to expect non-linear migration.

To assess whether a gradual temperature increase in the capillary could account for the observed discrepancy between recorded migration starting coordinates and reconstructed migration trajectory, a mathematical model for time dependence of temperature is presented for different cases. With this and a model for temperature dependence of mobility, the resulting deviation from linear migration will be estimated and compared to the observed values.

The temperature increases due to Joule heating, with the generated power P [W]

$$P = VI = \frac{V^2}{R} = \frac{V^2 A}{\rho L} \quad (2.23)$$

for voltage V [V], current I [A], resistance R [Ω], capillary cross sectional area A [m^2], specific resistance of the buffer ρ [Ωm], and capillary length L [m]. The sieving medium requires a minimal buffering power, which puts an upper limit on resistance. To develop an analytical model of errors due to changes in temperature, the following equations are used to find the time-dependent temperature $T(t)$ [K] in the capillary.

Joule heating:

$$P = \frac{\partial Q}{\partial t} = \dot{Q}_{HV} = VI \quad (2.24)$$

Newton's law of unforced cooling:

$$\frac{\partial Q}{\partial t} = \dot{Q}_{Co} = kS(T - T_\infty) \quad (2.25)$$

Heat capacity Q_{Ca} [J]:

$$Q_{Ca} = mcT \quad (2.26)$$

$$\dot{Q}_{Ca} = mc\dot{T} \quad (2.27)$$

with energy Q [J], time t [s], heat transfer coefficient k [$\text{W}/(\text{m}^2\text{K})$], surface area across which heat is transferred S [m^2], environment temperature T_∞ [K], unit mass m [kg], and specific heat capacity (per unit mass) c [$\text{J}/(\text{kgK})$]. Under the assumption that the temperature in the capillary is uniform throughout the capillary, and that the Joule heating is constant, the balance of energies gives:

$$\dot{Q}_{HV} - \dot{Q}_{Co} = \dot{Q}_{Ca} \quad (2.28)$$

leading to the differential equation for $T(t)$:

$$mc\dot{T} + kST - kST_\infty - VI = 0 \quad (2.29)$$

With $T(t = 0) = T_\infty$ as the boundary condition, this solves for:

$$T(t) = T_\infty + \frac{VI}{\alpha} \left(1 - e^{-\frac{\alpha}{\beta}t} \right) \quad \text{with } \alpha = kS \quad \text{and } \beta = mc \quad (2.30)$$

With appropriate coefficients (see Appendix A for derivation):

$$T(t) = 22 + 11.897 \left(1 - e^{-1.404t} \right) \quad (2.31)$$

Equation 2.31 describes a temperature that increases gradually until it converges to a steady-state value. In Equation 2.30, the $\frac{\alpha}{\beta}$ term describes how fast the steady-state temperature is reached. From thermodynamic tables the $\frac{\alpha}{\beta}$ term for the given system is calculated to be in the range of 1.4 s^{-1} (see Appendix A for details). With this rate constant, the capillary is fractions of a degree away from its steady-state temperature within

seconds. As this time scale is much smaller than the separation time scale (30–40 min), this model describes constant temperature of the buffer after a very short initial heating up phase.

Two approximations are employed to derive this result: the current I is assumed constant, and the environment temperature T_∞ is assumed constant. Measurements show that the environment temperature actually increases slightly in the course of a separation from 22 to 26 °C, due to heat dissipation from instrument components such as light source and power supplies. The current also increases gradually in the course of a separation, typically by about 6 % in the course of a separation (87 to 92 µA in 40 min). Finally, the temperature within the capillary may very well not be uniform, neither in axial nor in radial direction. For a more detailed study numerical simulations are necessary, however this goes beyond the scope of this thesis. For an approximate analysis of the impact of capillary temperature on protein migration trajectories, only the effect of increasing environment temperature will be considered.

The constant environment temperature (T_∞) is replaced with an increasing environment temperature ($T_0 + ht$), with a temperature-increase coefficient h [K/s]. This results in an additional term of $h \left(t - \frac{\beta}{\alpha} t \right)$ in Equation 2.29, the differential equation that is solved to find $T(t)$. The result describes a temperature profile similar to the one above, but with an overlaid gradual (linear) increase. The resulting equation for temperature is: (see Appendix A for derivation)

$$T(t) = T_0 + \left(\frac{VI}{\alpha} - \frac{\beta}{\alpha} h \right) \left(1 - e^{-\frac{\alpha}{\beta} t} \right) + ht \quad \text{with } \alpha = kS \quad \text{and } \beta = mc \quad (2.32)$$

With suitable coefficients:

$$T(t) = 22 + 11.896 \left(1 - e^{-1.404t} \right) + 0.001t \quad (2.33)$$

Although this model still uses some approximations, such as uniform temperature within the capillary at all times and constant Joule heating, it agrees well with the observable data (i.e. thermal increase in buffer conductivity). Temperature measurements at different positions along the capillary (shown in Figure 2.60) further confirm this as a suitable description of the course of temperature. Equation 2.33 will be used as the basis for estimating non-linear migration in the course of a separation with changing temperature.

Previous work has characterised how temperature influences the sieving effect in capillary gel electrophoresis of proteins.^[187] For different proteins between 10 and 100 kDa, the electrophoretic mobility μ increases with temperature. In a $(\ln \mu)$ vs. $(1/T)$ plot, a linear relationship is observed, with the slope varying for different sieving matrices and

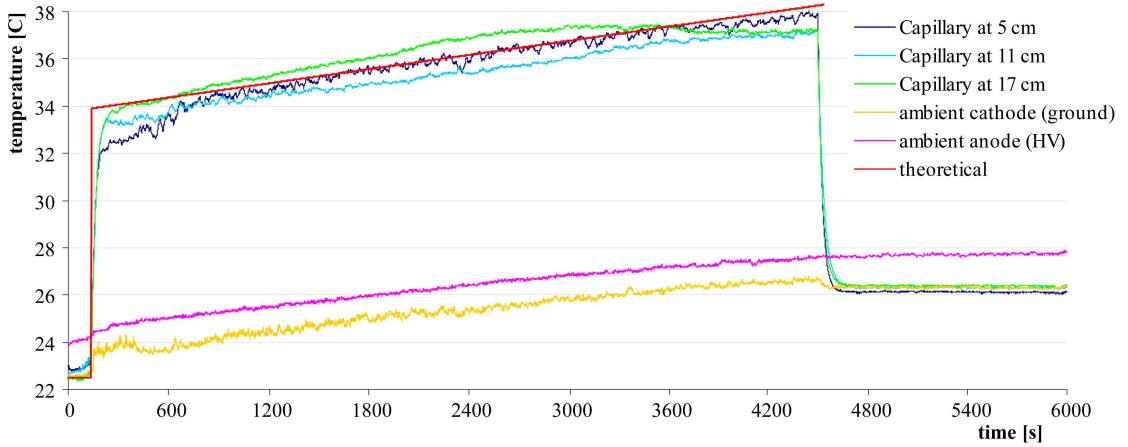


Figure 2.60: Calculated and measured temperature profiles. The red line indicates the calculated temperature profile (the starting temperature T_0 is adjusted to 22.5 °C). Green and blue lines represent different positions along the capillary, measured from the injection point. Orange and pink lines represent the ambient temperature at different regions of the instrument. Data are recorded with K-type thermocouples and an eight-channel data logger (Pico TC-08) at 1 Hz, on a standard capillary (200 mm separation length, total length 350 mm; 75 µm inner diameter, 375 µm outer diameter) under standard conditions (12.25 kV at anode, or 350 V/cm electric field strength; Beckman SDS gel buffer).

the intercepts varying for different molecular weights:

$$\mu(T) = Ae^{-\frac{B}{T}} \quad (2.34)$$

With the condition $\mu(T = T_0) = \mu_0$ we find:

$$A = \mu_0 e^{\frac{B}{T_0}} \quad (2.35)$$

$$\mu(T) = \mu_0 e^{\frac{B}{T_0}} e^{-\frac{B}{T}} \quad (2.36)$$

To find the temperature dependence of migration trajectories, Equation 2.33 for $T(t)$ is used to express the mobility μ as function of time:

$$\mu(t) = \mu_0 e^{\frac{B}{T_0}} e^{-\frac{B}{T_0 + 11.897(1 - e^{-1.404t}) + 0.001t}} \quad (2.37)$$

Although the $(1 - e^{-1.404t})$ term is already at 0.999 for $t = 5$ seconds, and could be considered negligible, it will be included for the sake of accuracy. With the definition of mobility ($\mu = v/E$ with velocity v [m/s] and electric field strength E [V/m], see Equation 1.17), the migration speed becomes:

$$v(t) = v_0 e^{B \left(\frac{1}{T_0} - \frac{1}{T_0 + 11.897(1 - e^{-1.404t}) + 0.001t} \right)} \quad (2.38)$$

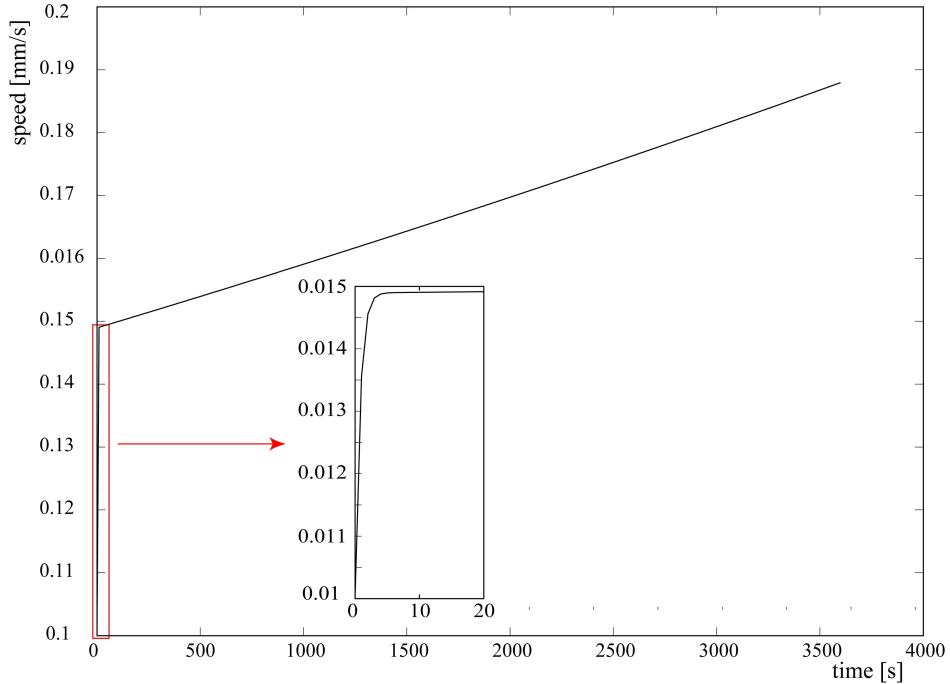


Figure 2.61: Migration speed as a function of time according to time-dependent temperature. The plotted speeds are for $T_0 = 295$ K, $v_0 = 0.1$ mm/s and $B = 1850$, which are typical values. All other conditions are standard ($E = 350$ V/cm, Beckman buffer, fused silica capillary with 75 μm inner diameter, 375 μm outer diameter).

In Figure 2.61 the resulting migration speeds in the course of a separation are shown. The migration trajectory $x(t)$ is found by integrating Equation 2.38 over t .

As no exact integration is available for the function in Equation 2.38, integration is done numerically with MATLAB. First, the temperature coefficient B , which defines how strongly the mobility changes with temperature, is determined.

In the literature B values in the range of 1000–3000 are reported, depending on matrix composition ($B = 2310$ for 15 % dextran 72 kDa matrix; $B = 1573$ for 3 % PEO 100 kDa matrix^[187]). Because the buffer used (Beckman SDS-MW gel buffer) is of undisclosed formulation, different values for B are compared to find the value that fits the observed data best. In Figure 2.62 trajectories with the same initial speed, but different B values are plotted. A value of $B = 1850$ is found to fit actual migration data best.

For a variety of initial migration speeds, the trajectories are plotted. Once the trajectory is given, the time a protein arrives at the detection window (at 200 mm) can be determined, along with the speed at which it migrates. The tracks that would be fitted from observation at the detection window are determined. The reconstructed starting point for these fitted tracks allows quantification of the error (the ‘starting position’ error as defined in Section 2.2.3). Figure 2.63 shows 10 migration trajectories with their fitted tracks.

By consulting a molecular weight-mobility calibration curve as measured in Section 2.3.1.3, the tracks with an arbitrary initial migration speed can be assigned an observed migration speed and from that a molecular weight.^[195] The errors associated with different molecular weights can then be determined. The results are shown in Figure 2.64. The

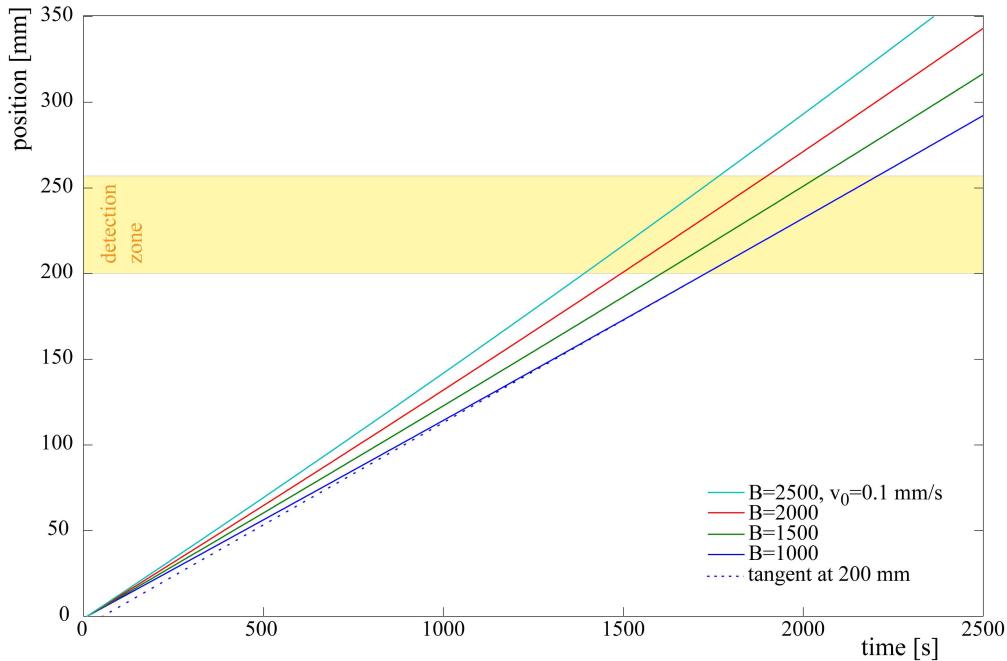


Figure 2.62: Migration trajectories for different temperature-mobility coefficients B . The tangent to one of the trajectories at 200 mm is indicated as a dashed line. All trajectories are for $T_0 = 295$ K, $v_0 = 0.1$ mm/s. All other conditions are standard ($E = 350$ V/cm, Beckman buffer, fused silica capillary with 75 μm inner diameter, 375 μm outer diameter).

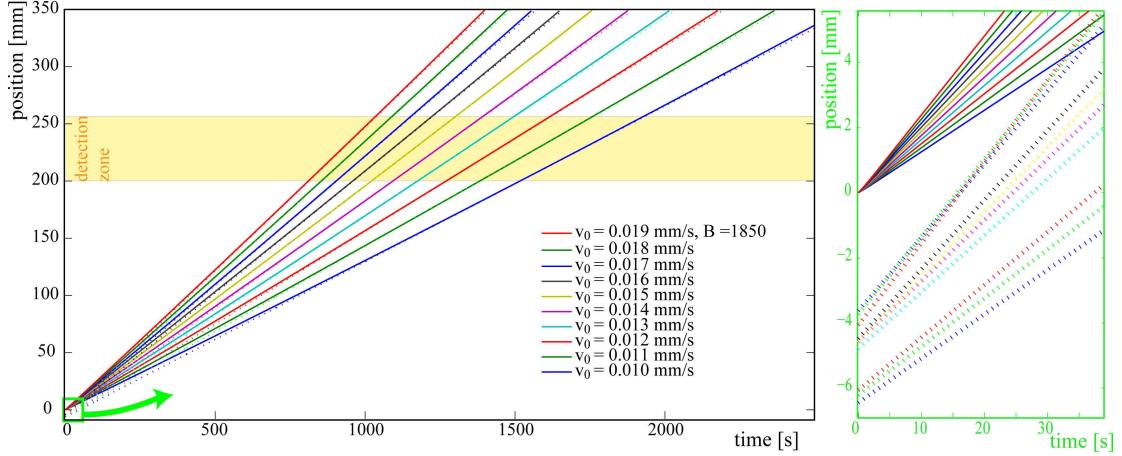


Figure 2.63: Migration trajectories for different initial migration speeds v_0 . The fitted tracks are indicated as a dashed line. The panel on the right shows an enlarged section of the vertex region. All trajectories are for $T_0 = 295$ K, $B = 1850$. All other conditions are standard ($E = 350$ V/cm, Beckman buffer, fused silica capillary with 75 μm inner diameter, 375 μm outer diameter).

unevenness of the thermal error curve is attributed to the resolution of the numerical model (1 second time steps), but a trend toward slightly larger errors for higher molecular weight can be recognised.

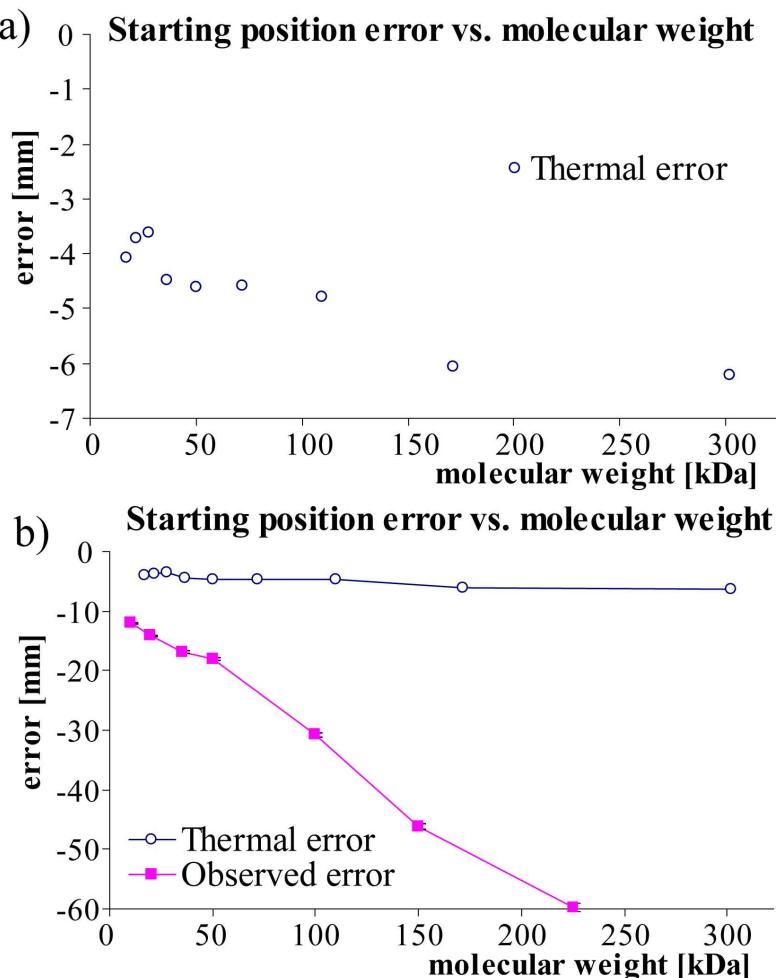


Figure 2.64: Errors for different molecular weights arising from non-linear migration due to temperature-dependent mobility ('thermal error'). a) Thermal errors. b) Thermal errors and actually observed ('starting position') errors of the reconstructed starting point. Thermal error model coefficients: $T_0 = 295$ K, $B = 1850$, $v_0 = 0.1\text{--}0.19$ mm/s. All other conditions are standard ($E = 350$ V/cm, Beckman buffer, fused silica capillary with 75 μm inner diameter, 375 μm outer diameter, detection at 200 mm).

Thermal errors are in the range of 3–7 mm for the chosen conditions. As seen in Figure 2.64b, the actually observed errors are around one order of magnitude greater. From this we conclude that thermal errors only make a minor contribution to the observed errors. For low molecular weight proteins thermal errors are of more relevance than for high molecular weight proteins. Non-linear behaviour due to thermal effects may be the reason why trajectories are never seen to go through the actual (recorded) coordinates of the start of separation, even for very good data quality and no otherwise observable deviation from a straight line track.

2.4.2 Errors due to background effects

As illustrated in Section 2.3.1.5, migration trajectories display kinks with peak widening thereafter. It was demonstrated that these trajectory bends coincided with a counterband, a background disturbance which slowly migrates from the anode to cathode, in the opposite direction as proteins. This discontinuous background is seen as steps in the electropherogram. Different steps of varying height but all the same migration speed can be distinguished.

By tracking the migration of steps, the starting position of different steps is found to be at different positions in the capillary. These starting positions lie between the last pixel of the PDA and the anodic buffer vial (Figure 2.27). The position where the counterbands intersect with a given protein band depends on capillary geometry, as well as separation voltage. Different buffers show similar counterbands under otherwise identical separation conditions. Notably buffers with high conductivity show the strongest counterbands.

To further characterise the behaviour of such counterbands, separations were performed on capillaries that were pre-run for 60 minutes, see Figure 2.65. It can be seen that if the separation voltage is applied for 60 minutes prior to protein injection and separation, the characteristic bends in the migration trajectories do not occur. From this it is concluded that the cause of counterbands is a characteristic of the system that takes longer than the typical separation time (40 minutes) to reach steady-state.

If there were some evidence that the counterbands all originate in the anodic vial, electrochemical modification of the buffer system could be a plausible explanation. Since this is not the case, thermal effects are investigated as a cause. The anodic vial is at high voltage, whereas the cathodic vial is grounded. If there were leakage current at

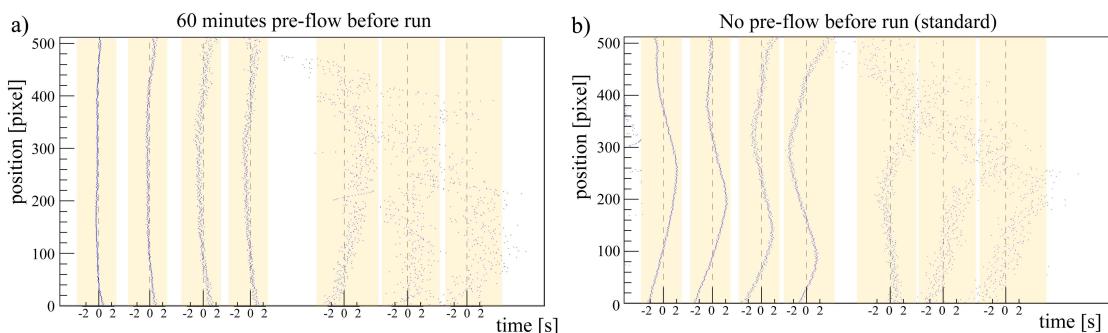


Figure 2.65: Track shape with and without pre-flow before separation. a) Residual plot of a separation with 60 minutes pre-flow prior to the separation. b) Residual plot of a standard separation without pre-flow. All other separation conditions are standard for both separations (200 mm separation length, 350 V/cm electric field strength, 75 μm inner diameter capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa).

the anodic vial, Joule heating would affect the vial and the inserted capillary end. Additionally, the anodic end of the capillary is in a section of the instrument with little air flow, so the heat transfer from capillary to environment could be significantly lower, causing excessive Joule heating in that capillary section. To assess this, the temperature at different positions along the capillary was measured with thermocouples.

In Figure 2.66a, the temperature in the course of a standard separation is recorded at three positions along the capillary between injection end and detection window. Alongside this the ambient temperature in different areas of the instrument is recorded, as well as the temperature near the electrode reservoir vials. It is noteworthy that the temperature at the high voltage electrode (anode) increases to a level above the ambient temperature, indicating non-uniform temperature within the capillary.

Following the observation that the capillary temperature at the anodic end of the capillary is significantly higher than at the cathodic end, cooling measures were installed and

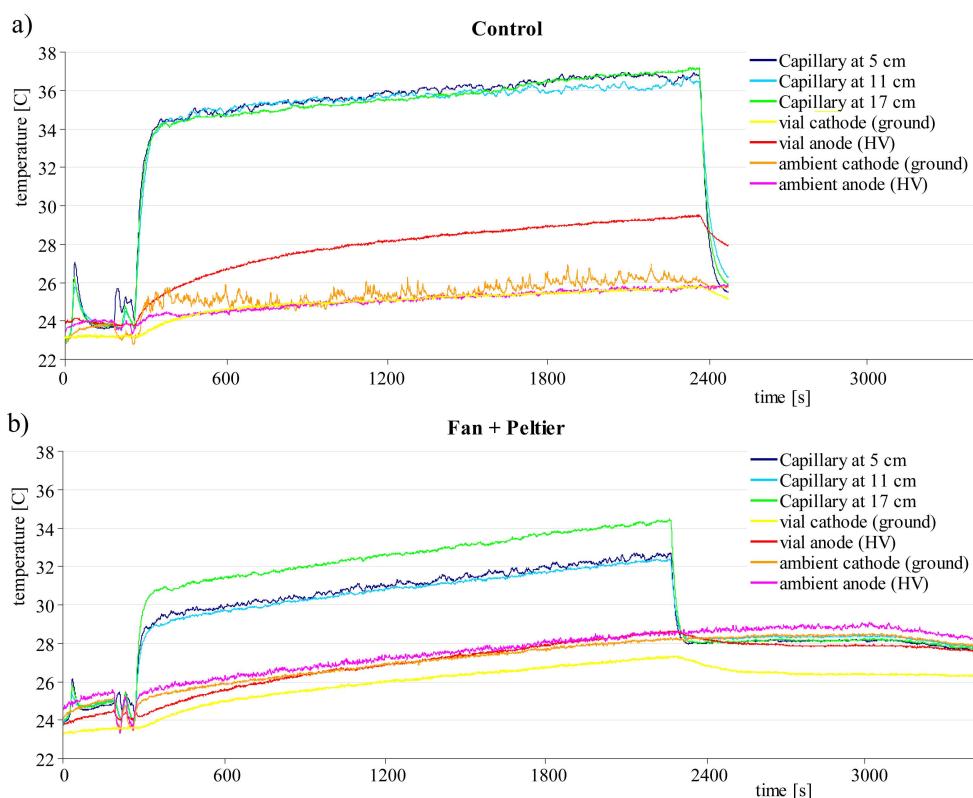


Figure 2.66: Temperature at different positions along the capillary in the course of a standard separation a) without and b) with additional cooling. Cooling is achieved with a 28.7 W Peltier element at the anodic vial holder, and a 1.4 W fan. The temperature is measured with K-type thermocouples at 5 cm, 11 cm, and 17 cm from the capillary end (cathodic sample injection end), as well as in close proximity to the electrode vials and the ambient temperature in the region of the capillary around both ends. All other separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μ m inner diameter capillary, Beckman SDS gel buffer), with the standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

separations under optimised anodic cooling conditions were performed. Figure 2.66b shows the development of temperature during the same standard separation, but with better cooling. Here the overall temperature differences are reduced, and the anodic electrode vial is maintained at the same temperature as ambient air.

In Figure 2.67 the residual plots of standard separations, with and without additional cooling measures, are shown. It is seen that with careful temperature control the large bends are suppressed and better linearity is achieved. From this it is concluded that what is seen as counterbands are actually temperature steps that run through the capillary. Upon encountering a temperature step, the migration speed of a band changes due to altered buffer sieving behaviour. This causes the observed kinks in the migration trajectory.

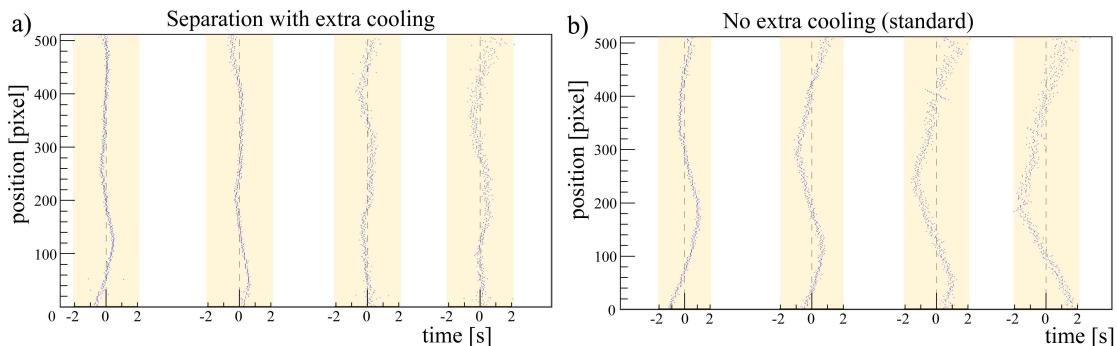


Figure 2.67: Track shape with and without additional cooling. a) Residual plot of standard separation with additional cooling. b) Residual plot of a standard separation (no additional cooling). For cooling a 28.7 W Peltier element is installed at the anodic vial holder, with 1.4 W fan. Residuals are shown for the 10, 20, 35, and 50 kDa tracks. All other separation conditions are same (200 mm separation length, 350 V/cm electric field strength, 75 μ m inner diameter capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

To confirm whether the migration of the counterbands is affected by electro-osmotic flow (EOF) or by thermal conductance, the EOF for otherwise standard condition is examined. This is done by filling the capillary with buffer as usual, but spiking the electrode vials with an EOF marker (1 mg/mL thiourea, which does not migrate electrophoretically because it is uncharged, but absorbs light at 214 nm). The result is shown in Figure 2.68, where it can be seen that there is indeed an electro-osmotic flow, from anode toward cathode (i.e. the opposite direction from protein migration), of comparable migration speed as the counterbands.

To summarise, migration trajectories of proteins display bends that are associated with non-uniform temperature distribution throughout the capillary. Data indicate that this

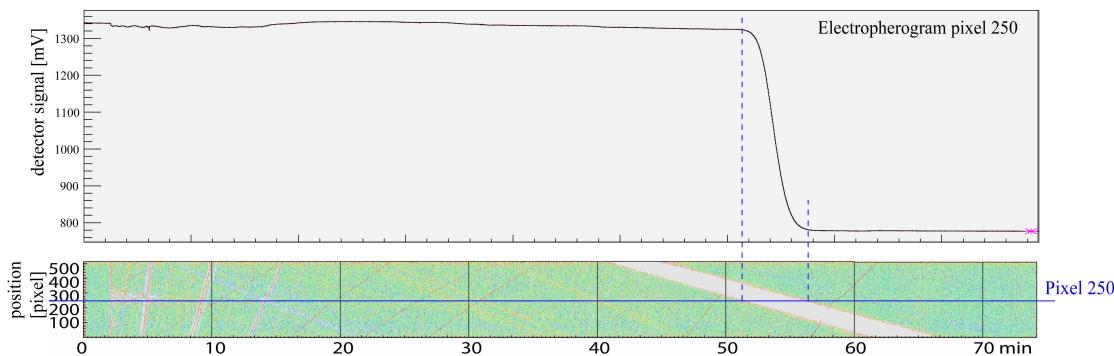


Figure 2.68: Electro-osmotic flow from anode. A strong drop in photocurrent is seen where run buffer from the anodic vial enters the detection window. All other conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μm inner diameter capillary, Beckman SDS gel buffer, no sample).

could be caused by buffer zones of different temperature that migrate with the electro-osmotic flow. This effect is dependent on instrumentation, and can be minimised by optimising temperature control.

2.4.3 Errors due to UV effects

As illustrated in the results characterising a standard separation (Section 2.3.1.6), protein bands are seen to accelerate as they progress through the detection window. This goes together with the development of a pronounced tail at the leading ('fast') flank of peaks. The peak asymmetry is used as an indicator for this phenomenon. It was shown that the tailing is suppressed when the capillary is shielded from irradiation with ultraviolet light at 214 nm.

Various studies have analysed capillary electropherogram peak shapes, in particular for capillary zone electrophoresis and DNA gel electrophoresis.^[180,181,218–220] Wall-analyte interactions are a common cause of tailing at the following peak flank. Peak symmetry can also be affected by background electrolyte conductivity and pH. Peak asymmetry can be caused by radial thermal gradients (causing mobility gradients) from Joule heating.^[200,201,213] Searching literature gives no indication that UV associated tailing has been observed before, which is not surprising because most instruments use single point illumination and detection.

Assuming UV irradiation affects the separation, different potential mechanisms are conceivable. Absorption of UV light could cause heating and expansion, although such expansion would be uniform across the detection window and not necessarily cause asymmetric band distortion. As mentioned above, if absorption of UV illumination were to cause a radial temperature gradient similar to Joule heating, but confined to the

illumination window, the changes in mobility might cause peak asymmetry and acceleration. Since in CGE separations mobility increases with temperature, this mechanism would cause a tail at the following flank of peaks, not at the leading flank.

Absorption of UV light could also cause degradation of the capillary contents. Either protein degradation, or degradation of the sieving matrix could cause acceleration. Protein degradation would cause breakage into smaller protein fragments, which migrate quicker and shift the bulk volume of the peak forward. Degradation of the sieving matrix would also cause proteins to accelerate, due to less matrix resistance. Matrix degradation does not explain peak asymmetry, as the matrix is stationary and uniform across the window, but it will not be ruled out. To explore whether protein degradation by UV irradiation is involved in causing the observed non-linear migration, further evidence is considered.

To study UV degradation mechanisms, the behaviour of proteins forced to migrate back and forth in the detection window is studied. For a standard separation of a certain number of proteins, the separation is allowed to progress until the proteins have passed beyond the detection window. At this point, the voltage polarity is reversed, forcing the proteins to migrate back through the detection window again.

So far it appears that whatever the mechanism of UV degradation, it is directed toward the anode, i.e. the driving force is electrophoretic. If this is correct, then after reversing polarity the leading tail of a peak will start the second pass at the following edge. The peak shape will progressively shift back to re-form a tail at the leading peak edge. Conversely, if the driving force is, for instance, pressure driven, then reversing polarity will have no effect on how the peak asymmetry develops, and in the second pass the tail will remain at the same peak flank. Figure 2.69 illustrates the two possible outcomes of the voltage switching experiment.

