3.16

$$p(\mathbf{t}|\alpha,\beta) = \int p(\mathbf{t}|\mathbf{w},\beta)\,p(\mathbf{w}|\alpha)\,d\mathbf{w}$$

From 3.10,

$$p(\mathbf{t}|\mathbf{w},\beta) = \left(\prod_{n=1}^{N} \mathcal{N}\left(t_n|\mathbf{w}^T\phi(\mathbf{x}_n),\beta^{-1}\right)\right) = \left(\prod_{n=1}^{N} \frac{1}{(2\pi\beta^{-1})^{1/2}} \exp\left\{-\frac{(t_n - \mathbf{w}^T\phi(\mathbf{x}_n))^2}{2\beta^{-1}}\right\}\right)$$

$$= \frac{1}{(2\pi\beta^{-1})^{N/2}} \exp\left\{-\frac{1}{2\beta^{-1}}\left(\sum_{n=1}^{N}(t_n - \mathbf{w}^T\phi(\mathbf{x}_n))^2\right)\right\}$$

$$= \frac{1}{(2\pi\beta^{-1})^{N/2}} \exp\left\{-\frac{1}{2\beta^{-1}}\sum_{n=1}^{N}\left(t_n^2 + (\mathbf{w}^T\phi(\mathbf{x}_n))^2 - 2t_n\mathbf{w}^T\phi(\mathbf{x}_n)\right)\right\}$$

Using intermediate results from exercise 3.7, this simplifies to:

$$= \frac{1}{(2\pi\beta^{-1})^{N/2}} \exp\left\{-\frac{1}{2\beta^{-1}}\left(\mathbf{t}^T\mathbf{t} + \mathbf{w}^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{w} - 2\mathbf{w}^T\mathbf{\Phi}^T\mathbf{t}\right)\right\}$$

$$= \frac{1}{(2\pi\beta^{-1})^{N/2}} \exp\left\{-\frac{1}{2\beta^{-1}}\left((\mathbf{t} - \mathbf{\Phi}\mathbf{w})^T(\mathbf{t} - \mathbf{\Phi}\mathbf{w})\right)\right\}$$

$$= \frac{1}{(2\pi)^{N/2}\ (\beta^{-1})^{N/2}} \exp\left\{-\frac{1}{2}\left((\mathbf{t} - \mathbf{\Phi}\mathbf{w})^T(\beta\mathbf{I})(\mathbf{t} - \mathbf{\Phi}\mathbf{w})\right)\right\}$$

$$= \frac{1}{(2\pi)^{N/2}\ |\beta^{-1}\mathbf{I}|^{1/2}} \exp\left\{-\frac{1}{2}\left((\mathbf{t} - \mathbf{\Phi}\mathbf{w})^T(\beta\mathbf{I})(\mathbf{t} - \mathbf{\Phi}\mathbf{w})\right)\right\}$$

$$= \mathcal{N}(\mathbf{t}|\mathbf{\Phi}\mathbf{w},\beta^{-1}\mathbf{I})$$

From 3.52,

$$p(\mathbf{w}|\alpha) = \mathcal{N}(\mathbf{w}|\mathbf{0},\alpha^{-1}\mathbf{I}) = \frac{1}{(2\pi)^{M/2}|\alpha^{-1}\mathbf{I}|^{1/2}} \exp\left\{-\frac{\mathbf{w}^T(\alpha\mathbf{I})\mathbf{w}}{2}\right\}$$

Applying 2.115,

$$p(\mathbf{t}|\alpha,\beta) = \int \mathcal{N}(\mathbf{t}|\mathbf{\Phi}\mathbf{w}, \beta^{-1}\mathbf{I})\,\mathcal{N}(\mathbf{w}|\mathbf{0}, \alpha^{-1}\mathbf{I})\,d\mathbf{w}$$

$$= \int \mathcal{N}(\mathbf{t}|\mathbf{\Phi}\mathbf{w}, \beta^{-1}\mathbf{I})\,\mathcal{N}(\mathbf{w}|\mathbf{0}, \alpha^{-1}\mathbf{I})\,d\mathbf{w}$$

$$= \mathcal{N}\left(\mathbf{t}|\mathbf{0}, \beta^{-1}\mathbf{I} + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T\right)$$

$$= \frac{1}{(2\pi)^{N/2}|(\beta^{-1}\mathbf{I} + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T)|^{1/2}} \exp\left(-\frac{\mathbf{t}^T(\beta^{-1}\mathbf{I} + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T)^{-1}\mathbf{t}}{2}\right)$$

Taking log, we get:

$$= -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|(\beta^{-1}\mathbf{I} + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T)| - \frac{\mathbf{t}^T(\beta^{-1}\mathbf{I} + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T)^{-1}\mathbf{t}}{2}$$

