HLA class I is most tightly linked to levels of tapasin compared with other antigen-processing proteins in glioblastoma

Camilla Thuring1, Elna Follin1,4, Linda Geironson1,4, Eva Freyhult2, Victoria Junghans1, Mikkel Harndahl3, Søren Buus3 and Kajsa M Paulsson *,1

1Immunology Section, Department of Experimental Medical Science, Lund University, SE-221 84 Lund, Sweden; 2Science for Life Laboratory, Bioinformatics Infrastructure for Life Sciences, Department of Medical Sciences, Cancer Pharmacology and Computational Medicine, Uppsala University, SE-751 05 Uppsala, Sweden and 3Department of Experimental Immunology, Institute of International Health, Immunology and Microbiology, DK-2200 Copenhagen, Denmark

Background: Tumour cells can evade the immune system by dysregulation of human leukocyte antigens (HLA-I). Low quantity and/or altered quality of HLA-I cell surface expression is the result of either HLA-I alterations or dysregulations of proteins of the antigen-processing machinery (APM). Tapasin is an APM protein dedicated to the maturation of HLA-I and dysregulation of tapasin has been linked to higher malignancy in several different tumours.

Methods: We studied the expression of APM components and HLA-I, as well as HLA-I tapasin-dependency profiles in glioblastoma tissues and corresponding cell lines.

Results: Tapasin displayed the strongest correlation to HLA-I heavy chain but also clustered with β2-microglobulin, transporter associated with antigen processing (TAP) and LMP. Moreover, tapasin also correlated to survival of glioblastoma patients. Some APM components, for example, TAP1/TAP2 and LMP2/LMP7, showed variable but coordinated expression, whereas ERAP1/ERAP2 displayed an imbalanced expression pattern. Furthermore, analysis of HLA-I profiles revealed variable tapasin dependence of HLA-I allomorphs in glioblastoma patients.

Conclusions: Expression of APM proteins is highly variable between glioblastomas. Tapasin stands out as the APM component strongest correlated to HLA-I expression and we proved that HLA-I profiles in glioblastoma patients include tapasin-dependent allomorphs. The level of tapasin was also correlated with patient survival time. Our results support the need for individualisation of immunotherapy protocols.

Human leukocyte antigens (HLA-I) present peptides from the endogenous pathway to CD8+ T cells, allowing cytotoxic responses against virally infected and tumour cells (Zinkernagel and Doherty, 1974). The assembly and maturation of HLA-I is promoted by the antigen-processing machinery (APM). Dysregulation of APM components, which in several tumour types result in downregulation of cell surface-expressed HLA-I molecules, may result in changed immunodominance hierarchies and decreased immunogenicity.

The HLA-I maturation, that is, folding of HLA-I heavy chain (HC), association with β2-microglobulin (β2m) and binding of a peptide of suitable length and motif takes place in the endoplasmic reticulum (ER). Peptides to be loaded onto HLA-I/β2m heterodimers are generated in the cytosol mainly by the proteasome and transported into the ER by the transporter associated with antigen processing (TAP) (Michalek et al, 1993; Neefjes et al, 1993). In the ER, the peptides may be further trimmed by the aminopeptidases ERAP1 and ERAP2 (Saveanu et al, 2005). Folding of the HLA-I
HC and its association with β2m is facilitated by ER chaperones including calnexin and calreticulin (Paulsson and Wang, 2004). The HLA-I HC/β2m dimer binds to tapasin and is incorporated into the peptide loading complex, which also consists of at least calreticulin, ERp57 and TAP (Van Hateren et al, 2010). Tapasin not only bridges HLA-I to TAP, but also promotes HLA-I maturation with both peptide- and HLA-I allomorph (allele-specific protein product) specificity (Peh et al, 1998; Roder et al, 2011; Geirong et al, 2013). Tapasin strongly influences HLA-I peptide presentation both in a qualitative and quantitative manner (Purcell et al, 2001). Although bound to tapasin, suboptimal peptides in the HLA-I binding cleft may be replaced in a peptide-editing process mediated by tapasin (Williams et al, 2002). ERp57 is involved in both early and late HLA-I maturation and the interaction of ERp57 with tapasin has been shown to enhance the peptide-editing function of tapasin (Wearsch and Cresswell, 2007).

Downregulation of HLA-I and APM components are common in tumours of different origin and are associated with worse prognosis in several malignancies as well as with the outcome of T-cell-based immunotherapy (Seliger et al, 2000; Dissemend et al, 2003; Seliger et al, 2003; Facoetti et al, 2005; Han et al, 2008; Mehta et al, 2008). Studies of a number of different tumours have focused on tapasin and have reported downregulation of tapasin to be linked to metastasis and poor outcome (Dissemend et al, 2003; Facoetti et al, 2005; Han et al, 2008; Jiang et al, 2010). In addition, tapasin has been found to be downregulated in, for example, renal cell carcinoma, prostate cancer, urothelial carcinoma, cervical carcinoma, head and neck squamous cell carcinoma and oesophageal squamous cell carcinoma (Seliger et al, 2003; Mehta et al, 2008; Bandoh et al, 2010; Cathro et al, 2010; Seliger et al, 2010; Sheyhidin et al, 2014). In astrocytic tumours, tapasin was associated, but not statistically proven to be correlated, with HLA-I downregulation or tumour grade. The same study showed that for the glioblastoma multiforme (GBM) lesions, 19 out of 24 were negative for tapasin in contrast to other APM proteins such as calnexin and calreticulin for which all lesions were positive (Facoetti et al, 2005). Glioblastoma multiforme, a grade IV astrocytic tumour, is an aggressive and highly invasive brain tumour. Treatments available today include surgical resection with subsequent radiotherapy and Temozolomide treatment (Nagasawa et al, 2012). Despite currently available treatments, the median survival following diagnosis is about 14 months, a prognosis that has remained virtually unchanged for the past decades (Wen and Kesari, 2008) and the development of complementary treatments (e.g., immunotherapy) is highly desirable.

The critical role of APM components in shaping the HLA-I presented peptide repertoire is well established but scarce information about HLA-I and APM component expression is available for malignant brain tumours. Here we have analysed the expression of APM components and HLA-I in GBM. Tapasin expression in particular was highly disregulated and significantly correlated to the expression of HLA-I HC and also to the survival time of patients. We propose that analysis of tapasin expression levels in combination with high-resolution HLA-I typing of patients could be used to improve personalised immunotherapy protocols for GBM patients.

