Table 1. Cyclic Penta Peptides Sequence, Global Minimum Energies and Bio-Activity Data

Training Compound #	Peptide Sequence	Global Minimum Energy[a] kcal / mol	Energy of 25th Conformer[a] kcal / mol	pIC_{50} [b]
FC08	L – Tyr – D – Arg – L – Arg – L – Nal – Gly	– 234.6435	– 232.04	6.260
FC28	L – Tyr – L – Arg – D – Arg – L – Nal – Gly	– 226.4976	– 224.24	4.959
FC48	D – Tyr – L – Arg – D – Arg – D – Nal – Gly	– 230.4091	– 228.20	4.854
FC68	D – Tyr – D – Arg – L – Arg – D – Nal – Gly	– 238.8565	– 235.95	8.810
FC130	L – Tyr – L – Arg – L – Arg – L – Nal – Gly	– 235.9761	– 232.94	7.252
FC131	D – Tyr – L – Arg – L – Arg – L – Nal – Gly	– 241.2344	– 237.04	9.495
FC132	L – Tyr – L – Arg – L – Arg – D – Nal – Gly	– 237.3278	– 232.66	6.268
FC133	D – Tyr – L – Arg – D – Arg – L – Nal – Gly	– 229.5455	– 227.22	6.260
FC134	D – Tyr – L – Arg – L – Arg – D – Nal – Gly	– 240.8301	– 237.90	6.252
FC135	L – Tyr – D – Arg – D – Arg – L – Nal – Gly	– 225.4785	– 223.16	6.268
FC136	L – Tyr – L – Arg – D – Arg – D – Nal – Gly	– 227.8612	– 224.17	4.921
FC137	D – Tyr – D – Arg – D – Arg – L – Nal – Gly	– 227.5789	– 225.76	6.268
FC138	L – Tyr – D – Arg – D – Arg – D – Nal – Gly	– 225.5718	– 223.81	5.268
FC139	D – Tyr – D – Arg – D – Arg – D – Nal – Gly	– 232.1411	– 229.55	5.260
FC91	L – Tyr – D – Arg – L – Arg – D – Nal – Gly	– 234.5048	– 231.71	6.252
FC92	D – Tyr – D – Arg – L – Arg – L – Nal – Gly	– 238.7297	– 234.09	8.979
Test Compound #	**Peptide Sequence**	**Global Minimum Energy[a] kcal / mol**	**Energy of 30th Conformer[a] kcal / mol**	**pIC_{50} [b]**
FC07	L – Arg – D – Tyr – L – Arg – L – Nal – Gly	– 236.2400	– 234.7500	6.301
FC10	L – Nal – D – Arg – L – Arg – L – Tyr – Gly	– 239.7200	– 236.4900	5.301
FC21	L – Tyr – L – Nal – D – Arg – L – Arg – Gly	– 230.4900	– 228.3600	5.268
FC27	L – Arg – L – Tyr – D – Arg – L – Nal – Gly	– 231.8300	– 229.7400	4.854
FC30	L – Nal – L – Arg – D – Arg – L – Tyr – Gly	– 234.0900	– 231.6600	4.886
FC47	D – Arg – L – Tyr – D – Arg – D – Nal – Gly	– 237.0200	– 231.8800	4.824
FC50	D – Nal – L – Arg – D – Arg – D – Tyr – Gly	– 240.6000	– 235.8000	4.921
FC67	D – Arg – D – Tyr – L – Arg – D – Nal – Gly	– 242.1500	– 239.2300	5.26
FC70	D – Nal – D – Arg – L – Arg – D – Tyr – Gly	– 243.6800	– 240.6100	6.26

(a) Computed with GB/SA water solvation in AMBER Force Field.
(b) IC_{50} values based on [^{125}I]SDF-1 binding to CXCR4.

The mol2 files, when read into Sybyl showed incorrect arginine N types, and incorrect amide bond types. These can be edited in Sybyl manually. Again, this repetitive task (for four hundred molecules) was achieved by a python script [43].

2.4 Alignment

It is well established that the success of a 3D-QSAR study is highly dependent on the bioactive conformer selection and alignment rule. The molecular database alignment in Sybyl is effected by "rigid-body rotations and translations only". The cross-validated r^2 (q^2) value is reported to be sensitive to the overall orientation of the superimposed molecules so much so that the value can vary as much as 0.5 q^2 units [19]. We therefore decided to use all of the conformers of the three most active compounds (FC131, FC92 and FC68). Each of these compounds had twenty five conformers, thus, we performed seventy five different molecular database alignments as described below. We chose the central cyclic pentapeptide backbone (fifteen atoms) along with the two hydrogen atoms of the Glycine residue, as the alignment rule for aligning the molecular databases (MDB). We created seventy five Sybyl MDBs, each including the twenty five conformers of each of the sixteen training set compounds. We accomplished this repetitive task by using a shell script [44]. The Sybyl MDB alignment protocol requires three input parameters (the MDB, the template molecule, and the common substructure query). The common substructure query contains the atoms and connectivity information of the substructure common between the template molecule and every molecule in the MDB. We used the seventy five different conformers of the three most active compounds (FC131, FC92 and FC68), as respective templates for aligning each of the respective MDBs. This repetitive task was automated by employing a Sybyl Programming Language (SPL) script [45].

2.5 CoMFA and PLS analyses

CoMFA analysis was performed using Sybyl, employing the Tripos force field to calculate the steric (van der waals) and electrostatic (coulombic) interactions. At all of the grid points, where the steric values exceeded 30 kcal/mol, the actual electrostatic value was replaced by the mean of

the non-excluded electrostatic values. This was done for all the compounds, so that the sterically excluded points would not contribute to the PLS calculations. The steric and electrostatic field cutoff values were set to 30 kcal/mol. Thus, at any grid point if the steric and electrostatic field values exceeded 30 kcal/mol then those values would by replaced by 30 kcal/mol. This ensured that the steric and electrostatic field values would reach, at most, a plateau closer to the center of any atom. The region settings option was set to "Create Automatically". The molecules of the aligned MDBs were placed in a cubic grid of 2 Å spacing, which extended 4 Å beyond every molecule in the training set. Further, the steric and electrostatic field energies would be calculated at all grid points using a sp^3 carbon as the steric probe atom and a +1.0 net charge as the electrostatic probe.

All cross-validated PLS analyses were performed using the Leave-One-Out (CV-LOO) approach. The scaling option was set to CoMFA default. The column filtering value was set to 2.0 kcal/mol. Thus, all of the CoMFA descriptor columns whose energy variance was less than 2.0 kcal/mol were omitted in the PLS calculations. The number of components to explore was set to 6.

2.6 Quasi-multi-way PLS analyses

The multi-way PLS method was developed by Bro *et al.* [46] and has been applied in the 3D-QSAR study of insecticidal neonicotinoid compounds [30]. Each dimension of the multi-way data corresponds to the compounds in training set, CoMFA field variables, conformations, and alignments. The conformers and alignments that gave the best correlation to the observed bioactivities were obtained from the solution of mathematical equations. We have mimicked the multi-way-PLS analyses by performing several sequential two-way PLS analyses. The repetitive task was automated by using a SPL script [47].

2.7 Final model selection and identification of bioactive conformer of test set compounds

The best alignments and best conformer set combination were identified based on the cross-validated q^2 values (>0.667) from the seventy five models. This gave eight candidate 3D-QSAR models (see Table 4) for the final model selection. Clearly, the model that most accurately predicts the bioactivities of the test set compounds would be the best model.

We created eight MDBs with the appropriate conformers of respective training set compounds (see Table 2), and computed the 3D-QSAR models by the non-validated PLS analyses. We then placed thirty conformers (within 1–5 kcal/mol range of the global minimum conformer) of every test set compound, in all of the eight molecular spreadsheets (MSS). To identify the bioactive conformers of the Test set compounds, we performed prediction of the bioactivities for all of the conformers for every test set compound, in all of the eight MSSs (candidate 3D-QSAR models).

3 Results and Discussion

3.1 Conformational search analysis

The Monte Carlo conformational search furnished thousands of unique conformations within a 12 kcal/mol energy range for most of the training set compounds. Out of these, there were several hundreds of conformers within a 5 kcal/mol range of the global minima. The water solvated global minimum energies and energies of the twenty-fifth conformers are presented in Table 1. All of the sixteen training set compound's energies fall within the 5 kcal/mol of their respective global minimas. The energies of conformers selected for the chosen eight 3D-QSAR models (q^2 > 0.667) are presented in Table 2. The global minimum energies and energies of conformers of the test compounds, selected for the 3D-QSAR model validation, are presented in Table 6. The difference in energies fall within the 5 kcal/mol energy range for each of the test set compounds.

3.2 3D-QSAR model development

The first generation of the seventy five MSSs contained twenty five conformers of each of the sixteen compounds (four hundred in total). The first set of PLS analyses was performed on all of the first generation of MSSs. We chose two conformers of every compound (thirty two in total) with the minimum and next-to-minimum residual values to give us the second generation of seventy five MSSs. The second set of PLS analyses was performed on the second generation of the seventy five MSSs. The statistical performance data of the seventy five different alignment-based models obtained is presented in Table 3. All of the seventy five models showed q^2 greater than 0.6. The best model displayed a q^2 of 0.863. The optimal number of components was 6 for sixty four models and 5 for the remaining eleven models. The standard error of prediction showed a maximum value of 1.07, with most of the models displaying values less than 1.0. We chose one conformer for every compound (sixteen in total) that displayed the least residual value of the two conformers. This gave the third generation of seventy five MSSs and the third set of PLS analyses was performed on this generation of seventy five MSSs. The statistical data of performance for all of the third generation models is presented in Table 4. The best q^2 found was 0.760. There were twenty eight models with q^2 greater than 0.6. The optimal number of components was 6 in nineteen models, 5 in eight models, and 3 in one model. The standard error of prediction was at most 1.13 with five models showing values less than 1.0. The fourth generation MSSs and models were selected from

Table 2. Conformers selected for QSAR models, their energies and performance of these models on Training data set.