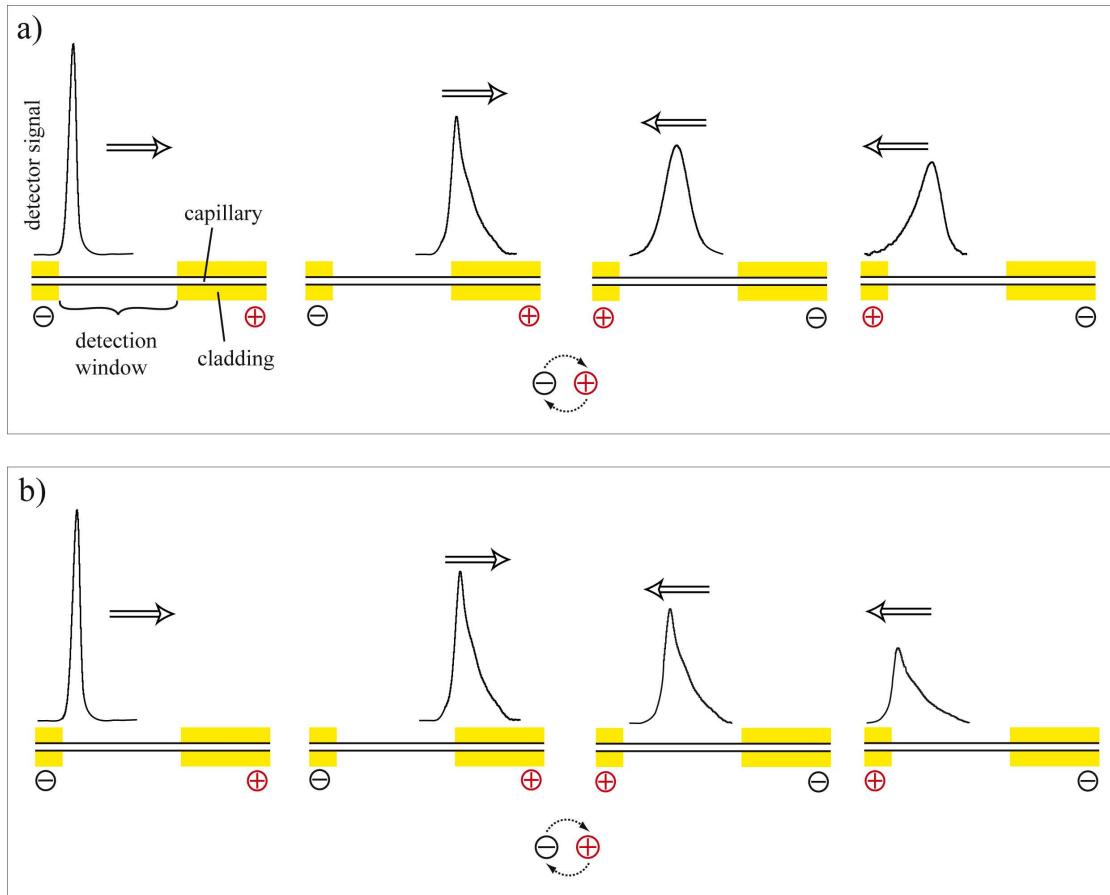


Figure 2.69: Schematic of anode directed peak asymmetry and electric field independent asymmetry. a) A leading tail develops as a peak progresses across the detection window. When the separation polarity is switched after the peak has passed, the tail re-enters the detection window at the following peak flank. Gradually, the bulk of the peak migrates forward again, creating a new, reversed, asymmetric peak. b) If the peak asymmetry is not related to electrophoresis, in the second pass the peak continues forming in the same manner, despite migrating in the opposite direction.

The results from voltage switching are shown in Figure 2.70. From the electropherograms, it can be observed that peak asymmetry is reversed with the change of separation polarity. This is quantified in the HWHM plot in Figure 2.71, where the peak shape from first and second pass of each band is compared.

From Figures 2.70 and 2.71 it is clear that in the second pass, what started as a following tail at the back of the peak returns to a symmetric peak and then shifts further to re-form tail at the leading ('fast') edge. This fits well with the tentative explanation of protein cleavage under UV irradiation: the shorter, quicker fragments that have accumulated at the front of a band suddenly find themselves at the back, and again migrate ahead and overtake larger proteins, resulting in the renewed formation of a leading tail shape.

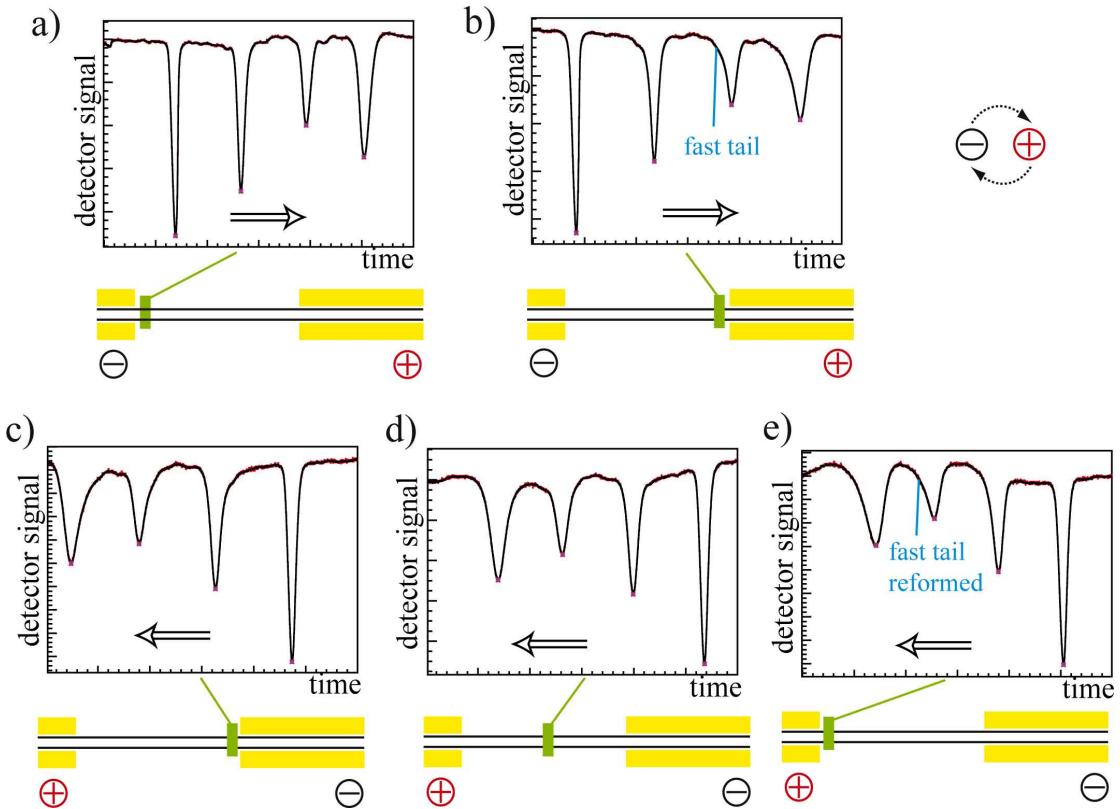


Figure 2.70: Electropherograms at different positions for first and second (reverse) pass of four protein peaks. a) Electropherogram of the first pixel in the detection window. b) Electropherogram of the last pixel in the detection window. c) Electropherogram of the last pixel after a voltage polarity switch. d) Electropherogram of the middle pixel in the detection window after a voltage polarity switch. e) Electropherogram of the first pixel in the detection window after a voltage polarity switch. Unlike the schematic in Figure 2.69, the x-axis indicates time. What was the last peak in the first pass is the first peak to re-enter the detection window, and electropherograms appear as mirror images.

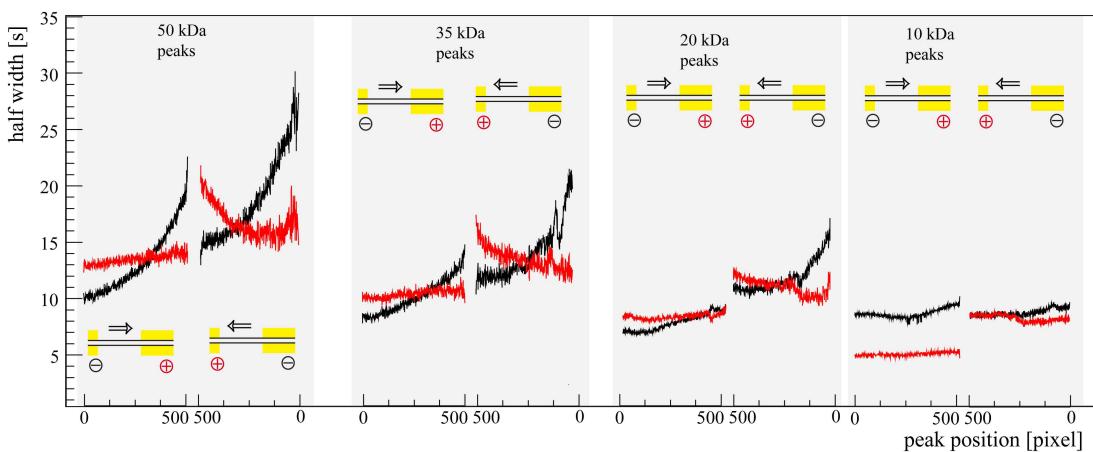


Figure 2.71: HWHM plot for the 4 bands seen in the electropherograms in Figure 2.70. To compare the peak shape from the first pass with the second pass, HWHM curves are positioned side by side, and not in the order of elution. Red colour indicates the slow, tailing half width. The leftmost group of curves is the first to re-enter the detection window, i.e. the last to pass the detection window before switching polarity (highest molecular weight).

To assess the role of the sieving matrix in peak asymmetry, it would be helpful to understand how migration progresses after UV exposure, in the absence of further illumination. Because the sieving matrix does not migrate (or only very slowly), after passing the detection window protein migration continues in a buffer environment that is unaffected by UV exposure. If buffer degradation is causing the peaks to deform, then once outside the illumination window, no further deformation is expected. If, on the other hand, protein degradation is causing peak deformation, then the fragments would continue to migrate farther away from the parent peak. In this case the peak shape would continue to change even in the absence of illumination.

A set of sliding window experiments was designed to determine post-exposure behaviour. A capillary with two detection windows was used. Initially the separation is observed in the first detection window, until at a given time the separation is interrupted and the capillary is shifted to allow detection in the second window. Typically proteins migrate for 18 cm before the first window, then undergo 5.1 cm illumination and detection, followed by another 7 cm unperturbed migration before entering the second detection window.

In Figure 2.72 the electropherograms of two peaks at different positions along the capillary are shown. Here the HWHM is not the ideal metric for observing the peak shapes,

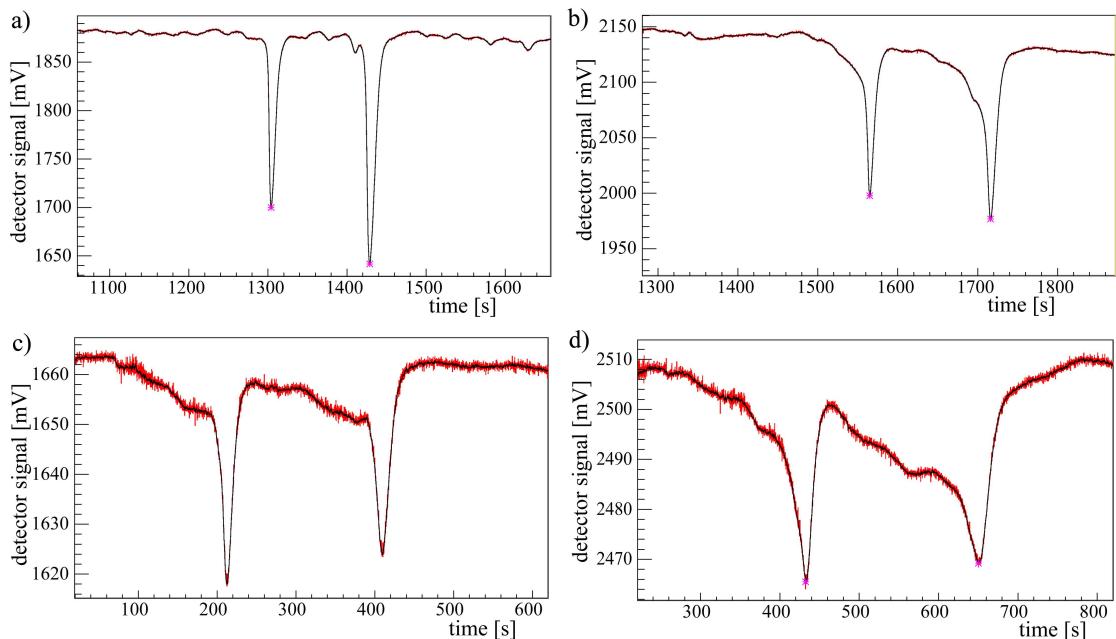


Figure 2.72: Sliding detection window: electropherograms of 35 kDa and 50 kDa peaks at different positions along the capillary. a) First window, pixel 10 (18.1cm). b) First window, pixel 500 (23 cm). c) Second window, pixel 10 (30.1 cm). d) Second window, pixel 450 (34.5 cm). Separation performed under otherwise standard conditions (350 V/cm electric field strength, 75 μm inner diameter capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL).

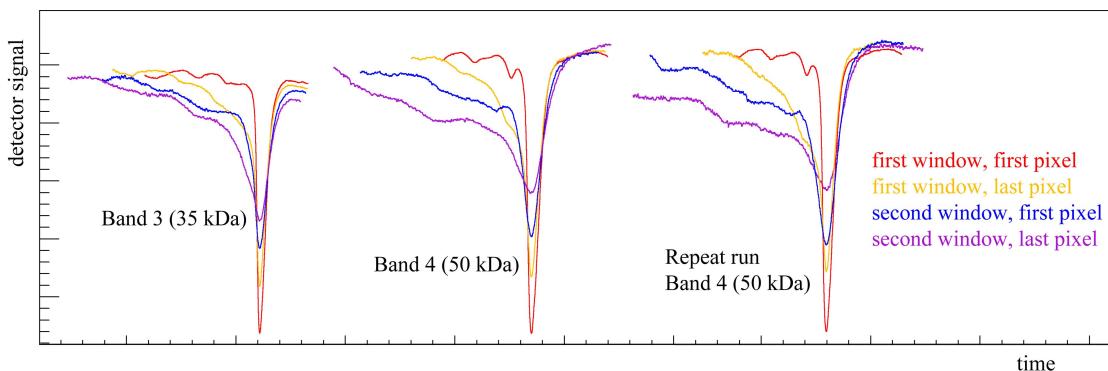


Figure 2.73: Sliding detection window: comparison of different peaks at different capillary positions. Detection windows at 18 and 30 cm. Separation performed under otherwise standard conditions (350 V/cm electric field strength, 75 μm inner diameter capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa).

because there is a distinct peak and tail, and, especially in the second window peaks, the tail does not affect the half widths. Comparing the peaks in Figures 2.72b and 2.72c, it appears that the tail has migrated independently of the main of the peak, and has spread out further. In Figure 2.73 the different peak electropherograms are overlaid for better comparison. In all peaks it can be seen that the leading tail has migrated ahead of the main peak, indicating that the damage sustained during the UV illumination is permanent damage to the proteins of that peak.

Having demonstrated that proteins undergo degradation under UV illumination, it remains unclear whether this is a buffer mediated phenomena, or direct photolysis of peptide bonds.

For protection of the high molecular weight sieving polymer in gel buffers, scavengers may be included in the gel buffer formulation, indeed even with the bulk polymer. Comparing different separation media, it is observed that separations in capillary gel electrophoresis (CGE) experience fronting/tailing, for all buffer formulations examined (Beckman SDS gel buffer, Groton Optima buffer, and custom prepared dextran-based buffers, see Section 2.3.4). This indicates that UV degradation occurs for different buffer formulations, and although it is particularly pronounced in some formulations, it occurs widely.

Unlike CGE separations, capillary zone electrophoresis (CZE) separations show symmetric peaks up until the end of the detection window. This is illustrated in Figure 2.74, where the HWHM toward front and back of peaks are plotted against the time coordinates of hits. In the left image the separation of five proteins in a CZE buffer is shown. The right image shows the separation results of five proteins in a commercial CGE buffer with a non-crosslinked, entangled polymer sieving matrix, where severe tailing is seen.

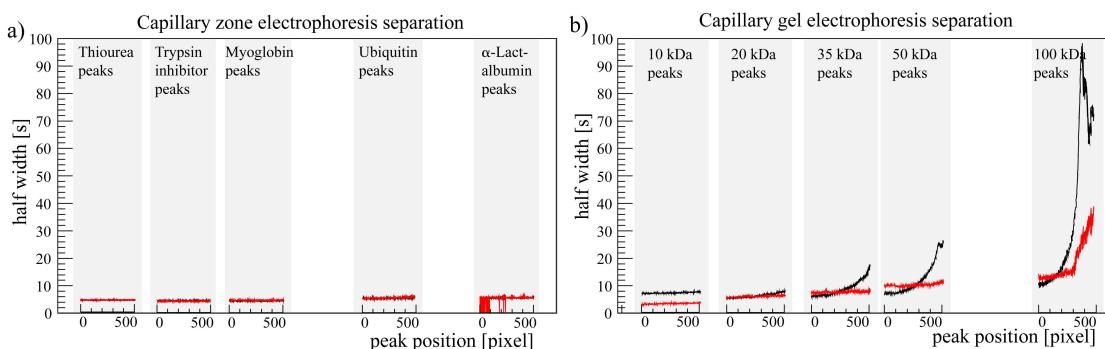


Figure 2.74: Comparison of peak asymmetry for different separations: a) capillary zone electrophoresis, b) capillary gel electrophoresis. Half width at half maximum toward fast (black) and slow (red) flanks of peaks are plotted against hit position. a) CZE, sample of five proteins with molecular weights 8–50 kDa, normal polarity separation in 42 mM sodium tetraborate, pH 9.3 (strong EOF). b) CGE, sample of five proteins with molecular weight 10–100 kDa, reverse polarity separation in commercial Beckman buffer (negligible EOF).

At a first glance it appears that proteins do not undergo degradation solely from UV irradiation, but degradation is related to the buffer environment. However, it must be taken in consideration that the mechanism of separation differs between CGE and CZE. In CZE the electrophoretic force depends on the charge-to-size ratio (Section 1.3.2). In CGE on the other hand, the driving force is same for all proteins, with a constant charge-to-mass ratio. Instead, the sieving effect of the matrix separates smaller, more mobile proteins from longer, less mobile proteins. Assuming protein cleavage, in CGE the fragments invariably migrate faster. In CZE the charge-to-size ratio might increase, decrease, or remain same, so fragments might migrate faster or slower, or at the same speed, and peak shape is not the anticipated indicator of degradation. Another consideration is that in CZE, proteins are not necessarily denatured, nor are disulphide bonds reduced. Therefore if peptide bonds were cleaved, conceivably the fragments could continue migration with the bulk of the protein and no changes would be observable. Finally, the mobility differences in CZE are far smaller than in CGE, so if protein fragmentation were to occur, it would be more difficult to observe than in CGE.

To determine whether protein degradation under UV illumination is related to CGE buffer composition, or if it is a universal, buffer-independent phenomenon, CZE experiments were designed to see if peak degradation can be observed under some circumstances. The CZE buffer used was a simple sodium tetraborate buffer (42 mM, pH 9.3) in water with 8 M urea and 2 % w/v CHAPS to ensure denaturing conditions (all from Sigma-Aldrich Co.). The previously introduced sliding window setup was used, with the first window for exposing one band to a high dose of UV illumination. After exposure of one band in the first window, the capillary is shifted to the second window and the peak shape is recorded. Because only one band received extra exposure, only one peak would

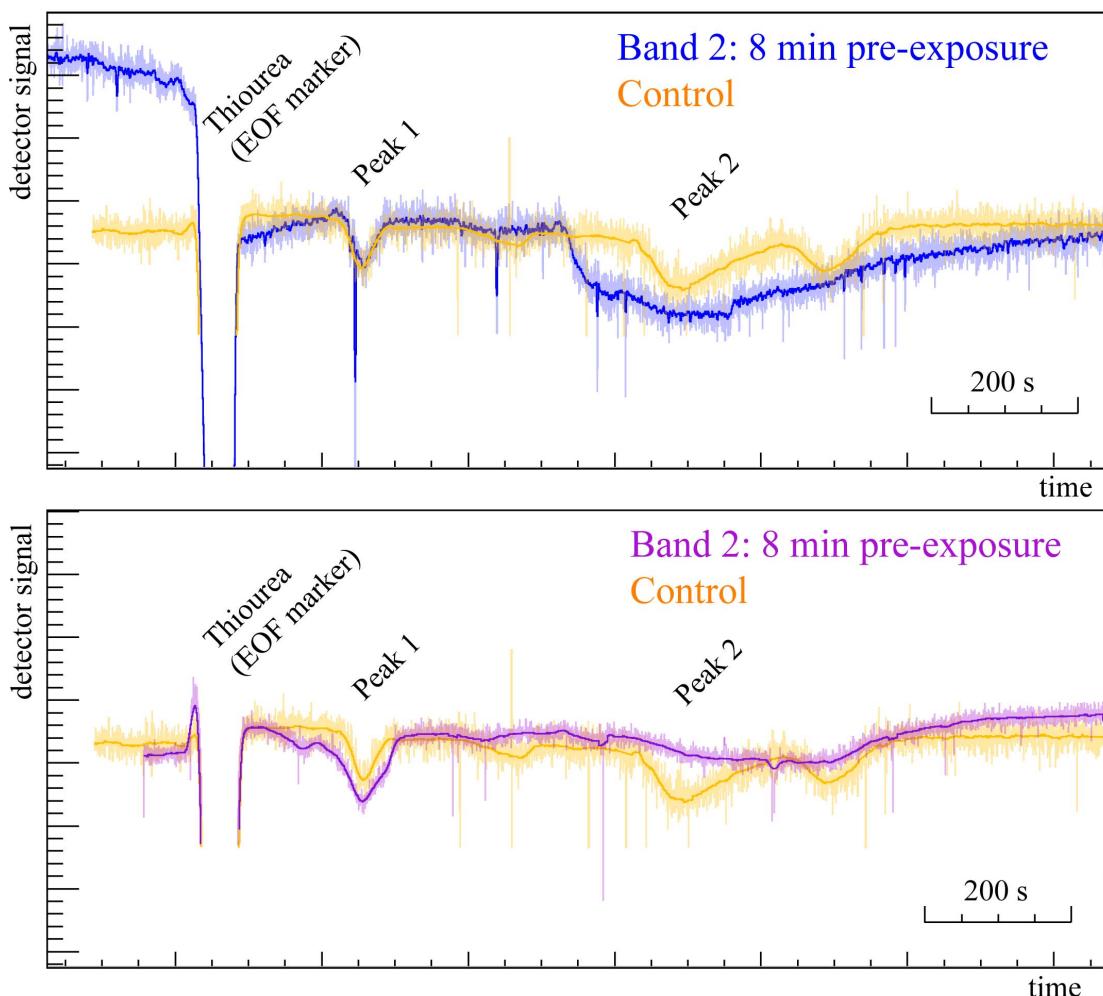


Figure 2.75: Comparison of peaks in denaturing CZE separation with and without pre-exposure of a peak. The upper and lower panel are repeat runs of the same experiment. The electropherogram of a control run with no pre-exposure (but otherwise identical conditions) is underlaid in orange, to illustrate changes in the pre-exposed peaks. Of the two prominent peaks, only one, the second peak, is pre-exposed. Conditions: normal polarity CZE separation in sodium tetraborate buffer (42 mM, pH 9.3) with 8 M urea and 2 % w/v CHAPS. 8 minutes pre-exposure of second peak at 15 cm, detection at 60 cm. Separation in 75 μ m inner diameter capillary at 230 V/cm electric field strength, under strong EOF conditions.

show degradation in the second window, allowing comparison to unexposed bands that underwent the same separation.

In Figure 2.75 the electropherograms of a separation with pre-exposure and a control separation is shown. As Figure 2.75 shows, the pre-exposed band (labelled as band 2) displays considerable peak degradation after an additional 45 cm migration distance, unlike the unexposed peak preceding it. This illustrates that under specific conditions, peak degradation can be shown to occur in CZE separations, indicating that protein fragmentation can be observed in a simple medium with minimal additives. This is

further supported by evidence found in the literature that peptide bonds can undergo photolysis around 214 nm in aqueous solution.^[77,221–226]

To summarise, UV illumination causes bands to accelerate and develop a tail at the leading front. If the polarity is switched, the peak tails shift back forward to the leading flank. After UV illumination, the leading tail continues to move farther away from the main peak. Peak degradation is not ordinarily observed in CZE, but can be induced under certain conditions. These observations are consistent with degradation and cleavage of proteins into smaller fragments which migrate with a higher mobility. The most plausible cause for proteins to undergo fragmentation is direct UV photolysis. Some evidence shows that sufficient UV irradiation can break peptide bonds.^[77,221–226]

2.4.4 Error optimisation and estimation

In a V2D separation, reconstructing the starting position of the CGE migration is crucial. In the light of minimising the error of this reconstructed starting position, the presented data can supply some optimal experimental conditions. This is summarised in Table 2.1. The values in Table 2.1 are applicable for separations of samples in the range of 10–225 kDa using Beckman buffer with the instrument described in Section 2.2.2.

Table 2.1: Optimal experimental conditions for minimal errors on the reconstructed starting position. For samples 10–225 kDa, using Beckman SDS gel buffer on the custom built instrument.

Experimental parameter	Optimal value
Hit threshold	100 x noise level
Detection length	> 20 mm
Electric field strength	300–400 V/cm
Separation length	10 cm
Capillary diameter	50 µm

With regards to instrumentation, the single most influential factor is temperature control. This affects how much proteins migration has deviated from a linear trajectory (thermal effects), as well as how tracks are fitted (background effects).

Further to the optimal experimental conditions, it is possible to derive estimators for the errors under arbitrary conditions. Conventional studies of errors in capillary electrophoresis generally describe the peak shapes, in particular the peak height and width, for different experimental conditions. This is the standard approach for a single detector, where the (average) migration speed is determined by the quotient of the separation distance divided by the migration time, i.e. the time of detection. In a V2D separation

by contrast, the separation distances vary for different proteins, and are not known *a priori*. Here the migration speed needs to be determined by tracking along multiple pixels. Hence the errors on migration speed are not necessarily dominated by the same factors.

One of the first observations is that the actual error on the reconstructed starting position is significantly greater than the error associated with the tracks (Section 2.3.1.2). As the track errors are related to peak shape parameters, similar to single detector data, this indicates that the observed reconstruction errors are heavily influenced by further factors.

In the foregoing discussion, three causes of non-linear migration and resulting errors were identified and characterised: thermal effects, background effects, and UV effects. Here we will develop an estimate of the contribution of each of these factors to the error on reconstructed starting position.

2.4.4.1 Estimate of thermal errors

The thermal error is caused by the separation buffer gradually heating up, causing proteins to migrate slower in the beginning and faster toward the end. This depends mostly on factors such as external heat sources (e.g. light source) and instrument cooling, and the duration of the separation. Although Joule heating (i.e. buffer resistance, separation voltage) causes the capillary temperature to increase, after the first couple of seconds of a separation a steady state is attained (for a narrow bore capillary in any case), and thereafter Joule heating hardly contributes to non-constant temperature conditions (and non-linear migration).

For the characterised system, the thermal error for a variety of proteins is small (3–7 mm) compared to the actually observed errors (10–60 mm). This means that in a first approximation the thermal error is independent of molecular weight, and adds an offset error. This explains why the tracks do indeed vertex reasonably well for the first few tracks, but not at the recorded vertex. For data from a standard separation, this offset is in the region of 13 mm, slightly greater than the thermal model in Section 2.4.1 predicts.

It is possible to define a function that describes the thermal error dependency on migration time, although this requires calibration data for different instrumentation. In

Figure 2.76 a linear function describes thermal errors with:

$$\epsilon_\Theta = -0.0042\tau - 0.2678 \quad (2.39)$$

where ϵ_Θ [mm] is the thermal error, and τ [s] is the migration time. With $\tau = L/(\mu E)$ this can be correlated with the more common experimental parameters mobility μ [mm²/(sV)], separation length L [mm], and electric field strength E [V/mm]. It is important to remember that Equation 2.39 is only an approximation. As shown in Section 2.4.1 the underlying function is more complicated.

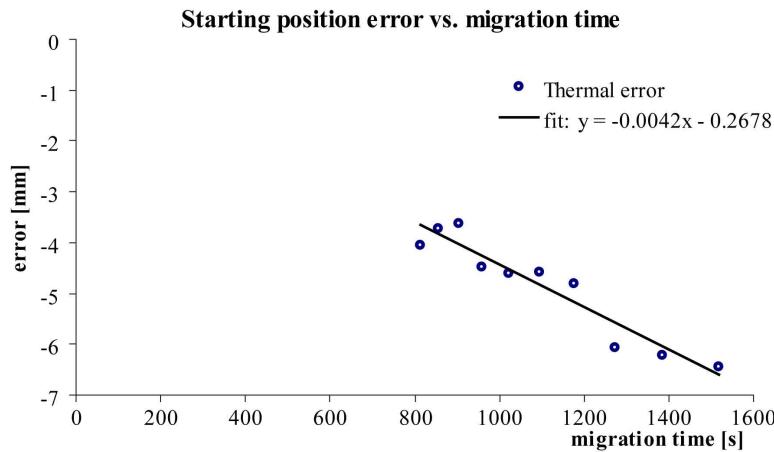


Figure 2.76: Errors for different migration times arising from thermal effects. Conditions: $T_0 = 295$ K, $B = 1850$, $v_0 = 0.1\text{--}0.19$ mm/s. All other conditions are standard ($E = 350$ V/cm, Beckman buffer, fused silica capillary with 75 µm inner diameter, 375 µm outer diameter, detection at 200 mm).

2.4.4.2 Estimate of errors from background effects

Errors from background effects (counterbands) are related to thermal errors in so far as they are also caused by a non-constant temperature, however they display quite different phenomena. Counterband errors are caused by distinct buffer zones of different migration behaviour. The buffer zones migrate against the proteins and originate from the anodic (high voltage) end of the capillary. Counterbands can be reduced by enforcing heat transport for a uniform temperature distribution. Counterband reduction brings about straighter tracks and smaller χ^2 values for better fits, but the influence on the reconstructed starting position error is more complicated. For instance, under some circumstances the counterband contributions balance out, and the overall influence is very small.

As shown in Section 2.3.2 where data from varying sub-sections of the detector were used for track fitting, depending on how much and which section of the track is included in

the analysis, the observed reconstructed starting position error can increase or decrease. Studying these errors gives an indication of how counterband-induced migration irregularities affect the errors. Apart from the last two tracks (which undergo particularly severe UV degradation), the range within which the reconstructed position error varies is same for all proteins. In the ‘increasing pixel subset’ data series, the variations lie within 12.1 ± 0.9 mm. In the ‘moving pixel subset’ data series, the variations lie within 19.4 ± 4.1 mm.

Although this gives a range within which errors due to counterbands may lie, the actual contribution of counterbands to the errors of an arbitrary track is still unknown. This factor depends very much on how exactly counterbands intersect with protein trajectories. In Figure 2.77 the first four trajectories are shown, along with their residuals, with each linear subsection marked in a different colour.

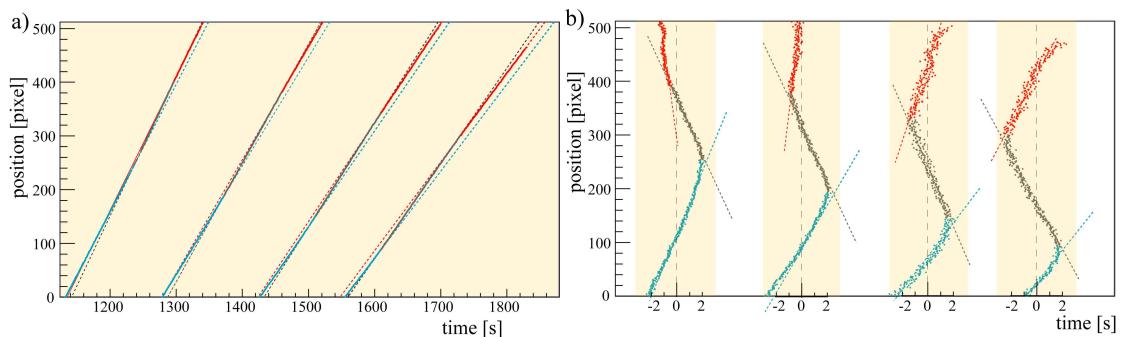


Figure 2.77: Protein migration trajectories: counterband subsections. a) Migration trajectories. b) Residual plots (trajectory as seen from fitted track). For each trajectory the first linear section is marked blue, the second section in grey, and the third section in red. Tracks fitted to the subsections are indicated with dashed lines. Separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μ m inner diameter capillary, Beckman SDS gel buffer). The four displayed tracks belong to 10, 20, 35, and 50 kDa proteins respectively.

It is possible analytically to describe the counterband migration trajectory, intercept the trajectories with an arbitrary protein migration track, and with that information describe how the protein migration trajectory changes. The occurrence of counterbands is highly dependent on experimental details such as how the capillary is clamped to a capillary holder, or fill levels of the electrode vials. Therefore it seems more useful to devote effort to ensuring robust separation conditions rather than attempting to develop a model that describes all these factors.

To give an approximation of the errors caused by counterbands, the ratio between the overall slope of trajectories in the detection window and the slope of the first trajectory subsection is determined. The first subsection merits special scrutiny because it has undergone the least counterband interference. The ratio of these slopes is constant for

proteins < 100 kDa, beyond which UV effects become dominant. For the tracks with molecular weight smaller than 100 kDa, this ratio is:

$$\mathbf{b_{S1}} = (1.052 \pm .005)\mathbf{b_{overall}} \quad (2.40)$$

with $\mathbf{b_{S1}}$ [s/mm] denoting the slope (inverse migration speed) of the track prior to counterband interference, and $\mathbf{b_{overall}}$ [s/mm] the slope of the track with counterband interference. With the corrected slope, using Equation 2.16 for error propagation, the contribution from counterbands on the reconstructed starting position error becomes:

$$\epsilon_{CB} = -\frac{\tau}{\mathbf{b_{overall}}} \left(1 - \frac{1}{1.052} \right) = -\frac{\tau}{\mathbf{b_{overall}}} 0.049 \quad (2.41)$$

where ϵ_{CB} [mm] is the error from counterband interference, and τ [s] is the migration time. With $\tau = L/(\mu E)$ this can be correlated with the mobility μ [mm²/(sV)], separation length L [mm], and electric field strength E [V/mm].

2.4.4.3 Estimate of errors from UV effects

The third characteristic error source, non-linear migration due to UV effects, is caused by UV illumination gradually generating protein fragments. As the smaller fragments begin to move ahead of the parent peak, the peak shape becomes asymmetric and the overall peak appears to accelerate. The migration trajectory becomes progressively steeper, and the fitted tracks deviate more and more from the true starting coordinates. Errors from UV effects depend on factors such as UV exposure, molecular weight, and sieving strength of the CGE buffer.

To quantify the effects of UV fragmentation, without the influence of other error sources, a Monte Carlo simulation was adapted.² Briefly, proteins of 10–225 kDa molecular weight are populated at a starting coordinate with a Gaussian position distribution. Next each protein is made to migrate according to a given molecular weight-mobility calibration function. Once the protein enters the illumination/detection window, in each time-step it absorbs photo-energy relative to its molecular weight, resulting in a probability for random fragmentation in that time step. The population of proteins is used to generate a set of data equivalent to photocurrent data from a photodiode array (based on UV absorption). These simulated data are then analysed the same as ordinary data. In Figure 2.78 the resulting simulated data are compared to real data.

²courtesy Phil Lewis, deltaDOT Ltd.