We simplify $(\beta^{-1}\mathbf{I} + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T)^{-1}$ using The Woodbury Identity (eqn 156 from The Matrix Cookbook), where

$$(\mathbf{A} + \mathbf{C}\mathbf{B}\mathbf{C}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B}^{-1} + \mathbf{C}^T\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{A}^{-1}$$

Here,

$$\mathbf{A} = \beta^{-1}\mathbf{I}$$
$$\mathbf{C} = \mathbf{\Phi}$$
$$\mathbf{B} = \alpha^{-1}\mathbf{I}$$

This gives us:

$$= (\beta^{-1}\mathbf{I})^{-1} - (\beta^{-1}\mathbf{I})^{-1}\mathbf{\Phi}((\alpha^{-1}\mathbf{I})^{-1} + \mathbf{\Phi}^T(\beta^{-1}\mathbf{I})^{-1}\mathbf{\Phi})^{-1}\mathbf{\Phi}^T(\beta^{-1}\mathbf{I})^{-1}$$

$$= (\beta^{-1}\mathbf{I})^{-1} - \beta\mathbf{\Phi}(\alpha\mathbf{I} + \beta\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\beta$$

$$= (\beta^{-1}\mathbf{I})^{-1} - \beta\mathbf{\Phi}\mathbf{A}^{-1}\mathbf{\Phi}^T\beta$$

Substituting this result in $\mathbf{t}^T(\beta^{-1}\mathbf{I} + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T)^{-1}\mathbf{t}$, we get:

$$\mathbf{t}^T \left((\beta^{-1}\mathbf{I})^{-1} - \beta\mathbf{\Phi}\mathbf{A}^{-1}\mathbf{\Phi}^T\beta\right)\mathbf{t}$$

$$= \beta\mathbf{t}^T\mathbf{t} - \beta\mathbf{t}^T\mathbf{\Phi}\mathbf{A}^{-1}\mathbf{\Phi}^T\beta\mathbf{t}$$

$$= \beta\mathbf{t}^T\mathbf{t} - \beta\mathbf{t}^T\mathbf{\Phi}\mathbf{A}^{-1}\mathbf{A}\mathbf{A}^{-1}\mathbf{\Phi}^T\beta\mathbf{t}$$

$$= \beta\mathbf{t}^T\mathbf{t} - (\beta\mathbf{t}^T\mathbf{\Phi}\mathbf{A}^{-1})\mathbf{A}(\mathbf{A}^{-1}\mathbf{\Phi}^T\beta\mathbf{t})$$

$$= \beta\mathbf{t}^T\mathbf{t} - (\beta\mathbf{A}^{-T}\mathbf{\Phi}^T\mathbf{t})^T\mathbf{A}(\beta\mathbf{A}^{-1}\mathbf{\Phi}^T\mathbf{t})$$

$$= \beta\mathbf{t}^T\mathbf{t} - (\beta\mathbf{A}^{-1}\mathbf{\Phi}^T\mathbf{t})^T\mathbf{A}(\beta\mathbf{A}^{-1}\mathbf{\Phi}^T\mathbf{t})$$

$$= \beta\mathbf{t}^T\mathbf{t} - \mathbf{m}_N^T\mathbf{A}\mathbf{m}_N$$

$$= \beta\mathbf{t}^T\mathbf{t} - \mathbf{m}_N^T(\alpha\mathbf{I} + \beta\mathbf{\Phi}^T\mathbf{\Phi})\mathbf{m}_N$$

$$= \beta\mathbf{t}^T\mathbf{t} - \alpha\mathbf{m}_N^T\mathbf{m}_N - \beta\mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N$$

$$= \beta(\mathbf{t}^T\mathbf{t} - \mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N) - \alpha\mathbf{m}_N^T\mathbf{m}_N$$

$$= \beta(\mathbf{t}^T\mathbf{t} + \mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N) - 2\beta\mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N - \alpha\mathbf{m}_N^T\mathbf{m}_N$$

$$= \beta(\mathbf{t}^T\mathbf{t} + \mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N - 2\mathbf{t}^T\mathbf{\Phi}\mathbf{m}_N + 2\mathbf{t}^T\mathbf{\Phi}\mathbf{m}_N) - 2\beta\mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N - \alpha\mathbf{m}_N^T\mathbf{m}_N$$

$$= \beta(\mathbf{t}^T\mathbf{t} + \mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N - 2\mathbf{t}^T\mathbf{\Phi}\mathbf{m}_N) - 2\beta\mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N + 2\beta\mathbf{t}^T\mathbf{\Phi}\mathbf{m}_N - \alpha\mathbf{m}_N^T\mathbf{m}_N$$

