

RESEARCH ARTICLE

A novel panel of protein biomarkers for predicting response to thalidomide-based therapy in newly diagnosed multiple myeloma patients

Rajesh Rajpal^{1,2,3*}, Paul Dowling^{1*}, Justine Meiller¹, Colin Clarke¹, William G. Murphy³, Robert O'Connor^{1,4}, Malcolm Kell⁵, Constantine Mitsiades⁶, Paul Richardson⁶, Kenneth C. Anderson⁶, Martin Clynes^{1**} and Peter O'Gorman^{1,2,5**}

¹ National Institute for Cellular Biotechnology (NICB), Dublin City University, Dublin, Ireland

² Mater Misericordiae University Hospital, Dublin, Ireland

³ Irish Blood Transfusion Services, Dublin, Ireland

⁴ School of Nursing, Dublin City University, Dublin, Ireland

⁵ Dublin Academic Medical Centre, University College Dublin, Dublin, Ireland

⁶ Dana-Farber Cancer Institute, Boston, MA, USA

Multiple myeloma (MM) is a heterogeneous group of disorders both genotypically and phenotypically. Response to thalidomide-based induction therapy in newly diagnosed patients varies significantly in published clinical trials. Proteomic analysis was performed on 39 newly diagnosed MM patients treated with a thalidomide-based regimen (22 responders; 17 non-responders) using immunodepletion, 2-D DIGE analysis and mass spectrometry. Zinc- α -2-glycoprotein (ZAG), vitamin D-binding protein (VDB), serum amyloid-A protein (SAA) and β -2-microglobulin (B2M) had statistically significant higher concentrations in non-responders compared to responders, while haptoglobin (Hp) had a lower concentration. ELISAs were used to validate the candidate protein biomarkers using unfractionated serum from 51 newly diagnosed MM patients (29 responders; 22 non-responders). Using logistic regression, the best possible area under the curve (AUC) was 0.96 using ZAG, VDB and SAA in combination. Leave-one-out-cross-validation (LOOCV) indicated an overall predictive accuracy of 84% with associated sensitivity and specificity values of 81.8 and 86.2%, respectively. Subsequently, 16 of 22 thalidomide-refractory patients successfully achieved complete response or very good partial response using second-line treatment suggesting that the biomarker profile is specific to thalidomide response rather than identifying patients with MM refractory to all therapies. Using a novel panel of predictive biomarkers, the feasibility of predicting response to thalidomide-based therapy in previously untreated MM has been demonstrated.

Received: August 6, 2010
Revised: October 8, 2010
Accepted: January 11, 2011

**Keywords:**

Biomarkers / Biomedicine / Multiple myeloma / Thalidomide

Correspondence: Dr. Peter O'Gorman, Dept of Hematology, Mater Misericordiae University Hospital, Eccles St, Dublin 7, Ireland

E-mail: pogorman@mater.ie

Fax: +353-1-8500962

Abbreviations: **AIC**, Akaike's Information Criterion; **B2M**, β -2-microglobulin; **BVA**, biological variation analysis; **CR**, complete response; **DE**, differentially expressed; **ISS**, International Staging System; **LOOCV**, leave-one-out-cross-validation; **LR**, logistic regression; **MM**, multiple myeloma; **Hp**, haptoglobin; **ROC**, receiver operating characteristic; **SAA**, serum amyloid A protein; **TTP**, time to progression; **VDB**, vitamin D-binding protein; **VGPR**, very good partial response; **ZAG**, zinc- α -2-glycoprotein

1 Introduction

Multiple myeloma (MM) is characterized by the proliferation of malignant plasma cells within a complex inter-dependent cellular network in the bone marrow (BM). Although MM remains an incurable blood cancer in the majority of patients, the development of novel therapies has dramatically increased response rates and long-term disease control is now attainable in a significant proportion of patients [1, 2].

*These authors contributed equally to this work.

**Joint senior authors.

A minority of centers has access to gene expression profiling and can stratify patients in terms of high- and low-risk disease, thus potentially individualizing treatment [3]. Molecular and cytogenetic analysis can identify high-risk groups but these facilities are not available in all centers. The more widely available International Staging System (ISS) allows clinicians to identify prognostic groups based on β -2-microglobulin (B2M) and serum albumin levels [4, 5]. However, despite major advances in our understanding of this complex disease, a relatively standard remission–induction therapeutic approach is typically used in patients in similar categories of age and performance status by the majority of groups [6, 7].

Thalidomide is an oral immunomodulatory [8] drug that has been shown to be highly active against MM, with response rates of 30–70% [9]. Recent long-term follow-up data indicate that patients exposed to thalidomide early in their treatment history obtain durable benefit [10]. Therefore, it is likely that thalidomide and its emerging analogues such as lenalidomide will remain a central component of anti-MM therapy in the future. Serious side effects observed with use of thalidomide in MM include thrombo-embolic disease and peripheral neuropathy. Neuropathy can be irreversible and may preclude the use of other potentially neuropathic therapies including bortezomib, thereby compromising survival.

In this study, we have focused specifically on identifying a biomarker profile that correlates with response rate to thalidomide-based regimens in newly diagnosed patients with MM. We have attempted to devise a proteomic profile using an accessible platform that will assist clinicians in individualizing treatment by identifying patients that will have a high likelihood of response to thalidomide-based therapy. If validated in prospective studies, this profile could allow the physician to choose an alternative therapy in patients predicted to be resistant to thalidomide-based treatment thereby avoiding ineffective, potentially costly treatment and exposure of the patient to unnecessary side effects.

2 Materials and methods

2.1 Patients and sample collection

Serum samples from 51 consecutive newly diagnosed MM patients who were having initial treatment with thalidomide-based regimens were analyzed. Samples were obtained at diagnosis and prior to commencement of therapy. The samples were collected according to standard phlebotomy procedures from consented patients. Ethical consent was granted from the Mater Misericordiae University Hospital, Dublin, Ireland, ethics committee. Ten milliliters of blood sample was collected into additive free blood tubes and was allowed to clot for 30 min to 1 h at room temperature. Samples were coded and transported on ice to the laboratory. The serum was denuded by pipette from the clot and placed

into a clean tube. The tubes were centrifuged at 400 relative centrifugal force (rcf) for 30 min at 4°C. Serum was aliquoted in the cryovial tubes, labeled and stored at –80°C until time of analysis. The time from sample procurement to storage at –80°C was less than 3 h. Each serum sample underwent no more than three freeze/thaw cycles prior to analysis.

2.2 Removal of high-abundance proteins from serum samples

Samples were prepared as outlined previously [11]. Briefly, diluted samples (Buffer A) were centrifuged to remove lipids and particulates. The Human Multiple Affinity Removal System (MARS) column was employed to remove 14 of the most abundant proteins (albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin (Hp), fibrinogen, α -2-macroglobulin, α -1-acid glycoprotein, IgM, apolipoprotein AI, apolipoprotein AII, complement C3 and transthyretin). For each sample, a low-abundance fraction was collected and concentrated using 5-kDa molecular weight cut-off spin concentrators (Agilent Technologies). The concentrated depleted serum samples were collected and immediately transferred and stored at –80°C until further analysis. The depletion protocol was found to be reproducible as demonstrated by Western blot analysis using an anti-albumin antibody, which confirmed the absence of albumin from the immunodepleted fraction (results not shown).