### MATERIALS AND METHODS

**Tumour tissues and cell lines.** Biopsies from 11 resected GBM tumours and cell lines derived from the same tumours were used. To protect patient anonymity, tumour samples were de-identified and registered in a database with a sample collection code. Cell lines were labelled with the abbreviation GCL (glioma cell line) and numbers 1–11. Corresponding codes for the tumour tissues registered in the database are: GCL1 (LU-94), GCL2 (LU-93), GCL3 (LU-96), GCL4 (LU-95), GCL5 (LU-200), GCL6 (LU-203), GCL7 (LU-10), GCL8 (LU-204), GCL9 (LU-37), GCL10 (LU-202) and GCL11 (LU-92). All patients were of age 50–69 years. The use of all material has been approved by the patients and the local ethical board (LU 307-98).

For peptide pulse experiments we used the transformed lymphoblastoid B cell lines 721.221 (tapasin proficient) and 721.220 (tapasin deficient), both transfected with HLA-A*02:01, a kind gift from Dr T Elliott, Southampton University, UK.

**Cell culture.** GCLs were cultured at 37 °C and 5% CO2 in R10 medium (RPMI1640 medium (GIBCO, Invitrogen, Thermo Scientific, Waltham, MA, USA) containing 10% FBS (HyClone, Thermo Scientific), 1 μM sodium pyruvate, 10 mM HEPES (both from GIBCO, Invitrogen, Thermo Scientific) and 50 μg ml-1 gentamicin (GIBCO, Invitrogen, Thermo Scientific). Cells were trypsinised with 0.25% Trypsin-EDTA (Invitrogen, Thermo Scientific) at 37 °C for 3–5 min. The GCLs were cultured over a time period of several months and expression levels of APM proteins and HLA-I (all stable over time) were analysed primarily with western blotting (WB) and flow cytometry. The cells were also closely monitored for changes in growth pattern and morphology. The 721.221 HLA-A*02:01 and 721.220 HLA-A*02:01 were maintained in R10 supplemented with 50 μM penicillin and 50 μg ml-1 streptomycin in 37 °C and 5% CO2.

**Western blotting.** Cells were lysed in 1% NP-40 Alternative (Calbiochem, San Diego, CA, USA), 1 mM PMSF (Sigma, Saint Louis, MO, USA) and 1× complete protease inhibitor cocktail (Roche, Basel, Schweiz). BCA Protein Assay Kit (Fierce, Thermo Scientific) was used to determine protein concentration. Proteins were separated on 10% SDS-polyacrylamide gels and transferred to nitrocellulose membranes (BioRad, Hercules, CA, USA). Equal loading was confirmed with MemCode staining (Thermo Scientific). Antibodies for β2m, BAP31, calnexin, calreticulin, ERp57, HLA-I, ERAP2, LMP2 and LMP7 were from Abcam (Cambridge, UK). β-COP1 was from Thermo Scientific, ERP1 was from Acris Antibodies (Herford, Germany), TAP1 was from Calbiochem, TAP2 was from MBL (Woburn, MA, USA) and anti-tapasin-antibody was generated as previously described (Roder et al, 2009). Nitrocellulose membranes were developed with Immun-Star WesternC Kit and molecular imager Chemidoc XRS system (BioRad) or manually (Amersham Hyperfilm, ECL, GE Healthcare Life Sciences, Buckinghamshire, UK). Quantification was made with BioRad’s imaging software Image Lab 2.0.1 using a complementary approach, where the quantification of total protein in each lane was used to correct the quantification of each WB-detected protein band after transfer to nitrocellulose membranes. The experiments were repeated at least three times with samples in duplicates. 721.221 HLA-A*02:01 was used as a positive control.

**Statistical analysis of WB analysis.** Two-way analysis of variance (ANOVA) was used to test the effect on the expression of APM components caused by cell lines, experimental time point and the interaction between these two variables (Table 1) (Kempthorne, 1979; Fox and Weisberg, 2011). The Bonferroni method was used to correct the P-threshold value for multiple tests. For correlation analyses, where values for one APM component are compared with another APM component, we calculated the Pearson’s correlation coefficient. The APM components were clustered using complete linkage hierarchical clustering and correlation distance measure (1 – |Pearson correlation|) (Hastie et al, 2009). Statistical analyses were performed in environment R (R Development Core Team) and GraphPad Prism 5.0a (La Jolla, CA, USA).

**Immunocytochemistry.** Cells grown on chamber slides were fixed and permeabilised with acetone before incubation with biotinylated or non-biotinylated W6/32, a conformation-specific antibody for HLA-I heterodimers (Abcam and American Type Culture Collection (Manassas, VA, USA), respectively). After wash, secondary
Table 1. Result from two-way ANOVA of analysed APM components

| APM component | GCLs | Experimental time points | Interaction |
|---------------|------|--------------------------|-------------|
| ERAP1         | 1.33*10⁻¹⁵ | 4.94*10⁻² | 5.70*10⁻² |
| LMP2          | 2.83*10⁻¹¹ | 5.31*10⁻⁸ | 4.00*10⁻⁴ |
| ERP57         | 2.13*10⁻¹⁰ | 3.72*10⁻⁶ | 6.61*10⁻⁴ |
| Tapasin       | 2.68*10⁻¹⁰ | 2.17*10⁻⁷ | 2.84*10⁻² |
| β₂m           | 8.74*10⁻¹⁰ | 8.10*10⁻⁶ | 3.00*10⁻¹ |
| LMP7          | 2.18*10⁻⁹  | 7.90*10⁻⁷ | 1.69*10⁻⁵ |
| HLA-I HC      | 3.27*10⁻⁸  | 2.36*10⁻⁵ | 5.23*10⁻³ |
| TAP1          | 5.01*10⁻⁷  | 1.77*10⁻³ | 2.79*10⁻⁴ |
| ERAP2         | 4.66*10⁻⁶  | 4.66*10⁻¹ | 5.92*10⁻⁵ |
| TAP2          | 3.21*10⁻⁵  | 2.42*10⁻³ | 2.65*10⁻² |
| Calnexin      | 1.24*10⁻³  | 1.51*10⁻³ | 5.49*10⁻³ |
| Calreticulin  | 2.65*10⁻³  | 1.13*10⁻³ | 5.14*10⁻¹ |
| β₂-COP1       | 1.94*10⁻²  | 3.27*10⁻² | 1.03*10⁻² |
| BAP31         | 5.92*10⁻¹¹ | 2.76*10⁻³ | 9.33*10⁻¹ |