Training Compound #	Actual Bio-Activity pIC_{50}	131_0			131_5			131_7			131_20		
		Conf. #	Energy[a] kcal / mol	Predicted Bio-Activity pIC_{50}	Conf. #	Energy[a] kcal / mol	Predicted Bio-Activity pIC_{50}	Conf. #	Energy[a] kcal / mol	Predicted Bio-Activity pIC_{50}	Conf. #	Energy[a] kcal / mol	Predicted Bio-Activity pIC_{50}
FC08	6.260	14	−232.65	6.215	14	−232.65	6.259	0	−234.79	6.255	14	−232.65	6.256
FC28	4.959	18	−224.50	4.965	18	−224.50	4.951	10	−224.78	4.932	4	−226.07	4.978
FC48	4.854	18	−228.39	4.894	15	−228.53	4.816	1	−229.96	4.854	1	−229.96	4.769
FC68	8.810	19	−236.21	8.811	2	−238.12	8.823	15	−236.53	8.774	15	−236.53	8.815
FC130	7.252	15	−234.02	7.242	5	−234.86	7.235	1	−235.83	7.228	3	−235.33	7.309
FC131	9.495	14	−237.43	9.473	11	−237.96	9.521	8	−238.44	9.432	11	−237.96	9.488
FC132	6.268	21	−232.86	6.256	9	−234.21	6.277	9	−234.21	6.286	8	−234.24	6.285
FC133	6.260	19	−227.54	6.250	14	−227.41	6.249	21	−227.30	6.257	12	−227.52	6.274
FC134	6.252	21	−238.15	6.311	18	−238.47	6.313	24	−237.90	6.214	21	−238.15	6.306
FC135	6.268	12	−223.90	6.285	12	−223.90	6.298	8	−224.19	6.221	4	−224.48	6.315
FC136	4.921	10	−225.19	4.877	16	−224.37	4.918	20	−224.34	4.987	10	−225.19	4.894
FC137	6.268	7	−226.63	6.268	10	−226.37	6.274	11	−226.23	6.283	8	−226.56	6.314
FC138	5.268	17	−224.38	5.259	23	−223.77	5.255	1	−225.36	5.264	17	−224.38	5.280
FC139	5.260	19	−229.74	5.274	19	−229.74	5.277	19	−229.74	5.269	15	−230.09	5.232
FC91	6.252	17	−232.24	6.238	10	−233.06	6.245	5	−233.70	6.262	5	−233.70	6.214
FC92	8.979	8	−235.00	9.007	9	−235.04	8.915	8	−235.00	9.107	8	−235.00	8.897

Training Compound #	Actual Bio-Activity pIC_{50}	68_19			68_2			92_23			92_7		
		Conf. #	Energy[a] kcal / mol	Predicted Bio-Activity pIC_{50}	Conf. #	Energy[a] kcal / mol	Predicted Bio-Activity pIC_{50}	Conf. #	Energy[a] kcal / mol	Predicted Bio-Activity pIC_{50}	Conf. #	Energy[a] kcal / mol	Predicted Bio-Activity pIC_{50}
FC08	6.260	0	−234.79	6.261	14	−232.65	6.280	14	−232.65	6.233	14	−232.65	6.239
FC28	4.959	18	−224.50	4.939	18	−224.50	4.878	10	−224.78	5.002	18	−224.50	4.947
FC48	4.854	18	−228.39	4.892	8	−228.89	4.910	12	−228.65	4.793	12	−228.65	4.866
FC68	8.810	21	−236.06	8.806	21	−236.06	8.777	15	−236.53	8.792	19	−236.21	8.795
FC130	7.252	7	−234.65	7.218	11	−234.70	7.333	7	−234.65	7.303	5	−234.86	7.269
FC131	9.495	11	−237.96	9.553	12	−237.69	9.399	14	−237.43	9.447	14	−237.43	9.416
FC132	6.268	13	−233.57	6.275	8	−234.24	6.283	8	−234.24	6.238	3	−235.23	6.258
FC133	6.260	8	−227.70	6.252	13	−227.42	6.263	2	−228.49	6.199	13	−227.42	6.295
FC134	6.252	4	−240.47	6.207	21	−238.15	6.141	24	−237.90	6.231	9	−239.94	6.231
FC135	6.268	2	−224.66	6.301	12	−223.90	6.218	2	−224.66	6.282	4	−224.48	6.281
FC136	4.921	21	−224.31	4.912	9	−225.58	4.842	21	−224.31	4.886	10	−225.19	4.908
FC137	6.268	11	−226.23	6.226	11	−226.23	6.293	11	−226.23	6.301	7	−226.63	6.237
FC138	5.268	1	−225.36	5.225	1	−225.36	5.351	23	−223.77	5.295	15	−224.58	5.268
FC139	5.260	21	−229.61	5.315	19	−229.74	5.281	15	−230.09	5.296	19	−229.74	5.268
FC91	6.252	8	−233.22	6.316	0	−234.62	6.317	10	−233.06	6.310	0	−234.62	6.276
FC92	8.979	8	−235.00	8.928	0	−238.60	9.059	8	−235.00	9.020	8	−235.00	9.071

Table 3. Statistical data of performance of different alignments for conformers with Minimum and Next to Minimum Residual Values

Alignment Name	q^2	ONC	S. E. in prediction	Alignment Name	q^2	ONC	S. E. in prediction	Alignment Name	q^2	ONC	S. E. in prediction
68_0	0.814	6	0.826511	92_0	0.723	6	0.987666	131_0	0.846	5	0.682582
68_1	0.761	6	0.936638	92_1	0.717	6	0.966680	131_1	0.723	6	0.860404
68_2	0.829	6	0.786740	92_2	0.798	5	0.888171	131_2	0.785	6	0.891423
68_3	0.734	6	1.020120	92_3	0.863	6	0.821501	131_3	0.801	6	0.889737
68_4	0.769	6	0.888725	92_4	0.697	6	1.002310	131_4	0.808	5	0.800617
68_5	0.752	6	0.914881	92_5	0.759	6	0.881225	131_5	0.815	6	0.706057
68_6	0.788	6	0.809620	92_6	0.775	6	0.942933	131_6	0.828	6	0.787380
68_7	0.804	6	0.779613	92_7	0.718	6	0.959982	131_7	0.735	6	0.930725
68_8	0.724	6	0.978718	92_8	0.806	6	0.941062	131_8	0.692	6	0.955353
68_9	0.711	6	0.985897	92_9	0.731	6	0.964540	131_9	0.777	6	0.908828
68_10	0.783	6	0.899782	92_10	0.799	6	0.875709	131_10	0.753	6	0.905429
68_11	0.767	6	0.928244	92_11	0.690	6	1.003070	131_11	0.777	5	0.876025
68_12	0.703	6	0.954935	92_12	0.830	6	0.885722	131_12	0.747	6	0.943408
68_13	0.767	6	0.886365	92_13	0.823	6	0.798505	131_13	0.653	6	1.031050
68_14	0.788	6	0.914451	92_14	0.758	6	0.900605	131_14	0.732	6	1.033780
68_15	0.865	6	0.802629	92_15	0.742	5	0.982651	131_15	0.736	6	0.868914
68_16	0.769	6	0.919262	92_16	0.685	6	1.076530	131_16	0.832	6	0.765060
68_17	0.747	5	0.892846	92_17	0.727	5	0.923361	131_17	0.758	6	0.907051
68_18	0.787	5	0.759792	92_18	0.765	6	0.890769	131_18	0.745	6	0.971360
68_19	0.769	5	0.828032	92_19	0.727	6	0.958452	131_19	0.728	6	0.906786
68_20	0.840	5	0.749218	92_20	0.789	6	0.859767	131_20	0.796	6	0.870979
68_21	0.800	6	0.820461	92_21	0.767	6	0.872157	131_21	0.810	6	0.723227
68_22	0.693	6	1.020230	92_22	0.780	5	0.820813	131_22	0.799	6	0.800829
68_23	0.752	6	0.916167	92_23	0.825	6	0.832032	131_23	0.711	6	0.993310
68_24	0.670	6	0.953167	92_24	0.690	6	0.961364	131_24	0.831	6	0.733437

ONC: Optimal number of components

the third generation models based on the cutoff q^2 value of 0.667. This gave us eight final candidate 3D-QSAR models. The eight final candidate models were based on the alignments done using the first (131_0), sixth (131_5), eighth (131_7), twenty-first (131_20) conformers of FC131, the third (68_2), twentieth (68_19) conformers of FC68, and the eighth (92_7), twenty-fourth (92_23) conformers of FC92 compounds. The conformer numbers of each of the sixteen compounds for each of the candidate models along with their energies and predicted activities are presented in Table 2. The final eight fifth generation MSSs were built with the selected conformer of each of the sixteen compounds and aligned based on the eight selected conformers (131_0, 131_5, 131_7, 131_20, 68_2, 68_19, 92_7 and 92_23). The cross-validated (CV-LOO) PLS analyses gave slightly varied statistical results, which are presented in Table 5. The q^2 values vary from 0.644 to 0.734, whereas the r^2 values vary from 0.998 to 1.0. The cross-validated standard error in prediction values range from 0.911 to 1.068, while the non-validated standard error in prediction values range from 0.036 to 0.082. The optimal number of components displayed is 6 for five models and 5 for three models. The F value, which is the ratio of r^2 to $(1.0-r^2)$, range from 954 to 4219. The predictive r^2, based on the prediction of nine test set compounds, range from 0.829 to 0.990. The predicted vs. observed bioactivities plot for the most predictive models (131_7, 68_2, and 92_7) are presented below Table 5. The prediction data of the nine test compounds for each of the selected 3D-QSAR models is presented in Table 6. The predicted vs. observed bioactivities plot for the nine test compounds in the most predictive models 131_7, 68_2, and 92_7 are presented below Table 6. On analyzing Table 5 and Table 6, it was evident that the 3D-QSAR model based on the alignment with the template molecule of compound FC68's third conformer (numbering of conformer starts with 0) from the global minimum conformer (which is 0.74 kcal/mol higher than the global minimum) furnished the most predictive 3D-QSAR model with q^2 of 0.734, and r^2 of 0.998.

3.3 CoMFA field interpretation

The alignment of all the sixteen compounds based on the alignment with template 68_2 is presented in Figure 1. The CoMFA field view based on electrostatics contribution of 80% for positive charge favored region (Blue), 20% for negative charge favored region (Red), and steric contribution of 80% for sterically favored region (green), 20% sterically disfavored region (yellow) is presented in Figure 2. The placement of most active analogs, FC131, FC92, and FC68 inside the CoMFA field shows that the positively charged arginines of all the compounds contribute to the bioactivity enhancement (Blue regions), whereas the bulky napthylalanine's napthyl ring contributes to enhanced

Table 4. Statistical data of performance of different alignments for conformers with LEAST Residual values of the QSAR models derived from the parent[a] QSAR models

Alignment Name	q^2	ONC[b]	S. E. in prediction	Alignment Name	q^2	ONC[b]	S. E. in prediction	Alignment Name	q^2	ONC[b]	S. E. in prediction
68_0	0.616	5	1.03669	92_0	0.342	3	1.289110	131_0	0.667	6	0.961062
68_1	0.562	6	1.132320	92_1	0.533	6	1.171450	131_1	0.635	6	1.093220
68_2	0.760	5	0.882714	92_2	0.222	6	1.450630	131_2	0.441	6	1.305870
68_3	0.497	6	1.291480	92_3	0.604	6	1.126800	131_3	0.324	6	1.446520
68_4	0.406	6	1.301930	92_4	0.497	6	1.255080	131_4	0.594	5	1.147870
68_5	0.642	6	1.042530	92_5	0.466	6	1.224660	131_5	0.732	6	0.920177
68_6	0.565	6	1.183300	92_6	0.664	6	1.047010	131_6	0.644	6	1.089570
68_7	0.660	6	1.030750	92_7	0.672	6	1.003900	131_7	0.706	6	0.995746
68_8	0.569	6	1.151640	92_8	0.487	5	1.352660	131_8	0.570	6	1.097350
68_9	0.544	6	1.187140	92_9	0.579	6	1.211430	131_9	0.417	6	1.265830
68_10	0.562	6	1.203240	92_10	0.539	6	1.228100	131_10	0.330	6	1.401430
68_11	0.519	6	1.293210	92_11	0.429	6	1.328670	131_11	0.432	6	1.328550
68_12	0.446	6	1.307170	92_12	0.569	5	1.137540	131_12	0.637	5	1.134420
68_13	0.648	6	1.070200	92_13	0.662	3	1.049500	131_13	0.607	5	1.082350
68_14	0.462	6	1.257070	92_14	0.625	6	1.062890	131_14	0.513	5	1.173510
68_15	0.576	6	1.080930	92_15	0.555	4	1.155510	131_15	0.13	6	1.534170
68_16	0.512	6	1.155140	92_16	0.484	6	1.325270	131_16	0.568	5	1.116240
68_17	0.646	6	1.070550	92_17	0.376	6	1.383080	131_17	0.481	6	1.374460
68_18	0.639	6	1.070720	92_18	0.560	6	1.143740	131_18	0.552	6	1.310520
68_19	0.687	5	1.002420	92_19	0.510	6	1.262410	131_19	0.558	6	1.206870
68_20	0.601	6	1.065120	92_20	0.540	5	1.235760	131_20	0.679	5	1.017180
68_21	0.630	6	1.068710	92_21	0.508	6	1.240920	131_21	0.614	5	1.103410
68_22	0.649	6	1.109490	92_22	0.609	6	1.090620	131_22	0.515	6	1.226910
68_23	0.438	5	1.270240	92_23	0.741	6	0.908023	131_23	0.469	4	1.237370
68_24	0.599	5	1.086680	92_24	0.428	6	1.307800	131_24	0.464	4	1.230530