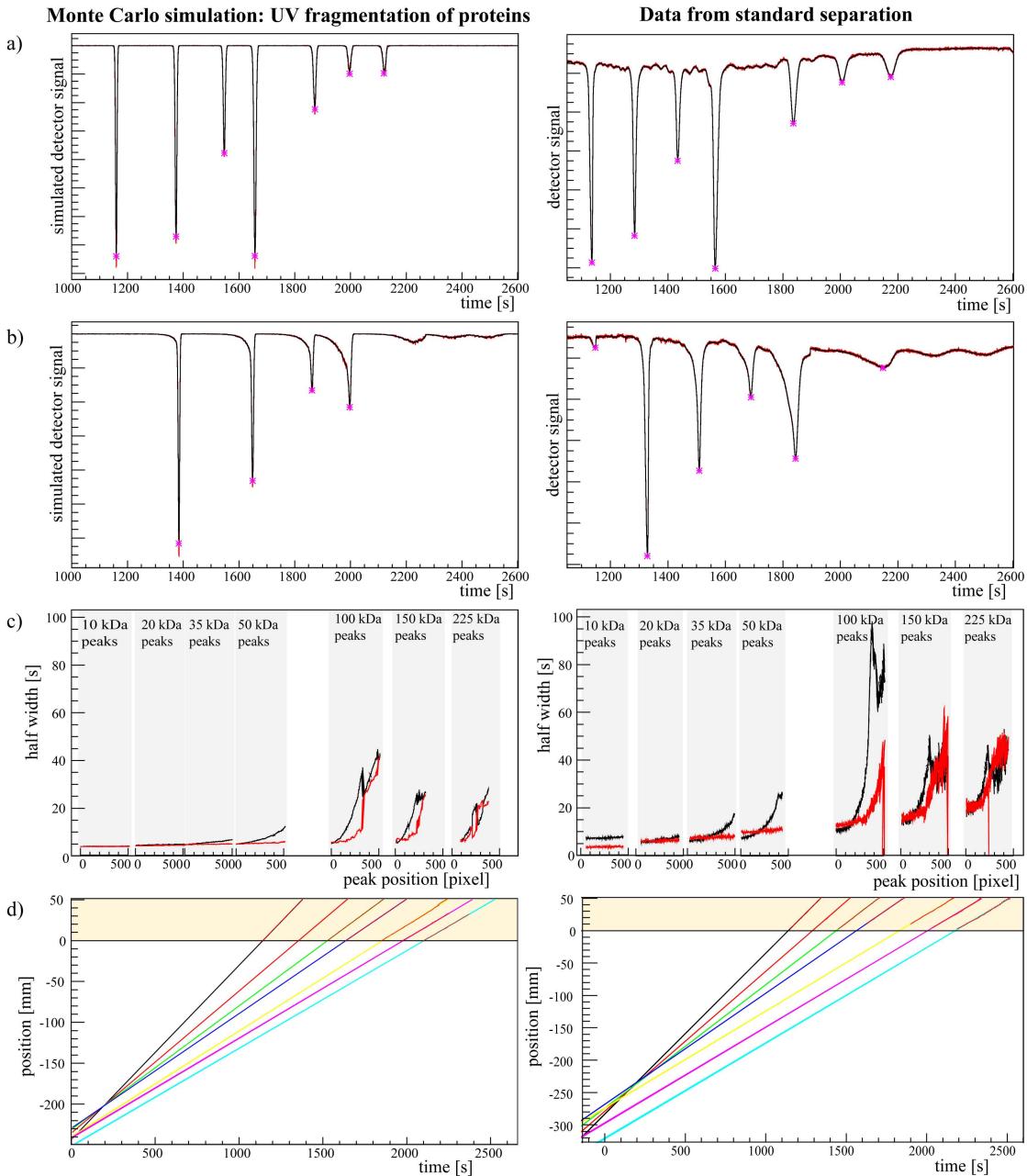


Figure 2.78: Monte Carlo simulation compared to real separation data. a) Electropherogram, first pixel. b) Electropherogram, last pixel. c) Half width at half maximum (peak asymmetry). d) Fitted tracks. All data are for identical conditions (200 mm separation distance, proteins molecular weight 10, 20, 35, 50, 100, 150, and 225 kDa, Beckman SDS gel buffer mobility calibration).

The Monte Carlo simulation is comparable to the real separation data, for which reason it is used to estimate the error contribution from UV mediated protein fragmentation. In Figure 2.79a the errors on the reconstructed starting coordinates are shown for proteins with different molecular weights. Notably, high molecular weight proteins suffer greater track distortion due to UV effects. This is caused in part by heavy proteins migrating more slowly, and thus receiving greater exposure; and in part by the size distribution of the fragments spanning a greater range (0.1–225 kDa vs. 0.1–10 kDa).

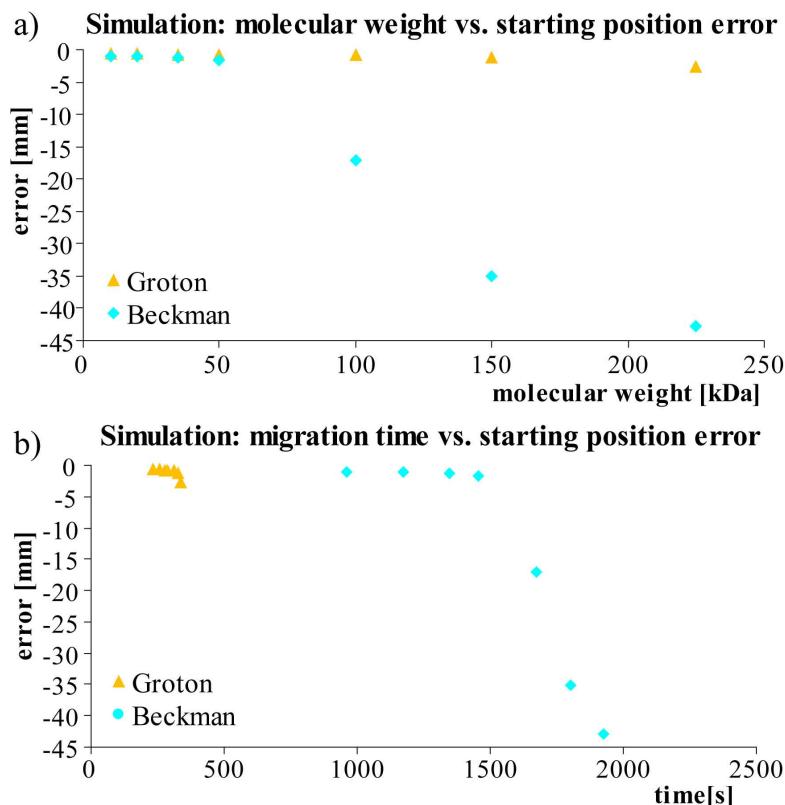


Figure 2.79: Monte Carlo simulation of UV induced fragmentation: errors on the reconstructed starting coordinates. a) Error vs. molecular weight, for two different buffers. b) Error vs. migration time, for two different buffers. All data are for identical conditions (200 mm separation distance, proteins molecular weight 10, 20, 35, 50, 100, 150, and 225 kDa), with Beckman SDS gel buffer (respectively Groton Optima gel buffer) mobility calibration.

If the migration calibration of another buffer, i.e. Groton buffer, is implemented, the UV effects are far less pronounced. This illustrates how the sieving strength of a buffer influences UV effects.

In Figure 2.79a, the error starts to increase significantly above a certain threshold. To study this the errors of the last four simulated tracks were evaluated for increasing pixel number (i.e. UV exposure time), see Figure 2.80.

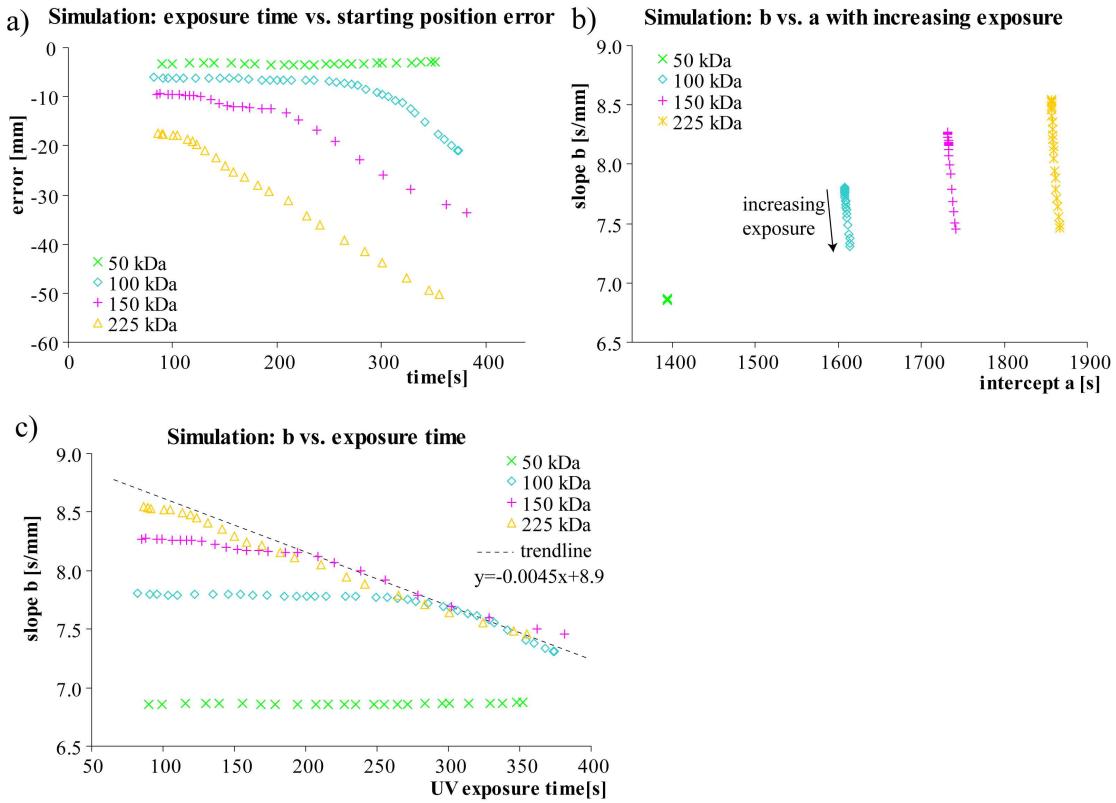


Figure 2.80: Monte Carlo simulation of UV induced fragmentation: increasing UV exposure. a) Error vs. UV exposure time for protein molecular weights 50, 100, 150, and 225 kDa. b) Development of track parameters **a** (intercept: migration time to first pixel) and **b** (slope: inverse migration velocity) with increasing UV exposure time. c) Track parameter **b** vs. UV exposure time. All data are for identical conditions (200 mm separation distance, protein molecular weight 50, 100, 150, and 225 kDa, Beckman SDS gel buffer mobility calibration).

In Figure 2.80a it is seen that the track error of a particular protein gradually increases, until a certain threshold is crossed, beyond which the error rapidly increases. In Figure 2.80b the two track parameters (intercept **a** and slope **b**) are plotted for increasing UV exposure. While the intercept only increases marginally, the slope of the tracks is strongly affected by UV exposure. In Figure 2.80c the slope of tracks of the same protein, but with increasing UV exposure is evaluated. Again a threshold-type behaviour is observed, where the slope **b** initially remains constant, and only starts to decrease above a certain threshold exposure.

Observing the 50 kDa protein in Figure 2.80, the slope (and error) do not vary greatly within the observation window. The next heavier protein (100 kDa) shows similar behaviour, up until a threshold, where the slope (**b** value) begins to drop off. It appears that the threshold depends on the molecular weight of the protein. For the 150 kDa protein a far lower exposure time induces the same behaviour, and for the heaviest protein (225 kDa) the exposure threshold is reached shortly after enough data are collected to

fit a track. This threshold behaviour is possibly due to a critical mass of fragmentation, before which the parent peak is still dominant.

In Figure 2.80c it is noteworthy that the \mathbf{b} values of different molecular weights form a common line, indicated with a dashed line in the figure. With increasing exposure \mathbf{b} values remain constant, until they intersect with this line, whereafter they decrease following the trendline. This observation is used to define a critical exposure, beyond which extra errors are introduced from changes of the track slope (\mathbf{b} value).

The critical exposure threshold δ^* [s] in dependence of the slope \mathbf{b}_0 [s/mm] (without UV exposure, or in approximation, the slope along the first couple of pixels) is:

$$\delta^* = \frac{8.9 - \mathbf{b}_0}{0.0045} \quad (2.42)$$

With a function describing the critical exposure threshold, an estimate for the errors below this threshold is developed. In Figure 2.79a the starting position error from simulated UV fragmentation data is plotted against the molecular weight. The proteins

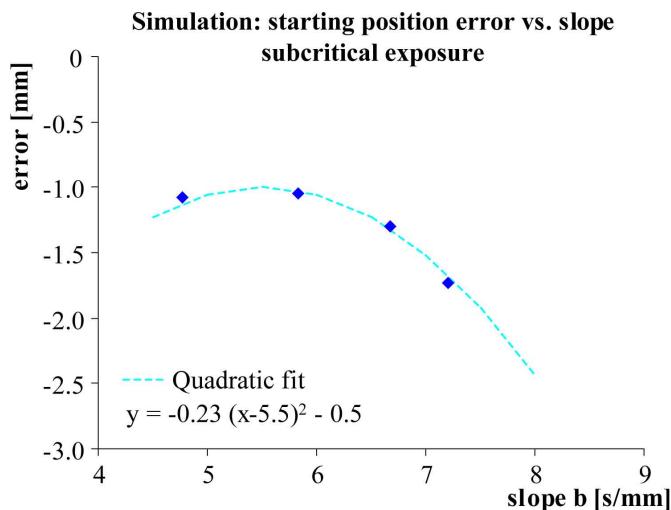


Figure 2.81: Monte Carlo simulation: function used to estimate errors from UV effects. Function describing the starting position errors of 10–50 kDa proteins (subcritical, baseline error $\epsilon_{baseline}$ [mm]) in dependence of the (measured) slope of the trajectory \mathbf{b}_0 .

10–50 kDa are below the critical exposure threshold, and are used to find a function describing the (subcritical) baseline error $\epsilon_{baseline}$ [mm], again correlated to the (measured) slope of the trajectory. Here a quadratic function describes the errors best, as illustrated in Figure 2.81.

$$\epsilon_{baseline} = -0.23(\mathbf{b}_0 - 5.5)^2 - 0.5 \quad (2.43)$$

Finally, an estimate for the errors above the critical exposure threshold is developed. The additional ‘supercritical’ error $\epsilon_{supercritical}$ [mm] is caused by the \mathbf{b} (track slope) values dropping off from the baseline \mathbf{b}_0 values, as observed in Figure 2.80c. Furthermore, the \mathbf{b} values for different molecular weights begin to follow the same function (dashed trendline in Figure 2.80c). The slope of this line is -0.0045 [mm]. Using this, the \mathbf{b} value at an arbitrary time after super critical exposure is:

$$\mathbf{b} = \mathbf{b}_0 - 0.0045(\delta^{total} - \delta^*) \quad (2.44)$$

where \mathbf{b}_0 [s/mm] is the track slope prior to UV exposure (or in approximation the track slope of the first 100 pixels), δ^{total} [s] is the total exposure time, and δ^* [s] is the critical exposure time, as expressed in Equation 2.42. With this expression for the slope, the additional error this contributes to the error in the reconstructed starting coordinate is found using $distance x = \frac{migrationtime \tau}{\mathbf{b}}$:

$$\epsilon_{supercritical} = x - x_0 = \frac{\tau}{\mathbf{b}} - \frac{\tau}{\mathbf{b}_0} \quad (2.45)$$

$$\begin{aligned} \epsilon_{supercritical} &= -\tau \left(\frac{1}{\mathbf{b}_0 - 0.0045(\delta^{total} - \delta^*)} - \frac{1}{\mathbf{b}_0} \right) \\ &= -\tau \left(\frac{1}{8.9 - 0.0045\delta^{total}} - \frac{1}{\mathbf{b}_0} \right) \end{aligned} \quad (2.46)$$

Equations 2.42, 2.43, and 2.46 are used to describe the critical exposure time δ^* [s], the error if the exposure is below the critical exposure $\epsilon_{baseline}$ [mm], and the additional error for exposure above the critical exposure $\epsilon_{supercritical}$ [mm]. The necessary values are the initial track slope (or for first 100 pixels) \mathbf{b}_0 [s/mm], the migration time to the first pixel τ [s], and the total exposure time δ^{total} [s].

Although these expressions are purely descriptive and fitted to a Monte Carlo fragmentation simulation, they describe the actual data behaviour well. For real data the numerical values in Equations 2.42, 2.43 and 2.46 are adjusted slightly, which describes actual (not simulated) data better:

$$\delta^* = \frac{8.1 - \mathbf{b}_0}{0.005} \quad (2.47)$$

$$\epsilon_{baseline} = -0.23(\mathbf{b}_0 - 4)^2 - 0.5 \quad (2.48)$$

$$\epsilon_{supercritical} = -\tau \left(\frac{1}{8.1 - 0.005\delta^{total}} - \frac{1}{\mathbf{b}_0} \right) \quad (2.49)$$

In Table 2.2 the estimates are summarised for errors on the reconstructed starting position from different causes (thermal, counterbands, UV effects). In Table 2.3 these estimates are applied to a set of standard data, and the resulting overall error estimates compared with the actual starting position errors.

Although there are minor differences between estimated and observed starting position errors (up to 3 mm overestimate), the standard deviation is 1.7 mm, which is a vast improvement over the uncorrected 10–70 mm error. Optimisation of the estimate functions could further improve the accuracy of the error estimates.

Table 2.2: Estimates of errors for different types of non-linear migration.

Error source	Estimate ¹
Non-linear migration from thermal effects	$\epsilon_\Theta = -0.0042\tau - 0.2678$
Counterband crossing	$\epsilon_{CB} = -\frac{\tau}{b_{\text{overall}}} 0.049$
UV degradation	<p>critical exposure $\delta^* = \frac{8.1 - b_0}{0.005}$</p> $\delta^{total} \leq \delta^* : \quad \epsilon_{UV} = -0.23(b_0 - 4)^2 - 0.5$ $\delta^{total} > \delta^* : \quad \epsilon_{UV} = -0.23(b_0 - 4)^2 - 0.5$ $- \tau \left(\frac{1}{8.1 - 0.005\delta^{total}} - \frac{1}{b_0} \right)$

¹For samples 10–225 kDa, using Beckman SDS gel buffer on custom built instrument.

With thermal error ϵ_Θ [mm], error caused by counterband interference ϵ_{CB} [mm], and error from UV effects ϵ_{UV} [mm]. τ [s] is the migration time to the first pixel, and can be substituted with $\tau = L/(\mu E)$ with separation length L [mm], mobility μ [mm²/(sV)], and electric field strength E [V/mm]. b_{overall} [s/mm] is the track slope, b_0 [s/mm] is the initial track slope (or for first 100 pixel), δ^{total} [s] the total exposure time.

Table 2.3: Comparison between observed starting position errors and estimated starting position errors for sample separation data¹.

MW	τ	$b_{overall}$	b_0	δ^{total}	δ^*	ϵ_Θ	ϵ_{CB}	ϵ_{UV}	$\epsilon_{estimated}$	$\epsilon_{observed}$
10	857	4.043	4.248	207	770	-3.87	-10.39	-0.51	-14.77	-12.06
20	1004	4.689	5.007	240	619	-4.48	-10.49	-0.73	-15.71	-14.14
35	1151	5.309	5.713	272	477	-5.10	-10.63	-1.18	-16.91	-16.88
50	1278	5.861	6.222	300	376	-5.64	-10.69	-1.64	-17.96	-18.09
100	1546	6.697	6.680	342	284	-6.78	-11.31	-12.71	-30.79	-30.91
150	1719	6.850	7.194	326	181	-7.49	-12.29	-29.58	-49.36	-50.89
225	1888	7.056	7.848	300	50	-8.20	-13.11	-49.36	-70.67	-67.55

¹Contributions are shown for thermal errors ϵ_Θ [mm], errors caused by counterband interference ϵ_{CB} [mm], and errors from UV effects ϵ_{UV} [mm]. τ [s] is the migration time to the first pixel, $b_{overall}$ [s/mm] is the track slope, b_0 [s/mm] is the initial track slope (or for first 100 pixel), δ^{total} [s] the total exposure time, and δ^* [s] the critical exposure time. Separation conditions are standard (200 mm separation length, 350 V/cm electric field strength, 75 μ m inner diameter capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL), using custom built instrument.

2.4.5 Expected performance of V2D separations

The data gathered so far can be used to generate a mock virtual 2D (V2D) separation, and quantify the expected performance of a V2D separation. Double injection data are used as the basis for the V2D mock-up, in order to include realistic band crossings and a variety of starting coordinates. In Figure 2.82 the double injection track plot along with its V2D mock-up is shown, with a conventional 2D gel image for comparison.

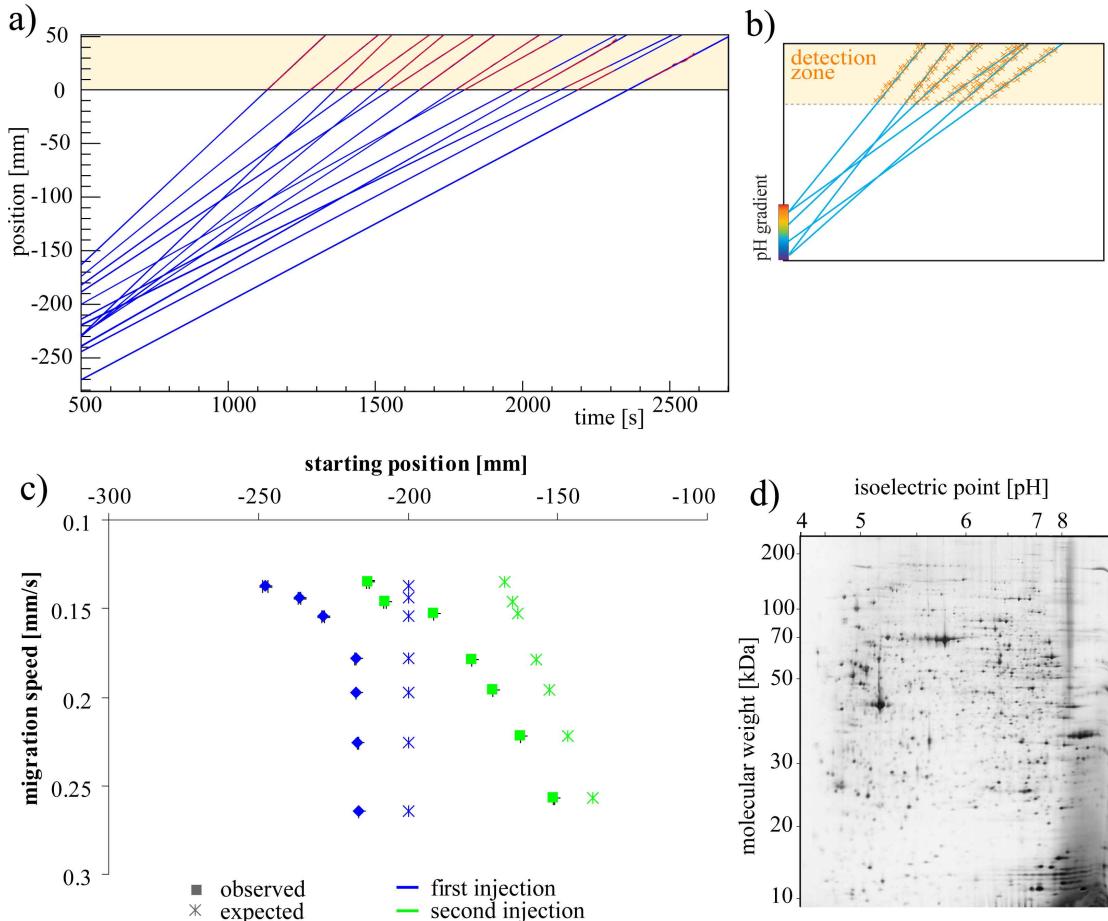


Figure 2.82: Mock V2D separation. a) Reconstructed migration tracks of the double injection separation used for a V2D mock-up. With observed trajectory (red), and fitted track (blue). b) Schematic illustration of migration tracks in a V2D separation. c) Mock V2D: starting coordinate (related to π) against migration speed (related to MW). The second injection time is used as the mock-up separation starting time. Asterisk symbols indicate expected values, full symbols observed values. Standard separation conditions apply (200 mm separation length, 350 V/cm electric field strength, 75 μm ID capillary, Beckman SDS gel buffer), with standard sample (7 proteins, 10–225 kDa, 3.2 mg/mL). Sample was injected twice with an interval of 4 minutes. d) 2D gel image with π against migration speed (human lymphoma, from ExPASy^[46]).

In Figure 2.82c the uncorrected reconstructed starting coordinates of tracks are shown. The errors are associated with the variance of the observed migration trajectory around the fitted track. Proteins of the first injection are marked in green, proteins of the second

injection in blue. Since this is a double injection, it is known that one set of proteins started at a given position (-200 mm in this case), and the other has migrated (for 4 minutes in this case) from that same starting position and is already farther along the capillary. This second group of proteins is expected to lie on a straight line (because the distance each protein travels in the same time is linearly proportional to its speed). In Figure 2.82c the expected positions of proteins are indicated with asterisk symbols. Clearly the reconstructed positions need correction before the data are of reasonable significance.

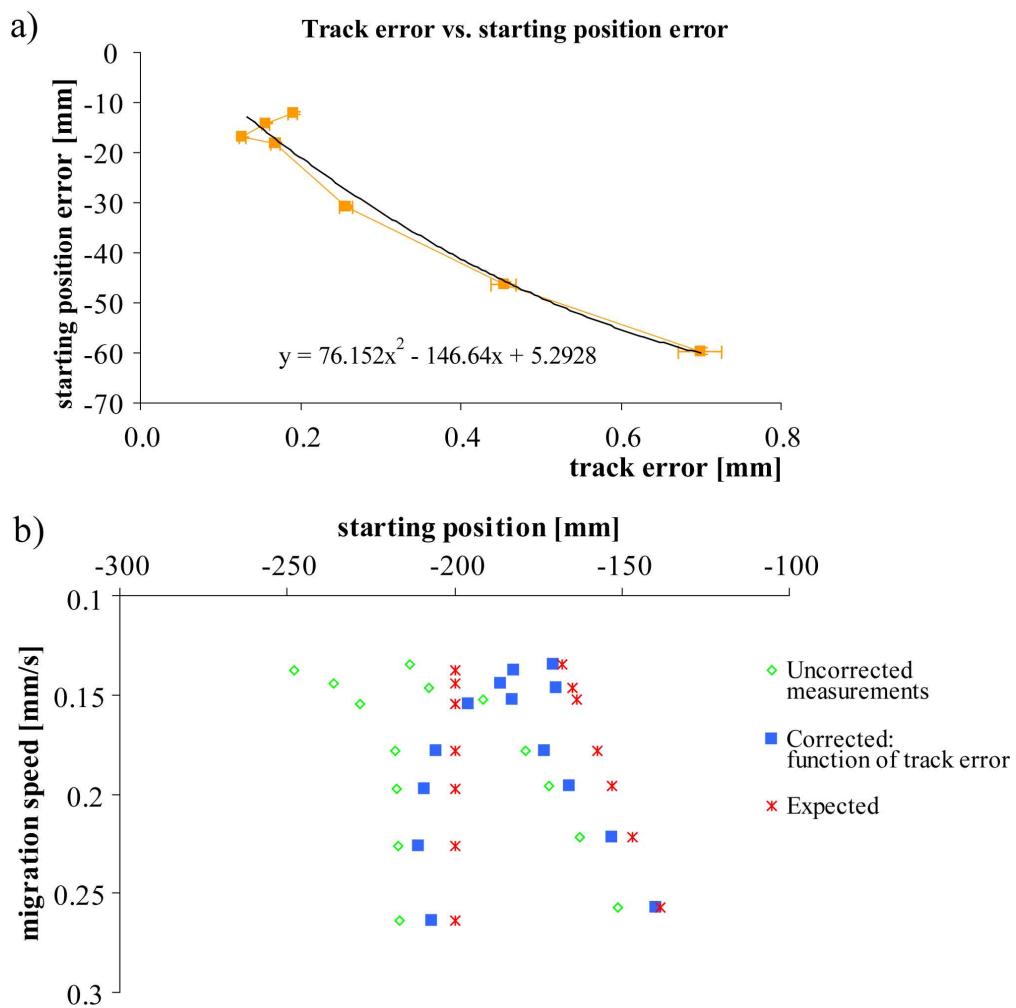


Figure 2.83: Calibration and correction of mock V2D separation. a) Calibration function (quadratic) for correlation of track error and starting position error, data from a standard (single injection) separation. b) Mock V2D with correction by track error calibration. Starting coordinates (related to pI) are plotted against migration speeds (related to MW). Red asterisk symbols indicate expected values, open green symbols uncorrected values, and blue solid symbols corrected values. Data is from a double injection separation (injection interval 4 minutes), and the second injection time is used as the separation starting time. All other separation conditions are standard, with standard sample (7 proteins, 10–225 kDa).

In Section 2.3.1.3 it was observed that the track error is loosely proportional (correlation coefficient 0.73) to the actual ('starting position') error. An error correction with the track error seems plausible because the track error contains influences from all error sources. Also, the track error is a value that is easily available for an arbitrary track. In Figure 2.83 the correction resulting from a linear correction function is seen. Comparing the uncorrected positions with the corrected ones, there is a noticeable improvement, however the inaccuracy is still considerable.

Instead of applying a correction that is correlated to the track errors, the more refined error estimates defined in Section 2.4.4 can be used. In Figure 2.84 the same mock-up V2D gel with error estimate corrections is seen.

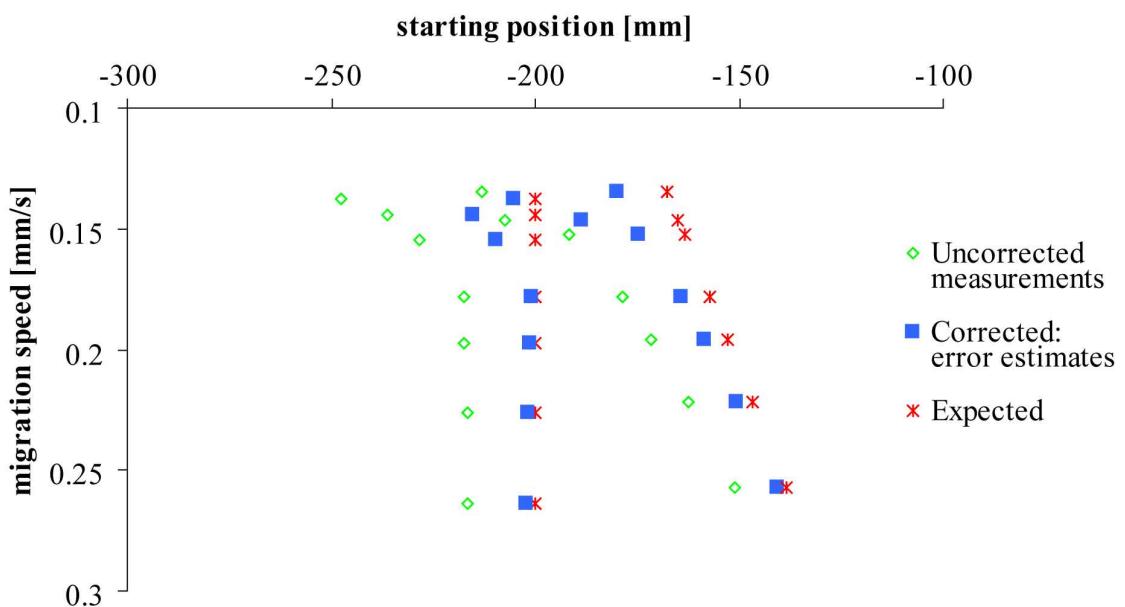


Figure 2.84: Correction of mock V2D separation with error estimates. Starting coordinates (related to pI) are plotted against migration speeds (related to MW). Red asterisk symbols indicate expected values, open green symbols uncorrected values, and blue solid symbols corrected values. Data is from a double injection separation (injection interval 4 minutes), and the second injection time is used as the mock-up separation starting time. All other separation conditions are standard, with standard sample (7 proteins, 10–225 kDa).

In Figure 2.85 the different corrections and their performance are compared. For lighter (faster) proteins (< 100 kDa) the correction with the error estimate produces quite satisfactory results, with the uncertainty of the starting positions in the range of 1–7 mm.

Assuming a pH gradient length of 50 mm, a 2 mm starting position error corresponds to a relative error of 4 %. If the pH gradient spans 7 pH units (pH 4–11), then the pI error is 0.28 pH units. Here up to 25 peaks can be separated, or 500 peaks per metre. These figures of merit are only for the first, reconstructed IEF dimension of the

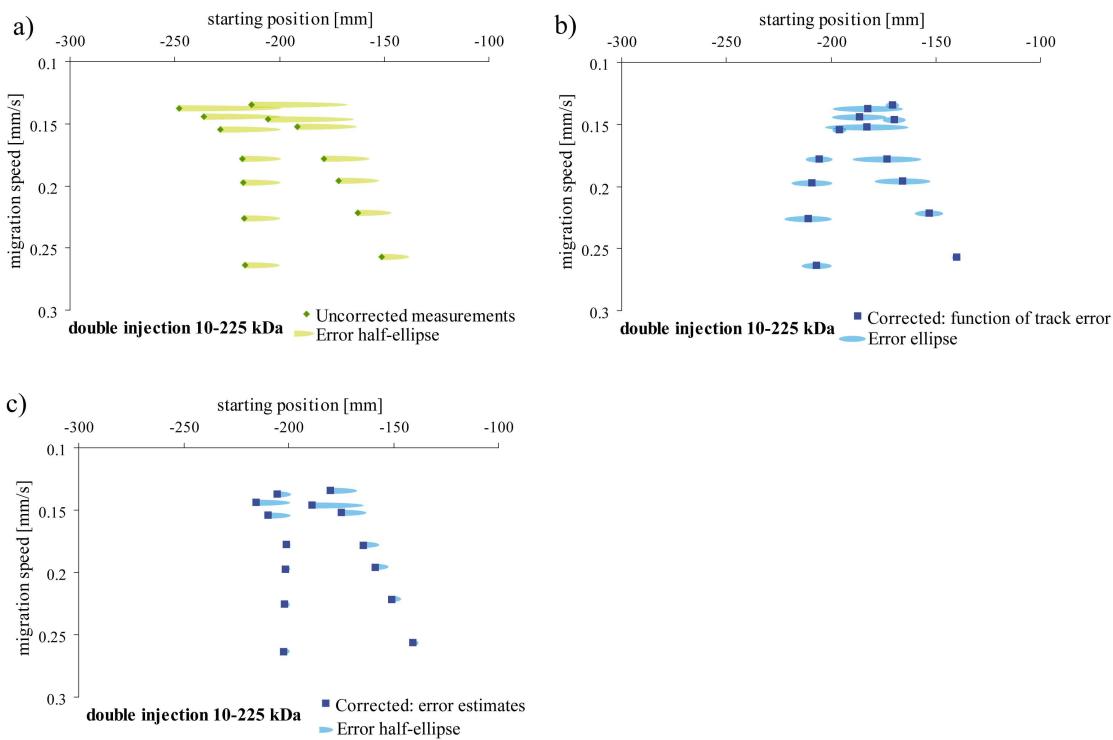


Figure 2.85: Comparison of different error corrections for V2D separations. a) Uncorrected mock V2D separation. b) Mock V2D separation with corrections as a function of track errors. c) Mock V2D separation with correction by the error estimates from Section 2.4.4. Starting coordinates (related to pI) are plotted against migration speeds (related to MW). Data is from a double injection separation (injection interval 4 minutes), and the second injection time is used as the mock-up separation starting time. All other separation conditions are standard, with standard sample (7 proteins, 10–225 kDa).

separation. For the second, gel electrophoresis dimension, the performance relates to the capillary gel electrophoresis separation. For a typical capillary gel electrophoresis separation under the conditions used here, the theoretical plate number is 24600 plates, or 123000 theoretical plates per metre. The height equivalent to one theoretical plate is 0.008 mm, and the resolution between two adjacent peaks is 3.9. The peak capacity is 39 peaks. For two orthogonal separations that are combined, the peak capacity is the product of the individual peak capacities. In this case the V2D separation has a peak capacity of 975 peaks. This is more than either separation individually, and more than a standard capillary gel electrophoresis separation can achieve under optimal conditions, but it is less than conventional two-dimensional gel electrophoresis. For longer separation lengths (e.g. 100 mm IEF, 600 mm CGE) higher peak capacities up to 5850 peaks can be obtained, which compares well to two-dimensional gel electrophoresis.