Abbreviations: ANOVA = analysis of variance; APM = antigen-processing machinery; β₂m = β₂-microglobulin; GCL = glioma cell line; HC = heavy chain. All the reported numbers are P-values.

antibodyAlexa 555 Streptavidin or Alexa 488, respectively, were applied and slides mounted with ProLong Gold Antifade Reagent with DAPI (Invitrogen, Thermo Scientific). Slides were analysed at room temperature (RT) with a Zeiss Axiocert 200M microscope (Carl Zeiss Microscopy, Göttingen, Germany), a Hamamatsu C4742-80-12AG camera/detector (Hamamatsu Photonics, Hamamatsu, Japan) and Volocity 5.0 software (Informer Technologies, Ashland, OR, USA) and FlowJo 9.5 (Tree Star Inc., Ashland, OR, USA) was used to analyse the images. Images are X = 0.31 μm, Y = 0.31 μm per pixel, and the image bit depth is millions of colours. The immunocytochemistry (ICC) staining was scored from 0 to 5 (with 5 being staining with highest intensity) independently by at least two scientists.

Flow cytometry. Cell surface expression of HLA-I was evaluated using 10 μg ml⁻¹ of HLA-I heterodimer-conformation-specific FITC-conjugated W6/32 antibody (Biolegend, San Diego, CA, USA, catalogue number 311404) for 30 min in the dark at 4°C. FITC-conjugated mouse IgG2a (Biolegend, catalogue number 400207) was used as isotype control. After incubation, cells were washed and filtered through a 60-μm nylon net filter before analysis. Samples were analysed on the BD LSR II (Becton Dickinson, San Jose, CA, USA) and FlowJo 9.5 (Tree Star Inc., Ashland, OR, USA) was used to analyse the data. To normalise, the results are shown as the ratio between the median fluorescence intensity (MFI) of cells stained with W6/32 and the MFI of isotype IgG2a-stained cells.

Immunohistochemical staining of tissues. Paraffin-embedded tissue sections were deparaffinised, rehydrated and stained using the MACH 2 system (Biocare Medical, Concord, CA, USA). Sections were stained using antibodies against HLA-I HC (Abcam (ab70328)) or tapasin (TO-3) and scanned with the ScanScope CS system (Aperio, Leica Microsystems, Wetzlar, Germany). The digital slides were analysed in ImageScope 7.01 (Aperio) and each pixel was analysed for positive staining and intensity (Thuring et al., 2014). For correlation analyses, the Pearson’s correlation coefficient was calculated and linear regression was done with tapasin, as well as age and gender of the patient as prognostic factors for survival. Areas with tumour grade lower than grade IV and necrotic areas as well were identified with the assistance of a pathologist and excluded from the analysis. Pearson’s test, a non-parametric (distribution-free) rank statistic test, was used to analyse possible monotonic relationships.

Peptide–HLA-I folding assay. Folding of HLA-I HC-β₂m dimers with graded amounts of peptide in the presence or the absence of tapasin_{1–87} (tpn_{1–87}) was monitored in a luminescence oxygen channeling immuno (LOCI) assay (AlphaScreen, Perkin Elmer, Waltham, MA, USA) (Roder et al., 2009). Recombinant HLA-I HCs, β₂m and tpn_{1–87} were generated as previously described (Ostergaard Pedersen et al., 2001; Leisner et al., 2008; Roder et al., 2009).

Peptide synthesis. The random peptide libraries (7–13 amino acids in length) were synthesised by Schaefer-N (Copenhagen, Denmark) and purified as previously described (Roder et al., 2011). NLVPVMATV (NLV), VLEETSVML (VLE) and GIYGAVIPL (GVI) used for the peptide pulse experiment were synthesised by ChinaPeptides Co., Ltd (Shanghai, China).

HLA typing. High-resolution HLA-I typing was performed by Genome Diagnostics B.V. (Nijmegen, The Netherlands) using sequence-based typing. In brief, DNA was purified from the 11 GCLs using the QIAamp DNA Mini Kit (Qiagen, Venlo, Netherlands). The DNA was sent off for PCR amplification and the full nucleotide sequences were determined and aligned to regularly updated databases containing all identified and verified allele sequences (IMGT/HLA sequence database).

Peptide pulse. The cells were starved 12–24 h, to increase the surface expression of HLA-I before they were pulsed with peptide. Briefly, the cells were washed twice in PBS and seeded in round-bottom 96-well plates at a density of 1 × 10⁶ cells per well and cultured at RT and atmospheric CO₂. The following day, the cells were pulsed with either 200 μM peptide or an equal volume PBS for 1 h in the presence of 10 μg ml⁻¹ Brefeldin A (BFA). After 1 h peptide pulse, the media was changed back to R10 media and cultured at 37°C and 5% CO₂ in the presence of 10 μg ml⁻¹ of BFA. The cells were collected at different time points and the surface expression was evaluated using flow cytometry.

Extensive variation in expression of APM components in GCLs. In a recent study we showed, for a set of GBMs (12 tumours), that the level of tapasin correlates with HLA-I expression and survival (Thuring et al., 2014). In this study we have done an extensive and detailed analysis of the expression levels of APM components in a unique set of cell lines (GCLs) derived from 11 GBMs (Figure 1). Despite highly variable levels, none of these GCLs showed a complete loss of expression of any of the APM components. However, GCL 5, 7 and 8 showed very low levels of β₂m (Figure 1A). A two-way ANOVA of the data showed significant variance between GCLs with P < 0.0012 for all the studied proteins except BAP31 (Table 1). The P-value threshold was divided with the number of test samples to get a P-threshold value corrected for multiple-tests. For TAP1, LMP2, LMP7, ERAP2 and ERP57, the P-value for the interaction term was significant (P < 0.0012).