(a) Parent QSAR Models are presented in Table 3.
(b) ONC: Optimal number of components

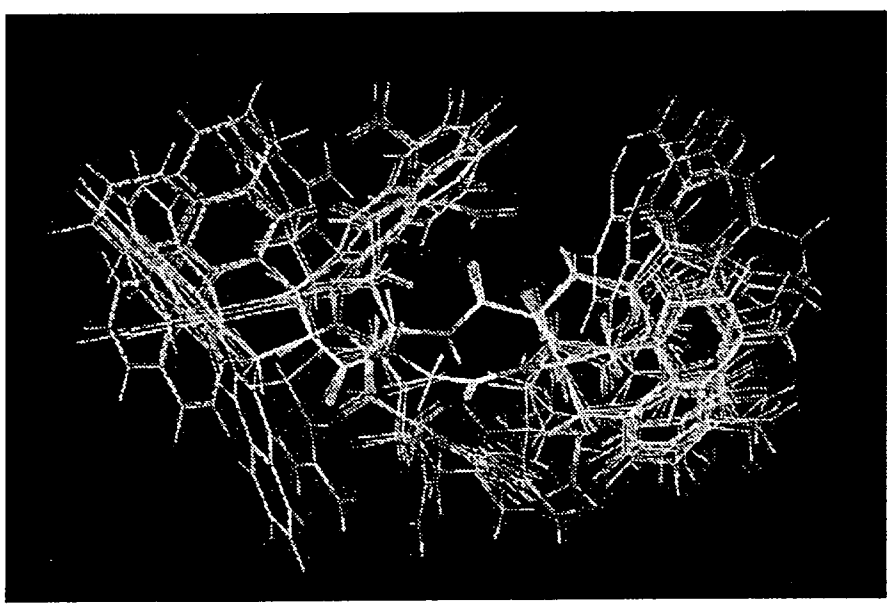
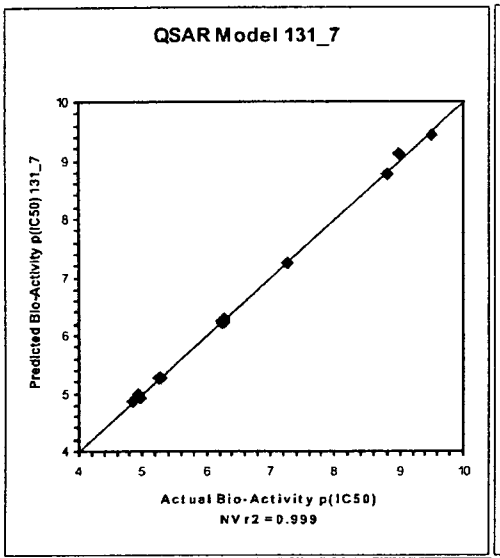


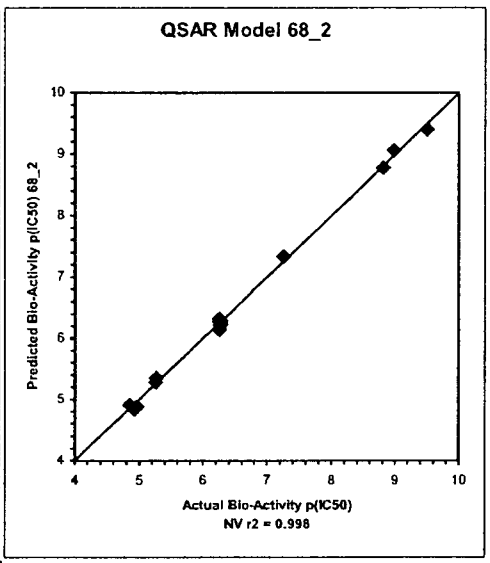
Figure 1. Alignment of the training set compounds along with the template molecule 68_2.

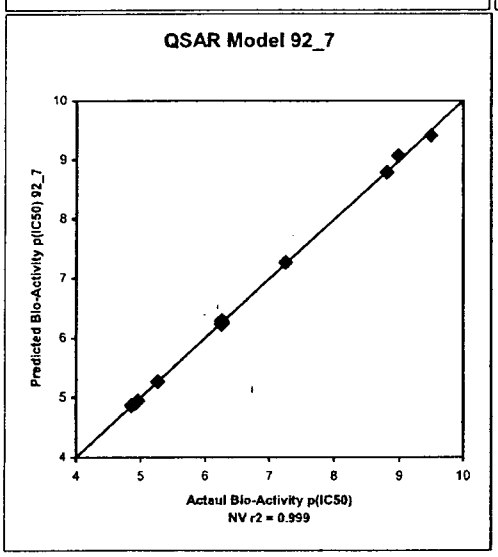
activity for FC131 and FC92 (green region). The reduced activity of FC68 could be attributed to the fact that the napthylalanine's napthyl ring gets placed outside the sterically favored region. The aromatic ring of the tyrosine residue of compounds FC131 and FC92 falls in the sterically favored region (green), while that of FC68 falls outside

Table 5. Statistical Data of selected eight QSAR Models

Model #	q^2	S_{cv}	ONC	NV r^2	S	F	Predictive r^2
131_0	0.691	0.938031	6	1.000	0.036	4219.058	0.900
131_5	0.665	1.033870	6	1.000	0.037	3956.214	0.829
131_7	0.714	0.992176	6	0.999	0.059	1521.240	0.951
131_20	0.680	1.041090	5	0.999	0.053	2242.042	0.928
68_19	0.644	1.068000	5	0.999	0.048	2775.759	0.918
68_2	0.734	0.914358	5	0.998	0.082	954.275	0.990
92_7	0.657	1.017910	6	0.999	0.047	2433.821	0.956
92_23	0.730	0.911822	6	0.999	0.054	1836.001	0.907



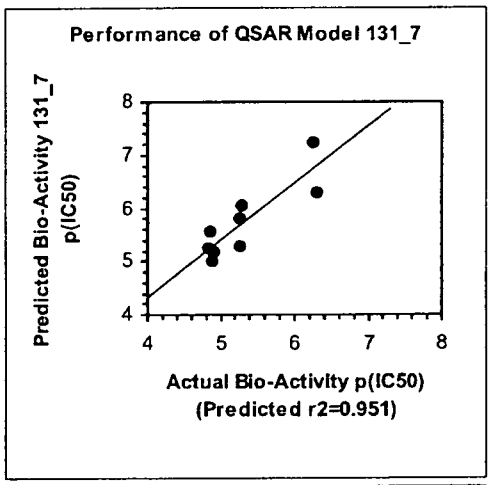


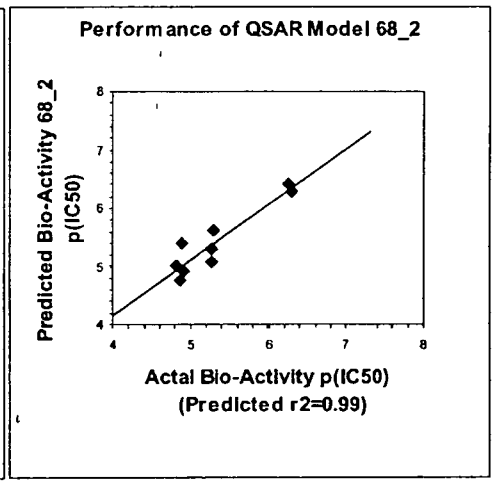


ONC : Optimal Number of Components.
Scv : Cross validated Standard Error of Prediction.
NV r2 : Non-Validated r2.
S : Standard Error of Prediction.
F : F-ratio is the ratio of r2 to (1.0−r2).

Table 6. Predictive Performance Of Data Of QSAR Models

Test Cmpd #	Conf #	Global Minimum Energy[a] kcal / mol	Energy of Conf.[a] kcal / mol	Actual Bio-Activity pIC_{50}	Calculated / Predicted Bio-Activity in various models							
					131_0	**131_5**	**131_7**	**131_20**	**68_19**	**68_2**	**92_23**	**92_7**
FC07	25	−236.24	−234.81	6.301	7.024	7.604	6.296	6.798	6.843	6.29	7.254	6.71
FC10	25	−239.72	−236.68	5.301	6.139	6.24	6.065	6.14	6.46	5.63	5.987	6.02
FC21	25	−230.49	−228.36	5.268	6.378	5.959	5.285	5.558	5.941	5.31	5.649	6.05
FC27	10	−231.83	−230.43	4.854	5.34	5.583	5.579	5.372	5.606	4.78	5.386	5.27
FC30	21	−234.09	−232.29	4.886	5.896	6.117	5.018	5.22	5.598	5.4	5.247	5.51
FC47	21	−237.02	−232.05	4.824	5.194	5.899	5.236	5.547	4.849	5.02	5.57	4.94
FC50	10	−240.60	−236.64	4.921	5.749	5.859	5.183	5.046	5.639	4.92	5.199	5.37
FC67	18	−242.15	−239.81	5.26	5.466	6.151	5.814	5.965	5.685	5.08	5.716	5.23
FC70	2	−243.68	−243.66	6.26	7.174	7.407	7.247	7.454	7.001	6.42	7.757	6.79





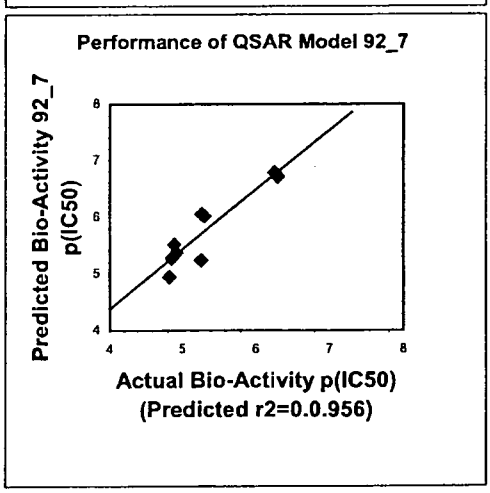


Figure 2. The CoMFA contours of the QSAR model based on 68_2 along with FC131, FC92, and FC68.

that region. The hydroxyl group of the tyrosine residue of FC131 falls in the negative-charge favored region (Red), whereas the tyrosine hydroxyl group of FC92 falls a little away from the negative-charge favored region. Since the side chain of the tyrosine reside of FC68 falls outside of the sterically favored region, the hydroxyl group gets positioned away from the negative-charge favored region (Red). The other negative-charge favored (Red), steric favored (green), and steric disfavored (yellow) regions where no part of FC131, FC92 or FC68 occupy, describe the bioactivities of the other moderately active compounds of the training set.