Particularly important factors for migration reconstruction are quick separation and rigorous temperature control. Microfluidic chips are especially favourable in both respects.

Table 2.4: Estimates for overall relative errors of a V2D separation under optimal experimental conditions.

V2D accuracy	
migration speed	< 1 %
starting position	4 %
peak area	3 %

Higher molecular weight proteins (100 kDa and over) suffer greater uncertainties regarding starting position, in the range of 1–2 cm. Presumably the correction could be improved for a given microfluidic chip, and with optimised instrumentation and buffers uncertainties in the range of millimetres should be possible, even for higher molecular weight proteins. This illustrates the importance of a relatively long pH gradient. Although from a microfluidic perspective a pH section as short as 1 cm is more obvious, from a tracking perspective short pH gradients result in poor sensitivity. In Table 2.4 the potential accuracy of a V2D separation under ideal conditions is summarised.

Chapter 3

V2D Separations

3.1 V2D separations in a two-buffer system

3.1.1 Introduction

In Section 1.5 the concept of virtual two-dimensional (V2D) separations was proposed in the form of collinear isoelectric focusing (IEF) and capillary gel electrophoresis (CGE), where the separation results of the individual separations are reconstructed by tracking. While IEF and GE are the most customary techniques for 2D protein electrophoresis separations, this approach requires the development of an elaborate microfluidic structure with an integrated pH gradient. A simpler way to combine two collinear separations is to start with a conventional capillary-based separation, but implement two different buffer sections in the same capillary.

In a separation with two buffer sections in the same capillary, a plug of protein mixture starts to migrate in the first section. Eventually a given protein passes through the interface into the second section, where its migration behaviour changes due to a change in the chemical conditions. After migration in the second section, the protein passes a detection window, where the migration trajectory is observed. From this track, the protein migration speed in the second buffer is known. Assuming the interface location is known (either by measurement or by prior positioning), the interface crossing can be reconstructed. With this and with information on injection time and position, it is possible to reconstruct the migration speed in the first section. From one set of measurements two different separations are resolved. This is illustrated in Figure 3.1.

The concept of changing the separation conditions mid-run has been proposed in earlier work, in particular dynamic change of pH.^[121,169,227] Although these reports include

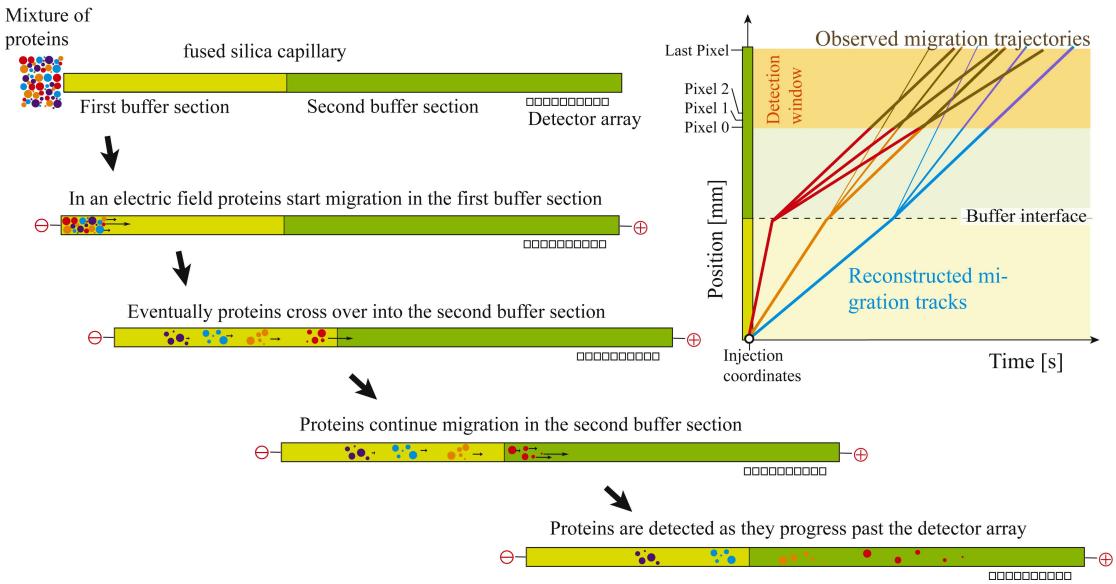


Figure 3.1: Schematic illustration of a V2D separation in two different buffers. After proteins migrate and separate in a first buffer section, they pass across an interface into a second buffer section. Here the migration behaviour changes due to a change in chemical conditions. After migration in the second section, the proteins pass a detection window, where the migration of each protein group is tracked. With this migration trajectory, the buffer interface crossing can be reconstructed. Given the injection coordinates, the migration speed in the first section is reconstructed.

schematic diagrams of the protein migration trajectories, these are exclusively for illustration. The aim of these separations is to improve performance by selectively shifting peaks that would otherwise co-migrate, and not to reconstruct separations under different conditions. By contrast, in this work the concept of tracking is applied to ‘two-condition’ separations. If the migration trajectory of proteins under different conditions is determined, it becomes possible to compare the migration change and gain additional information on the separating proteins.

While the capillary-based format is attractive in regards to simplicity, it does put some restrictions on the choice of separation modes. Since two buffers share the same capillary, they need to be compatible. Otherwise unwanted behaviour may interfere with the separations. Problems can arise from differences in electro-osmotic flow in the two sections, isotachophoretic behaviour if the background electrolyte varies,^[228,229] variation in the electric field strength if the buffer conductivities are mismatched,^[230] as well as migration of any charged buffer components such as SDS.^[40]

The two separations we chose to investigate in a model separation system are capillary zone electrophoresis (CZE) and sieving capillary zone electrophoresis. As described in Section 1.3.2, CZE defines an electrophoretic separation in a non-sieving matrix where the native charge of a protein (which depends on the protein’s pI relative to the buffer

pH) causes electrophoretic migration. Separation occurs according to the charge-to-size ratio of different proteins. In sieving CZE, the driving charge is still the native charge of a protein, but like in capillary gel electrophoresis, the buffer contains a sieving matrix. This affects small proteins less than large proteins, and contributes a size-dependent factor to the overall electrophoretic migration.^[203] Comparison of a protein's CZE mobility with its sieving CZE mobility then gives an indication of the protein's size.

Unlike in conventional 2D gel electrophoresis, the two separations are not orthogonal (independent), but somewhat linked. As previously mentioned, the conditions need to be compatible, resulting in a trade-off between compatibility and orthogonality.

One more issue needs consideration, namely how to control the position of the interface between the two different buffer sections. Either the position of the interface is defined, and does not vary in the course of the separation; or the interface moves in the course of the separation, and needs to be observed. Whether the interface position moves or is stationary depends on the electro-osmotic flow (EOF). Unless the EOF is zero, the interface is dragged along.

The EOF can be controlled with additives and surface treatments,^[210,231,232] and two segments of capillary with different filling can be coupled. Coupling of two capillary segments is likely to bring losses compared to an uninterrupted capillary, and the position of the buffer interface is not directly observable but must be assumed to be at the capillary joint.

To avoid these problems, the separation can be set up as normal polarity CZE, with strong EOF and sample injection at the anode end. If the anodic reservoir is changed to the second buffer after a given time, the EOF will draw the second buffer into the capillary and proteins migrating against the EOF will eventually migrate into the second buffer section. Both the interface and finally the proteins pass the detector near the cathode. This is illustrated in Figure 3.2. Here the interface position is directly observed and the problems arising from coupling are avoided.

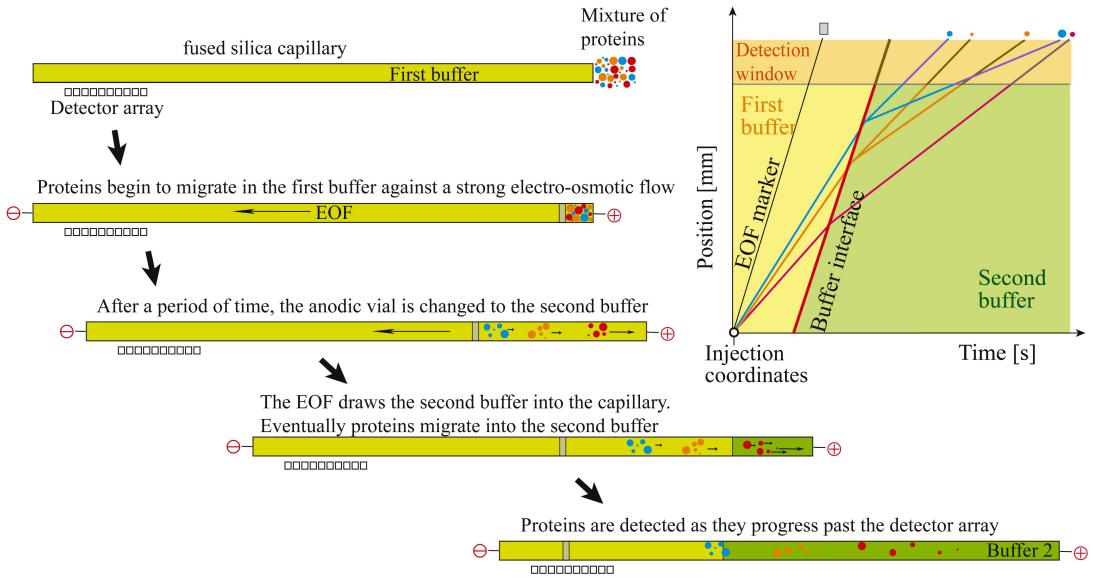


Figure 3.2: Schematic illustration of a V2D separation in two different buffers with a moving interface. After sample injection, the anodic reservoir is changed to the second buffer after a given time. The electro-osmotic flow (EOF) draws the second buffer into the capillary. Proteins migrating against the EOF eventually migrate into the second buffer section, where their migration behaviour changes due to different chemical conditions. Both the interface and finally the proteins pass the detector near the cathode. From the migration trajectory of the interface and a protein, the buffer interface crossing can be reconstructed. Given the injection coordinates, the migration speed in the first section is reconstructed.

3.1.2 Experimental methods

Instrumentation and data analysis: The instrumentation and data analysis package are the same as previously described in Chapter 2.2. Briefly, a custom built capillary electrophoresis system was used for standard measurements. Along an optical rail are positioned: a deuterium light source, imaging optics and wavelength filters, a capillary stage, and the detector. The capillary ends are immersed in vials with buffer and platinum electrodes that are connected to a high voltage power supply. The detector is a linear photodiode array with 2048 pixels, each $2.5\text{ mm} \times 25\text{ }\mu\text{m}$. Data are digitised and acquired at 10 Hz and the values of four adjacent pixels are averaged. This comes to the same effect as a detector with 512 pixels of $100\text{ }\mu\text{m}$ width each.

A set of custom offline analysis packages (ENIUS, deltaDOT) were used and adapted to requirements. To outline the data analysis process, the program begins by loading detector data from a file, then identifies so-called ‘hits’ (where something passes the detector), and finally fits tracks through the found hits. The slope and intercept for the fitted tracks are displayed for further use. Various operations can be performed with the data, for instance plotting the reconstructed migration trajectory.

Experimental procedures and reagents: Fused silica capillaries (PolyMicro Technologies) with 75 μm inner diameter and 375 μm outer diameter are used for all experiments. Standard separation length (to first pixel) is 60 cm, with a 5.5 cm imaging window and 75 cm total capillary length. Capillaries are subjected to a cleaning routine immediately prior to every separation (flush under pressure for 15 minutes with 1 M NaOH, 3 minutes ultrapure water, 3 minutes separation buffer, and 3 minutes separation buffer into cathode vial) as well as a special cleaning routine for new capillaries (5 min methanol, 3 min ultrapure water, 5 min 0.1 M HCl, 3 min ultrapure water, 15 min 1 M NaOH, 3 min ultrapure water, 3 minutes separation buffer, and 3 minutes separation buffer into cathode vial).

Standard CZE buffer is a sodium tetraborate buffer, 50 mM, pH 9.3 (from Sigma-Aldrich Co., 99.995 %). For sieving matrices, the following polymers were evaluated: polyethylene oxide, MW 5,000 kDa (Avocado Research Chemicals Ltd.); polyethylene oxide, MW 300 kDa (Alpha Aesar); dextran, MW 70 kDa (Sigma-Aldrich Co.); and dextran, MW 6 kDa (Sigma-Aldrich Co.). The composition of the standard sample, with 5 proteins and a marker for the electro-osmotic flow (all from Sigma-Aldrich Co.), is summarised in Table 3.1.

The sample is dissolved in the same sodium tetraborate solution as is used for the run buffer. The sample is injected electrokinetically (as described in Chapter 1.3.2), with 160 V/cm field strength for 5 seconds. Standard separations proceed at 160 V/cm. For the two-buffer separations, after 45 seconds of separation, a 5 second electrokinetic injection of only EOF marker in the run buffer precedes the second buffer section. This serves to mark the interface. The second buffer section is introduced simply by exchanging the reservoir vial at the anode, and letting the electro-osmotic flow draw it into the capillary.

Table 3.1: Standard sample for two-buffer V2D separations.

Protein	Molecular weight [kDa]	Isoelectric point	Concentration [mg/mL]
Ubiquitin	8.5	6.3	0.1
Myoglobin	17	7	0.3
α -Lactalbumin	14.2	4.3	0.3
Fetuin	48	3.5	0.6
Trypsin inhibitor	20	4.6	1.5
Thiourea	(EOF marker)	(uncharged)	0.03

Numerical simulation: To assess the effects of migration on the interface between the two buffer sections, a finite element simulation of the concentration dispersion by

hydrodynamic and electro-osmotic flow was performed using COMSOL Multiphysics software (from COMSOL AB., software formerly known as FEMLAB). The exact location of the interface is critical in reconstructing the migration trajectories. If the interface itself is not sharp, the uncertainty of the reconstruction will increase.

Within the COMSOL Multiphysics software, the following equation sets were combined: incompressible Navier-Stokes equations and conductive media direct current equations to describe the electro-osmotic flow (mmgfl and emdc, from the MEMS module), and convection and diffusion equations (chcd, from the chemical engineering module). The electro-osmotic mobility is defined as a function of concentration, and so is the viscosity. To prevent singularities, the functions are formulated with smoothing functions instead of step functions (hyperbolic tangent functions; see Appendix B for details).

To avoid the model becoming very large, only the buffer interface region in a 20 mm capillary section is simulated. As the interface itself migrates, this requires a moving reference system. The capillary is simplified to a two-dimensional model, because radial symmetry may be assumed. The very simple geometry (a long rectangle) makes a structured mesh of rectangles most suitable. The dimensions of the model for simulation are 20×10^{-3} m length and 75×10^{-6} m height (diameter), and the mesh rectangles are 10×10^{-6} m by 12.5×10^{-6} m. The initial conditions are a concentration distribution with a smoothed interfacial region of 50×10^{-6} m, less than the capillary diameter. Furthermore, there is no pressure at the capillary outlets, and the concentration at the outlets is constant.

Simulations are run over a time span of 900 seconds. For all functions and parameters the values that correspond to the actual system were measured and evaluated. Appendix B lists the values and functions that were applied. Figure 3.3 shows the model capillary with the initial concentration distribution.

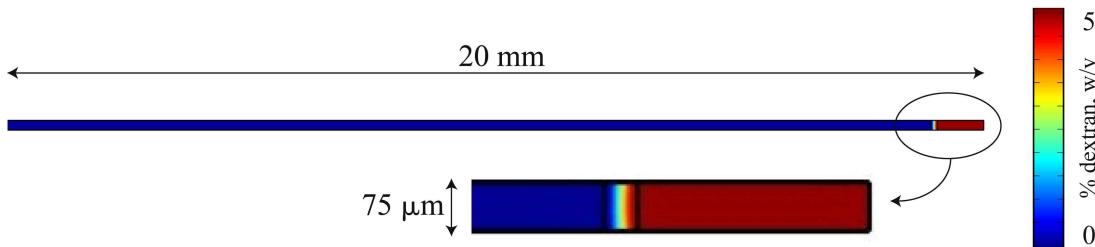


Figure 3.3: COMSOL model of the capillary with enlarged image of the buffer interface, initial conditions. The colourscale to the right indicates concentration [% w/v] of the sieving matrix component.

3.1.3 Results

First investigations were directed toward finding a sieving buffer that is compatible with the non-sieving capillary zone electrophoresis (CZE) buffer, and at the same time causes a change in protein mobility. The properties that are particularly important for compatibility are conductivity, viscosity, and electro-osmotic mobility. If the buffers have different conductivities, then the electric field strength is different in the two sections. Well-matched electro-osmotic flow (EOF) is important to ensure that the interface between the two sections is not unnecessarily spread out. A large difference in viscosity will cause a non-uniform flow profile.

As discussed in Chapter 1.3.2, the sieving property of a buffer is imparted by adding a soluble polymer that creates an entangled network with pores.^[69] For UV absorption detection, the buffer ought to be transparent at the detection wavelength, 214 nm here. Suitable polymers include polyethylene oxide and dextran, amongst others.^[40,88] These polymers are available in good purity and designated molecular weights.

In Figure 3.4a the change in mobility ('sieving response') of a sample protein, trypsin inhibitor, for different buffers is shown. In Figure 3.4b the EOF of these buffers is seen. The viscosity of PEO solutions becomes too great for capillary filling under pressure for solutions over 0.5 % (for 5,000 kDa MW), respectively 3 % (for 300 kDa MW). By comparing the two graphs in Figure 3.4 it is seen that PEO affects the EOF more strongly than dextran for the same sieving effect. For this reason dextran based buffers were chosen as a sieving system.

Theoretical models predict that the entanglement threshold (critical overlap concentration) for 70 kDa dextran is around 4.8 % (after Sartori *et al.*^[69]). A concentration of 5 % dextran, just above the entanglement threshold, was chosen, for at this concentration a change of mobility is observable, while minimising the difference of the EOF.

Addition of 5 % dextran decreases the conductivity of the buffer slightly. To adjust the buffer conductivities, the concentration of sodium tetraborate in the CZE buffer was reduced from 50 mM to 42 mM. Figure 3.5 shows the current for 5 % dextran in 50 mM sodium tetraborate buffer (referred to as 'sieving buffer' hereafter), the conductivity-adjusted 42 mM sodium tetraborate buffer (referred to as 'CZE buffer' hereafter), and also the conductivity of the 50 mM sodium tetraborate solution.

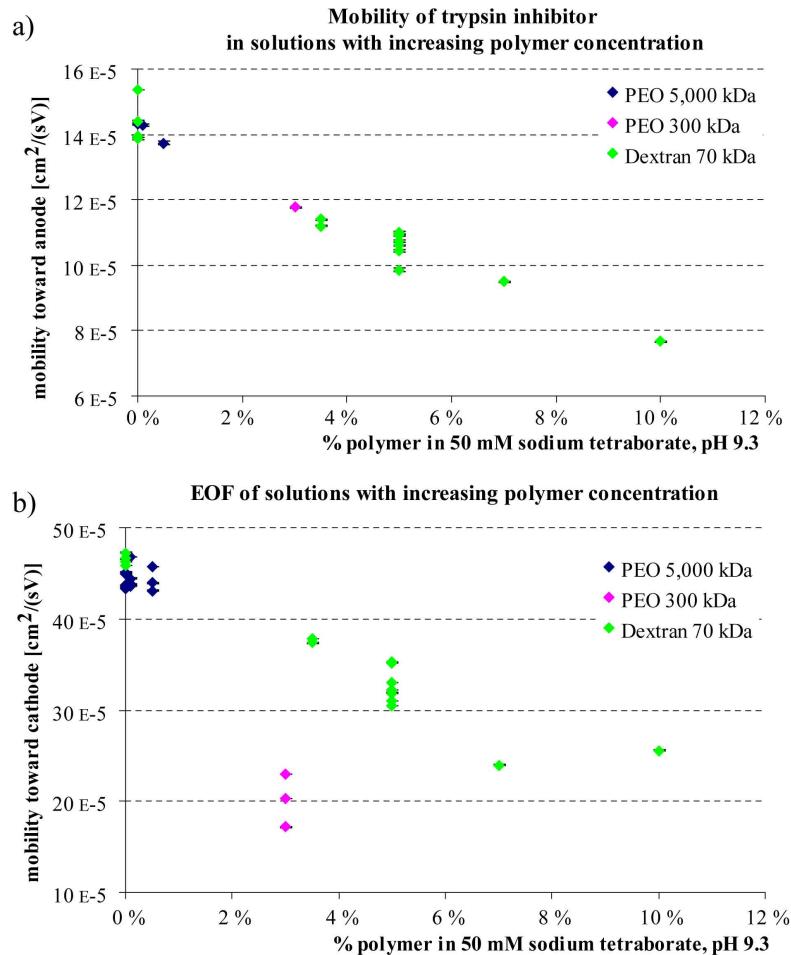


Figure 3.4: Effect of adding different polymers to CZE buffer solution. a) Mobility of a sample protein (trypsin inhibitor, MW 20 kDa) in the direction opposite to the EOF. With increasing polymer concentration, the mobility decreases. b) Electro-osmotic mobility vs. concentration for different polymers. With increasing addition of polymer to the solution, the EOF decreases. This is more pronounced for PEO than for dextran.

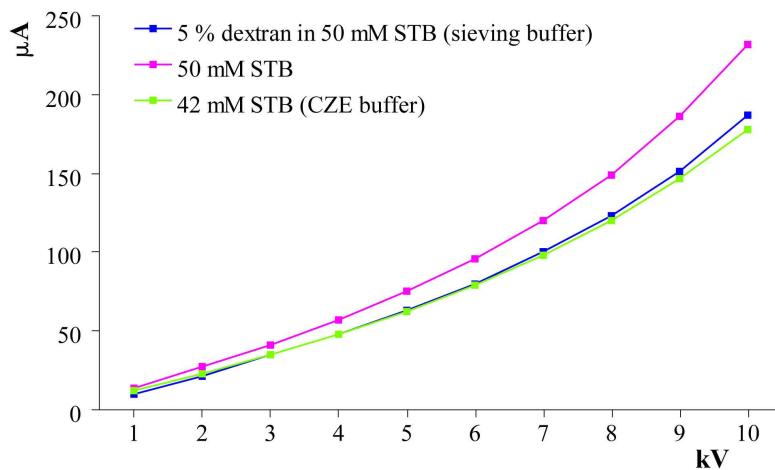


Figure 3.5: Current across a capillary (75 μ m inner diameter, 35 cm total length) with different buffer fillings. In the CZE buffer the concentration of the background electrolyte (sodium tetraborate) needs to be adjusted from 50 mM to 42 mM to ensure the conductivity is same as for the sieving buffer (5 % dextran in 50 mM sodium tetraborate).

With the buffer system characterised, a COMSOL simulation of the interface region was performed. The different physical constants are defined for a 5 % dextran (70 kDa) solution in 50 mM sodium tetraborate. Figure 3.7 shows how the interface develops over time, as well as a plot with concentration against position after 900 seconds. In Figure 3.7b the width of the 70 % concentration span is indicated. This is approximately 4 mm.

If the interface position assumed for trajectory reconstruction is incorrect by 4 mm, the resulting error in reconstructed migration in the unobserved buffer section can be characterised. This is illustrated in Figure 3.6 with data from a typical two-buffer reconstruction. Here the relative error of the reconstructed migration speed, as contributed by the interface uncertainty, is in the range of 1.3 %.

From the error analysis results previously presented, it is known that CZE tracks do not demonstrate significant non-linear migration due to UV effects, and background buffer migration effects that can cause non-linear migration are not observable either. Only thermal effects play much of a role, and these are of the same magnitude for most tracks. With regards to track reconstruction, the combination of CZE and sieving CZE is advantageous.

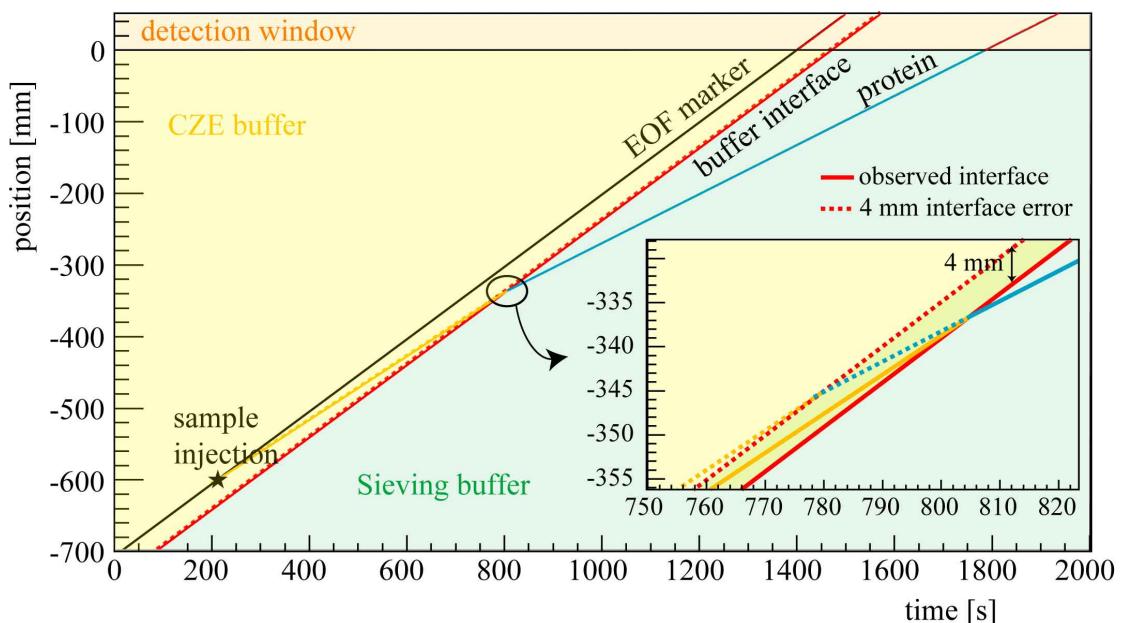


Figure 3.6: Reconstruction of the migration trajectory of a protein with an interface error of 4 mm. An inset shows an enlarged section of the interface crossing. Solid lines indicate the observed tracks, and dashed lines the tracks with a 4 mm interface error. Blue lines indicate the protein trajectory in the second buffer, red the buffer interface, and orange lines indicate the reconstructed track in the first buffer section. The data are from a typical two-buffer separation (EOF marker: thiourea, protein: α -lactalbumin). Separation conditions: CZE buffer 42 mM sodium tetraborate, sieving buffer 50 mM sodium tetraborate with 5 % dextran, separation distance 600 mm, $E = 160$ V/cm, sieving buffer introduced to anodic end 45 s after sample injection.

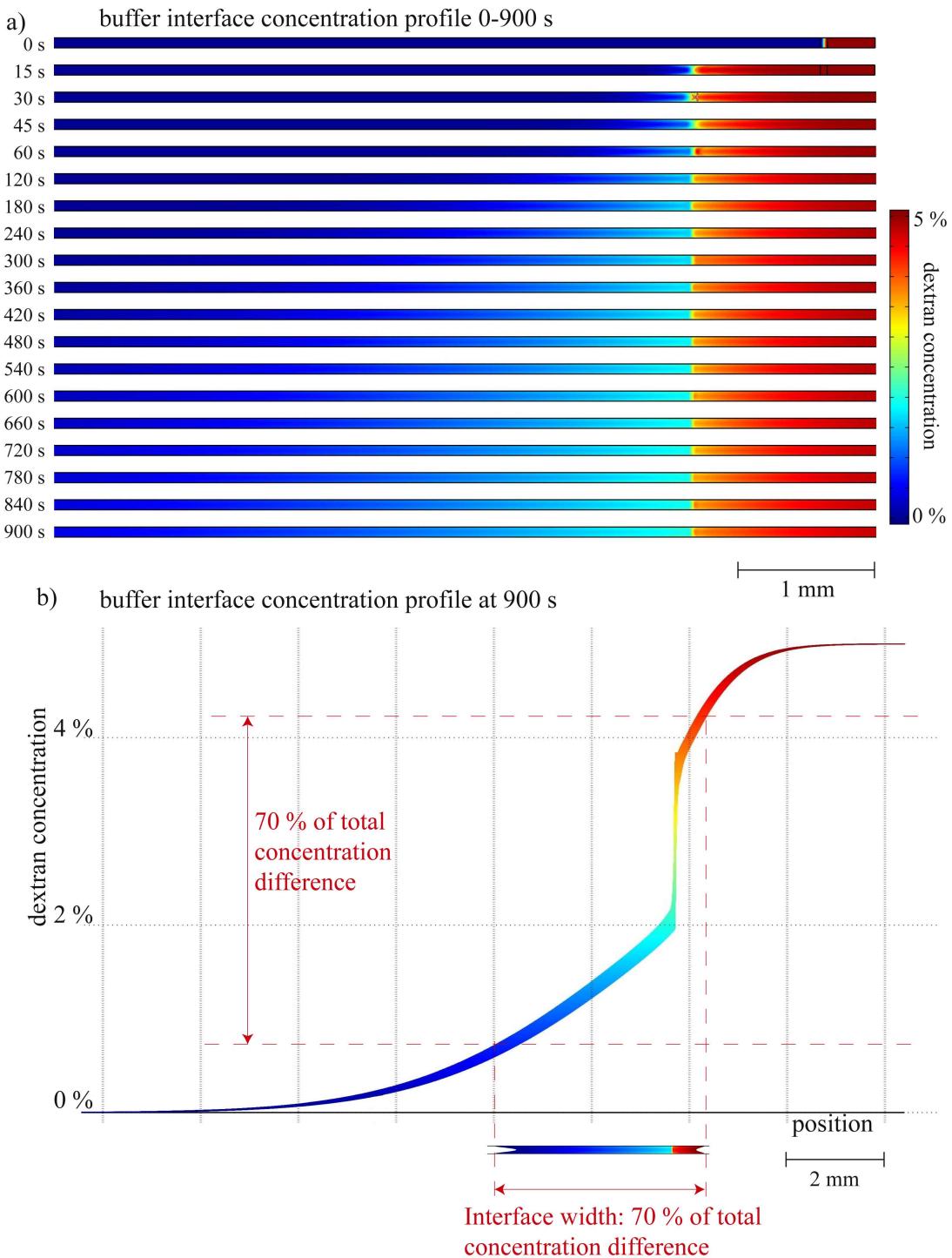


Figure 3.7: COMSOL simulation of interface region of two-buffer separation. a) Interface region at different times from 0 to 900 s. b) Concentration profile at 900 s with 70 % concentration difference as a measure for interface width. Conditions are for 5 % dextran at the right side (sieving buffer), 0 % dextran at left side (CZE buffer), with a background electrolyte of 50 mM sodium tetraborate.

With the confirmation that the chosen buffer system fulfils the requirements, a series of two-buffer separations was performed with a single protein. Figure 3.8 shows the mobilities in the two different buffer sections, whereby μ_{sieving} is directly observed and μ_{CZE} is reconstructed.

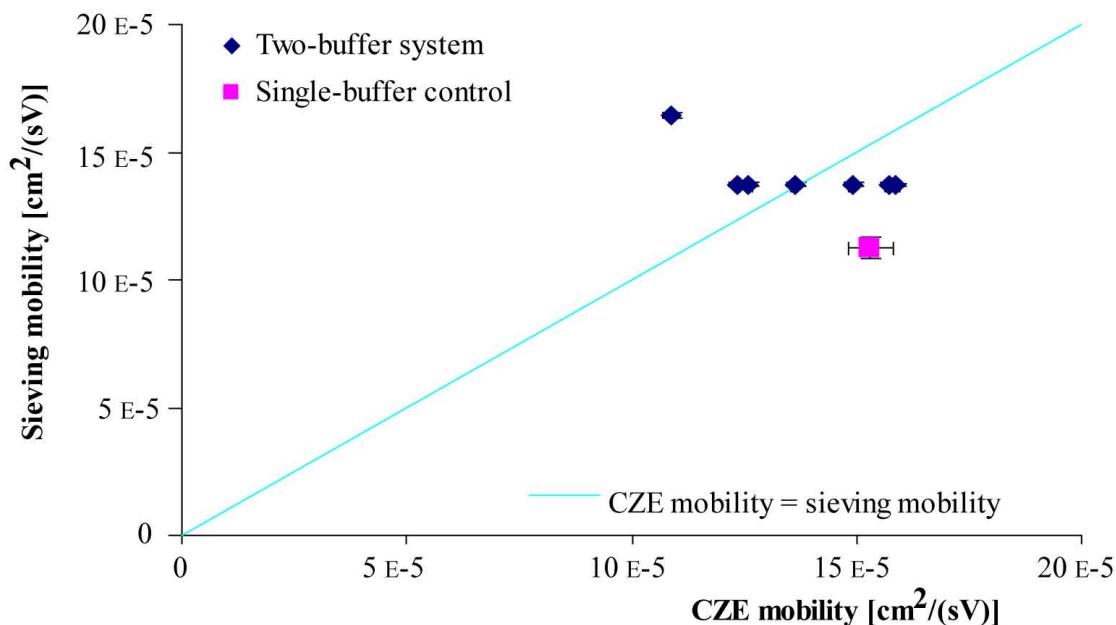


Figure 3.8: Repeat runs of a single protein in the two-buffer system. The mobility in the unobserved, CZE section is plotted against the mobility in the observed, sieving buffer section. The pink marker indicates mobility values obtained when measured in a single buffer system (averaged for four runs). The light blue line indicates where the mobility in both sections is identical. Protein: trypsin inhibitor (pI 4.6, MW 20 kDa). Separation conditions: CZE buffer 42 mM sodium tetraborate, sieving buffer 50 mM sodium tetraborate with 5 % dextran, separation distance 460 mm, $E = 160$ V/cm, sieving buffer introduction 120 seconds after sample injection. The error values from the two-buffer separation are obtained from the least squares fit to the detection data. The error value for the single-buffer mobility is obtained from the standard deviation of the four averaged runs.

Two observations can be made. The first is that the values for the observed (sieving) mobility show a small variation from run to run, whereas the reconstructed (CZE) mobilities vary significantly. The errors on the tracks are seen as error bars, indicating that it is not the reconstruction itself that causes the variability, but a more fundamental variability in the separation. This is likely to be the electro-osmotic flow, which is difficult to control in absolute terms and introduces considerable variations in separations. The second observation is that the values obtained from single-buffer systems are different to the values from two-buffer separations. Especially the observed sieving mobility is consistently higher in two-buffer separations.