Despite the significant interaction term for some of the APM proteins, we calculated the Pearson’s correlation coefficient for those that are likely to be expressed in a coordinated manner, for example, TAP1/TAP2 and LMP2/LMP7. The expression of TAP1/TAP2 (Pearson’s r = 0.93) was strongly correlated, as was LMP2/LMP7 (Pearson’s r = 0.92) (Supplementary Figure 1). Moreover, HLA-I HC expression correlated to the levels of β₂m (Pearson’s r = 0.67), tapasin (Pearson’s r = 0.87) and TAP2 (Pearson’s r = 0.62) (Figure 1C and Supplementary Figure 1). A clustering analysis of the APM components based on Pearson’s correlation coefficient showed that HLA-I HC, tapasin, TAP1, TAP2, LMP2, LMP7 and β₂m formed a cluster where all the proteins were more or less correlated to each other (Supplementary
Figure 1. The expression levels of APM components are highly variable in GCLs. (A) GCLs were lysed and equal amounts of protein were separated with SDS-PAGE and proteins were detected with immunoblotting. 721.221-HLA-A*02:01 cells were used as control. Representative raw data from three to six WB experiments is shown. (B) Densiometric analysis of WB data. Experiments have been repeated at least three times with samples in duplicate each time. Equal amounts of protein were separated with SDS-PAGE. Immunoblot detection was performed as indicated in Materials and Methods, with primary antibodies indicated in the diagrams respectively. Quantification of total protein in each lane was used to correct the quantification of each WB-detected protein band after transfer to nitrocellulose membranes. Error bars represent ± s.e.m. (C) Correlation analysis of HLA-I HC, β2m and tapasin. Values from densiometric analyses of immunoblotted proteins of the 11 GCLs are plotted against each other and Pearson’s correlation coefficient calculated.

GCLs express variable amounts of cell surface HLA-I. Owing to the dysregulations of APM protein levels, we wanted to assess the levels of assembled HLA-I on the cell surface of the GCLs. Immunocytochemistry staining of total HLA-I was done; however, when evaluated, only the cell surface staining was taken into account. The ICC indicated a high variation between GCLs in cell surface-expressed assembled HLA-I molecules (Figure 2A).

To further quantitate the cell surface-expressed HLA-I, we also did flow cytometry analysis (Figure 2B). This confirmed not only the high variation but also the hierarchy in HLA-I surface expression levels between the GCLs (Figure 2C). The levels of cell surface-expressed HLA-I molecules did not correlate to the expression levels of any of the analysed APM components (data not shown).
Tapasin is a prognostic marker for survival time of GBM patients. For a different set of samples (12 GBM tumours), we recently showed, by staining for tapasin and HLA-I HC with immunohistochemistry, that tapasin levels correlate with HLA-I HC expression and days of survival for patients (Thuring et al., 2014). Here we extended the analysis and included the 11 GBM tumours from which the GCLs used here were established. The IHC analysis showed extensive variation in the expression of tapasin and HLA-I among the GBMs (Figure 3A and B). Correlation analysis with calculated Pearson’s correlation coefficient showed that tapasin expression correlated with both HLA-I HC expression and days of survival in the extended set of 23 patients (Figure 3C and D). Linear regression further proved that tapasin is an independent factor for survival (Supplementary Table 2). This not only strengthens the results from our previous study but also firmly establishes tapasin as an important prognostic factor for GBM patients.

GCLs display HLA-I profiles with mixed high and low tapasin dependency allomorphs. Owing to the lack of correlation between tapasin expression and surface-expressed HLA-I molecules in the GCLs, we hypothesised that the HLA-I molecules presented at the cell surface could be less dependent on tapasin and thus be expressed at the cell surface even if tapasin is downregulated. To investigate this we performed high-resolution HLA-A and -B typing of each GCL and also, for comparison, of a healthy cohort of 31 individuals (Table 2 and Supplementary Table 1). To get the tapasin dependency for each allomorph, we used recombinant HLA-I heavy chains and studied the tapasin facilitation in a LOC5 assay (Roder et al., 2011). The studied allomorphs ranged from high to low tapasin dependency (Figure 4A). Analysis revealed that GCL 7 followed by GCL 5 had the lowest tapasin dependencies, and that all of the glioma patients expressed a mixture of HLA-I allomorphs with different tapasin dependencies (Figure 4B), which was also true for the healthy cohort (Figure 4C).

Lack of tapasin results in suboptimally loaded HLA-I on the cell surface. Tapasin was extensively dysregulated in the GCLs and GBMs, and was also the protein that had the strongest correlation to HLA-I HC expression; thus, we wanted to investigate tapasin’s effect on surface-expressed HLA-I molecules. The HLA-I typing revealed a high frequency of HLA-A*02:01, a very common allele in Caucasian populations and a molecule well studied in terms of tapasin dependence and peptide binding (Barber et al., 2001; Roder et al., 2011). In brief, 721.220 HLA-A*02:01 and 721.221 HLA-A*02:01 cells were incubated with peptides NLV, VLE or GIY, (predicted high affinities for HLA-A*02:01 (netMHCPan: 21, 150 and 47 nM, respectively)) and treated with BFA. Only one of these, GIY, generated an increased amount of W6/32 reactive HLA-I molecules on the tapasin-deficient 721.220 HLA-A*02:01 cells (Figure 5, upper panel).

DISCUSSION

Out of all the APM components analysed, tapasin had the strongest correlation with the expression of HLA-I HC (Figure 1C). The WB analysis also indicated a correlation between TAP1 and TAP2, as well as between LMP2 and LMP7, which could be expected, as their transcription is controlled by the same bidirectional promoter (Wright et al., 1995). Further, tapasin expression clustered with the TAP and LMP proteins, as well as HLA-I HC and β2m (Supplementary Figure S1). The correlated expression of these proteins suggests a regulatory mechanism that could cause simultaneous deficiency in the expression of genes at disperse loci. ERAPI and ERAP2 showed imbalanced expression, which has also been observed in tumours of different origin (Fruci et al., 2006, 2008). As ERAPI and 2 have different cleavage preferences, this may result in a shift of the presented antigenic peptide pool (Saveanu et al., 2005).
As expected, there was less variation in the expression of the more general chaperones calreticulin and calnexin, and the intracellular transport proteins β-COP1 and BAP31 (Figure 1 and Table 1). Similarly, ERp57 is involved in several different pathways. However, the levels among the GCLs varied and increased expression of ERp57 has been observed in transformed cells and has been linked to human diseases including cancer (Coe and Michalak, 2010). When analysing β-actin and GAPDH, we found that the levels varied a lot between the GCLs, which is well in line with several common housekeeping proteins not being stable in tumour cells (Ferguson et al., 2005; Waxman and Wurmbach, 2007). Although the transport protein BAP31 was expressed at equal levels (Figure 1), we decided to use a total protein stain to correct for equal loading of protein lysate.