3.4 Discussion

Our approach of quasi-4D-QSAR by incorporating several conformers and several alignments explicitly yields good, highly predictive QSAR models. The automated method of performing several two-way PLS analyses to identify the best conformers and the best alignment that gives the highest correlation between the CoMFA descriptors and bioactivity is demonstratively simple and efficient. This approximates the more formal 4D-QSAR and multi-way PLS analyses protocols quite effectively.

The observations made in this 3D-QSAR study are in agreement with those reported [47] in that, there is a correspondence between the total positive charge and anti-HIV activity of T140, the lead template for the cyclic pentapeptides used in this study. Our present CoMFA-QSAR study strongly suggests that the functional groups interacting with the CXCR4 receptor are from the side chains of the cyclic pentapeptides. From our 3D-QSAR study we conclude that 1) the strongest interactions for activity are contributed by the positioning of the positively charged arginine residues, 2) the intermediate interactions for binding affinity arise from the hydroxyl group of tyrosine residues' interaction with the receptor, and 3) the moderate interactions that augment the binding affinity come from the steric contributions of the napthyl ring on the napthylalanine residue and the benzene ring on the tyrosine residues.

4 Conclusion

We demonstrate the practical application of Macromodel and Sybyl for automated quasi-multi-way PLS analyses. Several highly predictive 3D-QSAR-candidate models have been developed. The automated approach is shown to furnish good 3D-QSAR models efficiently.

The 3D-QSAR model for these cyclic pentapeptide CXCR4 inhibitors corroborates the observation of the necessity of the positively charged arginine residues for the inhibition of CXCR4 activity by the cyclic pentapeptides. The electrostatic contributions of the hydroxyl group on tyrosine residue and the steric contribution of the napthylalanine and tyrosine residue are also found to contribute to the enhanced activity.

References

[1] H. K. Deng, R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. DiMarzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, N. R. Landau, *Nature* **1996**, *381*, 661 – 666.

[2] Y. Feng, C. C. Broder, P. E. Kennedy, E. A. Berger, *Science* **1996**, *272*, 872 – 877.

[3] A. Muller, B. Homey, H. Soto, N. Ge, D. Catron, M. E. Buchanan, T. McClanahan, E. Murphy, W. Yuan, S. N. Wagner, J. L. Barrera, A. Mohar, E. Verastegui, A. Zlotnik, *Nature* **2001**, *410*, 50 – 56.

[4] P. Matthys, S. Hatse, K. Vermeire, A. Wuyts, G. Bridger, G. W. Henson, E. De Clercq, A. Billiau, D. Schols, *J. Immunol.* **2001**, *167*, 4686 – 4692.

[5] M. Masuda, H. Nakashima, T. Ueda, H. Naba, R. Ikoma, A. Otaka, Y. Terakawa, H. Tamamura, T. Ibuka, T. Murakami, Y. Koyanagi, M. Waki, A. Matsumoto, N. Yamamoto, S. Funakoshi, N. Fujii, *Biochem. Biophys. Res. Commun.* **1992**, *189*, 845 – 850.

[6] H. Tamamura, Y. Xu, T. Hattori, X. Zhang, R. Arakaki, K. Kanbara, A. Omagari, A. Otaka, T. Ibuka, N. Yamamoto, H. Nakashima, N. Fujii, *Biochem. Biophys. Res. Commun.* **1998**, *253*, 877 – 882.

[7] N. Fujii, S. Oishi, K. Hiramatsu, T. Araki, S. Ueda, H. Tamamura, A. Otaka, S. Kusano, S. Terakubo, H. Nakashima, J. A. Broach, J. O. Trent, Z. Wang, S. C. Peiper, *Angew. Chem. Int. Ed.* **2003**, *42*, 3251 – 3253.

[8] C. A. Hansch, *Accts. Chem. Res.* **1969**, *2*, 232 – 239.

[9] R. D. Cramer III, D. E. Patterson, J. D. Bunce, *J. Am. Chem. Soc.* **1988**, *110*, 5959 – 5967.

[10] A. J. Hopfinger, B. J. Burke, *Molecular Shape Analysis: a formalism to quantitatively establish spatial molecular similarity in: Concepts and Applications of Molecular Similarity*, G. M. Maggiora (Ed.), Wiley and Sons, New York, **1990**, p. 173 – 209.

[11] S. Srivastava, Richardson, W. W., Bradley, M. P., Crippen, G. M., *Three-dimensional receptor modeling using distance geometry and Voronoi polyhedra*, ESCOM, Leiden, The Netherlands, **1993.**

[12] A. C. Good, S. J. Peterson, W. G. Richards, *J. Med. Chem.* **1993**, *36*, 2929 – 2937.

[13] G. Klebe, *Prospect. Drug Disc. Design* **1998**, *12*, 87 – 104.

[14] A. N. Jain, T. G. Dietterich, R. H. Lathrop, D. Chapman, R. E. Critchlow, B. E. Bauer, T. A. Webster, T. Lozanoperez, *J. Comput.-Aided Mol. Des.* **1194**, *8*, 635 – 652.

[15] A. M. Doweyko, *J. Math. Chem.* **1991**, *7*, 273 – 285.

[16] G. Folkers, A. Merz, D. Rognan, *CoMFA: scope and limitation in: 3D QSAR in Drug Design*, H. Kubinyi (Ed.), ESCOM, Leiden, The Netherlands, **1993**, pp. 583 – 618.

[17] K. H. Kim, *Comparative molecular field analysis (CoMFA) in: Molecular Similarity in Drug Design*, P. M. Dean (Ed.), Blackie Academic, Glasgow, UK, **1995**, pp. 291 – 331.

[18] G. Folkers, A. Merz, D. Rognan, *CoMFA as a tool for active site modeling in: Trends in QSAR Molecular Modeling 92*, ESCOM, Leiden, The Netherlands, **1993**, pp. 233 – 244.

[19] S. J. Cho, A. Tropsha, *J. Med. Chem* **1995**, *38*, 1060 – 1066.

[20] A. J. Hopfinger, S. Wang, J. S. Tokarski, B. Jin, M. Albuquerque, P. J. Madhav, C. Duraiswami, *J. Am. Chem. Soc.* **1997**, *119*, 10509 – 10524.

[21] K. Hasegawa, M. Arakawa, K. Funatsu, *Comp. Bio. Chem.* **2003**, *27*, 211 – 216.

[22] A. Vedani, D. R. McMasters, M. Dobler, *Quant. Struct.-Act. Relat.* **2000**, *19*, 149 – 161.

[23] M. Appell, W. J. Dunn III, M. E.A. Reith, L. Miller, J. L. Flippen-Anderson, *Bioorg. Med. Chem.* **2002**, *10*, 1197 – 1206.

[24] K. Hasegawa, M. Arakawa, K. Funatsu, K., *Chemom. Intell. Lab. Syst.* **2000**, *50*, 253 – 261.

[25] N. E. Jewell, D. B. Turner, P. Willett, G. J. Sexton, *J. Mol. Graph. Mod.* **2001**, *20*, 111 – 121.

[26] A. J. Tervo, T. H. Nyronen, T. Ronkko, A. Poso, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 807 – 816.

[27] W. Sippl. H.-D. Höltje, *J. Mol. Struct.* **2000**, *503*, 31 – 50.

[28] P. M. Kroonenberg, W. J. Dunn III, J. J.F. Commandeur, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 2025 – 2032.

[29] D. D. Robinson, P. D. Lyne, W. G. Richards, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 503 – 512.

[30] K. Hasegawa, M. Arakawa, K. Funatsu, *Chemom. Intell. Lab. Syst.* **1999**, *47*, 33 – 40.

[31] 1500 S. W. First Avenue, Suite 1180 Portland, OR 97201.

[32] 1699 Hanley Road St. Louis, MO USA.

[33] J. O. Trent, K. Napier, A. N. Lane (private communication; unpublished 3D coordinates of FC131 were determined by NMR).

[34] S. Guccione, A. M. Doweyko, H. Chen, G. U. Barretta, F. Balzano, *J. Comput.-Aided Mol. Des.* **2000**, *14*, 647 – 657.

[35] R. E. Taylor, J. Zajicek, *J. Org. Chem.* **1999**, *64*, 7224 – 7228.

[36] D. M. Ferguson, P. A. Kollman, *J. Comput. Chem.* **1991**, *12*, 620 – 626.

[37] E. Polak, G. Ribiere, *Rev. Fr. Inform. Rech. O.* **1969**, *16*, 35 – 45.

[38] G. Chang, W. C. Guida, W. C. Still, *J. Am. Chem. Soc.* **1989**, *111*, 4379 – 4386.

[39] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127 – 6129.

[40] J. B. Bhonsle, J. O. Trent, *Available upon request.*

[41] P. Walters, M. Stahl, Dolata Research Group, Department of Chemistry, University of Arizona, Tucson, AZ 85721, **1994.**

[42] J. B. Bhonsle, J. O. Trent, *Available upon request.*

[43] J. B. Bhonsle, J. O. Trent, *Available upon request.*

[44] J. B. Bhonsle, J. O. Trent, *Available upon request.*

[45] J. B. Bhonsle, J. O. Trent, *Available upon request.*

[46] R. Bro, *J. Chemom.* **1996**, *10*, 47 – 61.

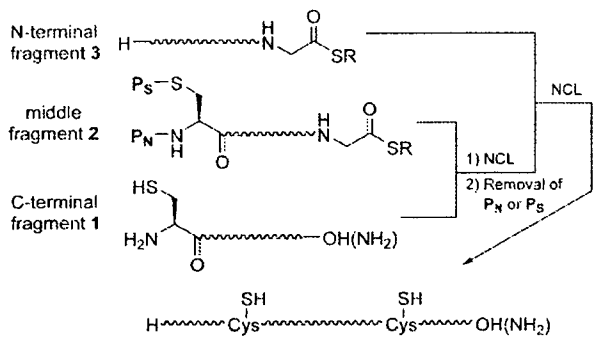
[47] H. Tamamura, A. Omagari, S. Oishi, T. Kanamoto, N. Yamamoto, S. C. Peiper, H. Nakashima, A. Otaka, N. Fujii, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2633 – 2637.