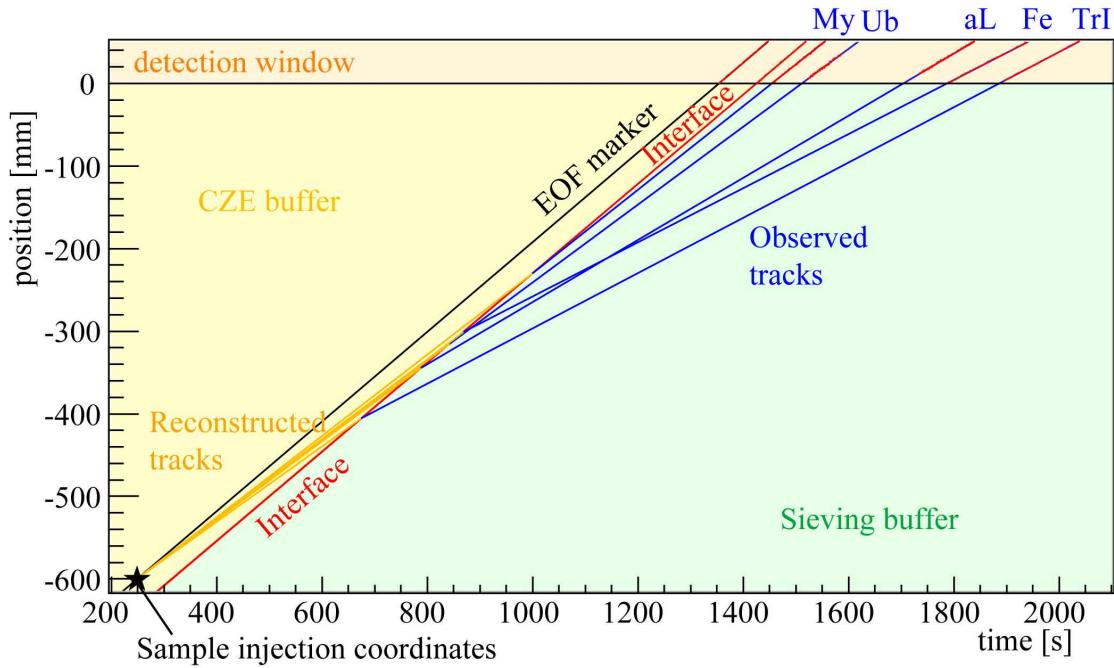


Figure 3.9: Separation of five proteins in the two-buffer system. Observed migration tracks of five proteins in blue, buffer interface in red, reconstructed migration trajectories in magenta, and EOF marker injected with sample in black. Proteins: My—myoglobin, Ub—ubiquitin, aL— α -lactalbumin, Fe—fetuin, TrI—trypsin inhibitor. Separation conditions: CZE buffer 42 mM sodium tetraborate, sieving buffer 50 mM sodium tetraborate with 5 % dextran, separation distance 600 mm, $E = 160$ V/cm, sieving buffer introduction 45 s after sample injection.

CZE separations can have relatively high run-to-run variations, because the EOF depends on the capillary surface charge, which can be influenced by many factors.^[40] Instead of comparing different runs of a single protein, a single run with numerous proteins will provide better consistency. In Figure 3.9 the reconstructed migration trajectories of five proteins in a two-buffer separation are shown.

In Figure 3.9 it can be seen that proteins appear to change migration speed at the interface in a characteristic manner. To ensure that the change in migration speed at the interface is not caused by systematic effects, control experiments were performed where the first and second buffer sections are identical. Figure 3.10 shows the tracks for such a control experiment. From Figure 3.10 it is concluded that the observed different migration behaviour in the two buffer sections is indeed a result of the different buffers.

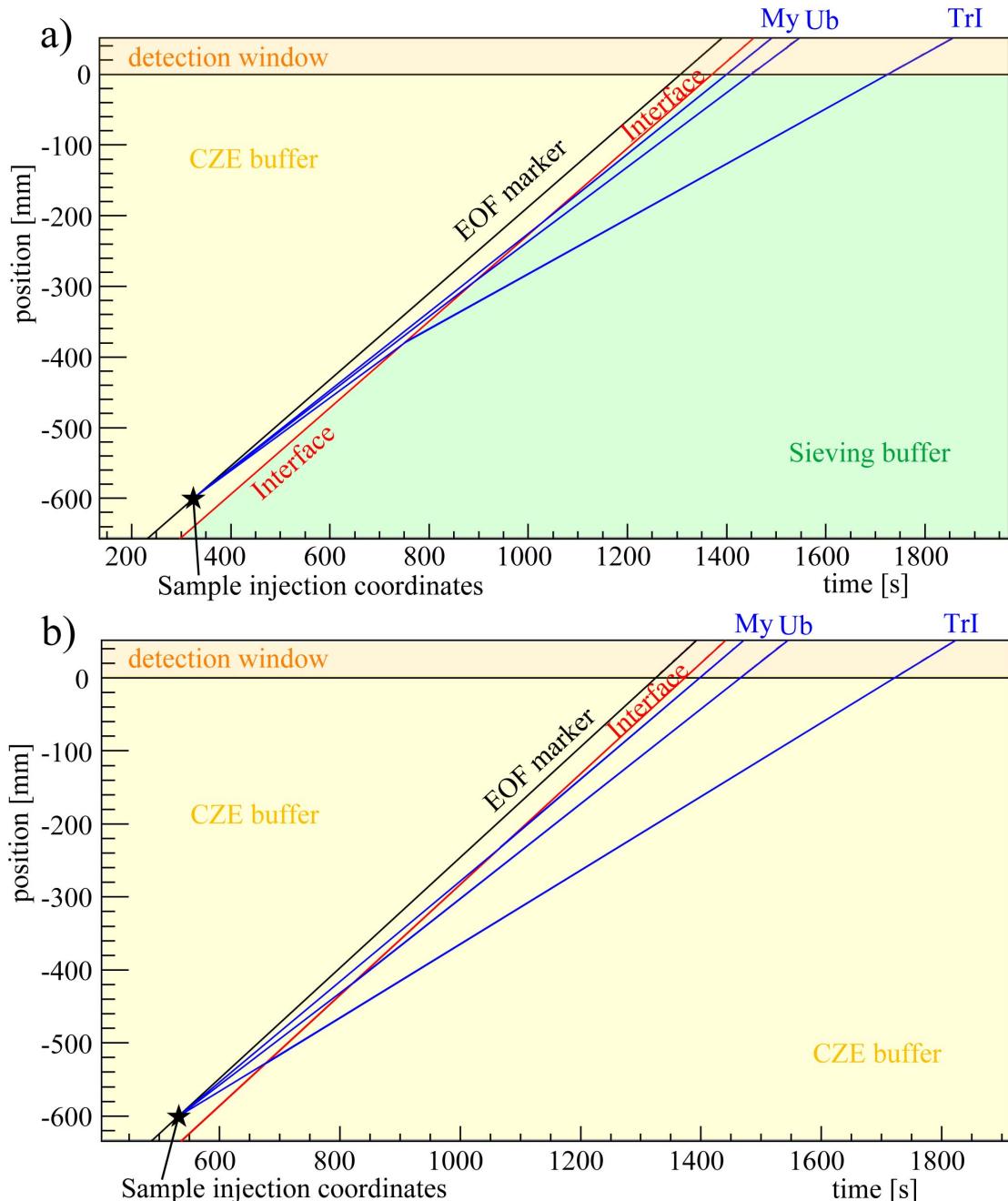


Figure 3.10: Two-buffer separation control experiment. a) Separation of three proteins in the CZE/sieving two-buffer system. b) Separation of the same three proteins in a CZE/CZE two-buffer system (with twice the same buffer). In the CZE/CZE two-buffer system no change of migration speed at the ‘interface’ is observed. Migration tracks of three proteins in blue, buffer interface in red, and EOF marker injected with sample in black. Proteins: My—myoglobin, Ub—ubiquitin, TrI—trypsin inhibitor. Separation conditions: CZE buffer 42 mM sodium tetraborate, sieving buffer 50 mM sodium tetraborate with 5 % dextran, separation distance 600 mm, $E = 160 \text{ V/cm}$. The second buffer is introduced 45 seconds after sample injection.

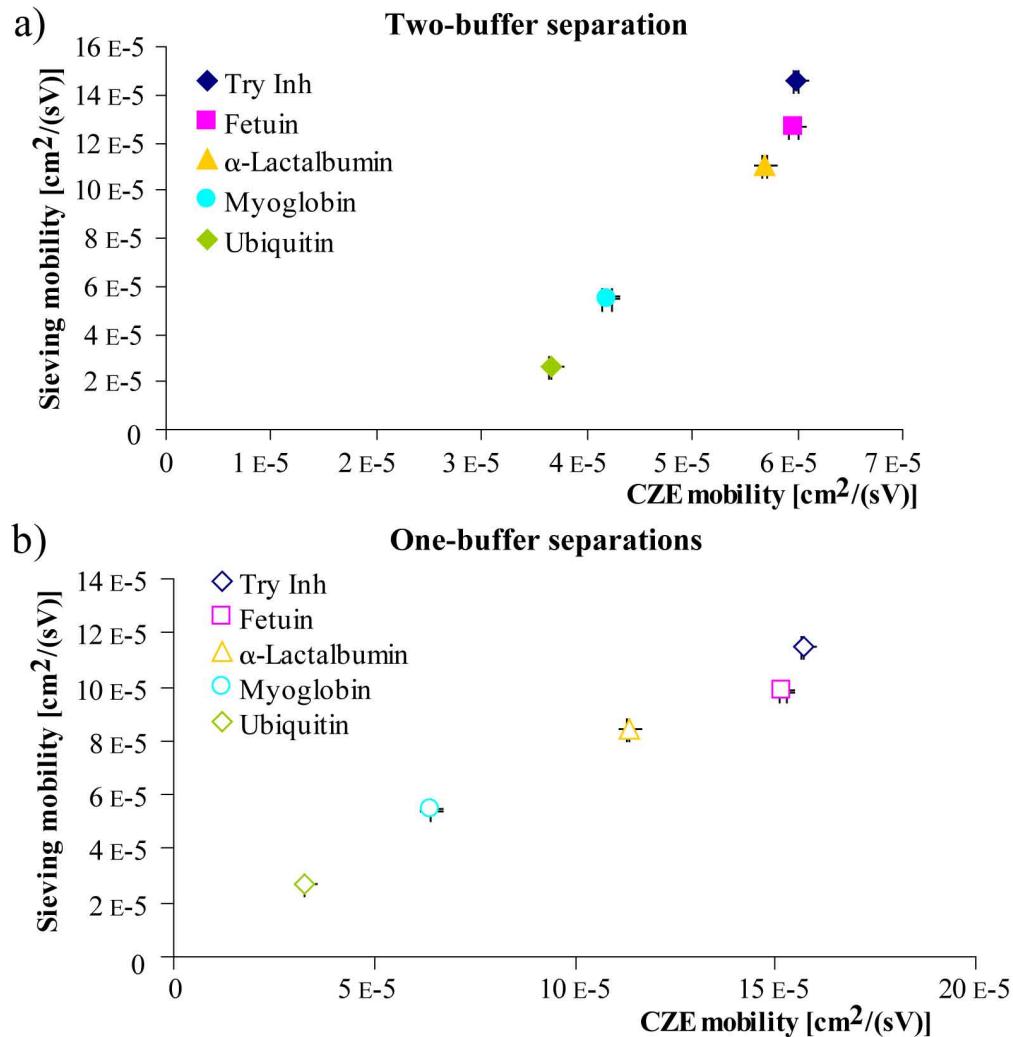


Figure 3.11: Two-buffer separation of a sample with five proteins: CZE mobility (first buffer section, reconstructed) vs. sieving mobility (second buffer section, observed). a) Two-buffer separation, and b) Mobilities measured independently in one-buffer systems. Proteins: myoglobin (pI 7, MW 17 kDa), ubiquitin (pI 6.3, MW 8.5 kDa), α -lactalbumin (pI 4.3, MW 14.2 kDa), fetuin (pI 3.5, MW 48 kDa), trypsin inhibitor (pI 4.6, MW 20 kDa). Separation conditions: CZE buffer 42 mM sodium tetraborate, sieving buffer 50 mM sodium tetraborate with 5 % dextran, separation distance 600 mm, $E = 160 \text{ V/cm}$, sieving buffer introduction 45 s after sample injection. Errors are obtained from the least squares fit of a line to detector data.

Figure 3.11 shows the mobility values for five proteins measured in a two-buffer system, along with the corresponding values measured in a single buffer system. In Figure 3.11, it appears that the mobilities of proteins display a similar pattern when measured in a two-buffer system or in a single buffer system. It is however remarkable that in the two-buffer system, the μ_{CZE} values are considerably smaller than when measured in the single buffer system. Indeed, where the value for μ_{sieving} is expected to be smaller than the corresponding μ_{CZE} value, it is observed to be consistently greater. Figure 3.12 shows repeats of the same measurement.

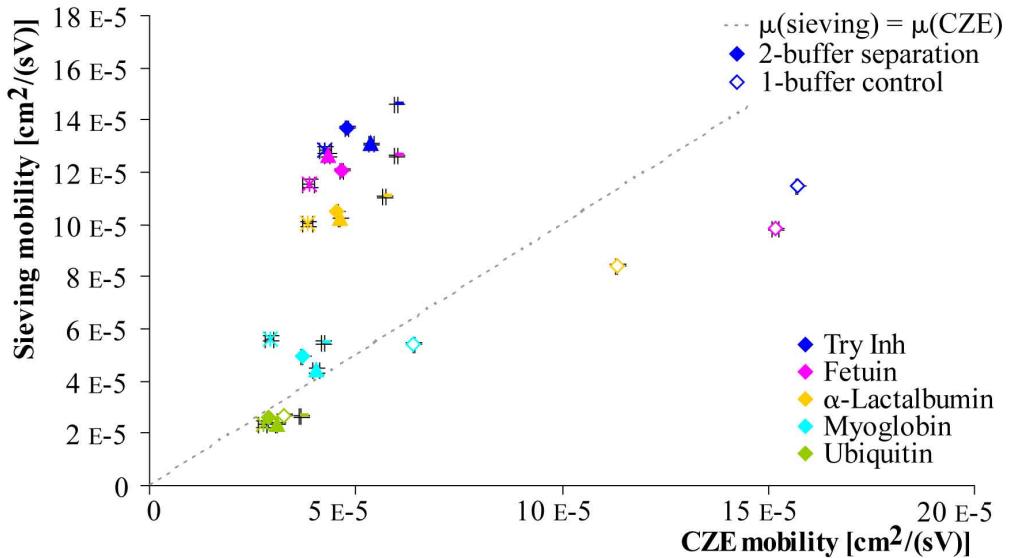


Figure 3.12: Repeat runs of two-buffer separation: CZE mobility (first buffer section, reconstructed) vs. sieving mobility (second buffer section, observed). Full symbols represent four repeat two-buffer separation, for each run a different symbol. Open symbols show mobilities measured in one-buffer systems. The line indicates where $\mu_{\text{sieving}} = \mu_{\text{CZE}}$. Proteins: myoglobin (pI 7, MW 17 kDa), ubiquitin (pI 6.3, MW 8.5 kDa), α -lactalbumin (pI 4.3, MW 14.2 kDa), fetuin (pI 3.5, MW 48 kDa), trypsin inhibitor (pI 4.6, MW 20 kDa). Separation conditions: CZE buffer 42 mM sodium tetraborate, sieving buffer 50 mM sodium tetraborate with 5 % dextran, separation distance 600 mm, $E = 160$ V/cm, sieving buffer introduction 45 s after sample injection. Errors are obtained from the least squares fit of a line to detector data.

The figures of merit for the separations seen in Figure 3.12 are found with Equations 1.2–1.14 in Chapter 1. The theoretical plate number for the first, unobserved, CZE separation is 2330 plates. The theoretical plate number for the second, observed sieving CZE separation is 2176 plates. The height equivalent to one theoretical plate is 0.17 mm for the first buffer section, and 0.09 mm for the second buffer section. The resolution of two adjacent peaks is 0.63. The peak capacity is 12.1 peaks for the first buffer section, and 11.7 peaks for the second buffer section. Because the separations are not orthogonal, the combined peak capacity is lower than the product of the individual numbers, 140.7 peaks. A conventional CZE separation is likely to perform slightly better than the individual sections in a two-buffer separation, but the combined separations are likely to perform slightly better than the conventional separation. The peak capacity of a two-buffer separation is lower than for conventional two-dimensional gel electrophoresis.

In Figure 3.12 the mobility pattern is seen to vary from run to run, and the only consistent feature is that the CZE mobility is consistently lower than expected from single-buffer data. Indeed, where the value for μ_{CZE} is expected to be greater than μ_{sieving} , for nearly all proteins the opposite is true. Although the tracking of proteins is successful, the dominating contribution to the different mobilities in the different buffer

sections is not related to properties of the proteins. Explanations for this behaviour will be discussed in the following section.

3.1.4 Discussion

Despite successfully reconstructing migration trajectories in different buffer sections, the observed protein mobility differences between the two buffers do not seem to be correlated with the protein size. It would appear that the difference in electro-osmotic flow completely dominates the change in migration at the interface. In other words, the change in protein mobility caused by the additional sieving effect is negligible compared to the change in electro-osmotic mobility. The separations in the two buffer sections are clearly far from orthogonal, and even when sieving mobility and CZE mobility are measured independently, only small differences can be seen between very large and very small proteins.

In CZE based separations, the mobilities are very low, and the migration speed is heavily influenced by the electro-osmotic flow. In the separations discussed, the migration speed of a protein is typically around 0.17 mm/s, while the EOF is typically 0.7 mm/s. Between the sieving and CZE buffers the change in protein migration is around 0.02–0.08 mm/s.

The EOF not being constant throughout the capillary is evidently problematic. Figure 3.13 illustrates the tracks of EOF markers injected at different points in the two-buffer system. Between the sieving and CZE buffers the change in electro-osmotic flow is around 0.25 m/s. This is an order of magnitude greater than the changes in protein migration that are of interest.

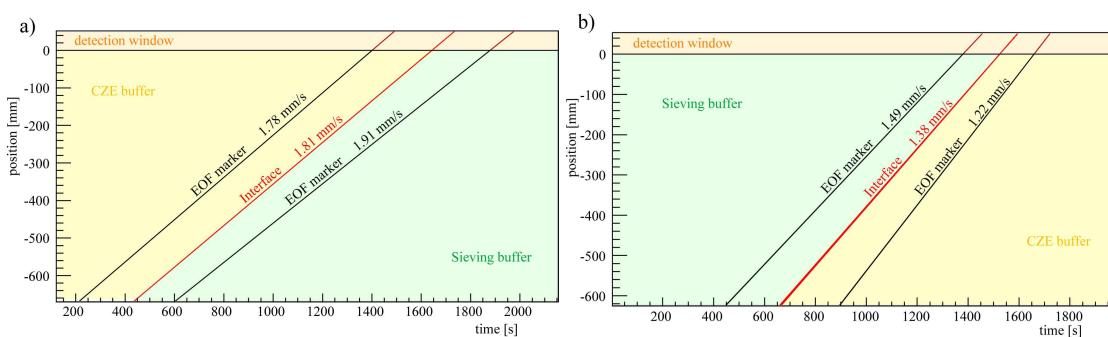


Figure 3.13: Electro-osmotic flow in different sections of the two-buffer system. a) First CZE buffer, second sieving buffer. b) First sieving buffer, second CZE buffer. The EOF differs in the first and second section by +7.3 % (a), respectively by -18.1 % (b). Separation conditions: CZE buffer 42 mM sodium tetraborate, sieving buffer 50 mM sodium tetraborate with 5 % dextran, separation distance 600 mm, $E = 160$ V/cm. EOF marker (thiourea 0.03 mg/mL in CZE buffer) was injected electrokinetically 3 minutes before and after the change of buffer, and at the buffer interface.

The problem of EOF dominating the migration change could be avoided by reducing the sieving component in the sieving buffer, for instance from 5 % dextran to 2 % dextran. As illustrated in Figure 3.4, this would reduce the EOF difference between the sieving buffer and the CZE buffer. This would also reduce the difference in protein mobility between sieving and CZE separation, and is therefore not desirable. Other buffer additives or surface treatments could be used to control the electro-osmotic flow, or different sieving matrices employed.

Some small improvements were achieved by adding 5 % dextran, molecular weight 6 kDa, to the CZE buffer. This concentration of dextran is below the entanglement threshold for this molecular weight. The EOF of the two buffers is more closely matched, but protein migration changes remain dominated by the change of EOF.

Finally it was concluded that in this separation system, the trade-off between orthogonality and compatibility does not leave enough margin to resolve protein-characteristic migration.

3.2 Microfluidics for V2D separations

3.2.1 Introduction

As described in Chapter 1.5, V2D separations combine isoelectric focusing (IEF) with capillary gel electrophoresis (CGE). Unlike conventional 2D separations, the two separations are performed collinearly, one after the other, and tracking of protein migration trajectories is used to reconstruct the first separation.

To implement a virtual two-dimensional (V2D) separation, it is necessary to integrate a pH gradient in a channel. Microfluidic structures offer far more configurational flexibility than conventional capillaries, and are therefore an ideal tool in this regard. Although pH gradients can be generated in free solution, here it is preferable for the pH gradient to be immobilised. Otherwise the pH generating additives would interfere with the second separation, as discussed in Chapter 1.3.

In Figure 3.14, a basic chip structure is seen, albeit without specific design elements for pH gradient genesis. In this section we will discuss suitable materials and microfabrication processes, characterise the microfluidic structures obtained, introduce different approaches to implementing immobilised pH gradients on-chip, and present the results of pH gradient fabrication on-chip.

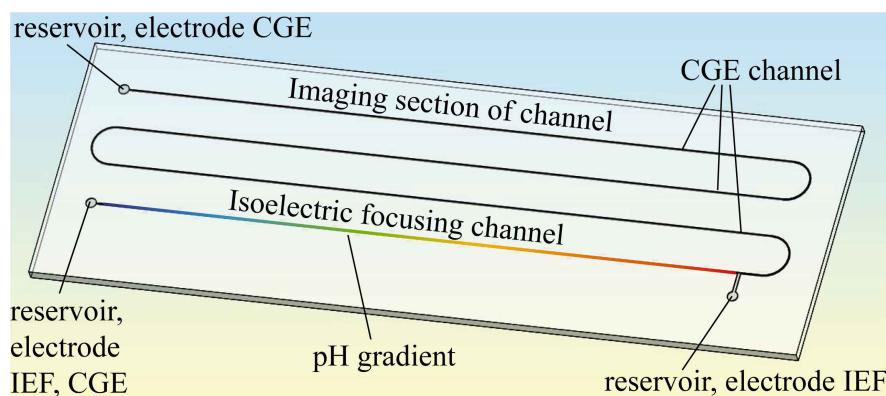


Figure 3.14: Basic layout of a V2D chip. In a first section, a pH gradient is integrated for isoelectric focusing. In the extension of the pH gradient channel, the second separation, capillary gel electrophoresis, is performed. At the end of the CGE channel, proteins are detected by UV absorption. Reservoirs are included to fill channels and to apply voltages. The chip dimensions are 25×75 mm.

3.2.2 Microfabrication

One of the first considerations relates to the choice of material, as this often dictates which fabrication techniques may be applied. As discussed in Chapter 1.5, the detection scheme is based on ultra-violet light absorption by proteins between 200 and 220 nm. This heavily influences the choice of material, since anything that is not transparent below 220 nm is unsuitable. This immediately excludes microfabrication favourites such as silicon and glass. Quartz and fused silica are transparent in this wavelength range.

Polymers are convenient from a fabrication perspective because they are generally easy to structure and inexpensive. Unfortunately, many polymers hardly transmit light below 300 nm. These include poly(methylene methacrylate) (PMMA), epoxy based polymers, polystyrenes, and polycarbonates.^[233] Another microfabrication favourite, poly(dimethyl siloxane) (PDMS), only transmits light down to around 250 nm. Finally a variety of cyclic olefin polymers and copolymers (COC) are available with different transparencies, some of them tolerably transparent below 220 nm. Their properties are illustrated in Figure 3.15, where the transmission between 200 and 300 nm is shown for a variety of polymeric films.

Some trials were performed with PDMS, but it was found that films of the necessary thinness are structurally not rigid enough. A related silicone, poly(methylhydrosiloxane) (PMHS), has better transparency,^[235] but trials showed it is unsuitable for casting and peeling, due to formation of a brittle surface layer. The materials found to be suitably transparent are fused silica, quartz, and COC.

For fabricating three-dimensional structures in quartz or fused silica, common procedures include etching in solutions of hydrofluoric acid, and powder blasting.^[236] Etching

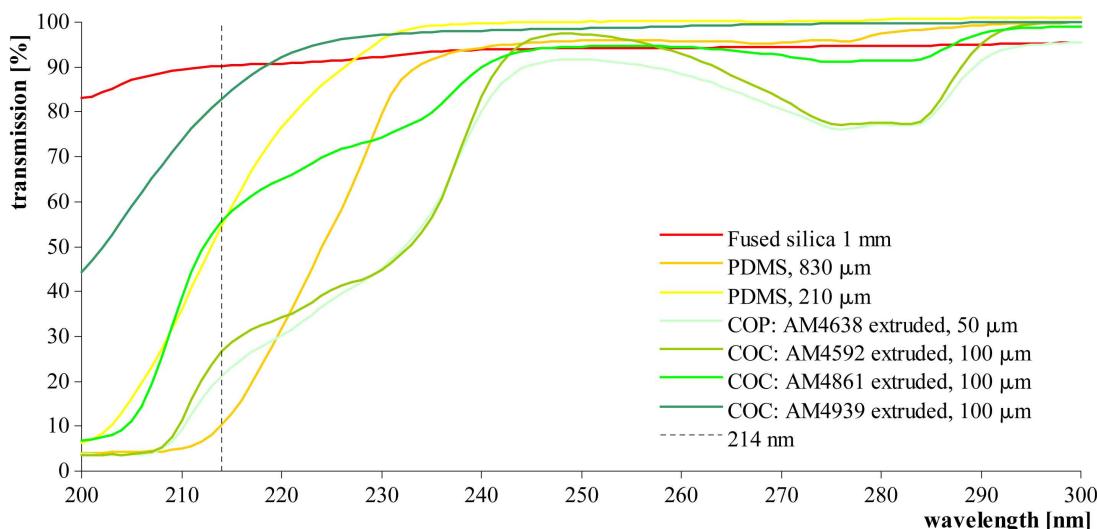


Figure 3.15: Transmission spectra between 200 and 300 nm for a variety of polymer films. PDMS (Dow Corning Corp., Sylgard 184) films are cast and polymerised according to standard protocols.^[234] Films of different grades of cyclic olefin polymers (COP) and cyclic olefin copolymers (COC) are obtained from Plitek LLC. The film thickness is measured to 10 μm accuracy with a micrometer. The wavelength of interest, 214 nm, is indicated with a dashed line.

with hydrofluoric acid produces shallow channels, and powder blasting causes rough surfaces. For these reasons neither technique is ideal for this application. Alternatively a sandwiching approach can be employed, where an easy-to-structure (polymeric), opaque material is sandwiched between two quartz slides. With regards to surface-dependent electro-osmotic behaviour of electrophoretic systems, this can be problematic. Furthermore, bonding of quartz to polymers to create sealed channels is not trivial.^[233] Instead, COC microfabrication was chosen. This fulfils the detection requirements and can be done with standard microfabrication equipment and without hazardous chemicals.

COC is an ethene-norbornene copolymer, as illustrated in Figure 3.16. Trade names of COC include Zeonor (Zeon Chemicals L.P.), Topas (TOPAS Advanced Polymers GmbH), and Apel (Mitsui Chemicals Inc.).^[237] Due to the variability of the copolymer's microstructure, COC is available in many different grades, with varying optical and thermal properties.

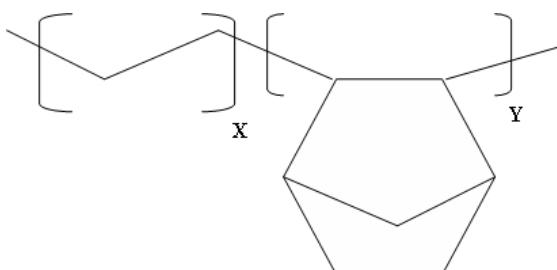


Figure 3.16: Structural formula of cyclic olefin copolymer (COC).

For structuring the high-UV-transparency COC, it was found that the originally intended process of dissolving COC pellets in toluene, casting a negative from a mold, and drying off the toluene affected the transparency of the COC. Even after extensive drying, residual toluene in the films caused poor transparency below 220 nm. This is shown in Figure 3.17.

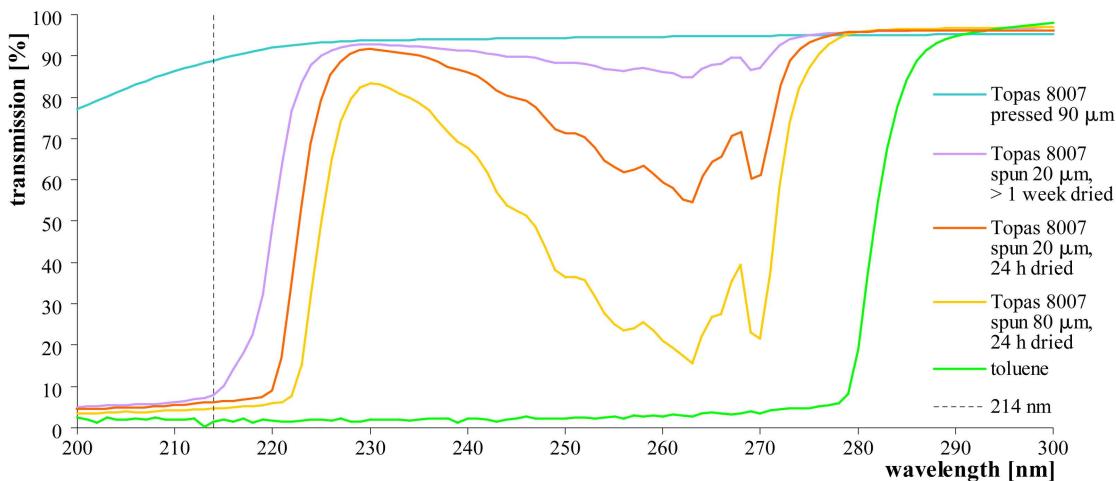


Figure 3.17: Transmission spectra of COC films with residual toluene. COC (Topas, 8007x4, granules) is dissolved 25 % w/v in toluene (BDH AnalaR), spin-coated onto glass slides, and dried at room temperature for at least 24 hours. Films of the same COC, but fabricated by hot pressing, have far higher transmission at 214 nm.

Since COC is a thermoplastic, embossing and other related techniques are an obvious alternative. A rapid prototyping process based on lamination was adapted, as illustrated in Figure 3.18.^[97] Briefly, a master structure is fabricated with SU-8 100 negative photoresist (MicroChem Corp.). COC films (Plitek AM4939) are laminated onto the master structure with a hot roll laminator (Albyco Photopro 33).

For sealing the channels, a second COC film is plasticised by wetting the surface with a hexane solution. Following plasticisation, the second film can be bonded to the first at a temperature slightly below the glass temperature, again with a hot roll laminator. This bonds the films without damaging the structure. Fabrication process details can be found in Appendix C.

The lamination-based microfabrication process allows quick fabrication of microfluidic structures that fulfil the minimum requirements for a V2D separation. This comes at the cost of channel cross section shape. In Figure 3.18c, the cross section of a channel is seen. The side walls are not vertical, but approach a catenary shape. This causes an enlargement of the channel compared to the master it was originally laminated onto. For a master width of 100 μm, the laminated channel is around 70 μm wide at the top, and around 350 μm wide at the bottom. Furthermore, the channel cross section is not

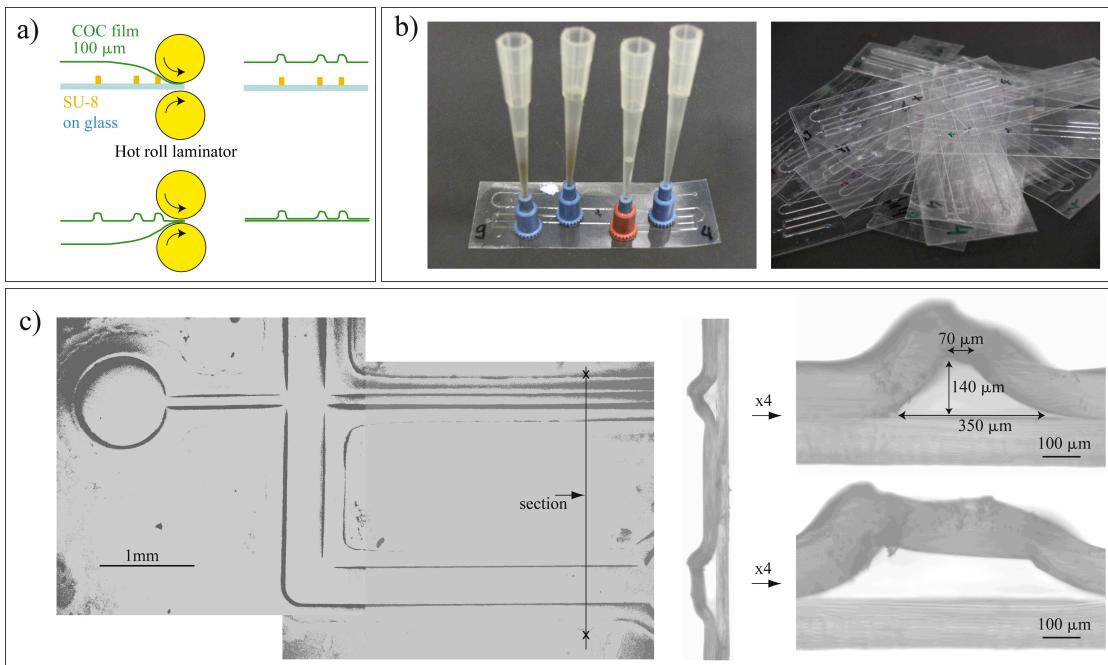


Figure 3.18: Microfabrication process: COC rapid prototyping with lamination.
a) Schematic of channel embossing and bonding of COC films. b) Images of COC microfluidic chips, each 25×75 mm in size. c) Micrographs of channels with cross section and enlarged micrograph of channel cross section. The SU-8 mold dimensions for the smaller channel are 100×100 µm, for the larger cross-channel 400×100 µm.

symmetrical, but appears to have suffered some distortion in the second pass through the laminator, when the sealing film is bonded on.

As discussed in Chapter 2.3.3.4, tracking becomes worse with increasing channel cross section area. For V2D separations, the cross section in Figure 3.18c is larger than ideal. For good master fidelity, hot embossing is a better technique.^[238] Despite these drawbacks, laminated chips are quick and easy to make and were therefore employed to establish a protocol for fabrication of a pH gradient on chip.

3.2.3 Fabrication of an immobilised pH gradient on-chip

With the UV transparent, rapidly fabricated chips characterised in the forgoing section, the next aim is to integrate a pH gradient that is immobilised in a channel section. The formation of immobilised pH gradients is a well established process for macroscale applications.^[59] For standard applications, acrylamide and bisacrylamide monomers are polymerised with ammonium persulfate as reaction initiator and tetramethylethyldiamine (TEMED) as catalyst, forming the matrix.^[239] The pH generating species, so-called Immobilines (from GE Healthcare), are copolymerised with acrylamide, as described in Chapter 1.3.

To form the gradient, the initial prepolymer aliquot is split in half, and one half is supplemented with a cocktail of acidic Immobilines, and the other half with basic Immobilines. These acidic and basic aliquots are poured to form a gradient, with exclusively acidic prepolymer on one side, and exclusively basic prepolymer on the other side, and a linear gradient of the two solutions in between. The gradient is formed with a gradient mixer, in which two chambers are joined and gradually flow out.^[59,60,65] The conventional gradient mixing technique cannot be applied directly to microfluidic chips. For microfluidic chips, the total gradient volume is in the range of microliters, which is far below the handling capabilities of a gradient mixer.

For generating immobilised pH gradients on-chip, two approaches have been reported so far: diffusion-based gradients of Immobilines in polyacrylamide,^[56] and surface functionalisation with focused carrier ampholytes.^[122–124] For surface functionalisation with carrier ampholytes, monolithic capillary columns are synthesised to increase the surface area. This approach has the advantage of very elegant pH gradient generation (the carrier ampholytes generate a pH gradient on application of an external electric field). A disadvantage is the number of reaction steps it takes to generate the monolith, prepare the surfaces, and link the carrier ampholytes to the surface. Considering monolith fabrication is not a routine technique, this reaction protocol may require significant optimisation and fine-tuning to implement.