2.3 Preparation for 2-D electrophoresis

Proteins from the immunodepleted serum samples were precipitated prior to labeling using a 2-D Cleanup Kit (Biorad). The protein pellets were resuspended in ice-cold DIGE lysis buffer (20 mM Tris, 7 M Urea, 2 M Thiourea, 4% CHAPS, pH 8.5). Protein quantification was performed using the Quick Start Bradford Protein Assay (Biorad) using BSA as a protein standard [11].

2.4 2-D DIGE

Depleted serum samples were labeled with *N*-hydroxy succinimidyl ester-derivatives of the cyanine dyes Cy2, Cy3 and Cy5 following a standard protocol [11]. Immobilized 24-cm linear pH gradient (IPG) strips, pH 3–11 NL, were rehydrated in rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 0.5% IPG Buffer, 50 mM DTT) overnight, according to the standard guidelines [11, 12]. Isoelectric focusing (IEF) was performed using an IPGphor apparatus (GE Healthcare) for a total of 40 kV/h at 20°C. Equilibrated IPG strips were placed on top of 12.5% uniform polyacrylamide gels cast between low-fluorescence glass plates. Strips were sealed on the top of the gels using a sealing solution (0.5% w/v low-melting-point agarose in running buffer containing bromphenol blue). Gels were run on the

Ettan Dalt 12 apparatus (GE Healthcare) using a two-step protocol, 2.5 W/gel for 30 min and then 100 W in total at 10°C until the dye front had run off the bottom of the gels.

2.5 DeCyder analysis

Gels were scanned using the Typhoon 9400 Variable Mode Imager (GE Healthcare) to generate images at the appropriate excitation and emission wavelengths from the Cy2-, Cy3- and Cy5-labeled samples. Subsequently, the gel images were cropped using ImageQuant software (GE Healthcare) and imported into Decyder 6.5 software. Decyder 6.5 was used to compare the immunodepleted serum from responders versus non-responders to generate lists of differentially expressed (DE) proteins [11, 13]. The differential in gel analysis (DIA) module was used to assign spot boundaries and to calculate parameters such as normalized spot volume. The BVA mode of DeCyder 6.5 was then used to match all paired image comparisons from difference in-gel analysis for a comparative cross gel statistical analysis. At this stage, operator intervention was required for more accurate matching. If the matching in an area required correction, the current matches were broken and remade with the appropriate spots [11, 12].

2.6 Spot digestion and identification by mass spectrometry

Proteins of interest were picked from preparative gels containing 400 µg of protein stained with Colloidal Coomassie Blue (CBB) stain (Sigma) using the Ettan Spot Picker robot (GE Healthcare). Tryptic digestions were performed on the proteins of interest according to standard protocols [11]. Samples were analyzed using a 1-D LC-MS/MS method run on an Ettan MDLC system (GE Healthcare) in high-throughput configuration directly connected to a Finnegan LTQ (Thermo Electron). Samples were initially concentrated and desalted on RPC trap columns (Zorbax 300SB C18, 0.3 mm × 5 mm, Agilent Technologies). Subsequently, peptides were separated on a nano-RPC column (Zorbax 300SB C18, 0.075 mm × 100 mm, Agilent Technologies) using a linear acetonitrile gradient from 0 to 65% ACN (Sigma) over 60 min. All buffers used for nano-LC separation contained 0.1% formic acid (Fluka) as the ion pairing reagent [14].

Identification of proteins was carried out using the Turbo-SEQUEST algorithm in the BioWorks 3.1 software package (Thermo Electron) and the Swiss-Prot human database (Swiss Institute of Bioinformatics, Geneva, Switzerland). Peptides were further evaluated using charge state versus cross-correlation number (XCORR). The following SEQUEST criteria were selected for high-confidence peptide identification: XCORR > 1.5 for singly charged ions, XCORR > 2.0 for doubly charged ions, and XCORR > 2.5 for triply charged ions, together with a minimum of two matched peptides for each protein [14, 15].

2.7 ELISA

ELISAs were used to confirm the differential expression of the five potential biomarkers discovered using 2-D DIGE analysis. ELISA-based validation was carried out using raw unfractionated serum samples from the original cohort of patients. Each sample was analyzed in duplicate using the following commercially available kits: for the measurement of human serum amyloid-A protein (SAA) (Invitrogen), serum Hp (AssayMax), zinc- α -2-glycoprotein (ZAG) (BioVendor), B2M and vitamin D-binding protein (VDB) (Immunodiagnostic) kits were used. The ELISA assays were performed according to each manufacturer's protocol and guidelines. The optical density (OD) was measured using a micro-plate reader (Bio-Tek) and the concentration of each protein in the serum samples was determined by comparing the OD of the samples against the respective standard curve provided by the kit.

2.8 Statistical analysis

DIGE gels were exported for image analysis using the BVA module of Decyder 6.5 software (GE Healthcare) for quantitation of protein abundance levels. Following confirmation of appropriate spot detection, matching, normalization and spot statistics were reviewed. The normalized volume of a spot was compared in all the gels between each group. Spots that were found to be statistically significant by univariate analysis (t -test ≤ 0.01) were selected for further investigation.

Multivariate logistic regression (LR) and receiver operating characteristic (ROC) curve analysis were carried out in the freely available statistical software R (<http://www.r-project.org/>). The ROC curves were used to interpret the utility of LR models for various combinations of the DE proteins. The probability of correct prediction for a given model was calculated from the ROC curve by determining the area under curve (AUC). Proteins and combinations of proteins returning the largest AUC values are deemed the most effective for the discrimination of responders from non-responders [16].

Akaike's Information Criterion (AIC) was also used as an additional factor to select the best combination of biomarkers. AIC is a commonly used measure to select between competing statistical models. The AIC is a trade off between the goodness of fit and model complexity, i.e. the number of parameters required (in this case proteins).

Sensitivity and specificity values were calculated for the best combination of biomarkers. Sensitivity is defined as the percentage of all non-responders who were refractory to a thalidomide-based treatment regime correctly identified as having this phenotype based on the panel of protein biomarkers (the true positive rate). Specificity is defined as the percentage of all responders who were sensitive to a thalidomide-based treatment regime correctly identified as

having this phenotype based on the panel of protein biomarkers (the true negative rate).

As an additional measure of the potential of these biomarkers to accurately predict response to thalidomide-based therapy, a commonly used internal validation technique known as leave-one-out-cross-validation (LOOCV) was performed. During the LOOCV procedure, data from a single observation are removed from the data set and the remaining samples are then utilized to construct a LR model. The “test” sample is presented to the trained model and the performance assessed; LOOCV continues until each observation is designated as the “test.” The average performance over the 51 tests is reported as the LOOCV accuracy.