$$= \beta||\mathbf{t} - \mathbf{\Phi}\mathbf{m}_N||^2 - 2\beta\mathbf{m}_N^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{m}_N + 2\beta\mathbf{t}^T\mathbf{\Phi}\mathbf{m}_N - \alpha\mathbf{m}_N^T\mathbf{m}_N$$

$$= \beta||\mathbf{t} - \mathbf{\Phi m}_N||^2 - 2\beta \mathbf{m}_N^T \mathbf{\Phi}^T \mathbf{\Phi m}_N + 2\beta \mathbf{m}_N^T \mathbf{\Phi}^T \mathbf{t} - \alpha \mathbf{m}_N^T \mathbf{m}_N$$

$$= \beta||\mathbf{t} - \mathbf{\Phi m}_N||^2 - 2\beta \mathbf{m}_N^T \mathbf{\Phi}^T \mathbf{\Phi m}_N + 2\beta \mathbf{m}_N^T \frac{1}{\beta} \mathbf{A m}_N - \alpha \mathbf{m}_N^T \mathbf{m}_N$$

$$= \beta||\mathbf{t} - \mathbf{\Phi m}_N||^2 - 2\beta \mathbf{m}_N^T \mathbf{\Phi}^T \mathbf{\Phi m}_N + 2\mathbf{m}_N^T \mathbf{A m}_N - \alpha \mathbf{m}_N^T \mathbf{m}_N$$

$$= \beta||\mathbf{t} - \mathbf{\Phi m}_N||^2 - 2\beta \mathbf{m}_N^T \mathbf{\Phi}^T \mathbf{\Phi m}_N + 2\mathbf{m}_N^T \mathbf{A m}_N - \alpha \mathbf{m}_N^T \mathbf{m}_N$$

$$= \beta||\mathbf{t} - \mathbf{\Phi m}_N||^2 - 2\beta \mathbf{m}_N^T \mathbf{\Phi}^T \mathbf{\Phi m}_N + 2\mathbf{m}_N^T (\alpha \mathbf{I} + \beta \mathbf{\Phi}^T \mathbf{\Phi}) \mathbf{m}_N - \alpha \mathbf{m}_N^T \mathbf{m}_N$$

$$= \beta||\mathbf{t} - \mathbf{\Phi m}_N||^2 - 2\beta \mathbf{m}_N^T \mathbf{\Phi}^T \mathbf{\Phi m}_N + \alpha \mathbf{m}_N^T \mathbf{m}_N + 2\beta \mathbf{m}_N^T \mathbf{\Phi}^T \mathbf{\Phi m}_N$$

$$= \beta||\mathbf{t} - \mathbf{\Phi m}_N||^2 + \alpha \mathbf{m}_N^T \mathbf{m}_N$$

Next, we simplify $|\beta^{-1}\mathbf{I} + \alpha^{-1}\mathbf{\Phi \Phi}^T|$:

$$= |\beta^{-1}\mathbf{I}_N + \alpha^{-1}\mathbf{\Phi \Phi}^T|$$

$$= |\beta^{-1}(\mathbf{I}_N + \alpha^{-1}\beta\mathbf{\Phi \Phi}^T)|$$

$$= \beta^{-N}|\mathbf{I}_N + \alpha^{-1}\beta\mathbf{\Phi \Phi}^T|$$

Using C.14 from Appendix C of Bishop, this becomes:

$$= \beta^{-N}|\mathbf{I}_M + \alpha^{-1}\beta\mathbf{\Phi}^T\mathbf{\Phi}|$$

$$= \beta^{-N}|\alpha^{-1}(\alpha\mathbf{I}_M + \beta\mathbf{\Phi}^T\mathbf{\Phi})|$$

$$= \beta^{-N}\alpha^{-M}|\alpha\mathbf{I}_M + \beta\mathbf{\Phi}^T\mathbf{\Phi}|$$

$$= \beta^{-N}\alpha^{-M}|\mathbf{A}|$$

Substituing all the results in the log, it becomes:

$$= -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(\beta^{-N}\alpha^{-M}|\mathbf{A}|) - \frac{(\beta||\mathbf{t} - \mathbf{\Phi}\mathbf{m}_N||^2 + \alpha\mathbf{m}_N^T\mathbf{m}_N)}{2}$$

$$= -\frac{N}{2}\ln(2\pi) + \frac{M}{2}\ln\alpha + \frac{N}{2}\ln\beta - \frac{1}{2}\ln|\mathbf{A}| - \frac{\beta}{2}||\mathbf{t} - \mathbf{\Phi}\mathbf{m}_N||^2 + \frac{\alpha}{2}\mathbf{m}_N^T\mathbf{m}_N$$

which is the same as 3.86.