The diffusion-based Immobiline gradients are closer to a well-established protocol, and the passive gradient generation is easily implemented. Figure 3.19 shows the chip layout employed for fabricating diffusion-based immobilised pH gradients. An important feature of the diffusion-based gradient is that the concentration at the ends must not be depleted, but maintained at the same level throughout the diffusion process. This is achieved by replenishing the ends with a constant flow from a cross-channel. Cross-channels are relatively wide to avoid influx to the gradient channel, and pressure levels are carefully adjusted to prevent pressure-driven flow through the gradient channel.

In conventional IPG fabrication, polymerisation of acrylamide and bisacrylamide is achieved with ammonium persulfate as a reaction initiator. For microfluidic applications, photopolymerisation is an attractive alternative. This can prevent unwanted polymerisation, for instance in access channels. Photopolymerisation can be achieved with commercial azo initiators such as VA-086 from Wako.^[56,240,241] The detailed composition and conditions of polyacrylamide polymerisation can be found in Appendix D.

Photopolymerisation of polyacrylamide in microfluidic COC channels successfully produces a gel structure, as confirmed in unsealed COC test channels. However, it was found that the polyacrylamide gel has very poor channel adhesion, and indeed is routinely

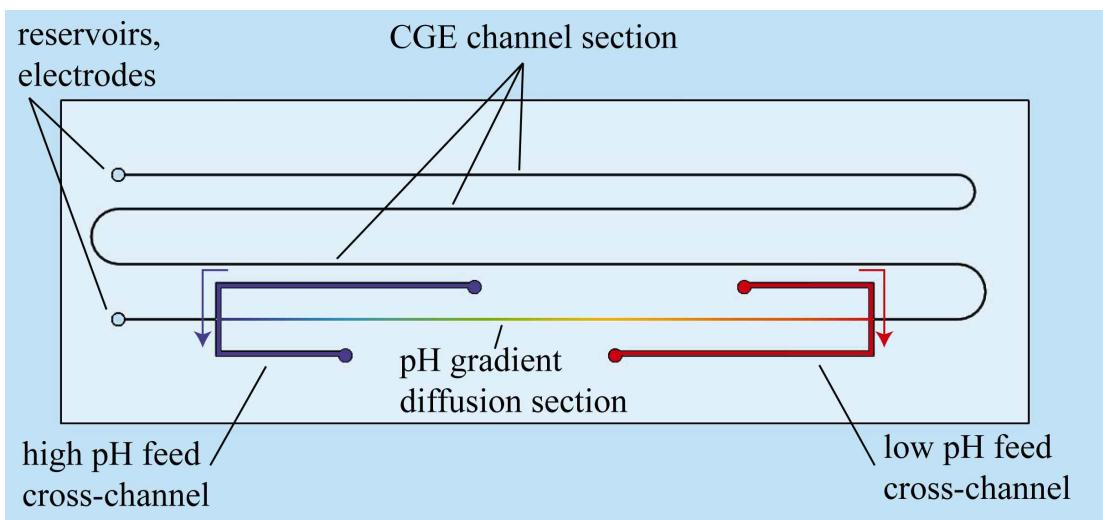


Figure 3.19: Microfluidic chip design for generating immobilised pH gradients by diffusion. Concentration at the ends of the gradient section is maintained constant with the feed from the cross-channels. The chip dimensions are 25 × 75 mm.

flushed out of the channel upon sample application. COC is known to have a low surface energy,^[237] which is undesirable for immobilising the polyacrylamide gel. To affect a bond between the gel and the channel, a photografting protocol was adapted.^[109,242,243] Briefly, a benzophenone solution is flushed through the channel, and exposed to UV light to form surface-bound radicals, as illustrated in Figure 3.20 (for details see Appendix D). Surface grafting of polyacrylamide then follows, by photopolymerisation as previously described.

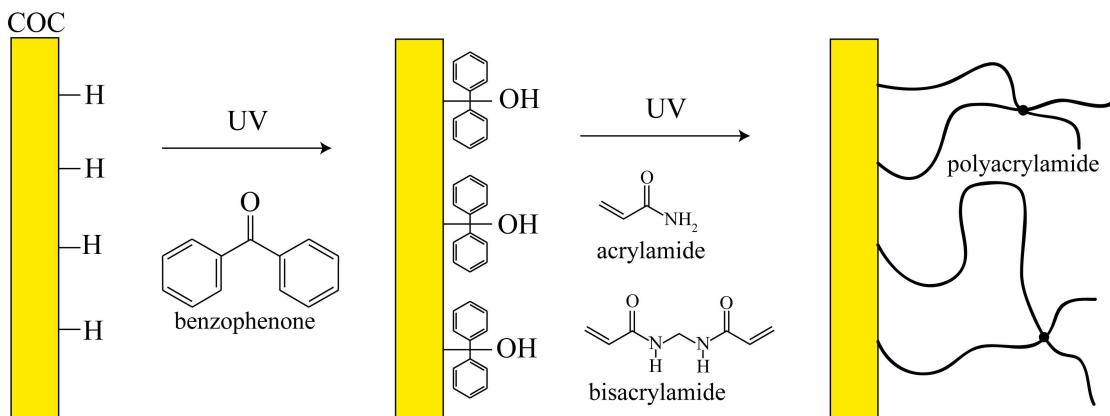


Figure 3.20: Photografting of polyacrylamide gel to COC channel. In an intermediate step, a benzophenone solution is irradiated to form surface-bound radicals. Adapted from^[243].

With satisfactory fabrication of polyacrylamide gels in COC microfluidic channels, fabrication of pH gradient gels was investigated. To visualise the gradient, dye gradients were formed. Figure 3.21 shows a dye gradient and the corresponding line scan intensity.

The line scan indicates that the gradient is approaching a linear gradient, in this case after 12 hours diffusion time.

With protocols in place for fabricating immobilised pH gradients in microfluidic COC channels, validation of isoelectric focusing on-chip was attempted. Prior to isoelectric focusing, the pH gradient was submitted to 60 minutes ‘electrophoretic flush’ at 30 V/cm, and 24 h rinsing of the cross-channels with ultrapure water. This is to remove unpolymerised acrylamide monomers and other residual contaminants in the pH gradient.

For isoelectric focusing, a high voltage power supply (Spellman MP10P24F) was controlled via a USB data acquisition module (Data Translation DT9800) using a custom-made script written in C. The electric field across the pH gradient was gradually ramped up to 280 V/cm, and held there for 120 minutes. This is comparable to the protocol suggested by manufacturers for commercial immobilised pH gradient strips.^[245]

Myoglobin (10 mg/mL in 50 mM sodium tetraborate) is used as a protein sample, because it has a vivid colour and is easily visualised. Figure 3.22 shows the results of myoglobin isoelectric focusing.

Although the myoglobin appears to have entered the immobilised pH gradient and collected in a general region of the pH gradient channel, this is far from a focused zone. Possible causes for such poor focusing and approaches for improvements are discussed in the next section.

3.2.4 Discussion and chip outlook

In the forgoing section it was demonstrated that pH gradients fabricated on-chip can perform isoelectric focusing. It was also noted that isoelectric focusing was not optimal, and the protein sample remains dispersed over a wide spatial range, even after excessive focusing. A number of causes are likely to contribute to this.

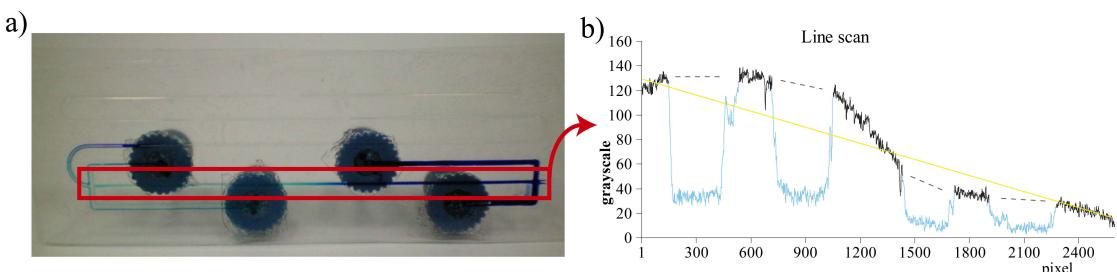


Figure 3.21: a) Dye gradient in microfluidic channel, after 12 h diffusion time. b) Line scan of intensity profile along gradient channel in a). In sections of the channel where the dye gradient is obstructed by nanoports, the line scan is tinted light blue, and a dashed line is drawn for guidance. In yellow a linear gradient is indicated. Line scan obtained with ImageJ software.^[244] The gradient channel is 60 mm long.

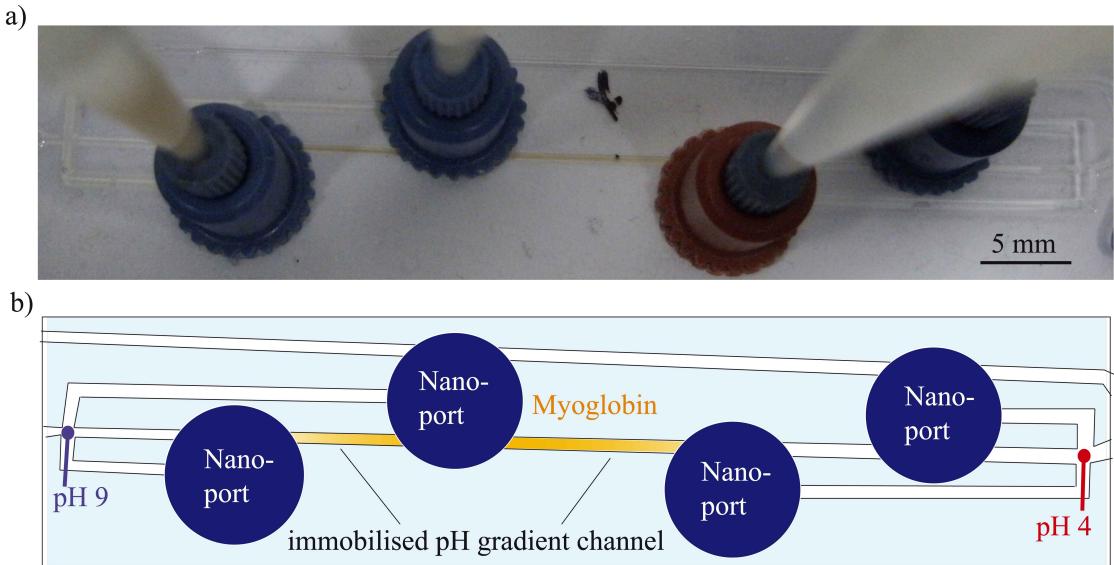


Figure 3.22: Isoelectric focusing of myoglobin in immobilised pH gradient on-chip. a) Photograph, and b) Schematic of photograph. The pH gradient spans from four to nine, and the isoelectric point of myoglobin is seven.

In conventional fabrication of immobilised pH gradients, the gel is thoroughly soaked in water to remove unpolymerised acrylamide, Immobilines, TEMED, and initiator.^[59] With a closed channel, this is not possible, and instead an ‘electrophoretic flush’ is applied. Although there have been reports of successful focusing after electrophoretic flushing,^[56] they were achieved in relatively short immobilised pH gradients of 5 mm, and supplemented by cross-channel flushing.

The chip design for V2D separations requires a rather longer pH gradient channel, since the reconstruction relative accuracy is better if proteins are spread farther apart. In the chip layout used for fabricating pH gradients, the gradient channel is 60 mm long. If cross-channel flushing is crucial for removing contaminants, it is likely that with longer channels this is insufficient.

The solution to this problem lies in a next generation of chip layout, where a dedicated flush-channel is implemented. This is also a requirement if proteins are to be ‘equilibrated’ with SDS after the isoelectric focusing and prior to capillary gel electrophoresis. In Figure 3.23, a chip design with a flush-channel adjacent to the pH gradient channel is seen. The flush-channel is connected to the pH gradient channel via narrow cross-channels, to enable access while affecting electric field distribution in the pH gradient channel only moderately.^[246]

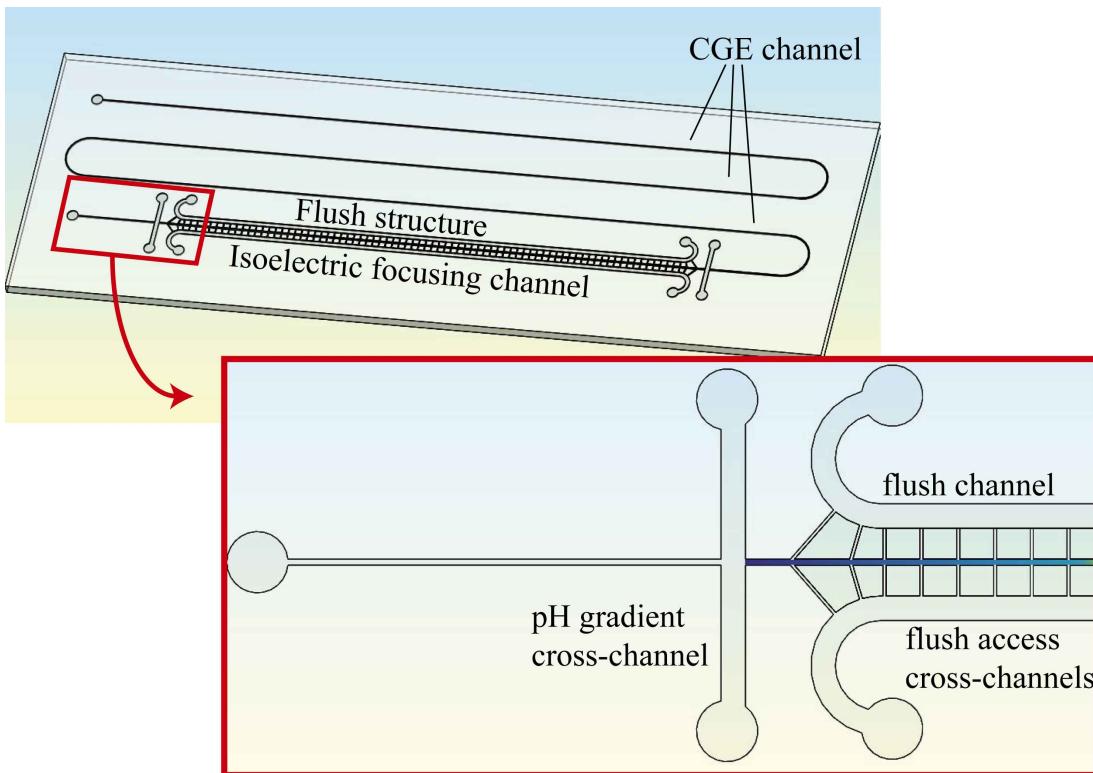


Figure 3.23: Microfluidic chip design with a flush-channel adjacent to the pH gradient channel. The flush-channel is connected to the pH gradient channel via narrow cross-channels, to enable rinsing of the gel after polymerisation and SDS treatment of focused proteins prior to the second (capillary gel electrophoresis) separation.

As mentioned above, previous success with isoelectric focusing in diffusion-based gradients was achieved in relatively short microfluidic channels. In this report, the channel cross section area is of a similar magnitude ($100 \times 100 \mu\text{m}$), but the channel length is an order of magnitude greater. Trials with dye already demonstrated that a linear gradient may not yet be formed within 12 hours. Using an estimate for steady-state diffusion time $t = L^2/(2D)$ with diffusion length $L = 60 \text{ mm}$, and diffusion constant of aqueous acrylamide monomer $D = 1.8 \times 10^{-10} \text{ m}^2/\text{s}$,^[241] we obtain a time of tens of days. Diffusion-based gradients are not ideal for the lengths intended here.

Alternative techniques to diffusion-based gradients may bring improvements. Passive generation of gradients has been addressed for microfluidic systems in the past. Approaches include micro-tunnel feed channels,^[117] ladder structures,^[118] cross-section dependent dispersion,^[247] designs with delay channels,^[119] and with channel networks.^[248] Some approaches are better suited for temporal gradients, i.e. they produce spatial gradients distributed along a very long channel section, which is useful for liquid chromatographic applications.^[119,120]

The generation of gradients with microfluidic networks has been well characterised, and is a promising approach for generating gradients in a channel as required for V2D

separations.^[249–252] Figure 3.24 shows the chip design for optimised isoelectric focusing on-chip, with a microchannel network for gradient genesis, and gel-rinse access channels. To improve channel cross section characteristics, the fabrication by hot embossing is preferred to lamination, and channel cross section dimensions are designed to be $50 \times 50 \mu\text{m}$, with a view toward tracking accuracy.

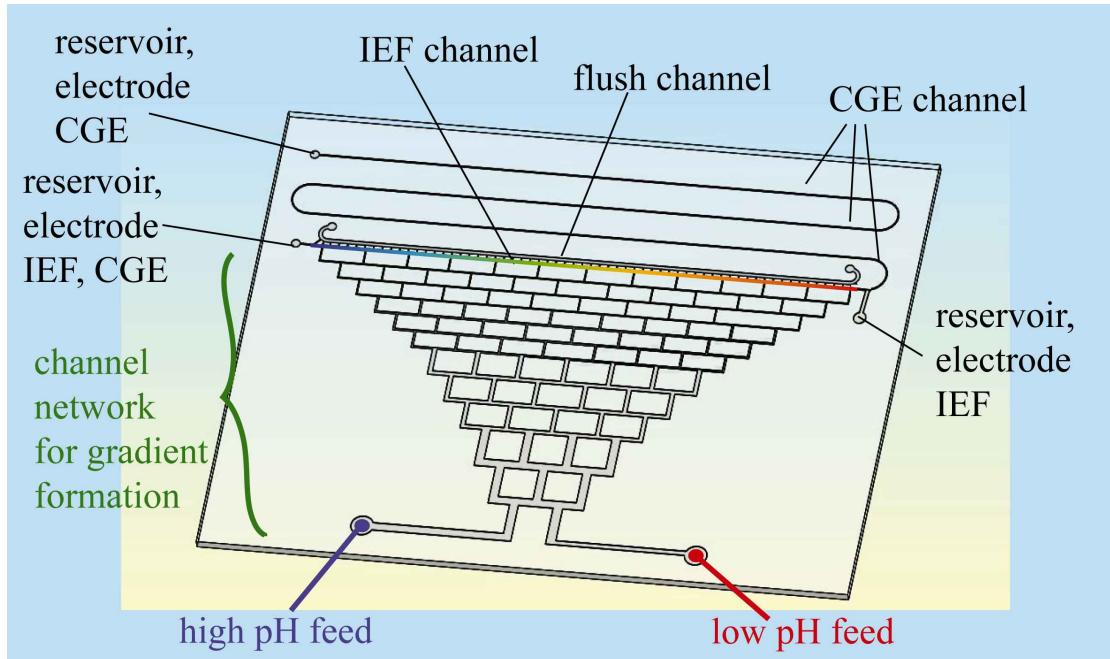


Figure 3.24: Chip design with a microfluidic channel network for forming a pH gradient. Two feed flows, one with high pH and the other with low pH, are successively split, rejoined, and mixed to generate a gradient in the channel for isoelectric focusing.

Chapter 4

Conclusion

4.1 Summary

The work in this thesis has been directed toward establishing virtual two-dimensional separations, where collinear separations are distinguished through the reconstruction of protein migration trajectories. Specifically, studies involved a theoretical analysis of the migration trajectory reconstruction process as well as software, instrument, and microfluidic device development. Significantly, for the first time, non-linear migration of proteins in capillary gel electrophoresis is characterised and quantified.

The migration behaviour of proteins in capillary gel electrophoresis has been characterised at a fundamental level. This is of crucial importance for virtual two-dimensional (V2D) separations, where reconstruction of the first separation is based upon migration trajectories of individual proteins. It was found that thermal effects caused by Joule heating in the course of electrophoresis have a significant impact on the migration trajectory, and hence on reconstruction accuracy. This affects both of the V2D dimensions, the molecular weight (defined by the migration speed or the slope in time-position space), and the pI of the protein (defined by the reconstructed starting coordinates). Accordingly, a variety of experimental settings were evaluated to find optimal conditions. Importantly, models describing non-linear migration effects have been developed and used to improve the accuracy of track reconstruction, allowing a ten-fold reduction of typical reconstruction errors. While thermal effects are highly dependent on experimental and instrumental factors, another more fundamental effect is described: UV induced fragmentation of proteins. Although non-linear migration trajectories are of particular relevance for V2D separations, UV effects may be of relevance for other protein electrophoresis applications.

The concept of protein track reconstruction is validated in a two-buffer separation system, where proteins migrate through sections with different chemical conditions. For the first time, the concept of migration trajectory reconstruction is applied to capillary electrophoresis separations of proteins under non-constant buffer conditions. Here protein migration reconstruction is successful, but the compatibility of the two buffers imposes severe limitations on the choice of separation conditions. It was found that the difference between the chemical conditions dominates and masks any potential difference in protein migration from one section to the other.

To implement V2D separations, a microfluidic system with integrated components for the two separations was developed. A protocol for rapid prototype fabrication was adapted to the requirements of the system. This included surface preparation and functionalisation. For the first separation, an immobilised pH gradient was fabricated within a cyclic olefin microfluidic chip. Isoelectric focusing of a protein sample on-chip was demonstrated.

Table 4.1 compares the performance of V2D separations and two-buffer separations with conventional separations. Based on the error analysis for the V2D reconstruction, a peak capacity comparable to two-dimensional gel electrophoresis is achievable. The advantage of the high sample load applicable to isoelectric focusing is combined with the detection performance of capillary gel electrophoresis. While two-buffer separations demonstrate the principle of separation reconstruction, they do not perform as well as the combination of isoelectric focusing with capillary electrophoresis.

Table 4.1: Comparison of performance metrics for V2D and conventional protein separation techniques.

	Capillary electrophoresis ¹	Isoelectric focusing ²	2D gel electrophoresis ³	2D liquid chromatography ⁴	V2D ⁵	2-Buffer ⁶
Peak capacity	65 ^[40]	130 ^[56]	2,000–10,000 ^[47]	100–500 ^[36]	5950	< 140
Dynamic range (orders of magnitude)	4 ^[58]	< 3 ^[59]	< 3 ^[36]	3 ^[36]	4	4
Sensitivity	400 pg ^[58]	100 pg ^[59]	100 pg ^[50]	20 pg ^[36]	400 pg	400 pg
Accuracy	1–10 % ^[58]	0.15 pH unit ^[60]	1–10 % ^[38]	< 1 Da ^[38]	4 %	2 %
Sample load	nL ^[58]	50–100 µg ^[50]	50–100 µg ^[50]	µL ^[36]	µg	nL

¹ Capillary gel electrophoresis, UV detection² In immobilised pH gradient strip, silver stain³ Silver stain⁴ Ion exchange-reverse phase LC, nanoLC coupled to nanoESI, detection by MS⁵ Reconstruction of IEF and CGE, from error analysis⁶ Reconstruction of sieving CZE and CZE

4.2 Outlook

With regards to the characterisation of non-linear migration, a Monte-Carlo simulation was applied to provide an estimate of the impact of UV induced protein fragmentation. A detailed quantitative study of UV absorbance and protein fragmentation during capillary electrophoresis would be of great merit, particularly as this appears to be fundamental behaviour and has not been previously reported. This could influence the design of future V2D separations, as UV exposure can be reduced with a trade-off in the volume of detection data. The other contributor to non-linear migration (thermal variation in space and time) is dependent on instrumentation and requires optimisation specific to the design.

Microfluidic devices for virtual two-dimensional separation need further optimisation before they can be applied in a robust way to V2D separations. As demonstrated and discussed in Chapter 3.2, isoelectric focusing results in a band of protein that is still widely dispersed. This is not suitable for V2D separations because the premise of the reconstruction is that the starting position is confined in a tight band. Optimisation of the microfluidic chip will include implementation of additional structures to assist in flushing the immobilised pH gradient gel, which will remove unpolymerised monomers and other interfering chemical residues. Rinse channels are also of critical importance for preparing the second, SDS-based capillary gel electrophoresis separation. Furthermore, it is desirable to design a long pH gradient section, as this improves the relative accuracy of the reconstruction. The implemented design for generating a gradient by diffusion is far better suited to short sections in the range of millimetres, because the diffusion time for distances of 5–10 cm become prohibitively long. An alternative design with a channel network for gradient genesis by flow splitting is expected to achieve a better, more linear gradient.

Once isoelectric focusing has been optimised, the remaining procedure for virtual two-dimensional separation must be defined. This will include rinsing the focused protein bands in the immobilised pH gradient gel with SDS and preparing the capillary gel electrophoresis channels. Rinsing needs to be adjusted to permit full SDS denaturing of proteins, while avoiding excessive diffusion. Preparing the CGE section of the channel is expected to be straightforward and broadly similar to conventional capillary gel electrophoresis. Optimisation of injection structures is of less relevance, since sample is focused in the immobilised pH gradient.

Other issues particular to on-chip protein electrophoresis pertain more to the V2D design. For example, as the surface properties of a microfluidic channel can dramatically influence the electro-osmotic flow, under some circumstances it is necessary to engineer

the surface.^[99,100,253] In capillary gel electrophoresis the sieving buffer usually contains surfactants (such as SDS), and the sieving constituents can provide additional dynamic coating, wherefore using polymeric channels might not have an adverse influence. Should the electro-osmotic flow prove problematic, the procedure for photografting described in Chapter 3.2 could easily be applied to the CGE section of the channel. Channel design is another important issue that may need consideration. This includes assessment of factors such as low dispersion turn geometry (to avoid the ‘race track effect’).^[254]

Following the results regarding optimal conditions for track reconstruction (Chapter 2.4.4), a different microfabrication process will be necessary to reduce the channel cross section area to around $50 \times 50 \mu\text{m}$. Instead of the lamination of cyclic olefin copolymer films, hot embossing is expected to be better suited to this problem. The requirements of the mold structure for hot embossing are different than for lamination, and possibly more sophisticated molds (than SU-8 on glass) will be necessary. Common problems include release of the mold from the polymer, especially if the side-walls are undercut, and mold robustness.

Despite all the potential pitfalls discussed above, it is not unlikely that a virtual two-dimensional separation can be demonstrated successfully without extensive optimisation efforts.

Bibliography

- [1] H. A. Harper, V. W. Rodwell, and P. A. Mayes, *Review of physiological chemistry*, 16th ed. Blackwell Scientific, 1977.
- [2] R. A. Freitas, *Nanomedicine*. Landes Bioscience, 1999, vol. 1, Basic capabilities.
- [3] *International Statistical Classification of Diseases and Related Health Problems*, World Health Organisation Std., Rev. 10, 2007.
- [4] (1948) Universal declaration of human rights. General Assembly of the United Nations. [Online]: www.un.org/en/documents/udhr/
- [5] J. Lederberg and A. T. McCray, “Ome sweet ’omics— a genealogical treasury of words,” *The Scientist*, vol. 15, no. 7, p. 8, 2001.
- [6] H. Winkler, *Verbreitung und Ursache der Parthenogenesis im Pflanzen- und Tierreiche*. Verlag Fischer, Jena, 1920.
- [7] W. Fiers, R. Contreras, F. Duerinck, G. Haegeman, D. Iserentant, J. Merregaert, W. Minjou, F. Molemans, A. Raeymaekers, A. Vandenberghe, G. Volckaert, and M. Ysebaert, “Complete nucleotide-sequence of bacteriophage MS2-RNA — primary and secondary structure of replicase gene,” *Nature*, vol. 260, no. 5551, pp. 500–507, 1976.
- [8] Human Genome Organisation. [Online]: <http://www.hugo-international.org/>
- [9] (2009) -Omes and -omics glossary & taxonomy. Cambridge Healthtech Institute. [Online]: <http://www.genomicglossaries.com/content/omes.asp>
- [10] B. Albers, *Molecular Biology of the Cell*, 5th ed. Garland Science, 2008.
- [11] J. B. German, B. D. Hammock, and S. M. Watkins, “Metabolomics: building on a century of biochemistry to guide human health,” *Metabolomics*, vol. 1, no. 1, pp. 3–9, 2005.
- [12] H. F. Lodish, *Molecular cell biology*, 6th ed. W.H. Freeman, 2008.
- [13] G. M. Cooper, *The cell: A molecular approach*, 4th ed. ASM Press, Sinauer Associates, 2007.
- [14] R. S. Tirumalai, K. C. Chan, D. A. Prieto, H. J. Issaq, T. P. Conrads, and T. D. Veenstra, “Characterization of the low molecular weight human serum proteome,” *Mol Cell Proteomics*, vol. 2, no. 10, pp. 1096–1103, 2003.
- [15] M. R. Wilkins, *Proteome research: Concepts, technology and application*. Springer, 2007.

- [16] R. M. Twyman, *Principles of proteomics*, ser. BIOS Scientific. Taylor and Francis, 2004.
- [17] M. Lomnytska and S. Souchelnytskyi, "Markers of breast and gynecological malignancies: The clinical approach of proteomics-based studies," *Proteomics - Clinical Applications*, vol. 1, no. 9, pp. 1090–1101, 2007.
- [18] L. Hartwell, D. Mankoff, A. Paulovich, S. Ramsey, and E. Swisher, "Cancer biomarkers: A systems approach," *Nat Biotech*, vol. 24, no. 8, pp. 905–908, 2006.
- [19] S. M. Hanash, S. J. Pitteri, and V. M. Faca, "Mining the plasma proteome for cancer biomarkers," *Nature*, vol. 452, no. 7187, pp. 571–579, 2008.
- [20] Any Test Franchising, Inc. [Online]: www.anylabtestnow.com
- [21] Monobind Inc. [Online]: www.monobind.com
- [22] C. A. Evans, A. Glen, C. L. Eaton, S. Larré, J. W. F. Catto, F. C. Hamdy, P. C. Wright, and I. Rehman, "Prostate cancer proteomics: The urgent need for clinically validated biomarkers," *Proteomics - Clinical Applications*, vol. 3, no. 2, pp. 197–212, 2009.
- [23] Y. Hu, A. Hosseini, J. S. K. Kauwe, J. Gross, N. J. Cairns, A. M. Goate, A. M. Fagan, R. R. Townsend, and D. M. Holtzman, "Identification and validation of novel CSF biomarkers for early stages of Alzheimer's disease," *Proteomics - Clinical Applications*, vol. 1, no. 11, pp. 1373–1384, 2007.
- [24] E. Camafeita, J. R. Lamas, E. Calvo, J. A. López, and B. Fernández-Gutiérrez, "Proteomics: New insights into rheumatic diseases," *Proteomics - Clinical Applications*, vol. 3, no. 2, pp. 226–241, 2009.
- [25] M. Latterich, M. Abramovitz, and B. Leyland-Jones, "Proteomics: New technologies and clinical applications," *European Journal of Cancer*, vol. 44, no. 18, Sp. Iss. SI, pp. 2737–2741, 2008.
- [26] L. L. Ilag, "Are genomic and proteomic technologies adding value to drug discovery?" *Targets*, vol. 1, no. 5, pp. 153–155, 2002.
- [27] (2009) Encarta world english dictionary. Microsoft. [Online]: encarta.msn.com
- [28] S. Ahuja, Ed., *Chromatography and Separation Science*, ser. Separation Science and Technology. Elsevier Science, 2003, vol. 4.
- [29] P. H. O'Farrell, "High resolution two-dimensional electrophoresis of proteins," *Journal of Biological Chemistry*, vol. 250, no. 10, pp. 4007–4021, 1975.
- [30] H. Zhu and M. Snyder, "Protein chip technology," *Current Opinion in Chemical Biology*, vol. 7, no. 1, pp. 55–63, 2003.
- [31] J. Seader and E. Henley, *Separation Process Principles*, 2nd ed. John Wiley and Sons, 2006.
- [32] C. Simpson, Ed., *Practical High Performance Liquid Chromatography*. Heyden and Son, 1979.
- [33] A. D. McNaught and A. Wilkinson, Eds., *Compendium of chemical terminology : IUPAC recommendations*. Blackwell Science, 1997.

- [34] B. Herbert and P. G. Righetti, "A turning point in proteome analysis: Sample prefractionation via multicompartment electrolyzers with isoelectric membranes," *Electrophoresis*, vol. 21, no. 17, pp. 3639–3648, 2000.
- [35] P. G. Righetti, A. Castagna, P. Antonioli, and E. Boschetti, "Prefractionation techniques in proteome analysis: The mining tools of the third millennium," *Electrophoresis*, vol. 26, no. 2, pp. 297–319, 2005.
- [36] R. Westermeier and T. Naven, *Proteomics in Practice*, second edition ed. Wiley-VCH, 2008.
- [37] J.-C. Sanchez, G. L. Corthals, and D. F. Hochstrasser, Eds., *Biomedical Applications of Proteomics*. Wiley-VCH, 2004.
- [38] M. Hamdan and P. G. Righetti, *Proteomics Today*, ser. Wiley-Interscience Series in Mass Spectrometry. Wiley-Interscience, 2005, vol. Protein Assessment and Biomarkers Using Mass Spectrometry, 2D Electrophoresis, and Microarray Technology.
- [39] A. L. Shapiro, E. V. nuela, and J. V. Maizel, "Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels," *Biochemical and Biophysical Research Communications*, vol. 28, no. 5, pp. 815–820, 1967.
- [40] R. Weinberger, *Practical Capillary Electrophoresis*. Academic Press Ltd, 1993.
- [41] P. G. Righetti, "Capillary electrophoretic analysis of proteins and peptides of biomedical and pharmacological interest," *Biopharmaceutics & Drug Disposition*, vol. 22, no. 7–8, pp. 337–351, 2001.
- [42] S. Babu, E. J. Song, S. M. E. Babar, M. H. Wi, and Y. S. Yoo, "Capillary electrophoresis at the omics level: Towards systems biology," *Electrophoresis*, vol. 27, no. 1, pp. 97–110, 2006.
- [43] *Mosby's Medical Dictionary*, 8th ed. Elsevier / Mosby Inc, 2009.
- [44] J. C. Giddings, "Concepts and comparisons in multidimensional separation," *Journal of High Resolution Chromatography*, vol. 10, no. 5, pp. 319–323, 1987.
- [45] B. D. Hames and D. Rickwood, Eds., *Gel electrophoresis of proteins: A practical approach*, second edition ed. Oxford University Press, 1998.
- [46] Expert Protein Analysis System (ExPASy). Swiss Institute of Bioinformatics. [Online]: www.expasy.ch
- [47] J. Taylor, N. L. Anderson, and N. G. Anderson, "Numerical measures of two-dimensional gel resolution and positional reproducibility," *Electrophoresis*, vol. 4, no. 5, pp. 338–346, 1983.
- [48] A. Abbott, "A post-genomic challenge: learning to read patterns of protein synthesis," *Nature*, vol. 402, no. 6763, pp. 715–720, 1999.
- [49] D. P. Clark, *Molecular Biology: Understanding the Genetic Revolution*. Elsevier Academic Press, 2005.
- [50] A. Görg, C. Obermaier, G. Boguth, A. Harder, B. Scheibe, R. Wildgruber, and W. Weiss, "The current state of two-dimensional electrophoresis with immobilized pH gradients," *Electrophoresis*, vol. 21, no. 6, pp. 1037–1053, 2000.