Despite the strong correlation between tapasin and HLA-I HC in total cell lysates, we could not detect any correlation between tapasin and the level of W6/32 reactive HLA-I molecules expressed on the cell surface. This might seem surprising, as a correlation between tapasin and surface-expressed HLA-I has been shown (Peh et al., 1998; Garbi et al., 2000; Grandea et al., 2000; Barber et al., 2001). However, those studies were made in models with tapasin either present or absent and otherwise intact APM, whereas the GCLs in our study all had multiple APM dysregulations. With tapasin present, the HLA-I HCs could, to a larger extent.

Table 2. High-resolution HLA-I typing of GCLs and mean tapasin facilitation

| GCL | HLA-A1 | HLA-A2 | HLA-B1 | HLA-B2 |
|-----|--------|--------|--------|--------|
|     | Allomorph | Mean Tpn fac ($r_B^{max}$) | Allomorph | Mean Tpn fac ($r_B^{max}$) | Allomorph | Mean Tpn fac ($r_B^{max}$) | Allomorph | Mean Tpn fac ($r_B^{max}$) |
| 1   | 02:01:01 | 1.7 | – | n.d. | 07:02 | 2.6 | – | n.d. |
| 2   | 01:01 | 1.9 | 02:01 | 1.7 | 08:01 | 2.6 | 51:01:01 | 2.1 |
| 3   | 01:01 | 1.9 | 02:01 | 1.7 | 08:01 | 2.6 | 57:01:01 | 1.6 |
| 4   | 02:01:01 | 1.7 | 68:02:01 | 1.8 | 15:01:01:01 | 2.0 | 44:02:01:01 | 10.6 |
| 5   | 24:02:01:01 | 1.9 | 68:02:01 | 1.8 | 15:01:01:01 | 2.0 | 40:01 | 1.6 |
| 6   | 03:01:01 | 2.5 | 68:02:01 | 1.8 | 14:01 | n.d. | 40:01:02 | 1.6 |
| 7   | 01:01 | 1.9 | 02:01 | 1.7 | 39:06 | 1.0 | 40:01 | 1.6 |
| 8   | 01:01:01:01 | 1.9 | – | n.d. | 08:01 | 2.6 | 15:01:01:01 | 2.0 |
| 9   | 03:01:01 | 24:02:01:01 | 1.9 | 07:02 | 2.6 | 15:01:01:01 | 2.0 |
| 10  | 01:01:01:01 | 1.9 | 30:04 | n.d. | 08:01 | 2.6 | 14:02:01 | 2.7 |
| 11  | 01:01:01:01 | 1.9 | 03:01:01 | 2.5 | 07:02 | 2.6 | 08:01:01 | 2.6 |

Abbreviations: fac = facilitation; GBM = glioblastoma multiforme; n.d. = not determined; $r_B^{max}$ = relative $B_{max}$; Tpn = Tapasin. Determination of tapasin dependency was not possible for all allomorphs present in GBMs (n.d.). Here ‘–’ refers to homozygocity, deletion or similar.

As expected, there was less variation in the expression of the more general chaperones calreticulin and calnexin, and the intracellular transport proteins β-COP1 and BAP31 (Figure 1 and Table 1). Similarly, ERp57 is involved in several different pathways. However, the levels among the GCLs varied and increased expression of ERp57 has been observed in transformed cells and has been linked to human diseases including cancer (Coe and Michalak, 2010). When analysing β-actin and GAPDH, we found that the levels varied a lot between the GCLs, which is well in line with several common housekeeping proteins not being stable in tumour cells (Ferguson et al., 2005; Waxman and Wurmbach, 2007). Although the transport protein BAP31 was expressed at equal levels (Figure 1), we decided to use a total protein stain to correct for equal loading of protein lysate.

Despite the strong correlation between tapasin and HLA-I HC in total cell lysates, we could not detect any correlation between tapasin and the level of W6/32 reactive HLA-I molecules expressed on the cell surface. This might seem surprising, as a correlation between tapasin and surface-expressed HLA-I has been shown (Peh et al., 1998; Garbi et al., 2000; Grandea et al., 2000; Barber et al., 2001). However, those studies were made in models with tapasin either present or absent and otherwise intact APM, whereas the GCLs in our study all had multiple APM dysregulations. With tapasin present, the HLA-I HCs could, to a larger extent.
extent, temporarily be rescued from degradation through tapasin’s retention/recycling function, but owing to the defects in other parts of the APM the HLA-I HCs could not reach a conformation stable enough to allow efficient presentation at the cell surface (Figure 6).

The effect of intact APM apart from tapasin deficiency was studied using the tapasin-deficient 721.220-HLA-A*02:01 cell line. On these cells we were able to increase the stability of surface-expressed HLA-A*02:01 molecules but only for one out of the three exogenously added peptides (Figure 5). Based on these data, we propose that the conformation of suboptimally loaded HLA-A*02:01 is different on those expressed at the cell surface compared with their conformation during maturation in the ER.

The relation between W6/32 reactive epitopes and tapasin levels likely varies between individuals, as the GBM patients, as well as the healthy cohort, expressed HLA-I allomorphs with a range of tapasin dependency (Figure 4B and C). HLA-I allomorphs with low tapasin dependency are suggested to have a higher intrinsic stability that allows them to be expressed on the cell surface even in the absence of tapasin (Geironson et al, 2013). The two GBM patients with lowest tapasin dependency were both in the group with less W6/32 surface expression but worth noting is that they also both had significantly downregulated tapasin expression (only 14% and 17%, respectively, of the level of the GCL with highest expression and 14% and 50%, respectively, of the level of the GBM with highest expression).

Although HLA-I allomorphs with low tapasin dependency can be presented on the cell surface in the absence of tapasin, there is still a strong correlation between the expression of tapasin and days of survival, which we have previously shown in a similar IHC study performed by our group (Thuring et al, 2014). We again analysed the GBM patients with an IHC method-based stainings being (I) weak positive, (II) positive and (III) strong positive. To study the effect of tapasin expression on survival, we here compared the material by dividing it into three equal sub-groups, which showed a significant correlation between tapasin expression and survival (Supplementary Figure 3). Moreover, when adding the data from the 11 GBMs to the previously analysed set of 12 GBMs, the correlation between tapasin and survival was strengthened.
Moreover, additional statistical analysis proved tapasin to be a prognostic factor of survival (Supplementary Table 2 and Supplementary Figure 3d).