3 Results

3.1 Sample set

The mean age of the patient group was 68 SD \pm 6.95 years (range 52–81 years); 27 were male and 24 female. Based on day-100 re-staging investigations and using the International Myeloma Working Group (IMWG) uniform response criteria for MM [17, 18], 29 responders and 22 non-responders to thalidomide-based therapy were identified. The mean age was 66 SD \pm 6.80 years (range 52–81 years) for responders and 70 SD \pm 6.69 years (range 57–79 years) for non-responders. Based on the ISS classification [19], 6 responding patients had stage I, 15 had stage II and 8 had stage III disease. Two non-responders had stage I, 11 had stage II and 9 had stage III disease. In the responders group, 24 patients were treated with thalidomide and dexamethasone (TD), 3 with thalidomide, cyclophosphamide and dexamethasone (CTD) and 2 with melphalan, prednisone and thalidomide (MPT). In the non-responder group, 17 patients received thalidomide and dexamethasone, 4 received thalidomide, cyclophosphamide and dexamethasone and 1 patient received melphalan, prednisone and thalidomide. Median follow-up was 18 months (range 9–27 months). In the responder group, 9 patients achieved complete response (CR) and 20 achieved very good partial response (VGPR). Five non-responders had Stable disease (SD) and the remaining 17 had Progressive disease (PD) (Table 1). Time to progression (TTP) was defined as the time from the start of thalidomide-based regimen to the date of documented progression. Of the 29 responders, 8 eventually relapsed with median TTP of 15 months (range 11–18 months). However, median TTP has not been calculated for this group as 21 patients remain in CR or VGPR.

3.2 Proteomic profiling

Proteins were precipitated from the low-abundance immunodepleted fraction, resuspended in lysis buffer, fluorescently

labeled and analyzed by 2-D DIGE using an internal standard design. This analysis was performed on 39 newly diagnosed MM patients (22 responders; 17 non-responders). Spot maps were generated and maps were aligned with a master spot map; relative abundance values were generated for each of 886 protein spots that were common to more than 90% of gels. Based on 2-D DIGE analysis, protein spots with a fold change of ≥ 1.25 in abundance level and a *t*-test of ≤ 0.01 were selected. Using these criteria, five individual DE protein spots were detected. Four proteins were increased in abundance level, and one was decreased in abundance level in thalidomide non-responders compared to responders. Figure 1 shows DeCyder analysis for B2M and Hp, indicating that these proteins were increased and decreased in abundance levels, respectively, between non-responders and responders to thalidomide-based therapy. Gel images and 3-D protein spot views for Hp and B2M are also displayed, demonstrating a clear difference in the abundance levels (Fig. 1).

3.3 Protein identification

Subsequently, these proteins of interest were identified by LC-MS/MS using an ion trap LTQ mass spectrometer and searched against the SWISS-PROT database using SEQUEST [20–22]. The serum Hp fragment, which was the only protein found to have a lower abundance level in non-responders compared to responders, was identified by LC-MS/MS resulting in nine matched peptides corresponding to a sequence coverage of 11.58% (Table 2). Proteins found to have higher abundance levels between non-responders and responders to thalidomide-based therapy were ZAG, VDB, SAA and B2M. These proteins were identified by LC-MS/MS resulting in 10, 20, 4, and 3 matched peptides, respectively, corresponding to a 37.63, 53.59, 48.36 and 35.29% sequence coverage, respectively (Table 2).

3.4 DeCyder ratios, ELISA data and ROC curve analysis

In this study, ELISA-based assays were used to measure the levels of the five candidate marker proteins in serum from thalidomide responders and non-responders (Supporting Information Table S1). The ELISA-based assays were performed on a larger cohort compared to the 2-D DIGE analysis, consisting of 51 consecutive MM patients (29 responders; 22 non-responders). The five DE protein concentrations were measured in duplicate for each patient. The box plots show the data for responders and non-responders (Fig. 2). The horizontal line within the boxes represents the median. The upper and lower box edges are the 1st and 3rd quartiles. The whiskers reach the nearest value within 1.5 times the inter-quartile range. The points outside the whiskers are considered outliers; however, no outlier value was removed from our analysis.

Table 1. Clinical details for the patients included in this study

Patients' age	Patients sex	Clinical "ISS" stage	Day-100 IMWG re-staging	Response to thalidomide	Follow-up (months)
Responders					
73	F	III	VGPR	R	17
66	M	II	CR	R	16
58	F	III	CR	R	16
67	M	II	VGPR	R	12
59	M	III	VGPR	R	18
67	M	III	CR	R	16
72	M	III	CR	R	13
60	M	II	CR	R	19
71	F	II	VGPR	R	13
69	M	I	VGPR	R	12
61	F	II	CR	R	19
74	F	I	VGPR	R	13
70	F	II	VGPR	R	14
61	F	I	VGPR	R	12
77	M	II	VGPR	R	14
63	M	I	CR	R	14
58	F	I	VGPR	R	15
59	F	I	VGPR	R	14
64	M	II	CR	R	12
63	F	II	VGPR	R	21
65	M	II	VGPR	R	8
70	M	III	VGPR	R	7
66	F	II	CR	R	8
80	M	II	VGPR	R	7
66	F	II	VGPR	R	10
81	F	III	VGPR	R	6
52	M	III	VGPR	R	11
64	M	II	VGPR	R	6
68	F	II	VGPR	R	6
Non-responders					
72	F	II	PD	NR	16
74	M	III	PD	NR	12
71	M	III	PD	NR	19
65	M	II	PD	NR	11
70	M	III	PD	NR	17
63	F	III	PD	NR	14
72	M	III	PD	NR	16
63	M	II	SD	NR	11
78	M	II	PD	NR	10
79	F	II	PD	NR	15
57	F	III	PD	NR	10
77	M	III	PD	NR	12
78	F	II	PD	NR	15
64	F	II	PD	NR	21
65	F	I	SD	NR	20
73	F	III	PD	NR	18
78	M	II	SD	NR	11
61	F	II	SD	NR	12
76	F	III	PD	NR	9
77	M	II	PD	NR	11
62	M	I	PD	NR	8
69	M	II	SD	NR	6

Clinical details include age, sex, ISS (International Staging System), day 100 restaging results based on IMWG uniform response criteria for MM. The last column shows duration of follow-up in months. Abbreviations: CR: complete response; VGPR: very good partial response; SD: stable disease; PD: progressive disease; IMWG: International Myeloma Working Group; R: responders to thalidomide-based therapy; NR: non-responders to thalidomide-based therapy; M: male; F: female.