DOI: 10.1002/cbic.200500272

Photolabile Protection for One-Pot Sequential Native Chemical Ligation

Satoshi Ueda,[a] Mizuno Fujita,[a] Hirokazu Tamamura,[a] Nobutaka Fujii,*[a] and Akira Otaka*[a, b]

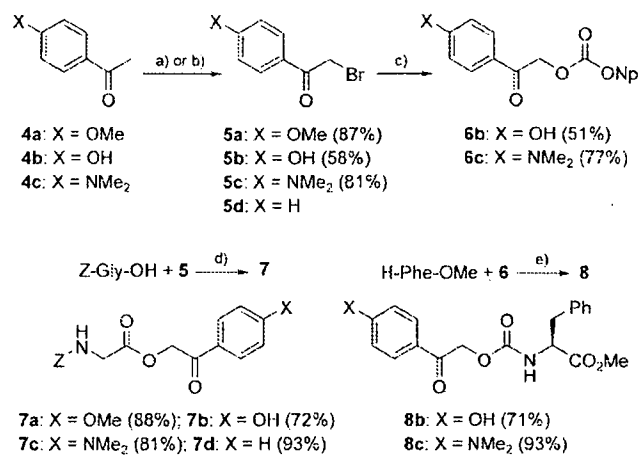
The strategy of native chemical ligation (NCL) has widespread application in the chemical synthesis of proteins.[1] This methodology features chemoselective amide formation between a thioester and an N-terminal cysteine through attack of the Cys sulfhydryl group on the thioester followed by intramolecular S-N acyl migration. Sequential NCL with more than one thioester fragment could potentially be an attractive method for constructing protein molecules (Scheme 1).[2,3] In this strategy, protection of the amino and/or sulfhydryl groups of the N-terminal Cys residue of the middle fragment **2** is needed to avoid the formation of undesired products such as cyclic peptides.[4] Recently, Kent's group reported an elegant one-pot sequential NCL approach for peptide synthesis in a purification-free manner using a thiazolidine derivative as N/S-protected Cys, which regenerates the Cys residues by the action of methoxyamine·hydrochloride.[5] The use of photolabile protecting groups in conjunction with acid- or base-sensitive groups provides a powerful means for selective manipulation of functional groups in synthetic chemistry.[6] Thus, we envisioned that a photolabile moiety should also serve as a useful protection for either the N^{α}-amino or the sulfhydryl group in a sequential NCL strategy, in which simple photoirradiation would probably allow consecutive NCL steps to be conducted in one-pot. Here we report the development of novel phenacyl-type photolabile protecting groups and their application to a purification-free one-pot sequential NCL strategy.



Scheme 1. Synthetic strategy for the preparation of peptides/proteins by using sequential native chemical ligation (NCL). Protection of the amino (P_N) or sulfhydryl (P_S) group in fragment **2** is necessary for the sequential NCL strategy.

Among several photoremovable groups, *o*-nitrobenzyl-type protecting groups have been shown to be useful in peptide chemistry.[7] However, such photodeprotection on an N-terminal Cys residue proceeds by a Norrish type II mechanism to release the N/S-unprotected Cys and an aldehyde, which then recombine to form an imine or thiazolidine moiety that fails to show the desired reactivity in the following NCL step. Photolabile phenacyl-type protection is a potential alternative to using nitrobenzyl groups due to their photosolvolytic character. Phenacyl-type groups with methoxy[8,9] or hydroxyl[10] substituents have received attention as photosensitive groups, in which the electron-donating substituents are responsible for the rapid release of parent molecules. This prompted us to examine the feasibility of -CO_2H or -NH_2 protection of phenacyl-type groups with a more electron-donating *p*-dimethylamino substituent and the applicability of such protecting groups to one-pot sequential NCL strategies.

First, we evaluated the usefulness of 4-(dimethylamino)phenacyl (Map) esters. The requisite carboxylates **7** were prepared as shown in Scheme 2. Photolysis of **7** was conducted in EtOH by using a 100 W high-pressure Hg lamp ($h\nu > 300$ nm) at a concentration of 1 mM. Deprotection profiles of **7** are summarized in Table 1. Photolysis of the novel Map ester **7c** proceeded efficiently to complete the release of Z-Gly-OH at a rate two times faster than 4-hydroxyphenacyl ester **7b**, which has great potential as a "photo-cage". Although the ester **7c** did not show sufficient stability toward 20% piperidine in DMF, this



Scheme 2. Syntheses of protecting agents **5** and **6** and protected amino acid derivatives **7** and **8**: a) $CuBr_2$, EtOAc, reflux for **5a** and **5b**; b) Br_2, conc. H_2SO_4, then $(EtO)_2P(O)H$, Et_3N for **5c**; c) HCO_2Na, EtOH, reflux, then *p*-nitrophenylchloroformate, pyridine, CH_2Cl_2; d) 1,8-diazabicyclo[5.4.0]undec-7-ene, benzene; e) Et_3N, DMF for **8b**; Et_3N, pyridine for **8c**.

[a] *S. Ueda, M. Fujita, Dr. H. Tamamura, Prof. N. Fujii, Prof. A. Otaka*
Graduate School of Pharmaceutical Sciences, Kyoto University
Sakyo-ku, Kyoto 606-8501 (Japan)
Fax: (+81) 75-753-4570
E-mail: nfujii@pharm.kyoto-u.ac.jp

[b] *Prof. A. Otaka*
Current address:
Graduate School of Pharmaceutical Sciences, The University of Tokushima
1-78-1 Shomachi, Tokushima 770-8505 (Japan)
Fax: (+81) 88-633-9505
E-mail: aotaka@ph.tokushima-u.ac.jp

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

Table 1. Photoinduced-deprotection of protected amino acid derivatives (**7** and **8**)

	Substrate (*p*-substituent)	Photolysis [*t*/min][a]	Yield [%][b]	$t^{1/2}$ [min][c]
1	**7a** (OMe)	90	22	n.d.[d]
2	**7b** (OH)	30	96	11
3	**7c** (NMe_2)	7	97	5.3
4	**7d** (H)	90	11	n.d.[d]
5	**8b** (OH)	70	93	19
6	**8c** (NMe_2)	20	95	7.5

[a] Each sample solution in EtOH (1 mM) was irradiated with a 100 W high-pressure Hg lamp ($h\nu > 300$ nm) at room temperature. [b] Determined by measurements of regenerated amino acid derivatives (Z-Gly-OH for **7** and H-Phe-OMe for **8**) by using reversed-phase HPLC analyses. [c] Determined by the disappearance of **7** or **8** by using HPLC analyses. [d] Not determined.

group could be used in the design of caged compounds as well as **7b**.

Next, we investigated the feasibility of the 4-(dimethylamino)phenacyloxycarbonyl (Mapoc) moiety for the protection of amines. To our knowledge, carbamates including the 4-hydroxyphenacyl unit have yet to be examined well for this purpose.[11] Phenylalanine-derived carbamates (**8b** and **8c**) were prepared according to the Scheme 2. Treatment of phenacyl bromides **5b** and **5c** with sodium formate followed by conversion to carbonate afforded phenacylcarbonates **6**. Treatment of H-Phe-OMe with **6** yielded *N*-protected derivatives **8**. Deprotection profiles of **8** under photoirradiation are also summarized in Table 1. Although the photolysis of **8b** or **8c** proceeded slowly as compared with that of the corresponding esters **7b** or **7c**, regenerated H-Phe-OMe was quantitatively recovered within 70 or 20 min, respectively. Mapoc carbamate **8c** also released the amine faster than **8b** did. Additionally, the Mapoc group was found to remain intact during acidic (TFA) or basic (20% piperidine in DMF) treatments commonly used in peptide chemistry. Furthermore, the half-life of the Map or Mapoc group in EtOH in normal daylight was 5.1 or 7.0 days, respectively; this shows that Map- or Mapoc-protected peptide can remain almost intact during usual experimental manipulations.

The utility of the Mapoc group in NCL was confirmed by its use in the syntheses of cyclic peptide **12** and a 32-residue human-brain natriuretic peptide (hBNP-32 **17**; Figures 1 and 2).[12] A "phototriggered intramolecular NCL" strategy was used for the preparation of **12**. As a model peptide, we chose a 12-mer peptide sequence (CSEFENEIIKQG) derived from the third extracellular loop region of the CXCR4-chemokine receptor.[13] The N^{α}-Mapoc thioester **9** was synthesized by Fmoc-based solid-phase protocols on a sulfonamide safety-catch linker.[14] The N^{α}-Mapoc group was introduced at the end cycle of the solid-phase protocol by using **6c** (5 equiv) in the presence of DIPEA (5 equiv). Completion of the protection with the Mapoc group was confirmed by a Kaiser ninhydrin test.[15] The protected resin was subjected to consecutive treatment for release of a thioester with iodoacetonitrile and with ethyl-3-mercaptopropionate to afford a fully protected thioester. TFA treatment of the protected thioester gave the N^{α}-Mapoc thioester **9**, which

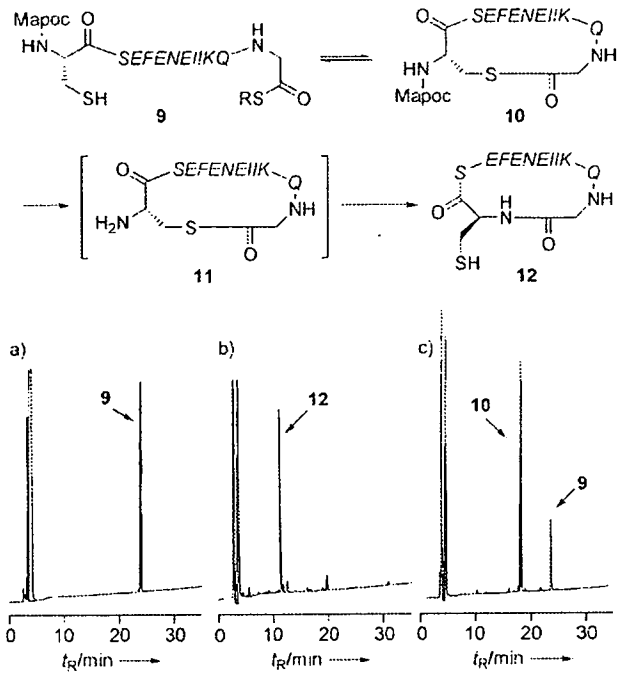


Figure 1. HPLC monitoring of phototriggered intramolecular NCL for the synthesis of cyclic peptide **12**: a) HPLC-purified **9**; b) phototriggered NCL ($t = 20$ min): **9** (0.2 mM) was photoirradiated at 25 °C in phosphate buffer (pH 7.6) containing guanidine-HCl (6 M) in the presence of 1% (*w*/*v*) sodium mercaptoethanesulfonate; c) Purified peptide **9** was dissolved in phosphate buffer (pH 7.6) containing guanidine-HCl (6 M) without photoirradiation. HPLC conditions: Cosmosil 5C_{18}ARII column (4.6×250 mm) with a linear gradient of MeCN/0.1% aq. TFA (20:80–80:20 over 50 min) at a flow rate of 1.0 mL min^{-1}, detection at 220 nm. Generally, compound **11** is thought to be a real intermediate; however, it is rapidly converted to compound **12**. Since we cannot detect compound **11** by HPLC, it is shown in brackets.

was dissolved in phosphate buffer (pH 7.6) containing 6 M guanidine hydrochloride and 1% (*w*/*v*) sodium mercaptoethanesulfonate at a concentration of 0.2 mM. Photolysis (> 300 nm, 20 min) triggered an intramolecular NCL reaction to yield the desired cyclic peptide **12** in high purity. Without UV irradiation, no ligation product was obtained and reversible thiol **9**⇌thioester **10** exchange was observed.