For the past years, the dogma that optimal peptides are those with high affinity for the HLA-I-binding cleft has been challenged (Assarsson et al., 2008; Roder et al., 2011; Geironson et al., 2012; Harndahl et al., 2012). It is evident that affinity is not the only factor that needs to be taken into account when searching for an optimal peptide but also the stability that the peptide induces in the HLA-I molecule (Geironson et al., 2012; Harndahl et al., 2012).

The dependency of tapasin is highly varying. Some HLA-I allomorphs, for example, HLA-B*44:02 are critically dependent on tapasin for maturation (Garbi et al., 2000; Purcell et al., 2001; Zarling et al., 2003; Geironson et al., 2013). Sieker et al., (2007, 2008) have shown that HLA-B*44:02 has a greater flexibility in the β2–helix in the peptide-binding cleft compared with the less tapasin-dependent HLA-B*44:05. In addition, exchange of alanine to leucine in position 2 of the melanoma antigen MART-127–35 increases the peptide affinity for the HLA-A*02:01 but renders the peptide:HLA-I complex worse in respect to its ability to bind to a T-cell receptor due to an increased flexibility in the β2-chain of the peptide-binding cleft (Insaidoo et al., 2011). Based on these findings, one could hypothesise that tapasin works to promote selection of peptides that reduce the flexibility of the peptide-binding cleft in the peptide:HLA-I complexes and thus induces more energetically favourable structures. The high variation in APM component expression and the pronounced correlation between HLA-I HC and tapasin levels shown here are together with the demonstrated function of tapasin to edit the HLA-I-presented peptide repertoire, change immune dominance hierarchies and exert quality control of HLA-I (Garbi et al., 2000; Purcell et al., 2001; Roder et al., 2011; Geironson et al., 2013), suggesting tapasin as a unique target molecule for manipulation of immune reactions.

The here presented work is another step towards improved molecular understanding of tumour evolution, immune evasion strategies and tumour immunology. The median survival, for GBM patients, has not changed during the last decades and it is evident that new treatments are highly desirable. Tapasin was here shown to be frequently dysregulated in GBMs and strongly linked to HLA-I expression, suggesting that histone deacetylase inhibitors such as trichostatin A (Setiadi et al., 2008) or IFNγ, either through gene transfer or as a recombinant protein, may upregulate both tapasin and other IFNγ-regulated APM proteins and improve HLA-I antigen presentation at the cell surface. The outcome for peptide vaccination in clinical trials has been very poor so far. Most commonly, the peptides are chosen based on high affinity and the patients have usually been heavily pre-treated with chemotherapy, which depletes the effector cells needed to elicit a tumoricidal immune response. It has previously been shown that directly targeting of TAP- and tapasin-independent peptides to the ER restores antigen presentation, and it was suggested that the critical factor was the ability of these peptides to be presented in the absence of tapasin (Aladin et al., 2007). A number of TAP-independent epitopes are presented in TAP-deficient cells by different allomorphs (Spies et al., 1992; Gil-Torregrosa et al., 1998; Yewdell et al., 1998; Kovjazin et al., 2011; Seidel et al., 2012), whereas the situation in tapasin-deficient cells is less dependent on peptide but more on the specific HLA-I allomorphs of the individual. Hence, HLA-I allomorph must always be considered in peptide-based vaccine strategies. In addition, TAP deficiency calls primarily for targeting of

**Figure 5. Exogenous peptide can be loaded on tapasin-deficient cells.** The three peptides tested were selected based on high-affinity peptides for HLA-A*02:01. Cells were pulsed at RT for 1 h with 0 μM (control) or 200 μM peptide in the presence of BFA. Peptide was removed and cells were transferred to 37 °C and treated with BFA for 0–4 h. Cells were stained with FITC-conjugated W6/32 and analysed with flow cytometry. Graphs show amount of surface HLA-I complexes present after 0 and 4 h of BFA treatment of tapasin-deficient 721.220 HLA-A*02:01 cells (upper panel) and tapasin-proficient 721.221 HLA-A*02:01 cells (lower panel).
TAP-independent tumour epitopes, whereas tapasin deficiency calls for an even more careful selection of HLA allomorph based on knowledge of the tapasin dependence. Importantly, deficiency in both tapasin and TAP likely block HLA-I maturation, even in the presence of TAP-independent peptides, because there will be too low local concentration of TAP-independent peptides and/or ER-retained semi-mature HLA-I molecules for efficient peptide loading. Choice of peptide should be made also considering factors such as co-stimulation, adjuvance and relieve of immune suppression. We propose that protocols involving HLA-I-binding antigenic peptides should be individualised and include high-resolution HLA-I typing along with determination of tapasin expression and dependency of allomorphs present in each patient. This would allow for a better selection of peptides of higher immunogenicity, as one could estimate which HLA-I allomorphs are likely to be presented at the cell surface. Alternatively, the total value of tapasin dependency of a HLA-I profile may be considered as a stratifying factor for inclusion or not in HLA-I-binding peptide-based immunotherapy.

Figure 6. Schematic figure of the effects of tapasin deficiencies and APM defects. The HLA-I and tapasin-deficient 721.220 cell line is often used as a model where an HLA-I allele of interest and tapasin is transfected into the cells to study the effect of tapasin on HLA-I surface expression. In these transfectants, the machinery of antigen processing is functional and the effect of tapasin is correlated with the surface expression of HLA-I, especially for HLA-I allomorphs with high tapasin dependency. In tumours, the APM is defect and thus the folding and maturation of HLA-I molecules are impaired even in the presence of tapasin. Tapasin can rescue misfolded HLA-I HCs from degradation through its retention/recycling function, but owing to defects in other parts of the APM many HLA-I molecules cannot reach a conformation making them stable enough to reach and efficiently present peptide on the cell surface. If there are multiple APM deficiencies including the level of tapasin expression, the maturation of HLA-I HCs are less successful, more HC is degraded and a higher proportion of unstable HLA-I molecules reach the cell surface.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Aladin F, Lautscham G, Humphries E, Coulson J, Blake N (2007) Targeting tumour cells with defects in the MHC Class I antigen processing pathway with CD8+ T cells specific for hydrophobic TAP- and Tapasin-independent peptides: the requirement for directed access into the ER. Cancer Immunol Immunother 56(8): 1143–1152.
Assarsson E, Bui HH, Sidney J, Zhang Q, Glenn I, Oseroff C, Mbawuike IN, Alexander J, Newman MJ, Grey H, Sette A (2008) Immune analysis of the repertoire of T-cell specificities for influenza A virus in humans. J Virol 82(24): 12241–12251.
Bandoh N, Ogino T, Katayama A, Takahara M, Katada A, Hayashi T, Harabuchi Y (2010) HLA class I antigen and transporter associated with antigen processing downregulation in metastatic lesions of head and neck squamous cell carcinoma as a marker of poor prognosis. Oncol Rep 23(4): 933–939.
APM dysregulations in glioblastoma multiforme