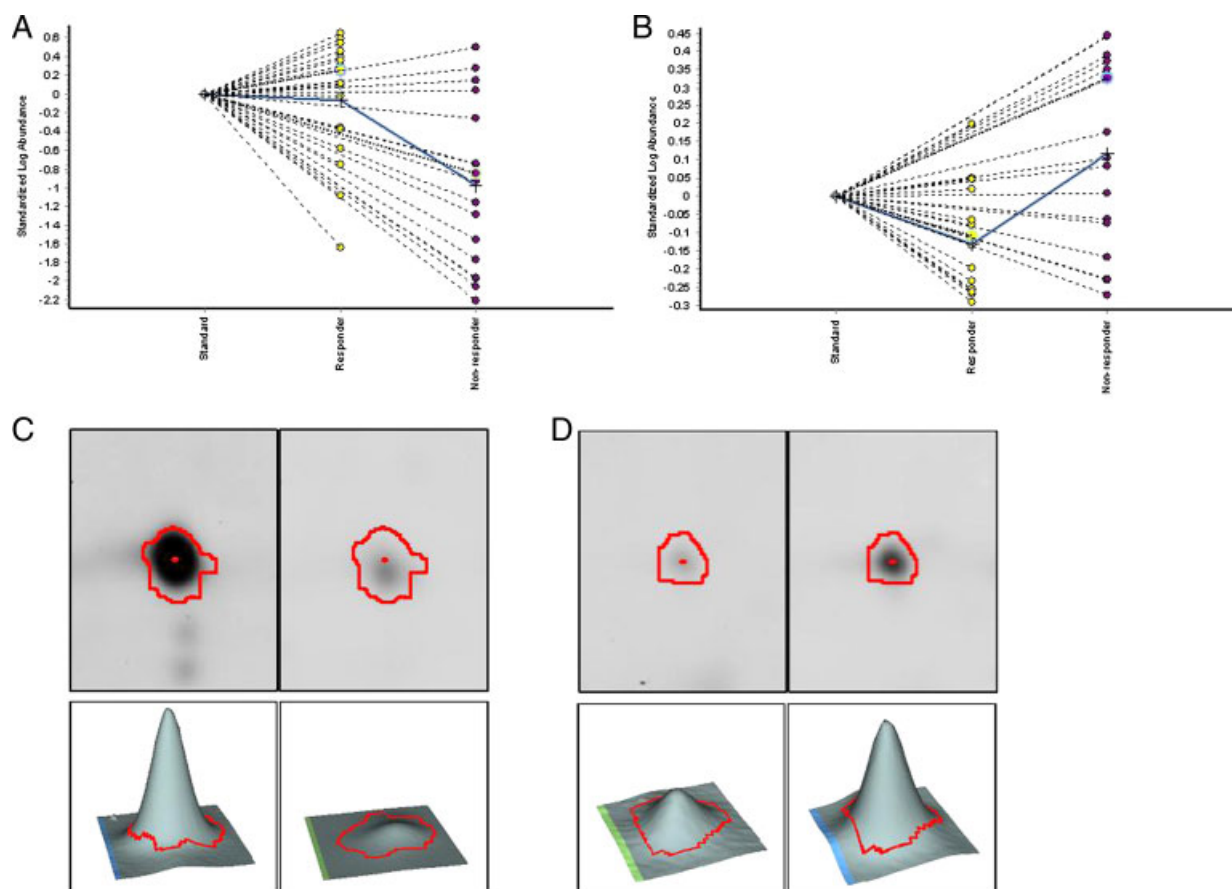


Figure 1. Statistical analysis of Hp and B2M respectively, using DeCyder BVA software. (A) and (B) display that these proteins were found to be 3.01-fold decreased and 1.96-fold increased, respectively, in the immunodepleted serum from non-responders compared to responders. (C) and (D) also show gel images and 3-D views for Hp and B2M, respectively, showing a clear change in expression levels. B2M: β -2-microglobulin, Hp: haptoglobin.

In clinical practice, it is rare that a chosen cut off point for a single analyte will achieve perfect discrimination between various groups of patients, and one has to select the best compromise between sensitivity and specificity by comparing the diagnostic performance of different tests or diagnostic criteria available. In this study, we assessed the suitability of a panel of proteins for potential clinical application using LR ROC curves and report the AUC of each test. ROC curves allow systematic analysis of the diagnostic performance of a test, a comparison of the performance of different tests, and the AUC provides a summary measure of the utility of the model [23, 24]. In our study, we determined the potential impact of the use of ZAG, SAA, VDB, B2M and Hp as single or combination biomarkers for distinguishing between responders and non-responders to thalidomide-based therapy in MM patients. The ELISA values for each of these five proteins were used to develop multivariate LR models and the subsequent ROC curves generated (Fig. 3) and the AUC determined.

2-D DIGE and ELISA analysis showed that ZAG had a 1.48- ($p = 0.000022$) and a 1.27-fold ($p = 0.00398$) increase in

abundance levels in non-responders compared to responders, respectively (Table 2), as determined by univariate analysis. The ROC curve generated from the ELISA data for ZAG yielded an AUC of 0.76 as an individual protein (Fig. 3). Results for SAA showed a 3.01-fold ($p = 0.006$) increase in abundance levels for non-responders compared to responders using 2-D DIGE protein profiling analysis. This result correlated strongly with data from the ELISA analysis, indicating a 3.80-fold ($p = 0.00016$) increase in SAA abundance levels in non-responders compared to responders (Table 2). ROC curves calculated from this ELISA data showed an AUC of 0.82 (Fig. 3), indicating excellent discriminatory power for this single protein. SAA is a sensitive marker of inflammation [25, 26]. In the 22 non-responder patients studied in this project, none had evidence of infection or fever at the time of sampling; out of 29 responder patients, 1 had an infection resulting in mild fever. Therefore, the SAA elevation in the non-responder patients appeared to be unrelated to infection or inflammation.

Hp is normally removed by the immunodepletion column; however, using 2-D DIGE analysis followed by

Table 2. Protein identities obtained by LC-MS/MS analysis

Protein name	MW (Da)	Matched peptides	% Coverage	XC-score	DeCyder ratio	DeCyder <i>p</i> -value	ELISA ratio	ELISA <i>p</i> -value
Vitamin D-binding protein (VDB)	52930	20	53.59	210.4	1.31	0.00044	1.28	0.02045
Haptoglobin fragment (Hp)	45177	9	11.58	76.3	-3.01	0.0017	-1.73	0.03241
Zinc- α -2-glycoprotein (ZAG)	33851	10	37.63	170.3	1.48	0.0000022	1.27	0.00398
β -2-microglobulin (B2M)	13706	3	35.29	46.3	1.96	0.0015	2	0.00118
Serum Amyloid A Protein (SAA)	13524	4	48.36	80.3	3.01	0.006	3.8	0.00016

Listed are the protein identities obtained by LC-MS/MS analysis, molecular weight (MW), number of matched peptides related to the protein, percentage coverage of the protein sequence identified, XC-Score (Sequest score), DeCyder ratio with associated *p*-value (immunodepleted serum) and ELISA ratio (NR/R) with associated *p*-value (unfractionated serum). The DeCyder ratio for Hp is the ~10 kDa fragment identified in the 2-D DIGE analysis while the ELISA ratio is based on the intact form of this protein.

LC-MS/MS, an ~10-kDa Hp fragment was identified [27, 28]. We suggest that this Hp fragment was not removed due to its size and non-interaction with the specific Hp antibody in the affinity column, and hence it was detected in the 2-D DIGE analysis. In the protein profiling analysis, the Hp fragment was found to be 3.01-fold ($p = 0.0015$) decreased in serum from non-responders compared to responders. As intact Hp was removed from this analysis because of the immunodepletion column, it was decided to investigate this protein using an ELISA-based assay approach. Data from the ELISA analysis for intact Hp showed a 1.73-fold ($p = 0.03241$) decrease in abundance levels in non-responders compared to responder. ROC curves were generated from the ELISA data, which showed an AUC of 0.64 (Fig. 3). The decrease in abundance of the Hp fragment seen in thalidomide non-responders compared to responders using 2-D DIGE analysis displayed a similar trend to that of the intact Hp detected by ELISA analysis in the unfractionated serum samples (Table 2).