Being encouraged by these results, we applied Mapoc protection to a one-pot peptide-chain assembly of hBNP-32 utilizing a sequential NCL strategy followed by disulfide-bond formation with DMSO. As shown in Scheme 3, hBNP-32 has two Gly-Cys sequences, and these Cys residues are connected to each other by a disulfide bond. Therefore, we prepared three peptide fragments consisting of a C-terminal segment **13**, a middle segment **14**, and an N-terminal segment **15** by Fmoc protocols. Synthesis of thioesters **14** and **15** was achieved by trithio-ortho ester methodology, in which a Gly trithio ester was anchored through a backbone amide linker (BAL) to a solid support.[16] Application of standard Fmoc protocols efficiently afforded the peptide thioesters. For the middle fragment **14**, the completed peptide resin with an unprotected N^{α}-amino group was treated with **6c** (5 equiv) in the presence of DIPEA (5 equiv). After 12 h reaction, a Kaiser test indicated that no free N^{α}-amino group remained on the resin. Treatment of

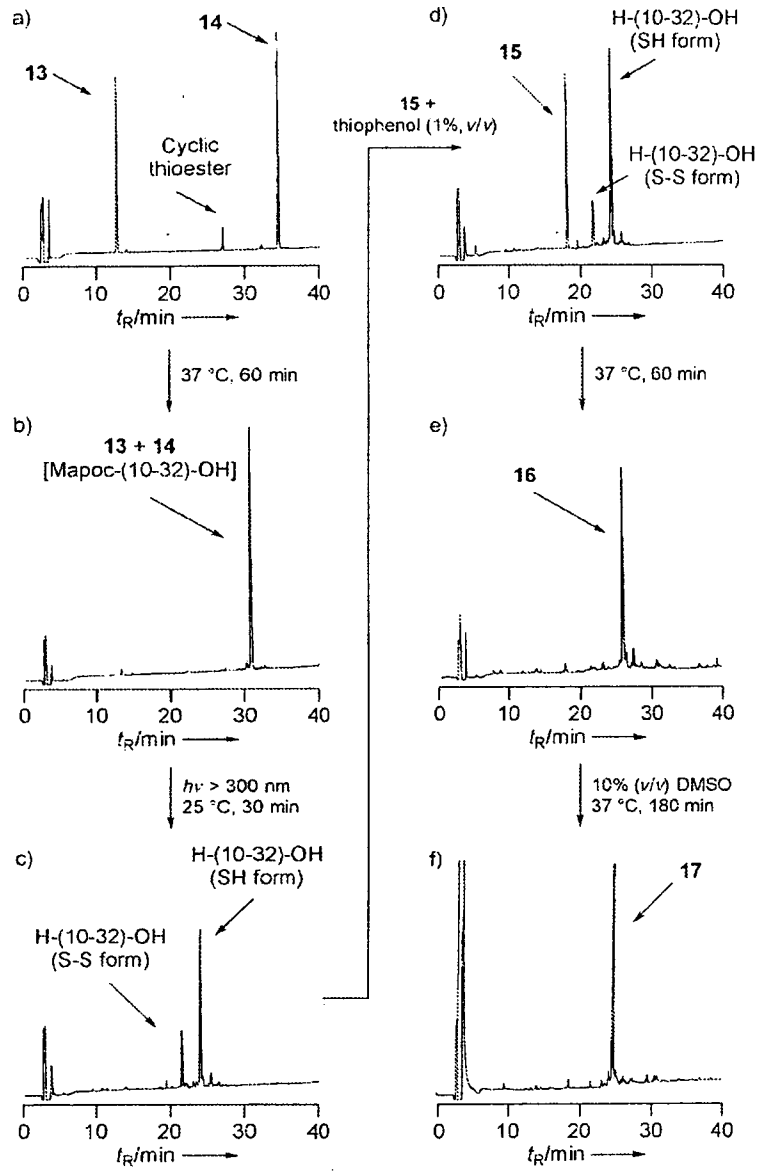
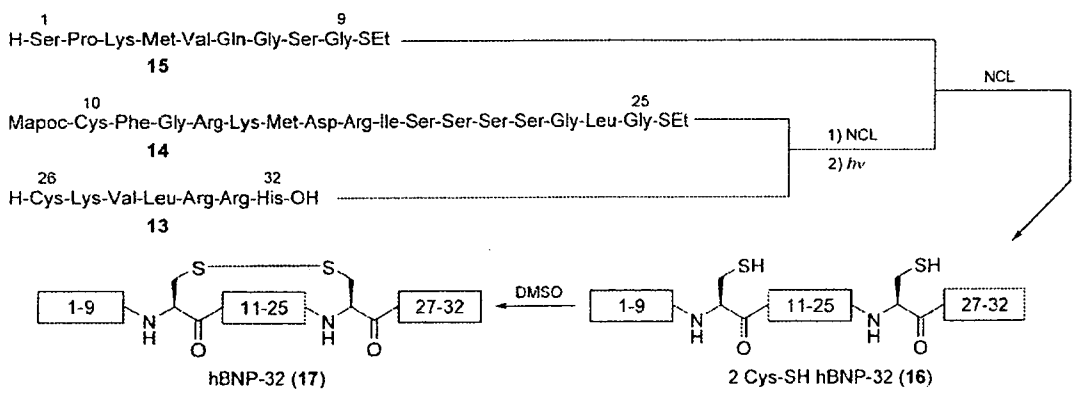


Figure 2. HPLC monitoring of one-pot sequential NCL reactions followed by disulfide-bond formation for the synthesis of hBNP-32 (**17**). a) The 1st NCL ($t = 0$ min): fragments **13** (1 mM) and **14** (1 mM) were ligated in phosphate buffer (pH 7.6) containing guanidine·HCl (6 M) in the presence of thiophenol (0.3%, *v/v*). b) The 1st NCL ($t = 60$ min). c) Photoinduced deprotection of the 1st NCL product ($t = 30$ min); the ligated product was photoirradiated ($h\nu > 300$ nm) in buffer for ligation. d) the 2nd NCL ($t = 0$ min): fragment **15** (1 mM) was coupled with the deprotected peptide in the presence of additional thiophenol (1%). e) the 2nd NCL ($t = 60$ min). f) formation of the disulfide bridge ($t = 180$ min); DMSO was added (10% *v/v*) to the ligation reaction mixture (0.33 mM peptide). HPLC conditions: Cosmosil $5C_{18}$ARII column (4.6×250 mm) with a linear gradient of MeCN/0.1% aq TFA (5:95–45:55 over 40 min) at a flow rate of 1.0 mL min^{-1}, detection at 220 nm.

the resulting resin with a TFA-based deprotection reagent yielded the N^{α}-Mapoc thioester **14**. Sequential NCL reactions were achieved in one pot without purification of the ligated intermediates (Figure 2). The initial ligation between the middle fragment **14** and the C-terminal fragment **13** was carried out in the presence of 0.3% (*v/v*) thiophenol in phosphate buffer (pH 7.6) containing 6 M guanidine hydrochloride at 37 °C for 1 h. The reaction proceeded quantitatively to yield the N^{α}-Mapoc ligated product (**13** + **14**, Figure 2a and b). Without purification, the coupled product was subjected to photoirradiation at 25 °C for 30 min (Figure 2c). This was followed by the second ligation with thioester **15** for 1 h in the presence of additional thiophenol (1% *v/v*) to afford the linear 2-Cys-SH hBNP-32 (**16**) with satisfactory purity (Figure 2d and e). Monitoring of the reaction progress by HPLC indicated that this sequence of reactions went to completion without any significant accompanying side reactions. The peptide solution was diluted threefold with phosphate buffer, followed by the addition of DMSO (10% *v/v*) to yield the disulfide-bridged hBNP-32 (**17**, Figure 2f).[17] HPLC purification of the crude material afforded purified **17** in 56% yield as calcu-



Scheme 3. Synthetic scheme for the preparation of hBNP-32 (**17**) by utilizing a one-pot sequential NCL followed by disulfide-bond formation with DMSO.

lated from peptide fragments employed in the sequential NCL reactions.

In this communication, we have presented novel 4-(dimethylamino)phenacyl-type photolabile protecting groups (Map and Mapoc) and demonstrated their unequivocal utility in peptide synthesis. In particular, the Mapoc unit represents a photosensitive protection group for amines that exhibits adequate stability against the acidic (TFA) and basic (20% piperidine/DMF) reagents commonly used in peptide chemistry. When used in conjunction with NCL reactions, the Mapoc group can be used for phototriggered intramolecular cyclizations. One-pot sequential NCL reactions can also be readily achieved. It is noteworthy that the one-pot synthesis presented herein could be conducted without either purification of ligated intermediates or readjustment of ligation conditions, such as pH. Finally, we believe the use of photolabile 4-(dimethylamino)phenacyl-type protection provides an efficient protocol for the synthesis of Cys-containing peptides/proteins and represents an alternative to Kent's thiazolidine-mediated approach. Extension of our methods to the preparation of proteins will be presented in due course.

Acknowledgements

We thank Dr. Terrence R. Burke, Jr., NCI, NIH, USA, for proofreading this manuscript. This research was supported in part by the 21st Century COE Program "Knowledge Information Infrastructure for Genome Science"; a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan; the Japan Society for the Promotion of Science (JSPS); and the Japan Health Science Foundation. S.U. is grateful for Research Fellowships from the JSPS for Young Scientists.

Keywords: native chemical ligation · peptides · proteins · synthesis design · thioesters

[1] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779; b) T. W. Muir, P. E. Dawson, S. B. H. Kent, *Methods Enzymol.* **1997**, *289*, 266–298; c) P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960; d) J. P. Tam, J. Xu, K. D. Eom, *Biopolymers*, **2001**, *60*, 194–205; e) S. B. H. Kent, *Curr. Opin. Biotechnol.* **2004**, *15*, 607–614.

[2] Examples of sequential NCL reactions: a) T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073; b) D. Bang, N. Chopra, S. B. H. Kent, *J. Am. Chem. Soc.* **2004**, *126*, 1377–1383.

[3] Example of tandem NCL reactions: K. D. Eom, Z. Miao, J.-L. Yang, J. P. Tam, *J. Am. Chem. Soc.* **2003**, *125*, 73–82.

[4] J. P. Tam, Y. I. Lu, *Protein Sci.* **1998**, *7*, 1583–1592.

[5] a) D. Bang, S. B. H. Kent, *Angew. Chem.* **2004**, *116*, 2588–2592; *Angew. Chem. Int. Ed.* **2004**, *43*, 2534–2538; b) D. Bang, S. B. H. Kent, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5014–5019.

[6] C. G. Bochet, *J. Chem. Soc. Perkin Trans. 1* **2002**, 125–142.

[7] S. P. A. Fodor, J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu, D. Solas, *Science* **1991**, *251*, 767–773.

[8] J. C. Sheehan, K. Umezawa, *J. Org. Chem.* **1973**, *38*, 3771–3774.

[9] For trimethoxyphenacyl groups as orthogonal safety-catch protection, see: A. Shaginian, M. Patel, M. Li, S. T. Flickinger, C. Kim, F. Cerrina, P. J. Belshaw, *J. Am. Chem. Soc.* **2004**, *126*, 16704–16705.

[10] R. S. Givens, J. F. W. Weber, P. G. Conrad II, G. Orosz, S. L. Donahue, S. A. Thayer, *J. Am. Chem. Soc.* **2000**, *122*, 2687–2697.

[11] G. Church, J.-M. Ferland, J. Gauthier, *Tetrahedron Lett.* **1989**, *30*, 1901–1904.

[12] Y. Kambayashi, K. Nakao, M. Mukoyama, Y. Saito, Y. Ogawa, S. Shiono, K. Inouye, N. Yoshida, H. Imura, *FEBS Lett.* **1990**, *259*, 341–345.

[13] a) J. O. Trent, Z. X. Wang, J. L. Murray, W. Shao, H. Tamamura, N. Fujii, S. C. Peiper, *J. Biol. Chem.* **2003**, *278*, 47136–47144; b) A. Otaka, S. Ueda, K. Tomita, Y. Yano, H. Tamamura, K. Matsuzaki, N. Fujii, *Chem. Commun.* **2004**, 1722–1723.