Barber LD, Howarth M, Bowes P, Elliott T (2001) The quantity of naturally processed peptides stably bound by HLA-A*0201 is significantly reduced in the absence of tapasin. *Tissue Antigens* 58(6): 363–368.

Cathro HP, Smolkin ME, Theodorescu D, Jo VY, Ferrone S, Frierson Jr, HF (2010) Relationship between HLA class I antigen processing machinery component expression and the clinicopathologic characteristics of bladder carcinomas. *Cancer Immunol Immunother* 59(3): 465–472.

Coe H, Michalak M (2010) ERp57, a multifunctional endoplasmic reticulum resident oxidoreductase. *Int J Biochem Cell Biol* 42(6): 796–799.

Dissemond J, Kochen T, Mors J, Weimann TK, Lindeke A, Goos M, Ferguson RE, Carroll HP, Harris A, Maher ER, Selby PJ, Banks RE (2005) Relationship between HLA class I antigen processing machinery component expression and intratumoral HLA class I molecules with preformed disulfide bonds. *Eur J Immunol* 35(10): 2986–2996.

Ferguson RE, Carroll HP, Harris A, Maher ER, Selby PJ, Banks RE (2005) Housekeeping proteins: a preliminary study illustrating some limitations as useful references in protein expression studies. *Proteomics* 5(2): 566–571.

Fox J, Weisberg S (2011) An R Companion to Applied Regression, 2nd edn. Sage.

Fruci D, Ferrari C, Luminizi MZ, Cunselo V, Gjoka E, Fraeli R, Sibilio L, Carroll O, Hattori A, van Endert PM, Giacomini P (2006) Expression of endoplasmic reticulum aminopeptidases in EBV-B cell lines from healthy donors and in leukemia/lymphoma, carcinoma, and melanoma cell lines. *J Immunol* 176(8): 4869–4879.

Fruci D, Giacomini P, Nicotra MR, Forloni R, Saveanu L, van Endert P, Natali PG (2008) Altered expression of endoplasmic reticulum aminopeptidases ERAP1 and ERAP2 in transformed non-lymphoid human tissues. *J Cell Physiol* 216(3): 742–749.

Garbi N, Tan P, Dierlh AD, Chambers BJ, Ljunggren HG, Blichert-Hammerling MJ (2000) Impaired immune responses and altered peptide repertoire in tapasin-deficient mice. *Nat Immunol* 1(3): 234–238.

Geirson R, Gorder, Paulsson K (2012) Stability of peptide-HLA-I complexes and tapasin folding facilitation tools to define immunogenic peptides. *FEBS Lett* 586(9): 1336–1343.

Geirson R, Lrnering M, Harndahl M, Rasmussen M, Buus S, Gorder, Paulsson K (2013) Tapasin facilitation of natural HLA-A and -B antigen processing. *Tissue Antigens* 82(4): 363–368.

Han LY, Fletcher MS, Urbaner DL, Mueller P, Landen CN, Kamat AA, Lin YG, Merritt WM, Spannuth WA, Deavers MT, De Geest K, Gershenson DM, Lutgendorf SK, Ferrone S, Sood AK (2008) HLA class I antigen processing machinery component expression and intratumoral T-Cell infiltrate as independent prognostic markers in ovarian carcinoma. *Clin Cancer Res* 14(11): 3372–3379.

Hastie T, Tibshirani R, Friedman J (2009) 3.13.12 Hierarchical clustering. *The Elements of Statistical Learning*, 2nd edn. Springer: New York.

Insaidoo FK, Borbulevych OY, Hossain M, Santhanagopolan SM, Baxter TK, Baker BM (2011) Loss of T cell antigen recognition arising from changes in peptide and major histocompatibility complex protein flexibility: implications for vaccine design. *J Biol Chem* 286(16): 40163–40173.

Jiang Q, Pan HY, Ye DX, Zhang P, Zhang LP, Zhang ZY (2011) Downregulation of tapasin expression in primary human oral squamous cell carcinoma: association with clinical outcome. *Tumour Biol* 31(5): 451–459.

Kemphoore O (1979) *The Design and Analysis of Experiments*.

Kovjazin R, Volovitz I, Daon Y, Vider-Shalit T, Azran R, Tsaban L, Carmon L, Ferrone S, Smolkin ME, Theodorescu D, Jo VY, Ferrone S, Tho R, Vanderbilt CM (2007) Signal peptides and trans-membrane regions are broadly immunogenic and have high CD8+ T cell epitope densities: Implications for vaccine development. *Mol Immunol* 48(8): 1009–1018.

Leisner C, Loeth N, Lambeth K, Justesen S, Sylvester-Hvid C, Schmidt EG, Claesson M, Buus S, Stryhn A (2008) One-pot, mix-and-read peptide-MHC tetramers. *PLoS One* 3(2): e1678.

Mehta AM, Jordanova ES, Kenter GG, Ferrone S, Fleuren GJ (2008) Association of antigen processing machinery and HLA class I defects with clinicopathological outcome in cervical carcinoma. *Cancer Immunol Immunother* 57(2): 197–206.

Michael MT, Grant EP, Grann C, Goldberg AL, Rock KL (1993) A role for the ubiquitin-dependent proteolytic pathway in HMC class I-restricted antigen presentation. *Nature* 363(6429): 552–554.