- [51] M. Ünlü, M. E. Morgan, and J. S. Minden, “Difference gel electrophoresis. A single gel method for detecting changes in protein extracts,” *Electrophoresis*, vol. 18, no. 11, pp. 2071–2077, 1997.
- [52] R. Marouga, S. David, and E. Hawkins, “The development of the DIGE system: 2D fluorescence difference gel analysis technology,” *Analytical and Bioanalytical Chemistry*, vol. 382, no. 3, pp. 669–678, 2005.
- [53] H. Wang and S. Hanash, “Multi-dimensional liquid phase based separations in proteomics,” *Journal of Chromatography B*, vol. 787, no. 1, pp. 11–18, 2003, proteomic Databases Part II.
- [54] C. R. Evans and J. Jorgenson, “Multidimensional LC-LC and LC-CE for high-resolution separations of biological molecules,” *Analytical and bioanalytical chemistry*, vol. 378, no. 8, pp. 1952–1961, 2004.
- [55] I. F. cois, K. Sandra, and P. Sandra, “Comprehensive liquid chromatography: Fundamental aspects and practical considerations—A review,” *Analytica Chimica Acta*, vol. 641, no. 1–2, pp. 14–31, 2009.
- [56] G. J. Sommer, A. K. Singh, and A. V. Hatch, “On-chip isoelectric focusing using photopolymerized immobilized pH gradients,” *Analytical Chemistry*, vol. 80, no. 9, pp. 3327–3333, 2008.
- [57] C. Wingren and C. A. Borrebaeck, “Antibody microarray analysis of directly labelled complex proteomes,” *Current Opinion in Biotechnology*, vol. 19, no. 1, pp. 55–61, 2008, analytical biotechnology.
- [58] S. Li, *Capillary electrophoresis: principles, practice, and applications*. Elsevier, 1992.
- [59] P. G. Righetti, *Immobilized pH gradients: Theory and methodology*, ser. Laboratory techniques in biochemistry and medical biology. Elsevier, 1990, vol. 20.
- [60] P. G. Righetti, “The Alpher, Bethe and Gamow of IEF, the alpha-Centaury of electrokinetic methodologies. Part II: Immobilized pH gradients,” *Electrophoresis*, vol. 28, no. 4, pp. 545–555, 2007.
- [61] M. Bier, J. Ostrem, and R. B. Marquez, “A new buffering system and its use in electrophoresis and isoelectric-focusing,” *Electrophoresis*, vol. 14, no. 10, pp. 1011–1018, 1993.
- [62] P. G. Righetti, M. Fazio, and C. Tonani, “Buffer isoelectric focusing revisited,” *Journal of Chromatography A*, vol. 440, pp. 367–377, 1988.
- [63] N. I. Tracy and C. F. Ivory, “Preparative isoelectric focusing of proteins using binary buffers in a vortex-stabilized, free-flow apparatus,” *Electrophoresis*, vol. 25, no. 12, pp. 1748–1757, 2004.
- [64] A. Tulp, D. Verwoerd, and A. A. M. Hart, “Density gradient isoelectric focusing of proteins in artificial pH gradients made up of binary mixtures of amphoteric buffers,” *Electrophoresis*, vol. 18, no. 5, pp. 767–773, 1997.

- [65] B. Bjellqvist, K. Ek, P. G. Righetti, E. Gianazza, A. Görg, R. Westermeier, and W. Postel, “Isoelectric-focusing in immobilized pH gradients—Principle, methodology and some applications,” *Journal of Biochemical and Biophysical Methods*, vol. 6, no. 4, pp. 317–339, 1982.
- [66] K. Salomon, D. S. Burgi, and J. C. Helmer, “Evaluation of fundamental properties of a silica capillary used for capillary electrophoresis,” *Journal of Chromatography A*, vol. 559, no. 1–2, pp. 69–80, 1991.
- [67] J. N. Israelachvili, *Intermolecular and surface forces*. Academic Press, 1985.
- [68] A. S. Cohen and B. L. Karger, “High-performance sodium dodecyl sulfate polyacrylamide gel capillary electrophoresis of peptides and proteins,” *Journal of Chromatography A*, vol. 397, pp. 409–417, 1987.
- [69] A. Sartori, V. Barbier, and J.-L. Viovy, “Sieving mechanisms in polymeric matrices,” *Electrophoresis*, vol. 24, no. 3, pp. 421–440, 2003.
- [70] G. J. M. Bruin, G. Stegeman, A. C. Van Asten, X. Xu, J. C. Kraak, and H. Poppe, “Optimization and evaluation of the performance of arrangements for UV detection in high-resolution separations using fused-silica capillaries,” *Journal of Chromatography A*, vol. 559, no. 1–2, pp. 163–181, 1991.
- [71] K. Kitagishi and Y. Sato, “Significant improvement of signal-to-noise ratio in capillary electrophoresis through optimization of aperture width for UV absorption detection,” *Electrophoresis*, vol. 22, no. 16, pp. 3395–3400, 2001.
- [72] G. P. Rozing, “Diode array detection,” in *High Resolution Separation and Analysis of Biological Macromolecules Part A: Fundamentals*, ser. Methods in Enzymology, B. L. Karger and W. S. Hancock, Eds. Academic Press, 1996, vol. 270, pp. 201 – 232.
- [73] X. Zhang, J. Stuart, and J. Sweedler, “Capillary electrophoresis with wavelength-resolved laser-induced fluorescence detection,” *Analytical and Bioanalytical Chemistry*, vol. 373, no. 6, pp. 332–343, 2002.
- [74] N. P. Beard and A. J. de Mello, “A polydimethylsiloxane/glass capillary electrophoresis microchip for the analysis of biogenic amines using indirect fluorescence detection,” *Electrophoresis*, vol. 23, no. 11, pp. 1722–1730, 2002.
- [75] S. M. Lunte, R. S. Martin, A. Gawron, N. Lacher, B. Fogarty, and F. Reagan, “Bioanalytical applications of microchip capillary electrophoresis with electrochemical detection.” *Abstracts Of Papers Of The American Chemical Society*, vol. 221, p. 214, 2001.
- [76] W. Yurong and C. Hengwu, “Microchip capillary electrophoresis with amperometric detection,” *Progress in Chemistry*, vol. 21, no. 1, pp. 200–209, 2009.
- [77] D. C. Neckers, “Photochemical reactions of natural macromolecules—Photoreactions of proteins,” *Journal of Chemical Education*, vol. 50, no. 3, pp. 164–168, 1973.
- [78] D. Janasek, J. Franzke, and A. Manz, “Scaling and the design of miniaturized chemical-analysis systems,” *Nature*, vol. 442, no. 7101, pp. 374–380, 2006.

- [79] A. Manz, N. Graber, and H. Widmer, “Miniaturized total chemical analysis systems: A novel concept for chemical sensing,” *Sensors and Actuators B: Chemical*, vol. 1, no. 1–6, pp. 244–248, 1990.
- [80] D. R. Reyes, D. Iossifidis, P.-A. Auroux, and A. Manz, “Micro total analysis systems. 1. Introduction, theory, and technology,” *Analytical Chemistry*, vol. 74, no. 12, pp. 2623–2636, 2002.
- [81] S. L. S. Freire and A. R. Wheeler, “Proteome-on-a-chip: Mirage, or on the horizon?” *Lab on a Chip*, vol. 6, no. 11, pp. 1415–1423, 2006.
- [82] C. J. Johnson, N. Zhukovsky, A. E. G. Cass, and J. M. Nagy, “Proteomics, nanotechnology and molecular diagnostics,” *Proteomics*, vol. 8, no. 4, pp. 715–730, 2008.
- [83] C. T. Culbertson and J. W. Jorgenson, “Lowering the UV absorbance detection limit and increasing the sensitivity of capillary electrophoresis using a dual linear photodiode array detector and signal averaging,” *Journal of Microcolumn Separations*, vol. 11, no. 9, pp. 652–662, 1999.
- [84] J. F. HASSARD and D. J. COLLING, “Molecular imaging,” USA Patent 6 613 210 B1, 2003.
- [85] “System and method,” USA Patent 7 497 934 B2, 2009.
- [86] “System and method,” USA Patent 7 425 252 B2, 2008.
- [87] F. Pereira, S. Hassard, A. Cass, and J. M. Nagy, “CZE of peptides using a multi-pixel detector array,” *Chimica Oggi-Chemistry Today*, vol. 26, no. 4, pp. 40–42, 2008.
- [88] F. Pereira, S. Hassard, J. Hassard, and A. de Mello, “CE of dsDNA in low-molecular-weight polyethylene oxide solutions,” *Electrophoresis*, vol. 30, no. 12, pp. 2100–2109, 2009.
- [89] L. Caldwell, “deltaDOT Background,” deltaDOT Ltd., Tech. Rep., 2005.
- [90] C. H. Ahn, J.-W. Choi, G. Beauchage, J. H. Nevin, J.-B. Lee, A. Puntambekar, and J. Y. Lee, “Disposable smart lab on a chip for point-of-care clinical diagnostics,” *Proceedings of the IEEE*, vol. 92, no. 1, pp. 154–173, 2004.
- [91] V. Dolnik and S. Liu, “Applications of capillary electrophoresis on microchip,” *Journal of Separation Science*, vol. 28, no. 15, pp. 1994–2009, 2005.
- [92] A. Manz, D. J. Harrison, E. M. J. Verpoorte, J. C. Fettinger, A. Paulus, H. Ludi, and H. M. Widmer, “Planar chips technology for miniaturization and integration of separation techniques into monitoring systems: Capillary electrophoresis on a chip,” *Journal of Chromatography A*, vol. 593, no. 1–2, pp. 253–258, 1992.
- [93] D. J. Harrison, A. Manz, Z. Fan, H. Luedi, and H. M. Widmer, “Capillary electrophoresis and sample injection systems integrated on a planar glass chip,” *Analytical Chemistry*, vol. 64, no. 17, pp. 1926–1932, 1992.
- [94] S.-W. Tsai, M. Loughran, H. Suzuki, and I. Karube, “Native and sodium dodecyl sulfate-capillary gel electrophoresis of proteins on a single microchip,” *Electrophoresis*, vol. 25, no. 3, pp. 494–501, 2004.

- [95] S. Yi, Y. C. Kwok, and N. Nam-Trung, “Low-pressure, high-temperature thermal bonding of polymeric microfluidic devices and their applications for electrophoretic separation,” *Journal of Micromechanics and Microengineering*, vol. 16, no. 8, p. 1681, 2006.
- [96] M. Agirregabiria, F. J. Blanco, J. Berganzo, A. Fullaondo, A. M. Zubiaga, K. Mayora, and J. M. Ruano-López, “SDS-CGE of proteins in microchannels made of SU-8 films,” *Electrophoresis*, vol. 27, no. 18, pp. 3627–3634, 2006.
- [97] D. Paul, A. Pallandre, S. Miserere, J. Weber, and J.-L. Viovy, “Lamination-based rapid prototyping of microfluidic devices using flexible thermoplastic substrates,” *Electrophoresis*, vol. 28, no. 7, pp. 1115–1122, 2007.
- [98] S. K. Mohanty, D. Kim, and D. J. Beebe, “Do-it-yourself microelectrophoresis chips with integrated sample recovery,” *Electrophoresis*, vol. 27, no. 19, pp. 3772–3778, 2006.
- [99] A. Pallandre, B. de Lambert, R. Attia, A. M. Jonas, and J.-L. Viovy, “Surface treatment and characterization: Perspectives to electrophoresis and lab-on-chips,” *Electrophoresis*, vol. 27, no. 3, pp. 584–610, 2006.
- [100] G.-B. Lee, C.-H. Lin, K.-H. Lee, and Y.-F. Lin, “On the surface modification of microchannels for microcapillary electrophoresis chips,” *Electrophoresis*, vol. 26, no. 24, pp. 4616–4624, 2005.
- [101] C. X. Zhang and A. Manz, “Narrow sample channel injectors for capillary electrophoresis on microchips,” *Analytical Chemistry*, vol. 73, no. 11, pp. 2656–2662, 2001.
- [102] L. Zhu, C. S. Lee, and D. L. DeVoe, “Integrated microfluidic UV absorbance detector with attomol-level sensitivity for BSA,” *Lab on a Chip*, vol. 6, no. 1, pp. 115–120, 2006.
- [103] N. P. Beard, C.-X. Zhang, and A. J. de Mello, “In-column field-amplified sample stacking of biogenic amines on microfabricated electrophoresis devices,” *Electrophoresis*, vol. 24, no. 4, pp. 732–739, 2003.
- [104] Caliper Life Sciences Inc. [Online]: www.caliperls.com
- [105] O. Hofmann, D. Che, K. A. Cruickshank, and U. R. Muller, “Adaptation of capillary isoelectric focusing to microchannels on a glass chip,” *Analytical Chemistry*, vol. 71, no. 3, pp. 678–686, 1999.
- [106] W. Tan, Z. H. Fan, C. X. Qiu, A. J. Ricco, and I. Gibbons, “Miniaturized capillary isoelectric focusing in plastic microfluidic devices,” *Electrophoresis*, vol. 23, no. 20, pp. 3638–3645, 2002.
- [107] J. Han and A. K. Singh, “Rapid protein separations in ultra-short microchannels: microchip sodium dodecyl sulfate-polyacrylamide gel electrophoresis and isoelectric focusing,” *Journal of Chromatography A*, vol. 1049, no. 1–2, pp. 205–209, 2004.
- [108] C. Das and Z. H. Fan, “Effects of separation length and voltage on isoelectric focusing in a plastic microfluidic device,” *Electrophoresis*, vol. 27, no. 18, pp. 3619–3626, 2006.

- [109] C. Li, Y. Yang, H. G. Craighead, and K. H. Lee, “Isoelectric focusing in cyclic olefin copolymer microfluidic channels coated by polyacrylamide using a UV photografting method,” *Electrophoresis*, vol. 26, no. 9, pp. 1800–1806, 2005.
- [110] Y. Li, D. L. DeVoe, and C. S. Lee, “Dynamic analyte introduction and focusing in plastic microfluidic devices for proteomic analysis,” *Electrophoresis*, vol. 24, no. 1–2, pp. 193–199, 2003.
- [111] H. Cui, K. Horiuchi, P. Dutta, and C. F. Ivory, “Multistage isoelectric focusing in a polymeric microfluidic chip,” *Analytical Chemistry*, vol. 77, no. 24, pp. 7878–7886, 2005.
- [112] Y. Xu, C.-X. Zhang, D. Janasek, and A. Manz, “Sub-second isoelectric focusing in free flow using a microfluidic device,” *Lab on a chip*, vol. 3, no. 4, pp. 224–227, 2003.
- [113] D. Kohlheyer, G. A. J. Besselink, S. Schlautmann, and R. B. M. Schasfoort, “Free-flow zone electrophoresis and isoelectric focusing using a microfabricated glass device with ion permeable membranes,” *Lab on a chip*, vol. 6, no. 3, pp. 374–380, 2006.
- [114] J. Chmelik, “Isoelectric focusing field-flow fractionation: Experimental study of the generation of pH gradient,” *Journal of Chromatography A*, vol. 539, no. 1, pp. 111–121, 1991.
- [115] Y. A. Song, S. Hsu, A. L. Stevens, and J. Han, “Continuous-flow pI-based sorting of proteins and peptides in a microfluidic chip using diffusion potential,” *Analytical Chemistry*, vol. 78, no. 11, pp. 3528–3536, 2006.
- [116] R. B. Schoch, A. Bertsch, and P. Renaud, “pH-controlled diffusion of proteins with different pI values across a nanochannel on a chip,” *Nano Lett.*, vol. 6, no. 3, pp. 543–547, 2006.
- [117] C.-W. Li, R. Chen, and M. Yang, “Generation of linear and non-linear concentration gradients along microfluidic channel by microtunnel controlled stepwise addition of sample solution,” *Lab on a Chip*, vol. 7, no. 10, pp. 1371–1373, 2007.
- [118] W. Saadi, S. Rhee, F. Lin, B. Vahidi, B. Chung, and N. Jeon, “Generation of stable concentration gradients in 2D and 3D environments using a microfluidic ladder chamber,” *Biomedical Microdevices*, vol. 9, no. 5, pp. 627–635, 2007.
- [119] R. A. Brennen, H. Yin, and K. P. Killeen, “Microfluidic gradient formation for nanoflow chip LC,” *Analytical Chemistry*, vol. 79, no. 24, pp. 9302–9309, 2007.
- [120] M. W. L. Watson, J. M. Mudrik, and A. R. Wheeler, “Gradient elution in microchannel electrochromatography,” *Analytical Chemistry*, vol. 81, no. 10, pp. 3851–3857, 2009.
- [121] W. E. Rae, J. E. Wong, and C. A. Lucy, “A simple means of generating pH gradients in capillary zone electrophoresis,” *Journal of Chromatography A*, vol. 781, no. 1–2, pp. 3–10, 1997.
- [122] G. Zhu, C. Yang, L. Zhang, Z. Liang, W. Zhang, and Y. Zhang, “Preparation and application of monolithic columns with narrow immobilized pH gradients,” *Talanta*, vol. 70, no. 1, pp. 2–6, 2006.

- [123] G. Zhu, C. Yang, L. Zhang, W. Zhang, and Y. Zhang, “Repeatedly usable immobilized pH gradient in a monolithic capillary column,” *Electrophoresis*, vol. 25, no. 12, pp. 1729–1734, 2004.
- [124] G. Zhu, H. Yuan, P. Zhao, L. Zhang, Z. Liang, W. Zhang, and Y. Zhang, “Macro-porous polyacrylamide-based monolithic column with immobilized pH gradient for protein analysis,” *Electrophoresis*, vol. 27, no. 18, pp. 3578–3583, 2006.
- [125] A. Tan, A. Pashkova, L. Zang, F. Foret, and B. L. Karger, “A miniaturized multichamber solution isoelectric focusing device for separation of protein digests,” *Electrophoresis*, vol. 23, no. 20, pp. 3599–3607, 2002.
- [126] G. Zilberstein, L. Korol, S. Bukshpan, and E. Baskin, “Parallel isoelectric focusing chip,” *Proteomics*, vol. 4, no. 9, pp. 2533–2540, 2004.
- [127] T. Huang and J. Pawliszyn, “Microfabrication of a tapered channel for isoelectric focusing with thermally generated pH gradient,” *Electrophoresis*, vol. 23, no. 20, pp. 3504–3510, 2002.
- [128] C. H. Lochmuller and C. S. Ronsick, “Isoelectric focusing with thermally formed, stepped-ramp pH gradients: Separation of human hemoglobin variants A and S,” *Analytica Chimica Acta*, vol. 249, no. 1, pp. 297–302, 1991.
- [129] J. Pawliszyn and J. Q. Wu, “Ampholyte-free isoelectric-focusing of proteins in cone-shaped capillaries,” *Journal of Microcolumn Separations*, vol. 5, no. 5, pp. 397–401, 1993.
- [130] X. H. Fang, M. Adams, and J. Pawliszyn, “A model of thermally generated pH gradients in tapered capillaries,” *Analyst*, vol. 124, no. 3, pp. 335–341, 1999.
- [131] T. Huang, X. Z. Wu, and J. Pawliszyn, “Capillary isoelectric focusing without carrier ampholytes,” *Analytical Chemistry*, vol. 72, no. 19, pp. 4758–4761, 2000.
- [132] C. R. Cabrera, B. Finlayson, and P. Yager, “Formation of natural pH gradients in a microfluidic device under flow conditions: Model and experimental validation,” *Analytical Chemistry*, vol. 73, no. 3, pp. 658–666, 2001.
- [133] C. F. Ivory, “A brief review of alternative electrofocusing techniques,” *Separation Science and Technology*, vol. 35, no. 11, pp. 1777–1793, 2000.
- [134] A. Griebel, “Integrated polymer chip for two-dimensional capillary gel electrophoresis,” *Lab on a chip*, vol. 4, no. 1, pp. 18–23, 2004.
- [135] N. Gottschlich, S. C. Jacobson, C. T. Culbertson, and J. M. Ramsey, “Two-dimensional electrochromatography/capillary electrophoresis on a microchip,” *Analytical Chemistry*, vol. 73, no. 11, pp. 2669–2674, 2001.
- [136] X. Chen, H. Wu, C. Mao, and G. M. Whitesides, “A prototype two-dimensional capillary electrophoresis system fabricated in poly(dimethylsiloxane),” *Analytical Chemistry*, vol. 74, no. 8, pp. 1772–1778, 2002.
- [137] C. Das, J. Zhang, N. D. Denslow, and Z. H. Fan, “Integration of isoelectric focusing with multi-channel gel electrophoresis by using microfluidic pseudo-valves,” *Lab on a Chip*, vol. 7, no. 12, pp. 1806–1812, 2007.

- [138] H. Becker, K. Lowack, and A. Manz, "Planar quartz chips with submicron channels for two-dimensional capillary electrophoresis applications," *Journal of Micromechanics and Microengineering*, vol. 8, no. 1, p. 24, 1998.
- [139] S. Yang, J. Liu, C. S. Lee, and D. L. DeVoe, "Microfluidic 2-D PAGE using multifunctional in situ polyacrylamide gels and discontinuous buffers," *Lab on a Chip*, vol. 9, no. 4, pp. 592–599, 2009.
- [140] J. Liu, S. Yang, C. S. Lee, and D. L. DeVoe, "Polyacrylamide gel plugs enabling 2-D microfluidic protein separations via isoelectric focusing and multiplexed sodium dodecyl sulfate gel electrophoresis," *Electrophoresis*, vol. 29, no. 11, pp. 2241–2250, 2008.
- [141] S.-W. Tsai, M. Loughran, and I. Karube, "Development of a microchip for 2-dimensional capillary electrophoresis," *Journal of Micromechanics and Microengineering*, vol. 14, no. 12, p. 1693, 2004.
- [142] Y. Li, J. S. Buch, F. Rosenberger, D. L. DeVoe, and C. S. Lee, "Integration of isoelectric focusing with parallel sodium dodecyl sulfate gel electrophoresis for multidimensional protein separations in a plastic microfluidic network," *Analytical Chemistry*, vol. 76, no. 3, pp. 742–748, 2004.
- [143] Y. M. Liu and J. V. Sweedler, "Two-dimensional separations: Capillary electrophoresis coupled to channel gel electrophoresis," *Analytical Chemistry*, vol. 68, no. 22, pp. 3928–3933, 1996.
- [144] C. A. Emrich, I. L. Medintz, W. K. Chu, and R. A. Mathies, "Microfabricated two-dimensional electrophoresis device for differential protein expression profiling," *Analytical Chemistry*, vol. 79, no. 19, pp. 7360–7366, 2007.
- [145] Z. Demianová, M. Shimmo, E. Pöysä, S. Franssila, and M. Baumann, "Toward an integrated microchip sized 2-D polyacrylamide slab gel electrophoresis device for proteomic analysis," *Electrophoresis*, vol. 28, no. 3, pp. 422–428, 2007.
- [146] K. Usui, A. Hiratsuka, K. Shiseki, Y. Maruo, T. Matsushima, K. Takahashi, Y. Unuma, K. Sakairi, I. Namatame, Y. Ogawa, and K. Yokoyama, "A self-contained polymeric 2-DE chip system for rapid and easy analysis," *Electrophoresis*, vol. 27, no. 18, pp. 3635–3642, 2006.
- [147] A. Xu, C. Sluszny, and E. S. Yeung, "Prototype for integrated two-dimensional gel electrophoresis for protein separation," *Journal of Chromatography A*, vol. 1087, no. 1–2, pp. 177–182, 2005.
- [148] A. Hiratsuka, H. Kinoshita, Y. Maruo, K. Takahashi, S. Akutsu, C. Hayashida, K. Sakairi, K. Usui, K. Shiseki, H. Inamochi, Y. Nakada, K. Yodoya, I. Namatame, Y. Unuma, M. Nakamura, K. Ueyama, Y. Ishii, K. Yano, and K. Yokoyama, "Fully automated two-dimensional electrophoresis system for high-throughput protein analysis," *Analytical Chemistry*, vol. 79, no. 15, pp. 5730–5739, 2007.
- [149] Y. C. Wang, M. H. Choi, and J. Han, "Two-dimensional protein separation with advanced sample and buffer isolation using microfluidic valves," *Analytical Chemistry*, vol. 76, no. 15, pp. 4426–4431, 2004.

- [150] A. E. Herr, J. I. Molho, K. A. Drouvalakis, J. C. Mikkelsen, P. J. Utz, J. G. Santiago, and T. W. Kenny, “On-chip coupling of isoelectric focusing and free solution electrophoresis for multidimensional separations,” *Analytical Chemistry*, vol. 75, no. 5, pp. 1180–1187, 2003.
- [151] H. Liu, C. Yang, Q. Yang, W. Zhang, and Y. Zhang, “On-line combination of capillary isoelectric focusing and capillary non-gel sieving electrophoresis using a hollow-fiber membrane interface: a novel two-dimensional separation system for proteins,” *Journal of Chromatography B*, vol. 817, no. 1, pp. 119–126, 2005.
- [152] D. Mohan and C. S. Lee, “On-line coupling of capillary isoelectric focusing with transient isotachophoresis-zone electrophoresis: A two-dimensional separation system for proteomics,” *Electrophoresis*, vol. 23, no. 18, pp. 3160–3167, 2002.
- [153] Y. Cong, L. Zhang, D. Tao, Y. Liang, W. Zhang, and Y. Zhang, “Miniaturized two-dimensional capillary electrophoresis on a microchip for analysis of the tryptic digest of proteins,” *Journal of Separation Science*, vol. 31, no. 3, pp. 588–594, 2008.
- [154] C. Yang, L. Zhang, H. Liu, W. Zhang, and Y. Zhang, “Two-dimensional capillary electrophoresis involving capillary isoelectric focusing and capillary zone electrophoresis,” *Journal of Chromatography A*, vol. 1018, no. 1, pp. 97–103, 2003.
- [155] J. D. Ramsey, S. C. Jacobson, C. T. Culbertson, and J. M. Ramsey, “High-efficiency, two-dimensional separations of protein digests on microfluidic devices,” *Analytical Chemistry*, vol. 75, no. 15, pp. 3758–3764, 2003.
- [156] H. Shadpour and S. A. Soper, “Two-dimensional electrophoretic separation of proteins using poly(methyl methacrylate) microchips,” *Analytical Chemistry*, vol. 78, no. 11, pp. 3519–3527, 2006.
- [157] M. M. Harwood, E. S. Christians, M. A. Fazal, and N. J. Dovichi, “Single-cell protein analysis of a single mouse embryo by two-dimensional capillary electrophoresis,” *Journal of Chromatography A*, vol. 1130, no. 2, pp. 190–194, 2006.
- [158] X. Chen, A. Fazal, and N. J. Dovichi, “CSE-MECC two-dimensional capillary electrophoresis analysis of proteins in the mouse tumor cell (AtT-20) homogenate,” *Talanta*, vol. 71, no. 5, pp. 1981–1985, 2007.
- [159] J. K. Osiri, H. Shadpour, S. Park, B. C. Snowden, Z.-Y. Chen, and S. A. Soper, “Generating high peak capacity 2-D maps of complex proteomes using PMMA microchip electrophoresis,” *Electrophoresis*, vol. 29, no. 24, pp. 4984–4992, 2008.
- [160] R. D. Rocklin, R. S. Ramsey, and J. M. Ramsey, “A microfabricated fluidic device for performing two-dimensional liquid-phase separations,” *Analytical Chemistry*, vol. 72, no. 21, pp. 5244–5249, 2000.
- [161] E. Sahlin, “Two-dimensional capillary electrophoresis using tangentially connected capillaries,” *Journal of Chromatography A*, vol. 1154, no. 1–2, pp. 454–459, 2007.
- [162] L. Sheng and J. Pawliszyn, “Comprehensive two dimensional separation based on coupling micellar electrokinetic chromatography with capillary isoelectric focusing,” *Analyst*, vol. 127, no. 9, pp. 1159–1163, 2002.

- [163] D. Kang and M. H. Moon, "Development of non-gel-based two-dimensional separation of intact proteins by an on-line hyphenation of capillary isoelectric focusing and hollow fiber flow field-flow fractionation," *Analytical Chemistry*, vol. 78, no. 16, pp. 5789–5798, 2006.
- [164] J. Chen, C. S. Lee, Y. Shen, R. D. Smith, and E. H. Baehrecke, "Integration of capillary isoelectric focusing with capillary reversed-phase liquid chromatography for two-dimensional proteomics separation," *Electrophoresis*, vol. 23, no. 18, pp. 3143–3148, 2002.
- [165] M. Hanna, C. Simpson, and D. Perrett, "Novel three-dimensional capillary electrophoresis system for complex and trace analysis," *Journal of Chromatography A*, vol. 894, no. 1–2, pp. 117–128, 2000.
- [166] D. C. Simpson and R. D. Smith, "Combining capillary electrophoresis with mass spectrometry for applications in proteomics," *Electrophoresis*, vol. 26, no. 7–8, pp. 1291–1305, 2005.
- [167] M. Fujita, W. Hattori, T. Sano, M. Baba, H. Someya, K. Miyazaki, K. Kamijo, K. Takahashi, and H. Kawaura, "High-throughput and high-resolution two dimensional mapping of pI and m/z using a microchip in a matrix-assisted laser desorption/ionization time-of-flight mass spectrometer," *Journal of Chromatography A*, vol. 1111, no. 2, pp. 200–205, 2006.
- [168] Z. Liu, T. Lemma, and J. Pawliszyn, "Capillary isoelectric focusing coupled with dynamic imaging detection: A one-dimensional separation for two-dimensional protein characterization," *Journal of Proteome Research*, vol. 5, no. 5, pp. 1246–1251, 2006.
- [169] W. L. Tseng, Y. W. Lin, and H. T. Chang, "Improved separation of microheterogeneities and isoforms of proteins by capillary electrophoresis using segmental filling with SDS and PEO in the background electrolyte," *Analytical Chemistry*, vol. 74, no. 18, pp. 4828–4834, 2002.
- [170] P. Sedláková, J. Svobodova, and I. Mikšík, "Capillary electrophoresis of peptides and proteins with plug of Pluronic gel," *Journal of Chromatography B*, vol. 839, no. 1–2, pp. 112–117, 2006.
- [171] "Method and apparatus for separating biological molecules," USA Patent 6 979 390 B2, 2005.
- [172] "Separation of biomolecules," World Intellectual Property Organization Patent 2007/11067 A1, 2007.
- [173] "Material separation device," World Intellectual Property Organization Patent 2 004 005 910 (A1), 2004.
- [174] J. Edgar, G. Milne, Y. Zhao, C. Pabbati, D. Lim, and D. Chiu, "Compartmentalization of chemically separated components into droplets," *Angewandte Chemie International Edition*, vol. 48, no. 15, pp. 2719–2722, 2009.
- [175] R. J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*, ser. The Manchester Physics Series. John Wiley and Sons, 1989.

- [176] D. Birkes and Y. Dodge, *Alternative methods of regression*, ser. Wiley Series in Probability and Statistics. Wiley-Interscience, 1993.
- [177] F. S. Acton, *Analysis of straight-line data*. New York: Dover, 1966.
- [178] M. Albin, P. D. Grossman, and S. E. Moring, “Sensitivity enhancement for capillary electrophoresis,” *Analytical Chemistry*, vol. 65, no. 10, pp. 489A–497A, 1993.
- [179] D. Wu, J. Qin, and B. Lin, “Electrophoretic separations on microfluidic chips,” *Journal of Chromatography A*, vol. 1184, no. 1–2, pp. 542–559, 2008.
- [180] S. Hjerten, S. Mohabbati, and D. Westerlund, “Influence of ignored and well-known zone distortions on the separation performance of proteins in capillary free zone electrophoresis with special reference to analysis in polyacrylamide-coated fused silica capillaries in various buffers: I. theoretical studies,” *Journal of Chromatography A*, vol. 1053, no. 1–2, pp. 181–199, 2004.
- [181] B. Gas, M. Stedrý, and E. Kenndler, “Peak broadening in capillary zone electrophoresis,” *Electrophoresis*, vol. 18, no. 12–13, pp. 2123–2133, 1997.
- [182] A. Manz, L. Bousse, A. Chow, T. B. Metha, A. Kopf-Sill, and J. W. Parce, “Synchronized cyclic capillary electrophoresis using channels arranged in a triangle and low voltages,” *Fresenius Journal of Analytical Chemistry*, vol. 371, no. 2, pp. 195–201, 2001.
- [183] Z. K. Shihabi, “Stacking in capillary zone electrophoresis,” *Journal of Chromatography A*, vol. 902, no. 1, pp. 107–117, 2000.
- [184] K. Swinney and D. J. Bornhop, “Detection in capillary electrophoresis,” *Electrophoresis*, vol. 21, no. 7, pp. 1239–1250, 2000.
- [185] K. B. Mogensen, H. Klank, and J. P. Kutter, “Recent developments in detection for microfluidic systems,” *Electrophoresis*, vol. 25, no. 21–22, pp. 3498–3512, 2004.
- [186] A. J. Pfeiffer, T. Mukherjee, and S. Hauan, “Design and optimization of compact microscale electrophoretic separation systems,” *Industrial & Engineering Chemistry Research*, vol. 43, no. 14, pp. 3539–3553, 2004.
- [187] A. Guttman, J. Horvath, and N. Cooke, “Influence of temperature on the sieving effect of different polymer matrixes in capillary SDS gel electrophoresis of proteins,” *Analytical Chemistry*, vol. 65, no. 3, pp. 199–203, 1993.
- [188] R. M. Latorre, J. Saurina, and S. Hernandez-Cassou, “Determination of amino acids in overlapped capillary electrophoresis peaks by means of partial least-squares regression,” *Journal of Chromatography A*, vol. 871, no. 1, pp. 331–340, 2000.
- [189] J. A. McReynolds, P. Edirisinghe, and S. A. Shippy, “Shah and sine convolution Fourier transform detection for microchannel electrophoresis with a charge coupled device,” *Analytical Chemistry*, vol. 74, no. 19, pp. 5063–5070, 2002.
- [190] H. J. Crabtree, M. U. Kopp, and A. Manz, “Shah convolution Fourier transform detection,” *Analytical Chemistry*, vol. 71, no. 11, pp. 2130–2138, 1999.

- [191] Y. C. Kwok and A. Manz, "Shah convolution differentiation Fourier transform for rear analysis in microchip capillary electrophoresis," *Journal of Chromatography A*, vol. 924, no. 1–2, pp. 177–186, 2001.
- [192] Y. C. Kwok and A. Manz, "Characterisation of Shah convolution Fourier transform detection," *The Analyst*, vol. 126, no. 10, pp. 1640–1644, 2001.
- [193] Y. C. Kwok and A. Manz, "Shah convolution Fourier transform detection: Multiple-sample injection technique," *Electrophoresis*, vol. 22, no. 2, pp. 222–229, 2001.
- [194] J. A. McReynolds, L. Gao, J. Barber-Singh, and S. A. Shippy, "Hadamard transform CE-UV detection for biological samples," *Journal of Separation Science*, vol. 28, no. 2, pp. 128–136, 2005.
- [195] *ProteomeLab SDS-MW Analysis Guide*, Beckman Coulter, Inc., January 2004, p/N 390963AB.
- [196] L. Ullman and A. Signer, *C++ Programming*. Peachpit Press, 2006.
- [197] Root v5.12. CERN. [Online]: root.cern.ch
- [198] L. A. Currie, "Nomenclature in evaluation of analytical methods including detection and quantification capabilities: (IUPAC recommendations 1995)," *Analytica Chimica Acta*, vol. 391, no. 2, pp. 105–126, 1999.
- [199] A. Guttman, P. Shieh, D. Hoang, J. Horvath, and N. Cooke, "Effect of operational variables on the separation of proteins by capillary sodium dodecyl sulfate-gel electrophoresis," *Electrophoresis*, vol. 15, no. 2, pp. 221–224, 1994.
- [200] X. Xuan and D. Li, "Band-broadening in capillary zone electrophoresis with axial temperature gradients," *Electrophoresis*, vol. 26, no. 1, pp. 166–175, 2005.
- [201] X. Xuan and D. Li, "Analytical study of Joule heating effects on electrokinetic transportation in capillary electrophoresis," *Journal of Chromatography A*, vol. 1064, no. 2, pp. 227–237, 2005.
- [202] K. Ganzler, K. S. Greve, A. S. Cohen, B. L. Karger, A. Guttman, and N. C. Cooke, "High-performance capillary electrophoresis of SDS-protein complexes using UV-transparent polymer networks," *Analytical Chemistry*, vol. 64, no. 22, pp. 2665–2671, 1992.
- [203] W. L. Tseng and H. T. Chang, "On-line concentration and separation of proteins by capillary electrophoresis using polymer solutions," *Analytical Chemistry*, vol. 72, no. 20, pp. 4805–4811, 2000.
- [204] K. Benedek and A. Guttman, "Ultra-fast high-performance capillary sodium dodecyl sulfate gel electrophoresis of proteins," *Journal of Chromatography A*, vol. 680, no. 2, pp. 375–381, 1994.
- [205] K. Benedek and S. Thiede, "High-performance capillary electrophoresis of proteins using sodium dodecyl sulfate-poly(ethylene oxide)," *Journal of Chromatography A*, vol. 676, no. 1, pp. 209–217, 1994.
- [206] T. Takagi, "Capillary electrophoresis in presence of sodium dodecyl sulfate and a sieving medium," *Electrophoresis*, vol. 18, no. 12–13, pp. 2239–2242, 1997.