2-D DIGE data and ELISA results showed a 1.96- ($p = 0.0015$) and 2.00-fold ($p = 0.00118$) increase, respectively, in the abundance level of B2M from non-responders compared to responders. The ROC curve generated from this data yielded an AUC of 0.87, indicating excellent discriminatory power for this protein (Fig. 3). The VDB protein level from 2-D DIGE analysis and ELISA data showed a 1.31 ($p = 0.00044$) and 1.28 ($p = 0.02045$) fold increase in abundance levels, respectively, from non-responders compared to responders (Table 2). ROC curves were generated from the ELISA data and showed an AUC of 0.70 (Fig. 3).

3.5 Logistic regression with multiple proteins

Initially, LR was used to develop predictive models for each individual DE protein (Fig. 3). The best predictive single proteins from the LR model were B2M and SAA, with AUC values of 0.87 and 0.82, respectively. The remaining single protein model had AUC values less than 0.8, indicating poorer predictive ability (Fig. 3). The predictive capability of

models developed based upon combinations of proteins was also assessed. LR models were constructed and ROC analysis carried out on all possible permutations of the DE proteins. The three most successful combinations of biomarkers from this analysis for predicting response to thalidomide-based therapy were found to be Hp, SAA, VDB (AUC = 0.96, LOOCV = 84.31, AIC = 35.26), ZAG, B2M, SAA, VDB (AUC = 0.96, LOOCV = 84.31, AIC = 37.02) and B2M, SAA, VDB (AUC = 0.94, LOOCV = 84.31, AIC = 36.59). The combination of Hp, SAA, VDB was found to be the most successful based on the values from the AUC, LOOCV and AIC analyses as shown (Fig. 4). The sensitivity and specificity of this model was 81.81 and 86.20%, respectively.

In this analysis, CRP level at diagnosis was also assessed but did not improve the predictive capability of this model. CRP was found to have a mean of $10.4 \pm 17.7 \mu\text{g/mL}$ in responders compared to $12.1 \pm 23.6 \mu\text{g/mL}$ in non-responders ($p = 0.770749339$). CRP was found to have an AUC of 0.4, indicating no discriminatory power for predicting response to thalidomide-based therapy.

No correlation between clinical parameters such as age and gender and response to thalidomide was found. Also, individual clinical parameters were combined with the model to determine if classification accuracy could be improved. However, no clinical parameter improved classification accuracy for thalidomide response.

3.6 Calculation of non-response probability

The equation shown here, allows for the calculation of the probability (p) of a patient being a non-responder based on the serum concentrations of ZAG, Hp and VDB as determined using ELISAs. Each protein concentration in $\mu\text{g/mL}$ is first multiplied by the regression coefficient (derived from the fitted model) as per the equation below. If the resulting value of p is below 0.5, response to thalidomide is predicted and if the value is above 0.5, non-response to thalidomide is predicted. As additional patient data is accumulated from future randomized clinical trials, the

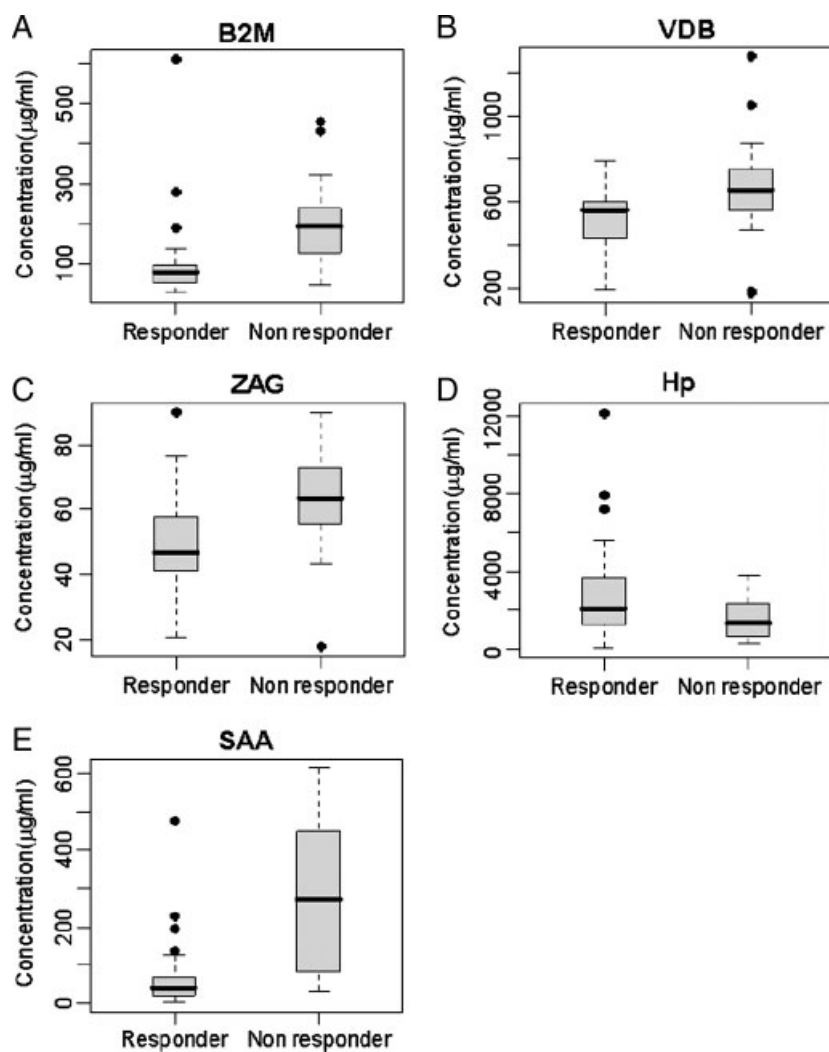


Figure 2. Displayed are the concentrations for the five differentially expressed proteins, obtained in duplicate for each patient using ELISAs. The box plots show the data for responder and non-responder patients. The horizontal lines within the boxes represent the median. The upper and lower box edges are the 1st and 3rd quartiles. The whiskers reach the nearest value within 1.5 times the inter quartile range. The points outside the whiskers are considered outliers. (A) B2M: β -2-microglobulin; (B) VDB: vitamin D-binding protein; (C) ZAG: zinc- α -2-glycoprotein; (D) Hp: haptoglobin; (E) SAA: serum amyloid-A protein.

values of the regression coefficients may be altered (Supporting Information Table S2).

$$p = \frac{1}{1 + e^{-((-0.001037 \times Hp) + (0.017653 \times SAA) + (0.013311 \times VDB) - 8.603880)}}$$

4 Discussion

Clinical trials of thalidomide have shown variable results [29, 30]. We believe that part of the reason for this is the application of a standardized approach to therapy for a heterogeneous group of patients. There is a significant risk of major acute adverse events (National Cancer Institute Common Toxicity Criteria Grade III–IV) including thrombosis (20%), infection (12%) and irreversible neuropathy (8%) associated with thalidomide-based therapy [31, 32], which can be both cumulative as well as dose related [33]. Furthermore, the development of severe neuropathy can

preclude the future use of other neurotoxic agents including bortezomib. Therefore, it is clearly desirable to develop means to predict responders to thalidomide in order to maximize the benefit to risk ratio in MM patients.