Nagasawa DT, Chow F, Yew A, Kim W, Cremers N, Yang I (2012) Temozolomide and other potential agents for the treatment of glioblastoma multiforme. *Neuro Oncol* 14(3): 307–322ix.

Neefjes JJ, Momburg F, Hammerling GJ (1993) Selective and ATP-dependent translocation of peptides by the MHC-encoded transporter. *Science* 261(5122): 769–771.

Ostergaard Pedersen L, Nissen NH, Hansen NJ, Nielsen LL, Lauenmoller SL, Blicher T, Nansen A, Sylvester-Hvid C, Thomsen AR, Buus S (2001) Efficient assembly of recombinant major histocompatibility complex I molecules with preformed disulfide bonds. *Eur J Immunol* 31(10): 2986–2996.

Paulsson KM, Wang P (2004) Quality control of MHC class I maturation. *Faseb J* 18(1): 31–38.

Pelt CA, Burrows SR, Barnden M, Khanna R, Cresswell P, Moss DJ, McCloskey J (1998) HLA-B7-restricted antigen presentation in the absence of tapasin reveals polymorphism in mechanisms of HLA class I peptide loading. *Immunol* 8(5): 531–542.

Porcelli AW, Gorman JJ, Garcia-Peydro M, Paradela A, Burrows SR, Talbo GH, Laham N, Pelt CA, Reynolds EC, Lopez D, Castro JA, McCloskey J (2001) Quantitative and qualitative influences of tapasin on the class I peptide repertoire. *Immunol* 166(2): 1016–1027.

Roder G, Geirson R, Darabi A, Harndahl M, Schafer-Nielsen C, Skjoldi K, Buus S, Paulsson K (2009) The outermost N-terminal region of tapasin facilitates folding of major histocompatibility complex class I. *Eur J Immunol* 39(10): 2602–2609.

Roder G, Geirson R, Lrasmussen M, Harndahl M, Buus S, Paulsson K (2011) Tapasin discriminates peptide-human leukocyte antigen-A*0201 complex forms with natural ligands. *J Biol Chem* 286(23): 20547–20557.

Saveau L, Carroll O, Lindo V, Del Val M, Lopez D, Lepelletier Y, Greer F, Schomburg L, Fruci D, Niedermann G, van Endert PM (2005) Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nat Immunol* 6(7): 689–697.

Seidel U, Oliveira CC, Lampen MH, Hall T (2012) A novel category of antigens enabling CTL immunity to tumor escape variants: Cinderella antigens. *Cancer Immunol Immunother* 61(11): 221–225.

Seliger B, Atkins D, Bock M, Ritzi U, Ferrone S, Huber C, Storkel S (2003) Characterization of human lymphocyte antigen class I antigen-processing machinery defects in renal cell carcinoma lesions with special emphasis on transporter-associated with antigen-processing down-regulation. *Clin Cancer Res* 9(5): 1712–1727.

Seliger B, Maeruer MJ, Ferrone S (2000) Antigen-processing machinery breakdown and tumor growth. *Immunol Today* 21(9): 455–464.

Seliger B, Stoehr O, Handke D, Mueller A, Ferrone S, Willich B, Tannapfel A, Hofstaedter F, Hartmann A (2010) Association of HLA class I antigen abnormalities with disease progression and early recurrence in prostate cancer. *Cancer Immunol Immunother* 59(4): 529–540.

Setiadi AF, Omulius K, David MD, Seipp RP, Hartikainen J, Gopaul R, Choi KB, Jefferies WA (2008) Epigenetic changes within the promoter regions of antigen processing machinery family genes in Kazakh primary esophagal squamous cell carcinoma. *Asian Pac J Cancer Prev* 9(15): 10299–10306.

Sieker F, Springer S, Zacharias M (2007) Comparative molecular dynamics analysis of tapasin-dependent and -independent MHC class I alleles. *Protein Sci* 16(2): 299–308.

Sieker F, Straatsma TP, Springer S, Zacharias M (2008) Differential tapasin dependence of MHC class I molecules correlates with conformational changes upon peptide dissociation: a molecular dynamics simulation study. *Mol Immunol* 45(14): 3714–3722.

Spies T, Cerundolo V, Colonna M, Cresswell P, Townsend A, DeMars R (1992) Presentation of viral antigen by MHC class I molecules is
dependent on a putative peptide transporter heterodimer. Nature 355(6361): 644–646.

Thuring C, Geironson L, Paulsson K (2014) Tapasin and human leukocyte antigen class I dysregulation correlates with survival in glioblastoma multiforme. Anticancer Agents Med Chem 14(8): 1101–1109.

Van Hateren A, James E, Bailey A, Phillips A, Dalchau N, Elliott T (2010) The cell biology of major histocompatibility complex class I assembly: towards a molecular understanding. Tissue Antigens 76(4): 259–275.

Waxman S, Wurmbach E (2007) De-regulation of common housekeeping genes in hepatocellular carcinoma. BMC Genomics 8: 243.

Wearsch PA, Cresswell P (2007) Selective loading of high-affinity peptides onto major histocompatibility complex class I molecules by the tapasin-ERp57 heterodimer. Nat Immunol 8(8): 873–881.

Wen PY, Kesari S (2008) Malignant gliomas in adults. N Engl J Med 359(5): 492–507.

Williams AP, Peh CA, Purcell AW, McCluskey J, Elliott T (2002) Optimization of the MHC class I peptide cargo is dependent on tapasin. Immunity 16(4): 509–520.

Wright KL, White LC, Kelly A, Beck S, Trowsdale J, Ting JP (1995) Coordinate regulation of the human TAP1 and LMP2 genes from a shared bidirectional promoter. J Exp Med 181(4): 1459–1471.

Yewdell JW, Snyder HL, Bacik I, Anton LC, Deng Y, Behrens TW, Bachi T, Bennink JR (1998) TAP-independent delivery of antigenic peptides to the endoplasmic reticulum: therapeutic potential and insights into TAP-dependent antigen processing. J Immunother 21(2): 127–131.

Zarling AL, Luckey CJ, Marto JA, White FM, Brame CJ, Evans AM, Lehner PJ, Cresswell P, Shabanowitz J, Hunt DF, Engelhard VH (2003) Tapasin is a facilitator, not an editor, of class I MHC peptide binding. J Immunol 171(10): 5287–5295.

Zinkernagel RM, Doherty PC (1974) Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature 248(5450): 701–702.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)