[14] a) R. Ingenito, E. Bianchi, D. Fattori, A. Pessi, *J. Am. Chem. Soc.* **1999**, *121*, 11369–11374; b) Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi, *J. Am. Chem. Soc.* **1999**, *121*, 11684–11689.

[15] E. Kaiser, R. L. Colescott, C. D. Bassinger, P. I. Cook, *Anal. Biochem.* **1970**, *34*, 595–598.

[16] J. Brask, F. Albericio, K. J. Jensen, *Org. Lett.* **2003**, *5*, 2951–2953.

[17] a) A. Otaka, T. Koide, A. Shide, N. Fujii, *Tetrahedron Lett.* **1991**, *32*, 1223–1226; b) J. P. Tam, C. Wu, W. Liu, J. Zhang, *J. Am. Chem. Soc.* **1991**, *113*, 6657–6662.

Received: June 29, 2005
Published online on October 5, 2005

European Urology

European Urology 48 (2005) 1025–1030

Translational Research - From Lab to Clinic

New HIV-Drug Inhibits In Vitro Bladder Cancer Migration and Invasion

M. Retz[a], S.S. Sidhu[a,*], J. Lehmann[b], H. Tamamura[c], N. Fujii[c], C. Basbaum[a]

[a]Biomolecular Sciences Program, Cardiovascular Research Institute, Department of Anatomy, UCSF, San Francisco, CA 94143-0452, USA
[b]Department of Urology and Pediatric Urology, Saarland University, Homburg/Saar, Germany
[c]Department of Bio-organic Medicinal Chemistry, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto, Japan

Received 5 May 2005; accepted 20 July 2005

Available online 15 August 2005

Abstract

Objective: The CXCR4/CXCL12 axis appears crucial in the metastasis of bladder cancer. Our aim was to evaluate the potency of the CXCR4 antagonist, 4F-benzoyl-TE14011 (4F-bTE), as an anti-metastatic drug in this disease. In this study, we assessed the ability of 4F-bTE to inhibit tumor cell motility, invasion through extracellular matrix (ECM), matrix metalloproteinase (MMP) secretion and cytoskeletal responses to chemokine.

Methods: To assess the degree to which cells could migrate and invade ECM under various conditions, we used TCCSUP bladder cancer cells in a Boyden chamber system. To monitor actin polymerization, we stained cells on chamber slides with AlexaFluor 594 phalloidin. To measure matrix-metalloproteinase-2 and -9 (MMP) activity, we used gelatin zymography. To assess the effects of the CXCR4 antagonist 4F-bTE on each of the above parameters, we exposed bladder cancer cells either to chemokine CXCL12, alone, or to both CXCL12 and 4F-bTE. We also monitored cells for apoptotic and necrotic changes during drug treatment.

Results: The CXCR4 antagonist 4F-bTE markedly decreased CXCL12-induced bladder cancer cell migration and ECM invasion in Boyden chamber assays. The antagonist also blocked chemokine-induced actin polymerization as well as the induction of MMP-2 and MMP-9 in these cells.

Conclusion: The CXCR4 antagonist 4F-bTE has the potential to inhibit expression of the metastatic phenotype and may provide therapeutic value to patients.

Keywords: Chemokines; Bladder cancer; Migration; Invasion; CXCR4 antagonist; 4F-benzoyl-TE14011; MMP

1. Introduction

Pelvic lymph nodes are the primary sites for bladder cancer metastasis, followed by widespread metastasis to lung, liver and bone marrow [1]. Chemokines and their receptors have been shown to play pivotal roles in organ-specific metastasis [2]. They also control angiogenesis, B-cell lymphopoiesis and myelopoiesis [3,4]. The interaction of soluble chemokines with their specific, transmembrane G-protein-coupled receptors mediates their biological effects. To date, 18 chemokine receptors (CCR1-10, CXCR1-6, XCR1 and CX_3CR1) have been identified in humans [5]. Recent evidence indicates that tumor cells express distinct, tumor-specific, non-random patterns of chemokine receptors. Intriguingly, the specific chemokine ligands for these receptors exhibit peak expression in lymph nodes, lung, liver and bone marrow, the most common sites of cancer metastasis [2].

Previously, we identified CXCR4 as the only chemokine receptor in bladder tumor cells whose expression levels correlated with tumor progression [6]. Our

* Corresponding author. Present address: University of California, San Francisco, Anatomy/CVRI, 513 Parnassus Ave, HSW1330A, San Francisco, CA 4143, USA. Tel. +1 415 476 4514; Fax: +1 415 514 3933/4764845.

E-mail address: sidhu@itsa.ucsf.edu (S.S. Sidhu).

0302-2838/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.eururo.2005.07.016

studies also showed that the corresponding chemokine ligand CXCL12 had a strong chemoattractant effect on bladder cancer cells. Consistent with its role in invasive processes, the CXCR4/CXCL12 axis is also functionally linked to actin polymerization, which is responsible for lamellipodia formation, and secretion of collagenolytic matrix-metalloproteinases (MMPs), which are believed to facilitate movement through the ECM [6]. Our findings on bladder cancer cells are in agreement with other recent studies, which have shown that the CXCR4/CXCL12 axis is required for metastasis in a variety of tumor types [2,7,8]. In particular, Müller and coworkers observed that neutralizing antibodies directed against the chemokine receptor CXCR4 potently reduced the number of metastases deriving from mouse xenografts of human breast tumors [2]. Based on our results, the CXCR4/CXCL12 axis also appears to play a pivotal role in the metastasis of bladder cancer and is a potential therapeutic target.

Prior to the discovery of CXCR4's involvement in cancer metastasis, CXCR4 had been identified as an important co-receptor for the human immunodeficiency virus (HIV-1). CXCR4 is the major co-receptor for the entry of T-cell-tropic HIV-1 (X4-HIV-1) [9]. To date, several CXCR4 antagonists have been developed as HIV-entry inhibitors. Tamamura and co-workers synthesized the polyphemusin T140, a 14-residue peptide containing a disulfide bridge. Although T140 exhibited strong anti-HIV activity, it lacked serum stability. A new generation of T140 derivatives, including 4F-benzoyl-TE14011 (4F-bTE), were developed and characterized by both high anti-HIV activity and high serum-stability. The new compounds also exhibit low toxicity [10].

Our aim was to evaluate the potency of the small molecule CXCR4 antagonist 4F-bTE as an anti-metastatic drug in bladder cancer. As described below, we assessed the ability of this compound to inhibit tumor cell motility, invasion through extracellular matrix, MMP secretion and chemokine-induced cytoskeletal reorganization.

2. Materials and methods

2.1. CXCR4 antagonist

The polyphemusin II peptide derivative 4F-bTE was synthesized as described previously [10].

2.2. Cell line

The undifferentiated bladder cancer cell line TCCSUP was originally obtained from the American Type Culture Collection (Manassas, VA). TCCSUP was cultured in 75 cm^2 flasks in RPMI 1640 medium supplemented with 10% heat-inactivated fetal bovine serum (FBS), 10 U/ml penicillin and 10 mg/ml streptavidin (Sigma, St. Louis, MO).

2.3. In vitro cell migration and invasion assays

Migration and invasion assays were performed in a Boyden chamber system as described previously [6]. Briefly, chemotactic migration of bladder cancer cells (TCCSUP) was investigated using 24-well cell-culture chambers containing inserts with 8-μm pores (BD PharMingen, San Diego, CA). Invasion assay chambers also contained a reconstituted extracellular matrix membrane overlying the insert (Matrigel with 8-μm pore membrane, BD PharMingen). Cells (5×10^4 cells/500 μl) were suspended in chemotaxis buffer (DMEM/0.1% BSA/ 12 mM HEPES) and were added to the upper chamber of the assay wells. Migration and invasion assays were performed in presence of a chemokine CXCL12 gradient (R&D Systems, Minneapolis, MN) established by placing chemokine (100 ng/ml) in the lower chamber. Control wells contained buffer alone. To investigate the inhibitory effect of the CXCR4 antagonist 4F-bTE, some bladder cancer cells were pre-incubated with 100 nM 4F-bTE for 1 h at 37 °C prior to placing them in the upper chamber. All experiments were repeated three times with triplicate samples.

2.4. Apoptosis/cell death assay

Apoptotic cells and necrotic cells were quantified using the Annexin V-FITC and propidium iodide (PI) detection kit (BD PharMingen, San Diego, CA) following standard protocols. Annexin V-FITC is a marker of apoptotic cells and the entry of PI into un-permeabilized cells acts as a marker for cells undergoing necrosis. TCCSUP cells were treated with 100 nM 4F-bTE for 24 h at 37 °C. Untreated TCCSUP cells served as negative control. Some cells were incubated with 100 μM etoposide as a positive control to induce apoptosis. Cells were adjusted to a concentration of 1×10^5 cells/100 ul in PBS. Each sample was stained with 5 μl Annexin V-FITC and 5 μl propidium iodide. Cells were analysed on a FACSscan (Becton Dickinson, Mountain View, CA) and flow cytometry data were analysed using Cellquest software (Becton Dickinson, Mountain View, CA).

2.5. Actin polymerization assay

To assess the effect of 4F-bTE on actin polymerization, TCCSUP cells were seeded on Labtech-eight-well-chamber slides pre-coated with 0.01% poly-L-lysine solution (Sigma, St. Louis, MO). Cells were serum-starved overnight and then incubated with 100 ng/ml CXCL12 (R&D Systems, Minneapolis, MN) ± 100 nM 4F-bTE for 30 min at 37 °C to stimulate actin polymerization. After 30 min, cells were fixed for 2 min (4 °C) in 3% paraformaldehyde in PBS, permeabilized for 5 min with 0.2% Triton X-100, incubated with 5 mU/ml AlexaFluor 594 phalloidin (Molecular Probes, INC., Eugene, OR) for 1 h, washed with PBS and mounted in Vectashield aqueous mounting medium containing DAPI (Vector Laboratories, Burlingame, CA). Fluorescence microscopy was performed using a Nikon Eclipse, E600 microscope.

2.6. Gelatin zymography

To investigate the effects of 4F-bTE on MMP activity, TCCSUP bladder cancer cells were cultured with 100 ng/ml CXCL12 (R&D Systems, Minneapolis, MN) ± 100 nM 4F-bTE under serum-free conditions. After 24 h at 37 °C, cell supernatants were collected and samples were resolved on 10% SDS-PAGE containing 0.3% gelatin. Following electrophoresis, gels were incubated in buffer containing 2.5% Triton X-100, 50 mM Tris-HCl, 6.5 mM $CaCl_2$, 5 μM $ZnCl_2$

and 0.5 g/l NaN_3 (Sigma, St. Louis, MO) for 48 h at 37 °C. Subsequently, gels were fixed and stained with 0.1% Coomassie Brilliant Blue (Biorad, Hercules, CA) to reveal bands of enzymatic activity.

2.7. Statistical analysis

Since most variables did not comply with normal Gaussian distribution as determined by the Kolmogoroff-Smirnoff test, a non-parametric Mann-Whitney-*U*-test or Kruskal-Wallis-test was applied in order to compare independent samples of continuous variables [11–14]. All *p*-Values were based on two-sided tests and the threshold to accept statistical significance was set at the alpha level 0.05. Analyses were performed with the statistical software package SPSS version 10.0 (SPSS Inc., Chicago, IL).