The mechanism of action of thalidomide and its analogues is mediated through modulation of the cells of the microenvironment with anti-inflammatory and immunomodulatory effects [8, 34]. These include downregulation of stromal cell adhesion molecules, inhibition of secretion of angiogenic and survival growth factors, pleiotropic immunomodulatory effects including inhibition of tumor necrosis factor- α (TNF- α) secretion, stimulation of IL2 and interferon-gamma (IFN γ). Cellular targets include lymphocytes, monocytes and NK cells [35]. In patients with an activated immune system, thalidomide inhibits tumor necrosis factor- α secretion from monocytes [36] and downregulates secretion of stimulatory cytokines including IL1, IL6, IL12 and GM-CSF [37]. Conversely, thalidomide has been shown to co-stimulate CD4+T cells by CD28 phosphorylation. In vitro, immunomodulatory effects of thali-

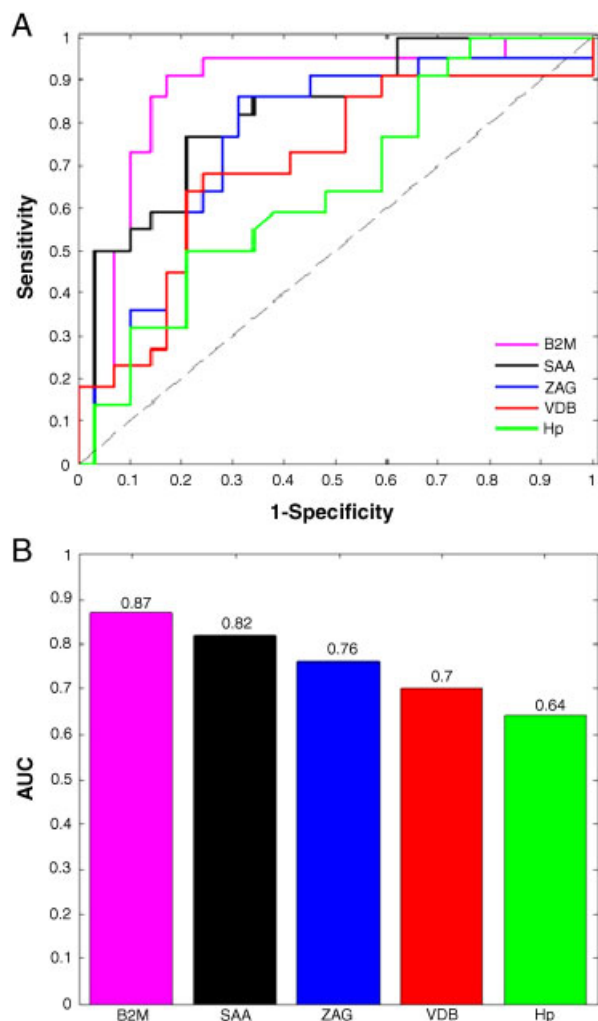


Figure 3. Logistic regression analysis used to develop a predictive model for each individual differentially expressed protein. The performance of the models was assessed using ROC curves (A), and the AUC for each individual protein are shown (B). The best predictive ability for logistic regression model for single proteins was for B2M and SAA, with AUC values of 0.87 and 0.82, respectively. ZAG: zinc- α -2-glycoprotein; VDB: vitamin D-binding protein; SAA: serum amyloid-A protein; B2M: β -2-microglobulin; Hp: haptoglobin.

domide and its analogues therefore vary depending on the level of activation of the patient's immune system. Using 2-D DIGE and MS technologies, we have previously demonstrated novel biomarkers for the early detection of squamous carcinoma of lung [11] and proteins associated with tumor progression in squamous cell carcinoma of the head and neck region [14]. We have adopted a similar scientific approach to the analysis of serum from newly diagnosed MM patients. Interestingly, ZAG, VDB, SAA, B2M and Hp were the five individual proteins with statistically different concentrations between responders and non-responders. A combination of protein biomarkers

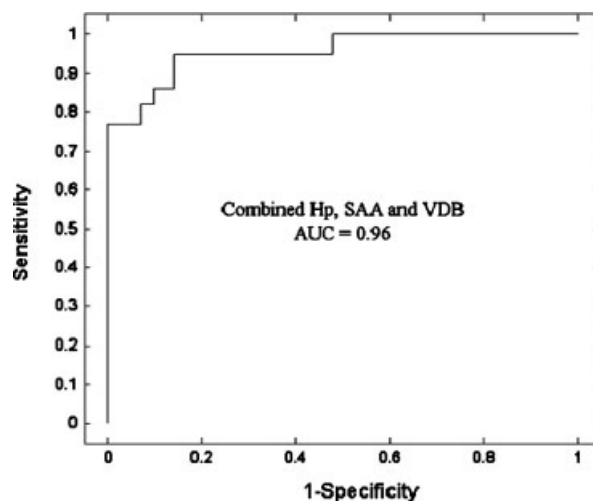


Figure 4. ROC curve analysis using a combination of Hp, SAA and VDB. The best possible AUC was found with combined Hp, SAA and VDB, which yielded an AUC of 0.96 indicating excellent discriminatory power. Hp: haptoglobin; SAA: serum amyloid-A protein; VDB: vitamin D-binding protein.

consisting of serum proteins ZAG, SAA and VDB yielded an AUC of 0.96, indicating an outstanding predictive capability.

B2M is an 11-kD protein associated with the outer membrane of many cells including lymphocytes. It is the small subunit of the MHC class I molecule and is present in small amounts in serum, CSF and urine of normal people. In patients with MM, B2M has been confirmed as a highly significant prognostic factor in combination with albumin as part of the ISS. ZAG is an adipokine with structural organization and folding similar to MHC class I antigen-presenting molecules and therefore may have a role in the immune response [38]. SAA is an apolipoprotein synthesized in response to inflammatory cytokines released by activated monocytes as well as macrophages, and also has a role in fibrosis [39, 40]. Its persistently high concentration is also a prerequisite for the development of AA Amyloidosis [41]. VDB protein is a glycoprotein that functions as the main carrier for vitamin D and its hydroxylated metabolites in the plasma. It is also a precursor for the principal macrophage-activating factor (MAF) [42, 43]. VDB is present on the surface of several cell types including T and B lymphocytes. Hp is a glycoprotein mainly secreted by hepatocytes in response to a variety of stimuli mainly as an acute phase protein and has been shown to have angiogenic effects [44–46].

In vitro, it has been shown that thalidomide analogues can have variable effects depending on the immunological status of the patient [35, 47]. Interestingly, four of the five proteins with increased concentration in non-responders to thalidomide are acute phase reactants. However, Hp and CRP, two significant acute phase proteins, were not found to be of higher abundance in the non-responder group so it is unclear if the response is related to acute phase or immune response.

Thalidomide has potent anti-angiogenic effects mediated by AKT signaling and secretion of angiogenic cytokines [48]. The observation that Hp, a pro-angiogenic protein, has a higher concentration in responders to thalidomide supports the idea that inhibition of angiogenesis is an important part of disease response to thalidomide.