- [207] M. R. Karim, J. C. Janson, and T. Takagi, "Size-dependent separation of proteins in the presence of sodium dodecyl-sulfate and dextran in capillary electrophoresis—Effect of molecular-weight of dextran," *Electrophoresis*, vol. 15, no. 12, pp. 1531–1534, 1994.
- [208] T. Takagi and M. R. Karim, "A new mode of size-dependent separation of proteins by capillary electrophoresis in presence of sodium dodecyl sulfate and concentrated oligomeric dextran," *Electrophoresis*, vol. 16, no. 1, pp. 1463–1467, 1995.
- [209] Reagents for protein molecular weight determination analysis- CE system. Groton Biosystems. [Online]: http://www.grotonbiosystems.com/products/consumables_reagents.cfm
- [210] D. Corradini, "Buffer additives other than the surfactant sodium dodecyl sulfate for protein separations by capillary electrophoresis," *Journal of Chromatography B: Biomedical Sciences and Applications*, vol. 699, no. 1–2, pp. 221–256, 1997.
- [211] C. J. Evenhuis and P. R. Haddad, "Joule heating effects and the experimental determination of temperature during CE," *Electrophoresis*, vol. 30, no. 5, pp. 897–909, 2009.
- [212] C. J. Evenhuis, R. M. Guijt, M. Macka, P. J. Marriott, and P. R. Haddad, "Temperature profiles and heat dissipation in capillary electrophoresis," *Analytical Chemistry*, vol. 78, no. 8, pp. 2684–2693, 2006.
- [213] W. A. Gobie and C. F. Ivory, "Thermal model of capillary electrophoresis and a method for counteracting thermal band broadening," *Journal of Chromatography A*, vol. 516, no. 1, pp. 191–210, 1990.
- [214] J. Knox, "Thermal effects and band spreading in capillary electro-separation," *Chromatographia*, vol. 26, no. 1, pp. 329–337, 1988.
- [215] J. Knox and K. McCormack, "Temperature effects in capillary electrophoresis. 1: Internal capillary temperature and effect upon performance," *Chromatographia*, vol. 38, no. 3, pp. 207–214, 1994.
- [216] J. Knox and K. McCormack, "Temperature effects in capillary electrophoresis 2: Some theoretical calculations and predictions," *Chromatographia*, vol. 38, no. 3, pp. 215–221, 1994.
- [217] K. O. Voss, H. P. Roos, and N. J. Dovichi, "The effect of temperature oscillations on DNA sequencing by capillary electrophoresis," *Analytical Chemistry*, vol. 73, no. 6, pp. 1345–1349, 2001.
- [218] P. Gebauer, P. Borecka, and P. Bocek, "Predicting peak symmetry in capillary zone electrophoresis. Background electrolytes with two co-ions: Schizophrenic zone broadening and the role of system peaks," *Analytical Chemistry*, vol. 70, no. 16, pp. 3397–3406, 1998.
- [219] Z. Deyl, I. Mikšík, J. Charvátová, and A. Eckhardt, "Comparison of the electrophoretic separation of proteins in capillaries with different inner diameter," *Journal of Chromatography A*, vol. 1013, no. 1–2, pp. 233–238, 2003.
- [220] Y. Sun, Y. Kwok, and N.-T. Nguyen, "Modeling and experimental characterization of peak tailing in DNA gel electrophoresis," *Microfluidics and Nanofluidics*, vol. 3, no. 3, pp. 323–332, 2007.

- [221] W. D. Wilson and J. F. Foster, "Photolysis accompanying peptide absorption in proteins: Electrophoresis and optical rotary dispersion studies," *Biophysical Journal*, vol. 12, no. 6, pp. 609–624, 1972.
- [222] R. R. Hill, J. D. Coyle, D. Birch, E. Dawe, G. E. Jeffs, D. Randall, I. Stec, and T. M. Stevenson, "Photochemistry of dipeptides in aqueous solution," *Journal of the American Chemical Society*, vol. 113, no. 5, pp. 1805–1817, 1991.
- [223] Y. A. Repeyev, E. V. Khoroshilova, and D. N. Nikogosyan, "212.8 nm laser photolysis of aromatic and aliphatic amino acids and related peptides," *Journal of Photochemistry and Photobiology B: Biology*, vol. 12, no. 3, pp. 259–274, 1992.
- [224] Y. F. Lion, M. Kuwabara, and P. Riesz, "UV photolysis of aqueous solutions of aliphatic peptides. an ESR and spin-trapping study," *The Journal of Physical Chemistry*, vol. 84, no. 25, pp. 3378–3384, 1980.
- [225] H. Görner and D. N. Nikogosyan, "Indirect 248 nm 20 ns photolysis of aliphatic amino acids in aqueous solution," *Journal of Photochemistry and Photobiology B: Biology*, vol. 39, no. 1, pp. 84–89, 1997.
- [226] M. Mulcahy, J. G. McInerney, D. N. Nikogosyan, and H. Goerner, "193 nm photolysis of aromatic and aliphatic dipeptides in aqueous solution: Dependence of decomposition quantum yield on the amino acid sequence," *Biological Chemistry*, vol. 381, no. 12, pp. 1259–1262, 2000.
- [227] F. Foret, S. Fanali, and P. Bocek, "Applicability of dynamic change of pH in the capillary zone electrophoresis of proteins," *Journal of Chromatography A*, vol. 516, no. 1, pp. 219–222, 1990.
- [228] J. Petr, V. Maier, J. Horáková, J. Scaronevcík, and Z. Stránský, "Capillary isotachophoresis from the student point of view—Images and the reality," *Journal of Separation Science*, vol. 29, no. 18, pp. 2705–2715, 2006.
- [229] F. Foret, K. Kleparník, P. Gebauer, and P. Bocek, "Ionic boundaries in biological capillary electrophoresis," *Journal of Chromatography A*, vol. 1053, no. 1–2, pp. 43–57, 2004.
- [230] S. Devasenathipathy, R. Bharadwaj, and J. Santiago, "Investigation of internal pressure gradients generated in electrokinetic flows with axial conductivity gradients," *Experiments in Fluids*, vol. 43, no. 6, pp. 959–967, 2007.
- [231] Ultratrol dynamic pre-coatings. Target Discovery, Inc. [Online]: www.targetdiscovery.com
- [232] C. Yang and Z. E. Rassi, "Capillary zone electrophoresis of proteins with fused-silica capillaries having polymers and surfactants adsorbed onto surfactant moieties previously covalently bound to the capillary column surface," *Electrophoresis*, vol. 19, no. 13, pp. 2278–2284, 1998.
- [233] O. Geschke, H. Klank, and P. Telleman, Eds., *Microsystem engineering of lab-on-a-chip devices*. Wiley-VCH, 2004.
- [234] M. Krogh and P. Asberg, *My Little Guide to Soft Lithography (or Soft Lithography for Dummies)*, Linköping University, 2003.

- [235] S. J. Lee, M. Goedert, M. T. Matyska, E. M. Ghandehari, M. Vijay, and J. J. Pesek, “Polymethylhydrosiloxane (PMHS) as a functional material for microfluidic chips,” *Journal of Micromechanics and Microengineering*, vol. 18, no. 2, p. 025026, 2008.
- [236] R. E. Oosterbroek, Ed., *Lab-on-a-chip : miniaturized systems for (bio)chemical analysis and synthesis*. Elsevier B.V., 2003.
- [237] J. Y. Shin, J. Y. Park, C. Liu, J. He, and S. C. Kim, “Chemical structure and physical properties of cyclic olefin copolymers,” International Union of Pure and Applied Chemistry (IUPAC), Tech. Rep., 2005.
- [238] C. K. Fredrickson, Z. Xia, C. Das, R. Ferguson, F. T. Tavares, and Z. H. Fan, “Effects of fabrication process parameters on the properties of cyclic olefin copolymer microfluidic devices,” *Microelectromechanical Systems, Journal of*, vol. 15, no. 5, pp. 1060–1068, 2006.
- [239] P. Menter, “Acrylamide polymerization—A practical approach,” Bio-Rad Laboratories, Tech. Rep. 1156.
- [240] A. V. Hatch, A. E. Herr, D. J. Throckmorton, J. S. Brennan, and A. K. Singh, “Integrated preconcentration SDS-PAGE of proteins in microchips using photopatterned cross-linked polyacrylamide gels,” *Analytical Chemistry*, vol. 78, no. 14, pp. 4976–4984, 2006.
- [241] C. T. Lo, D. J. Throckmorton, A. K. Singh, and A. E. Herr, “Photopolymerized diffusion-defined polyacrylamide gradient gels for on-chip protein sizing,” *Lab on a Chip*, vol. 8, no. 8, pp. 1273–1279, 2008.
- [242] T. Rohr, D. F. Ogletree, F. Svec, and J. M. J. Fréchet, “Surface functionalization of thermoplastic polymers for the fabrication of microfluidic devices by photoinitiated grafting,” *Advanced Functional Materials*, vol. 13, no. 4, pp. 264–270, 2003.
- [243] T. B. Stachowiak, D. A. Mair, T. G. Holden, L. J. Lee, F. Svec, and J. M. J. Fréchet, “Hydrophilic surface modification of cyclic olefin copolymer microfluidic chips using sequential photografting,” *Journal of Separation Science*, vol. 30, no. 7, pp. 1088–1093, 2007.
- [244] ImageJ. National Institutes of Health, USA. [Online]: <http://rsbweb.nih.gov/ij/>
- [245] *2-D Electrophoresis: Principles and Methods*, GE Healthcare Amersham Biosciences.
- [246] M. A. Lerch, M. D. Hoffman, and S. C. Jacobson, “Influence of channel position on sample confinement in two-dimensional planar microfluidic devices,” *Lab on a Chip*, vol. 8, no. 2, pp. 316–322, 2008.
- [247] J. Goulpeau, B. Lonetti, D. Trouchet, A. Ajdari, and P. Tabeling, “Building up longitudinal concentration gradients in shallow microchannels.” *Lab on a Chip*, vol. 7, no. 9, pp. 1154–1161, 2007.
- [248] N. L. Jeon, S. K. W. Dertinger, D. T. Chiu, I. S. Choi, A. D. Stroock, and G. M. Whitesides, “Generation of solution and surface gradients using microfluidic systems,” *Langmuir*, vol. 16, no. 22, pp. 8311–8316, 2000.

- [249] S. K. W. Dertinger, D. T. Chiu, N. L. Jeon, and G. M. Whitesides, “Generation of gradients having complex shapes using microfluidic networks,” *Analytical Chemistry*, vol. 73, no. 6, pp. 1240–1246, 2001.
- [250] D. Amarie, J. A. Glazier, and S. C. Jacobson, “Compact microfluidic structures for generating spatial and temporal gradients,” *Analytical Chemistry*, vol. 79, no. 24, pp. 9471–9477, 2007.
- [251] D. Irimia, D. A. Geba, and M. Toner, “Universal microfluidic gradient generator,” *Analytical Chemistry*, vol. 78, no. 10, pp. 3472–3477, 2006.
- [252] K. Lee, C. Kim, B. Ahn, R. Panchapakesan, A. R. Full, L. Nordee, J. Y. Kang, and K. W. Oh, “Generalized serial dilution module for monotonic and arbitrary microfluidic gradient generators.” *Lab on a Chip*, vol. 9, no. 5, pp. 709–717, 2009.
- [253] J. Liu and M. L. Lee, “Permanent surface modification of polymeric capillary electrophoresis microchips for protein and peptide analysis,” *Electrophoresis*, vol. 27, no. 18, pp. 3533–3546, 2006.
- [254] S. Hardt and F. Schönfeld, Eds., *Microfluidic Technologies for Miniaturized Analysis Systems*, ser. MEMS Reference Shelf. Springer, 2007.
- [255] V. S. M. und Physiklehrer, Ed., *Formeln und Tafeln Mathematik – Physik*. Orell Füssli, 1992.
- [256] D. R. Lide, Ed., *CRC handbook of chemistry and physics*. CRC, 2008.
- [257] “NanoTM SU-8 negative tone photoresist formulations 50–100,” MicroChem Corp., Tech. Rep., 2002.

Appendix A

Derivation of thermal model

A.1 Differential equations for temperature-time function

Assumptions:

- Uniform temperature throughout the capillary (in axial and in radial direction).
- Constant Joule heating (constant electric current and voltage).
- Constant environment temperature.

Joule heating:

$$P = \frac{\partial Q}{\partial t} = \dot{Q}_{HV} = VI \quad (\text{A.1})$$

with power P [W], energy Q [J], time t [s], voltage V [V], and current I [A].

Newton's law of cooling:

$$\frac{\partial Q}{\partial t} = \dot{Q}_{Co} = kS(T - T_{\infty}) \quad (\text{A.2})$$

with heat transfer coefficient k [W/(m²K)], surface area across which heat is transferred S [m²], capillary temperature T [K], and environment temperature T_{∞} [K].

Heat capacity Q_{Ca} [J]:

$$\begin{aligned} Q_{Ca} &= mcT \\ \dot{Q}_{Ca} &= mc\dot{T} \end{aligned} \quad (\text{A.3})$$

with mass m [kg], and specific heat capacity c [J/(kgK)].

Substituting

$$kS = \alpha$$

$$mc = \beta \quad (\text{A.4})$$

Equations A.1, A.2, and A.3 become:

$$\begin{aligned} \dot{Q}_{HV} &= VI \\ \dot{Q}_{Co} &= \alpha(T - T_\infty) \\ \dot{Q}_{Ca} &= \beta\dot{T} \end{aligned} \quad (\text{A.5})$$

A balance of energies gives:

$$\dot{Q}_{HV} - \dot{Q}_{Co} - \dot{Q}_{Ca} = 0 \quad (\text{A.6})$$

Applying Equations A.5 to the balance of energies leads to a differential equation for $T(t)$:

$$\begin{aligned} VI - \beta\dot{T} - \alpha T + \alpha T_\infty &= 0 \\ \beta\dot{T} + \alpha T &= \alpha T_\infty + VI \\ \dot{T} + \frac{\alpha}{\beta}T &= \frac{\alpha T_\infty + VI}{\beta} \end{aligned} \quad (\text{A.7})$$

The differential equation A.7 is solved with the following equations:^[255]

$$\begin{aligned} \text{for } y' + u(x)y &= v(x) \\ y &= G(x)e^{-U(x)} \end{aligned} \quad (\text{A.8})$$

$$\text{with } U(x) = \int u(x) \, dx \quad (\text{A.9})$$

$$G(x) = \int v(x)e^{U(x)} \, dx \quad (\text{A.10})$$

Applying this to Equation A.7, the differential equation for $T(t)$, we obtain:

$$\begin{aligned} U(t) &= \int \frac{\alpha}{\beta} dt = \frac{\alpha}{\beta}t + C_1 \\ G(t) &= \int \frac{\alpha T_\infty + VI}{\beta} e^{\frac{\alpha}{\beta}t+C_1} dt = \frac{\beta}{\alpha} \frac{\alpha T_\infty + VI}{\beta} e^{\frac{\alpha}{\beta}t+C_1} + C_2 \\ &= \left(T_\infty + \frac{VI}{\alpha} \right) e^{\frac{\alpha}{\beta}t+C_1} + C_2 \\ T(t) &= \left[\left(T_\infty + \frac{VI}{\alpha} \right) e^{\frac{\alpha}{\beta}t+C_1} + C_2 \right] e^{-\frac{\alpha}{\beta}t-C_1} \\ &= \left(T_\infty + \frac{VI}{\alpha} \right) + C e^{-\frac{\alpha}{\beta}t} \end{aligned} \quad (\text{A.11})$$

The integration constant C is found by applying the condition $T(t = 0) = T_\infty$:

$$\begin{aligned} T_\infty &= T_\infty + \frac{VI}{\alpha} + C \\ C &= -\frac{VI}{\alpha} \end{aligned} \quad (\text{A.12})$$

With Equation A.12 the equation describing the temperature in dependence of time becomes:

$$T(t) = T_\infty + \frac{VI}{\alpha} \left(1 - e^{-\frac{\alpha}{\beta}t}\right) \quad (\text{A.13})$$

Case: non-constant environment temperature

In the temperature model developed above, it was assumed that the environment temperature T_∞ is constant. In the following a more realistic model with gradually increasing environment temperature is developed. The environment temperature will be replaced with:

$$T_\infty = T_0 + ht \quad (\text{A.14})$$

where T_0 [K] is the initial environment temperature, and h [K/s] is a temperature-increase coefficient. Equation A.14 describes a linear increase in environment temperature with time, starting from an initial value T_0 . Replacing T_∞ in the energy equations A.6 gives:

$$\begin{aligned} \dot{Q}_{HV} &= VI \\ \dot{Q}_{Ca} &= \beta \dot{T} \\ \dot{Q}_{Co} &= \alpha(T - T_\infty) = \alpha(T - T_0 - ht) \end{aligned} \quad (\text{A.15})$$

Again Equations A.15 are applied to a balance of energies $\dot{Q}_{HV} - \dot{Q}_{Ca} - \dot{Q}_{Co} = 0$, leading to a differential equation for $T(t)$:

$$\begin{aligned} VI - \beta \dot{T} - \alpha T + \alpha T_0 + \alpha h t &= 0 \\ \beta \dot{T} + \alpha T &= VI + \alpha T_0 + \alpha h t \\ \dot{T} + \frac{\alpha}{\beta} T &= \frac{\alpha T_0 + VI}{\beta} + \frac{\alpha h}{\beta} t \end{aligned} \quad (\text{A.16})$$

The differential equation A.16 is solved with the same approach as above:

$$\begin{aligned} \text{for } y' + u(x)y &= v(x) \\ y &= G(x)e^{-U(x)} \\ \text{with } U(x) &= \int u(x) \, dx \\ G(x) &= \int v(x)e^{U(x)} \, dx \end{aligned}$$

The solution to Equation A.16 is obtained with:

$$\begin{aligned} U(t) &= \int \frac{\alpha}{\beta} dt = \frac{\alpha}{\beta}t + C_1 \\ G(t) &= \int \left(\frac{\alpha T_0 + VI}{\beta} + \frac{\alpha h}{\beta}t \right) e^{\frac{\alpha}{\beta}t+C_1} \, dt \\ &= \int \left[\left(\frac{\alpha T_0 + VI}{\beta} \right) e^{\frac{\alpha}{\beta}t+C_1} + \frac{\alpha h}{\beta}te^{\frac{\alpha}{\beta}t+C_1} \right] \, dt \\ &= \left(T_0 + \frac{VI}{\alpha} \right) e^{\frac{\alpha}{\beta}t+C_1} + \frac{\alpha h}{\beta}e^{\frac{\alpha}{\beta}t+C_1} \left(\frac{\beta}{\alpha}t - \frac{\beta^2}{\alpha^2} \right) + C_2 \\ &= e^{\frac{\alpha}{\beta}t+C_1} \left(T_0 + \frac{VI}{\alpha} + ht - \frac{\beta}{\alpha}h \right) + C_2 \\ T(t) &= \left[e^{\frac{\alpha}{\beta}t+C_1} \left(T_0 + \frac{VI}{\alpha} + ht - \frac{\beta}{\alpha}h \right) + C_2 \right] e^{-\frac{\alpha}{\beta}t-C_1} \\ &= \left(T_0 + \frac{VI}{\alpha} + ht - \frac{\beta}{\alpha}h \right) + Ce^{-\frac{\alpha}{\beta}t} \end{aligned} \tag{A.17}$$

(A.18)

Again the integration constant C is found by applying the condition $T(t = 0) = T_0$:

$$\begin{aligned} T_0 &= T_0 + \frac{VI}{\alpha} - \frac{\beta}{\alpha}h + C \\ C &= -\frac{VI}{\alpha} + \frac{\beta}{\alpha}h \end{aligned} \tag{A.19}$$

With Equation A.19 the equation describing the temperature in dependence of time becomes:

$$T(t) = T_0 + \left(\frac{VI}{\alpha} - \frac{\beta}{\alpha}h \right) \left(1 - e^{-\frac{\alpha}{\beta}t} \right) + ht \tag{A.20}$$

A.2 Coefficients for temperature-time function

Joule heating $\dot{Q}_{HV} = VI$

$$V = 12.25 \text{ kV} = 12250 \text{ V}$$

$$I = 90 \mu\text{A} = 9 \times 10^{-5} \text{ A}$$

$$VI = 1.1025 \text{ J/s}$$

Heat capacity $Q_{Ca} = \beta T$

$$\begin{aligned} \beta &= mc = m_{\text{buffer}} c_{\text{buffer}} + m_{\text{capillary}} c_{\text{capillary}} \\ m_{\text{buffer}} c_{\text{buffer}} &= V_{\text{buffer}} \rho_{\text{buffer}} c_{\text{buffer}} \\ &= \pi \frac{d_i^2}{4} L \rho_{\text{buffer}} c_{\text{buffer}} \\ m_{\text{capillary}} c_{\text{capillary}} &= V_{\text{capillary}} \rho_{\text{capillary}} c_{\text{capillary}} \\ &= \pi \frac{(d_a^2 - d_i^2)}{4} L \rho_{\text{capillary}} c_{\text{capillary}} \end{aligned}$$

where m [kg] is the mass, c [J/(kgK)] the specific heat capacity (per unit mass), V [m^3] the volume, ρ [kg/ m^3] the density, d_i [m] the inner capillary diameter, d_a [m] the outer capillary diameter, and L [m] the capillary length.

	d [m]	L [m]	ρ [kg/ m^3]	c [J/(kgK)]
buffer (water)	75×10^{-6}	0.35	1000 ^[255]	4182 ^[255]
capillary (quartz)	375×10^{-6}	0.35	2660 ^[255]	710 ^[255]

$$\begin{aligned} m_{\text{buffer}} c_{\text{buffer}} &= \pi \frac{(75 \times 10^{-6})^2}{4} \times 0.35 \times 1000 \times 4182 \\ &= 0.0065 \\ m_{\text{capillary}} c_{\text{capillary}} &= \pi \frac{(375 \times 10^{-6})^2 - (75 \times 10^{-6})^2}{4} \times 0.35 \times 2660 \times 710 \\ &= 0.0595 \\ \beta &= 0.660 \frac{\text{J}}{\text{K}} \end{aligned}$$

Newton's law of cooling ($\dot{Q}_{Co} = \alpha(T - T_\infty)$):

$$\begin{aligned}\alpha &= kS \\ S &= \pi d_a L = \pi \times 375 \times 10^{-6} \times 0.35 \\ &= 4.123 \times 10^{-4} \text{ m}^2\end{aligned}$$

with heat transfer coefficient k [W/(m²K)], surface area across which heat is transferred S [m²], capillary temperature T [K], environment temperature T_∞ [K], outer capillary diameter d_a [m], and capillary length L [m].

The heat transfer coefficient k for n different layers of material, each with thickness x_n , is found with:

$$\frac{1}{k_{\text{total}}} = \sum_n \frac{x_n}{K_n}$$

The x and K values are:^[255]

	quartz	polyimide	air
thickness x_n [m]	150×10^{-6}	20×10^{-6}	100×10^{-6}
unit thickness heat transfer coefficient K_n [W/(mK)]	1.3	0.12	0.024

$$k_{\text{total}} = 244.785 \text{ W/(m}^2\text{K)}$$

$$\alpha = 0.093 \text{ W/K}$$

Increasing environment temperature $T_\infty = T_0 + ht$

T_0 [K] is the initial environment temperature, and h [K/s] is a temperature increase coefficient. The temperature increase coefficient is found with two measurement values:

$$\begin{aligned}T_\infty(t = 0) &= 295 \text{ K} \\ T_\infty(t = 60 \text{ min}) &= 299 \text{ K} \\ \rightarrow h &= 0.001 \text{ K/s} \\ T_\infty(t) &= 295 + 0.001t\end{aligned}$$

With the calculated coefficients the temperature equations can be formulated with numerical values.

In the case of **constant environment temperature**:

$$\begin{aligned} T(t) &= T_{\infty} + \frac{VI}{\alpha} \left(1 - e^{-\frac{\alpha}{\beta}t} \right) \\ \frac{VI}{\alpha} &= 11.897 \text{ K} \\ \frac{\alpha}{\beta} &= 1.404 \text{ 1/s} \\ T(t) &= 295 + 11.897 \left(1 - e^{-1.404t} \right) \end{aligned} \quad (\text{A.21})$$

The initial temperature can also be written in °C instead of K (295 K ≈ 22 °C).

In the case of **increasing environment temperature**:

$$\begin{aligned} T(t) &= T_0 + \left(\frac{VI}{\alpha} - \frac{\beta}{\alpha} h \right) \left(1 - e^{-\frac{\alpha}{\beta}t} \right) + ht \\ \frac{VI}{\alpha} - \frac{\beta}{\alpha} h &= 11.896 \text{ K} \\ \frac{\alpha}{\beta} &= 1.404 \text{ 1/s} \\ h &= 0.001 \text{ K/s} \\ T(t) &= 295 + 11.896 \left(1 - e^{-1.404t} \right) + 0.001t \end{aligned} \quad (\text{A.22})$$

Appendix B

COMSOL conditions and parameters

Concentration c [% w/v] dextran 70 kDa in 50 mM sodium tetraborate

Viscosity $\eta(c)$ [sPa]:

$$\eta(c) = 4e-3 \tanh(2c - 5) + 5e-3$$

$$\eta(c = 0) = 1e-3; \quad \eta(c = 5) = 9e-3 \quad (\text{values from measurement})$$

Electro-osmotic mobility $\mu(c)$ [$\text{m}^2/(\text{sV})$]:

$$\mu(c) = -4.3352e-9 \tanh(2c - 5) + 3.9528e-8$$

$$\mu(-\infty \rightarrow 1.5) = 4.385e-8; \quad \mu(3.6 \rightarrow \infty) = 3.519e-8 \quad (\text{values from measurement})$$

Constants:

Density $\rho = 1e3$ [kg/m^3] (for dilute aqueous system)

Electrical conductivity $\sigma = 6.2e-9$ [S/m] (from measurements)

Diffusion coefficient $D = 4e-11$ [m^2/s] (for dextran 70 kDa^[256])

Initial conditions, and boundary conditions: see B.1.

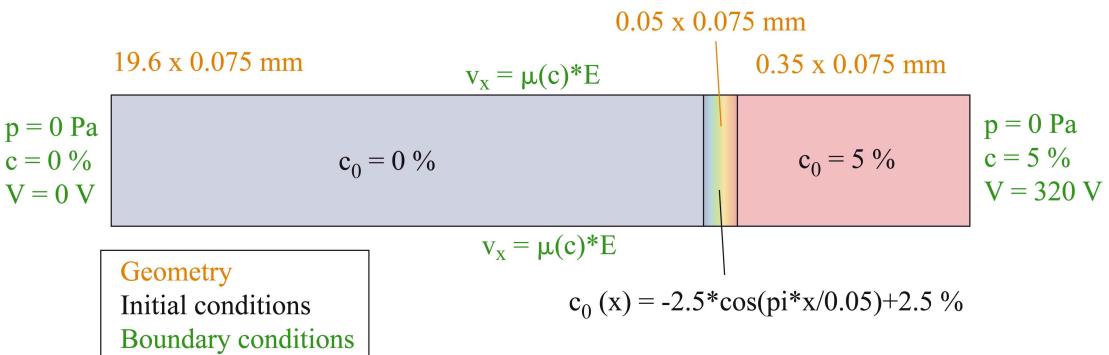


Figure B.1: Geometry, initial conditions, and boundary conditions for COMSOL model of interface region of two buffers co-migrating (not to scale).

Appendix C

Cyclic olefin copolymer microfabrication process

SU-8 mold Photolithography process

1. Clean float glass slides (VWR International LLC, 25 × 75 × 1 mm) for 10 minutes in an ultrasonic bath (Ultrawave Ltd.) in acetone (VWR International LLC, HiPer-Solve CHROMANORM), followed by 10 minutes in an ultrasonic bath in ultrapure water (Purite Ltd., > 18 MΩ), then at least 24 h in sulphuric acid (Riedel-de Han, 95-97 %). Finally slides are rinsed four times with ultrapure water and baked for 45 minutes at 200 °C (Clifton, hot plate) for surface dehydration.
2. Spin coat SU-8 100 (MicroChem Corp.) using the following spin procedure: ramp up to 500 rpm at 100 rpm/s, hold for 30 seconds; ramp up to 2000 rpm at 300 rpm/s, and hold for 40 seconds before deceleration (Brewer Science, Inc., Cee 200 Coat-Bake system). Nominal SU-8 thickness is 125 µm.^[257]
3. Soft bake (Brewer Science, Inc., Cee 200 Coat-Bake system) with the following temperature profile: ramp up to 65 °C, hold for 5 minutes, ramp up to 95 °C, hold for 60 minutes, and gradually cool down (ca. 4 hours) to room temperature.
4. Expose the SU-8 by flood exposure for 500 seconds (Tamarack Scientific Co., Inc., PRX 200/350 collimated exposure system). This creates an adhesion layer of SU-8 on the glass slide.
5. Post exposure bake with the following temperature profile: ramp up to 65 °C, hold for 10 minutes, ramp up to 95 °C, hold for 40 minutes, and gradually cool down (ca. 4 hours) to room temperature.
6. Spin coat a second SU-8 100 layer using the following spin procedure: ramp up to 500 rpm at 100 rpm/s, hold for 30 seconds; ramp up to 2000 rpm at 300 rpm/s, and hold for 40 seconds before deceleration.

7. Soft bake with the following temperature profile: ramp up to 65 °C, hold for 5 minutes, ramp up to 95 °C, hold for 60 minutes, and gradually cool down (ca. 4 hours) to room temperature.
8. Expose the top SU-8 layer for 500 seconds with a mask for photolithography (photoemulsion transparency, plotted from AutoCAD drawing at Technical Graphics Ltd.). The mask design is shown in Figure C.1.
9. Post exposure bake with the following temperature profile: ramp up to 65 °C, hold for 10 minutes, ramp up to 95 °C, hold for 40 minutes, and gradually cool down (ca. 4 hours) to room temperature.
10. Develop structure in Microposit EC solvent (Shipley Europe Ltd.) under agitation and visual monitoring. Stop development with isopropanol alcohol (VWR International LLC, GPR Rectapur), and dry under nitrogen stream.

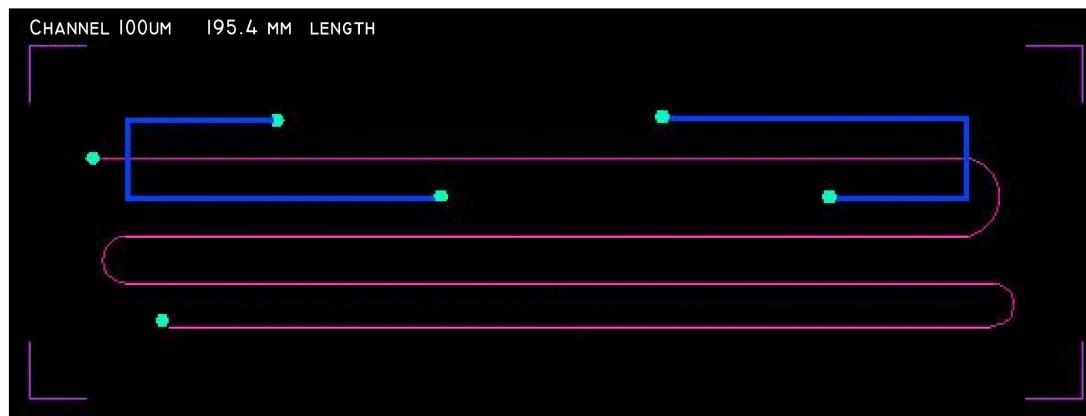


Figure C.1: Photolithographic mask design for SU-8 master for embossing microfluidic channels in COC by lamination. Dimensions of framed area: 25 × 75 mm.

Cyclic olefin copolymer (COC) microfluidic chip Lamination process COC films from Plitek (AM4939, Topas 8007x4 extruded to 100 μm thickness)

1. Emboss pattern from SU-8 master with hot roll laminator (Albyco Photopro 33) at lamination speed 1, lamination temperature 130 °C. Allow to cool before detaching the COC film.
2. Cut access holes at reservoirs manually with 0.7 mm diameter syringe needle (Terumo Europe N.V., Neolus).
3. Plasticise the surface of a second COC film with 7 % hexadecane (Sigma-Aldrich Co., 99 %) in isopropanol alcohol (VWR International LLC, GPR Rectapur) by spin coating the solution at 2000 rpm for 60 seconds.

4. Immediately bond the plasticised film to the embossed film by a single pass through the hot roll laminator at speed 1, 80 °C.
5. Glue on Nanoport reservoirs (IDEX Corp, previously Upchurch Scientific Inc.) with Loctite polyolefin primer 770 and Loctite 406 cyanoacrylate adhesive.

Appendix D

Photopolymerisation of polyacrylamide on-chip

Prepare acrylamide solution For 2 mL of acrylamide mix:

- 8 mg VA-068 (final concentration 0.2 % w/v), from Wako Pure Chemical Industries Inc.
- 330.3 μ L acrylamide/bisacrylamide solution 30 %, 2.6 % C (final concentration 5 % T, 2.6 % C), from Sigma-Aldrich Co.
- 1 μ L TEMED from Fluka Biochemika
- 1665.7 μ L ultrapure water (Purite Ltd., $> 18 \text{ M}\Omega$)

Prepare Immobiline mix For 1 mL final IPG solution:^[59]

Immobiline species ¹	pH 4 mix [μ L]	pH 9 mix [μ L]
pK 3.6	55.24	9.82
pK 4.6	15.69	28.28
pK 6.2	15.46	24.01
pK 7.0	1.49	19.76
pK 8.5	16.68	4.71
pK 9.3	14.76	44.17
total volume	119.3	130.7
add acrylamide solution for 1 mL IPG solution	880.7	869.3

¹Immobilines from GE Healthcare, 10 mL solution 2M

Immobiline mix is prepared weekly in aliquots, acrylamide solution is prepared fresh daily. IPG solution is degassed for 60 minutes prior to use.

Prepare COC channel surface for photografting

1. Flush channel with 5 % benzophenone (Fluka Analytical) in methanol (BDH AnalaR) for 5 minutes.
2. UV exposure (Tamarack Scientific Co., Inc., PRX 200/350 collimated exposure system) for 6 minutes.
3. Flush channel with methanol for 5 minutes.

Prepare and polymerise diffusion gradient

1. Flush channel with acrylamide solution (without Immobilines) for 60 minutes.
2. Fill reservoirs with pH 4, respectively pH 9, IPG solution, adjust levels and allow to diffuse 15 hours.
3. Expose IPG channel section with a slit mask for 45 minutes (Tamarack Scientific Co., Inc., PRX 200/350 collimated exposure system). Allow polymerisation to continue for > 6 hours prior to further use.