We considered the possibility that the non-responders were simply patients with disease that would be refractory to all therapies. Of 22 non-responders to thalidomide, 16 successfully achieved CR/VGPR using other therapies including bortezomib and lenalidomide (Supporting Information Table S3). This suggests that the biomarker profile is predictive of thalidomide response rather than that the non-responders were simply patients with MM refractory to all therapies (Supporting Information Table S3). Furthermore, these data suggest that, had the non-responsive patients received an alternative to thalidomide as induction therapy, a significant response could have been obtained. The ISS, based on serum B2M and albumin levels, is a predictor of overall survival in MM patients [17, 49]. The endpoint in our study was response to thalidomide in contrast to survival (as applied to the ISS) and we found no correlation between ISS values and response rate to thalidomide. Moreover, incorporation of albumin, B2M and CRP into the Hp/VDB/SAA panel did not improve predictive ability (data not shown).

In conclusion, accurate prediction of sensitivity to thalidomide will help progress toward the goal of personalized medicine in MM patients. We have identified a predictive biomarker panel that has the potential to allow thalidomide-based therapy to be directed to patients with sensitive disease resulting in higher response rates. Avoidance of thalidomide use in patients with unresponsive disease will decrease unnecessary potential side effects including thrombo-embolic disease and irreversible peripheral neuropathy. Better response rates and avoidance of therapy and disease-related complications will increase patient quality of life, reduce cost to health care providers and may impact on survival. However, 2-D DIGE analysis is expensive, labor intensive and requires complex equipment and advanced analytical skills, limiting its applicability to only a minority of clinical centers. ELISA tests are inexpensive, robust, sensitive, specific and widely used in hematopathology laboratories attached to clinical units. For this reason, we validated the results from the 2-D DIGE using commercially available ELISA kits. Validation of this proteomic profile using ELISA in large-scale prospective trials of patients receiving thalidomide in the setting of induction, relapse and post-transplant maintenance is required to further define its clinical utility and to expand this paradigm to other agents.

The authors thank Dr. P. Murphy and Dr. G. Crotty for providing "clinical samples." The Mater Myeloma Cancer Research Fund, Irish Blood Transfusion Services, Irish Higher

Education Authority (HEA) PRTL Cycle 4 and Enterprise Ireland for the funding support.

The authors have declared no conflict of interest.

5 References

- [1] Barlogie, B., Shaughnessy, J., Zangari, M., Tricot, G., High-dose therapy and immunomodulatory drugs in multiple myeloma. *Semin. Oncol.* 2002, 29, 26–33.
- [2] Rajkumar, S. V., Kyle, R. A., Multiple myeloma: diagnosis and treatment. *Mayo Clin. Proc.* 2005, 80, 1371–1382.
- [3] Burington, B., Barlogie, B., Zhan, F., Crowley, J., Shaughnessy, J. D., Jr., Tumor cell gene expression changes following short-term in vivo exposure to single agent chemotherapeutics are related to survival in multiple myeloma. *Clin. Cancer Res.* 2008, 14, 4821–4829.
- [4] Scudla, V., Zemanova, M., Minarik, J., Bacovsky, J. et al., International prognostic index (IPI)—a critical comparison with five multiple myeloma staging systems in the group of 270 patients treated by conventional chemotherapy. *Neoplasma* 2006, 53, 277–284.
- [5] Tichy, M., Maisnar, V., Palicka, V., Friedecky, B. et al., International Staging System required standardization of biochemical laboratory testing in multiple myeloma. *Neoplasma* 2006, 53, 492–494.
- [6] Anderson, K. C., Alsina, M., Bensinger, W., Biermann, J. S. et al., Multiple myeloma. Clinical practice guidelines in oncology. *J. Natl. Compr. Cancer Netw.* 2007, 5, 118–147.
- [7] Anderson, K. C., Hamblin, T. J., Traynor, A., Management of multiple myeloma today. *Semin. Hematol.* 1999, 36, 3–8.
- [8] Vallet, S., Palumbo, A., Raje, N., Boccadoro, M., Anderson, K. C., Thalidomide and lenalidomide: Mechanism-based potential drug combinations. *Leuk. Lymphoma* 2008, 49, 1238–1245.
- [9] Thalidomide: new indication. For elderly myeloma patients: some improvement in first-line treatment. *Prescrire Int.* 2009, 18, 49–52.
- [10] Barlogie, B., van Rhee, F., Shaughnessy, J. D., Jr., Epstein, J. et al., Seven-year median time to progression with thalidomide for smoldering myeloma: partial response identifies subset requiring earlier salvage therapy for symptomatic disease. *Blood* 2008, 112, 3122–3125.
- [11] Dowling, P., O'Driscoll, L., Meleady, P., Henry, M. et al., 2-D difference gel electrophoresis of the lung squamous cell carcinoma versus normal sera demonstrates consistent alterations in the levels of ten specific proteins. *Electrophoresis* 2007, 28, 4302–4310.
- [12] Nagalla, S. R., Canick, J. A., Jacob, T., Schneider, K. A. et al., Proteomic analysis of maternal serum in down syndrome: identification of novel protein biomarkers. *J. Proteome Res.* 2007, 6, 1245–1257.
- [13] Huang, J. T., Wang, L., Prabakaran, S., Wengenroth, M. et al., Independent protein-profiling studies show a decrease in

- apolipoprotein A1 levels in schizophrenia CSF, brain and peripheral tissues. *Mol. Psychiatry* 2008, *13*, 1118–1128.
- [14] Dowling, P., Wormald, R., Meleady, P., Henry, M. et al., Analysis of the saliva proteome from patients with head and neck squamous cell carcinoma reveals differences in abundance levels of proteins associated with tumour progression and metastasis. *J. Proteomics* 2008, *71*, 168–175.
- [15] Pendarvis, K., Kumar, R., Burgess, S. C., Nanduri, B., An automated proteomic data analysis workflow for mass spectrometry. *BMC Bioinformatics* 2009, *11*, S17.
- [16] Fu, A. Z., Cantor, S. B., Kattan, M. W., Use of nomograms for personalized decision-analytic recommendations. *Med. Decision Making* 2009, *2*, 267–274.
- [17] Rajkumar, S. V., Buadi, F., Multiple myeloma: new staging systems for diagnosis, prognosis and response evaluation. *Best Pract. Res. Clin. Haematol.* 2007, *20*, 665–680.
- [18] Durie, B. G., Harousseau, J. L., Miguel, J. S., Blade, J. et al., International uniform response criteria for multiple myeloma. *Leukemia* 2006, *20*, 1467–1473.
- [19] Hotta, T., [Classification, staging and prognostic indices for multiple myeloma]. *Nippon. Rinsho.* 2007, *65*, 2161–2166.
- [20] Brizard, J. P., Ramos, J., Robert, A., Lafitte, D. et al., Identification of proteomic changes during human liver development by 2D-DIGE and mass spectrometry. *J. Hepatol.* 2009, *51*, 114–126.
- [21] Martinez-Gomariz, M., Hernaez, M. L., Gutierrez, D., Ximenez-Embun, P., Prestamo, G., Proteomic analysis by two-dimensional differential gel electrophoresis (2D DIGE) of a high-pressure effect in *Bacillus cereus*. *J. Agric. Food Chem.* 2009, *57*, 3543–3549.
- [22] Wu, C., Luo, Z., Chen, X., Wu, C. et al., Two-dimensional differential in-gel electrophoresis for identification of gastric cancer-specific protein markers. *Oncol. Rep.* 2009, *21*, 1429–1437.
- [23] Ralhan, R., Desouza, L. V., Matta, A., Chandra Tripathi, S. et al., iTRAQ-multidimensional liquid chromatography and tandem mass spectrometry-based identification of potential biomarkers of oral epithelial dysplasia and novel networks between inflammation and premalignancy. *J. Proteome Res.* 2009, *8*, 300–309.
- [24] Marazia, S., Barnabei, L., De Caterina, R., Receiver operating characteristic (ROC) curves and the definition of threshold levels to diagnose coronary artery disease on electrocardiographic stress testing. Part II: the use of ROC curves in the choice of electrocardiographic stress test markers of ischaemia. *J. Cardiovasc. Med. (Hagerstown)* 2008, *9*, 22–31.
- [25] Saleh, N., Olausson, A., Nilsson, T., Hansson, L. O., Tornvall, P., Comparison of effects of a thrombin-based femoral artery closure device with those of a mechanical compression device on serum C-reactive protein and amyloid A after percutaneous coronary intervention. *Am. J. Cardiol.* 2003, *92*, 324–327.
- [26] Jousilahti, P., Salomaa, V., Hakala, K., Rasi, V. et al., The association of sensitive systemic inflammation markers with bronchial asthma. *Ann. Allergy Asthma Immunol.* 2002, *89*, 381–385.
- [27] Rompp, A., Dekker, L., Taban, I., Jenster, G. et al., Identification of leptomeningeal metastasis-related proteins in cerebrospinal fluid of patients with breast cancer by a combination of MALDI-TOF, MALDI-FTICR and nanoLC-FTICR MS. *Proteomics* 2007, *7*, 474–481.
- [28] Heo, S. H., Lee, S. J., Ryoo, H. M., Park, J. Y., Cho, J. Y., Identification of putative serum glycoprotein biomarkers for human lung adenocarcinoma by multilectin affinity chromatography and LC-MS/MS. *Proteomics* 2007, *7*, 4292–4302.
- [29] Warzocha, K., Kraj, M., Poglod, R., Szpila, T. et al., Efficacy and safety of thalidomide in the treatment of multiple myeloma. *Acta Pol. Pharm.* 2008, *65*, 771–774.
- [30] Ludwig, H., Hajek, R., Tothova, E., Drach, J. et al., Thalidomide-dexamethasone compared with melphalan-prednisolone in elderly patients with multiple myeloma. *Blood* 2009, *113*, 3435–3442.
- [31] Facon, T., Mary, J. Y., Hulin, C., Benboubker, L. et al., Melphalan and prednisone plus thalidomide versus melphalan and prednisone alone or reduced-intensity autologous stem cell transplantation in elderly patients with multiple myeloma (IFM 99-06): a randomised trial. *Lancet* 2007, *370*, 1209–1218.
- [32] Zangari, M., Barlogie, B., Anaissie, E., Saghafifar, F. et al., Deep vein thrombosis in patients with multiple myeloma treated with thalidomide and chemotherapy: effects of prophylactic and therapeutic anticoagulation. *Br. J. Haematol.* 2004, *126*, 715–721.
- [33] Richardson, P., Schlossman, R., Jagannath, S., Alsina, M. et al., Thalidomide for patients with relapsed multiple myeloma after high-dose chemotherapy and stem cell transplantation: results of an open-label multicenter phase 2 study of efficacy, toxicity, and biological activity. *Mayo Clin. Proc.* 2004, *79*, 875–882.
- [34] Gockel, H. R., Luger, A., Heidemann, J., Schmidt, M. et al., Thalidomide induces apoptosis in human monocytes by using a cytochrome c-dependent pathway. *J. Immunol.* 2004, *172*, 5103–5109.
- [35] Anderson, K. C., Lenalidomide and thalidomide: mechanisms of action—similarities and differences. *Semin. Hematol.* 2005, *42*, S3–8.
- [36] Zidi, I., Mestiri, S., Bartegi, A., Amor, N. B., TNF-alpha and its inhibitors in cancer. *Med. Oncol.* 2009, *2*, 185–198.
- [37] Bartlett, J. B., Michael, A., Clarke, I. A., Dredge, K. et al., Phase I study to determine the safety, tolerability and immunostimulatory activity of thalidomide analogue CC-5013 in patients with metastatic malignant melanoma and other advanced cancers. *Br. J. Cancer* 2004, *90*, 955–961.
- [38] Hassan, M. I., Waheed, A., Yadav, S., Singh, T. P., Ahmad, F., Zinc alpha 2-glycoprotein: a multidisciplinary protein. *Mol. Cancer Res.* 2008, *6*, 892–906.
- [39] Lee, H. Y., Kim, M. K., Park, K. S., Shin, E. H. et al., Serum amyloid A induces contrary immune responses via formyl peptide receptor-like 1 in human monocytes. *Mol. Pharmacol.* 2006, *70*, 241–248.
- [40] Han, C. Y., Subramanian, S., Chan, C. K., Omer, M. et al., Adipocyte-derived serum amyloid A3 and hyaluronan play

- a role in monocyte recruitment and adhesion. *Diabetes* 2007, *56*, 2260–2273.
- [41] Yamada, T., [Serum amyloid A (SAA)–pathogenicity and implication of appearance in plasma]. *Rinsho. Byori*. 2006, *54*, 509–512.
- [42] Schneider, G. B., Grecco, K. J., Safadi, F. F., Popoff, S. N., The anabolic effects of vitamin D-binding protein-macrophage activating factor (DBP-MAF) and a novel small peptide on bone. *Crit. Rev. Eukaryot. Gene Expr.* 2003, *13*, 277–284.
- [43] Rehder, D. S., Nelson, R. W., Borges, C. R., Glycosylation status of vitamin D binding protein in cancer patients. *Protein Sci.* 2009, *18*, 2036–2042.
- [44] Abdullah, M., Schultz, H., Kahler, D., Branscheid, D. et al., Expression of the acute phase protein haptoglobin in human lung cancer and tumor-free lung tissues. *Pathol. Res. Pract.* 2009, *205*, 639–647.
- [45] Cid, M. C., Grant, D. S., Hoffman, G. S., Auerbach, R. et al., Identification of haptoglobin as an angiogenic factor in sera from patients with systemic vasculitis. *J. Clin. Invest.* 1993, *91*, 977–985.
- [46] de Kleijn, D. P., Smeets, M. B., Kemmeren, P. P., Lim, S. K. et al., Acute-phase protein haptoglobin is a cell migration factor involved in arterial restructuring. *FASEB J.* 2002, *16*, 1123–1125.
- [47] Sze, D. M., Chan, G. C., Supplements for immune enhancement in hematologic malignancies. *Hematol. Am. Soc. Hematol. Educ. Program* 2009, 313–319.
- [48] Komorowski, J., Jerczynska, H., Siejka, A., Baranska, P. et al., Effect of thalidomide affecting VEGF secretion, cell migration, adhesion and capillary tube formation of human endothelial EA.hy 926 cells. *Life Sci.* 2006, *78*, 2558–2563.
- [49] Greipp, P. R., San Miguel, J., Durie, B. G., Crowley, J. J. et al., International staging system for multiple myeloma. *J. Clin. Oncol.* 2005, *23*, 3412–3420.