3. Results

3.1. Effect of 4F-bTE on CXCL12-induced migration of bladder cancer cells

Chemotaxis of the bladder cancer cell line TCCSUP was examined along a CXCL12 chemokine gradient using 8 μm pore membranes in a modified Boyden chamber system. Optimal chemotactic response to CXCL12 was observed at a concentration of 100 ng/ml, which induced a 2.5 fold increase ($p = 0.0006$) in migration after 24 h compared to unstimulated cells. To determine whether there was an inhibitory effect of the CXCR4 antagonist, some TCCSUP cells were pre-incubated with 4F-bTE at various concentrations (1 h, 37 °C). 4F-bTE inhibited CXCL12-induced chemotaxis in a dose-dependent manner. At a concentration of 100 nM, the compound reduced CXCL12-induced migration to background levels ($p = 0.0059$; Fig. 1).

3.2. Effect of 4F-bTE on CXCL12-induced invasion

Invasion through ECM was assayed in Boyden chambers overlaid with reconstituted extracellular matrix (Matrigel). The presence of 100 ng/ml CXCL12 in the lower chamber increased the number of invading cells significantly over 24 h ($p = 0.0037$). The ability of the CXCR4 antagonist to inhibit invasion was evident from experiments in which we pre-treated the cells with various concentrations of 4F-bTE (1 h, 37 °C). Invasion was completely blocked by 4F-bTE at 100 nM ($p = 0.0074$; Fig. 2).

3.3. Low cytotoxicity of 4F-bTE

4F-bTE is thought to specifically bind to the CXCR4 receptor. In order to rule out that drug induced inhibition of tumor motility and invasion was due to general toxicity, we performed an annexin V and propidium iodide assay. Annexin staining is an early sign of apoptosis. Propidium iodide staining in non-permeabilized cells is a reflection of ongoing necrosis. As shown in Fig. 3, no differences were observed between

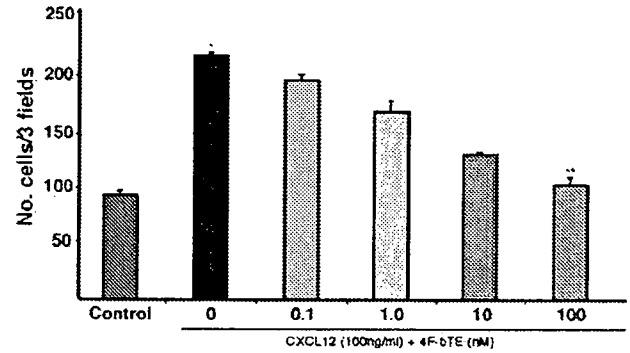


Fig. 1. Effect of CXCR4 antagonist on CXCL12-stimulated migration. Chemotaxis of TCCSUP bladder cancer cells was investigated using Boyden chambers with 8 μm pore membrane inserts. Chemokine CXCL12 (100 ng/ml) was added to the lower chamber. Cell suspensions (5×10^4 TCCSUP cells/500 μl) were added to the upper chamber. Bladder cancer cells were pre-incubated (1 h, 37 °C) with various concentration of 4F-bTE (0.1 nM–100 nM). The CXCL12-stimulated migration was inhibited by 4F-bTE and completely eradicated at a concentration of 100 nM. * indicates a significant difference between control and CXCL12 ($p = 0.0006$) and ** indicates a significant difference between CXCL12 and 100 nM 4F-bTE + CXCL12 ($p = 0.0059$).

untreated TCCSUP cells and cells treated with 100 nM 4F-bTE suggesting low cytotoxicity (Fig. 3). As a positive control for the apoptosis/cell death assay, TCCSUP cells were treated with 100 μM etoposide for 24 h (data not shown) (Table 1).

3.4. Effect of 4F-bTE on CXCL12-induced actin polymerization

Extensive cytoskeletal reorganization involving actin polymerization is a prerequisite for migration

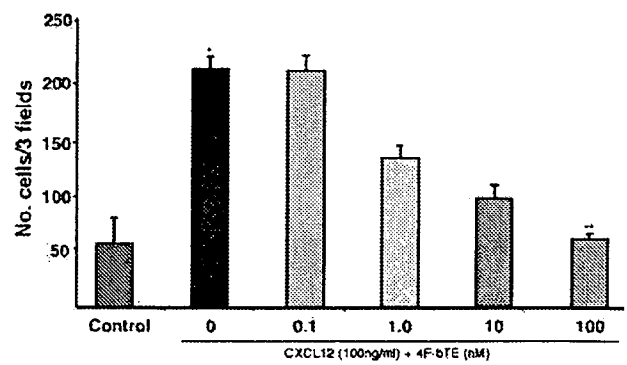


Fig. 2. Effect of CXCR4 antagonist on CXCL12-stimulated invasion. As in Fig. 1 except the transwell insert was coated with ECM proteins (Matrigel) prior to the assay. CXCL12 (100 ng/ml) was placed in the lower chamber of the transwell plate. Bladder cancer cells were preincubated (1 h, 37 °C) with various concentrations of 4F-bTE (0.1 nM–100 nM). Invasion through the ECM membrane was completely inhibited by 100 nM 4F-bTE. * indicates a significant difference between control and CXCL12 ($p = 0.0037$) and ** indicates a significant difference between CXCL12 and 100 nM 4F-bTE plus CXCL12 ($p = 0.0074$).

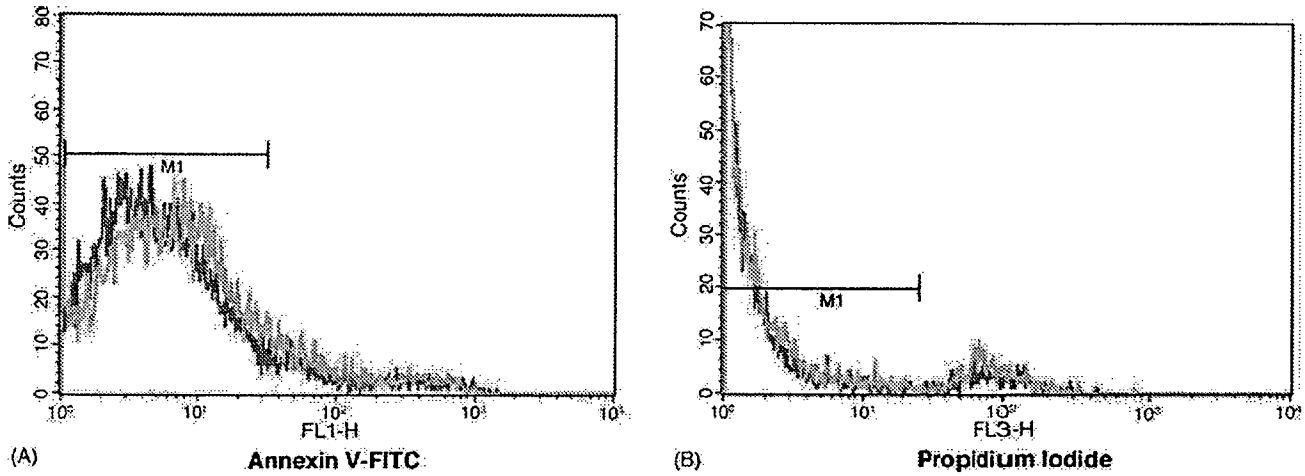


Fig. 3. Low cytotoxicity of 4F-bTE. FACS analysis of annexin-V (A) and propidium iodide (B) staining of TCCSUP cells, untreated or treated with 4F-bTE (100 nM) for 24 h. No significant differences were observed between untreated and 4F-bTE treated cells indicating low cytotoxicity. Key: red - untreated TCCSUP cells; green - 4F-bTE.

Table 1
Percentage of cells staining for annexin-V or propidium iodide

Treatment	Concentration	Annexin-V (positive stained cells %)	Propidium iodide (positive stained cells %)
Untreated cells	–	11.28 ± 6.5	2.74 ± 1.28
4F-benzoyl-TE14011	100 nM	14.63 ± 7.8	4.49 ± 1.39

No statistically significant differences were observed between untreated and 4F-benzoyl-TE14011 treated cells indicating low cytotoxicity.

of tumor cells [15]. AlexaFluor 594 phalloidin staining of CXCL12-treated tumor cells showed a distinct cytoskeletal redistribution of F-actin stress fibers and lamellipodia formation, beginning 10 min after CXCL12 exposure. Treatment of bladder cancer cells with 100 nM 4F-bTE (30 min, 37 °C) inhibited CXCL12-induced actin polymerization and lamellipodia formation (Fig. 4).

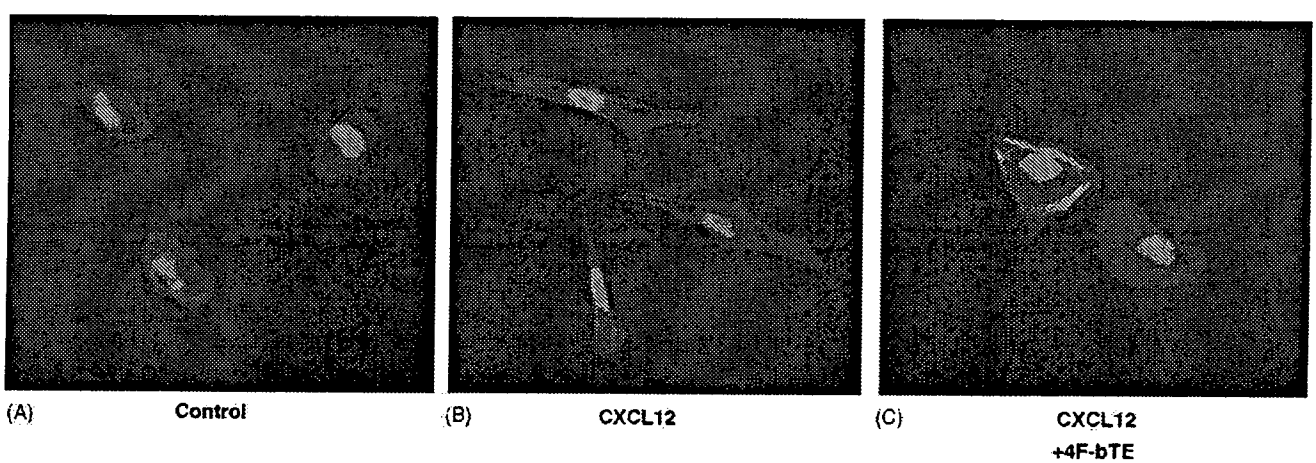


Fig. 4. Effect of CXCR4 antagonist on CXCL12-induced actin polymerization. Untreated TCCSUP bladder tumor cells contained few cytoplasmatic stress fibers and no obvious lamellipodia formation (A). TCCSUP bladder tumor cells were stimulated with 100 ng/ml CXCL12. Actin cytoskeleton was visualized by Alexa Fluor 594 phalloidin and fluorescence microscopy. CXCL12-stimulated cells showed a distinct cytoskeletal redistribution of F-actin stress fibers and lamellipodia formation 30 min after exposure to CXCL-12 (B). Co-incubation of bladder cancer cells with 100 nM 4F-bTE during CXCL12 exposure inhibited the CXCL12-induced actin redistribution and lamellipodia formation (C).

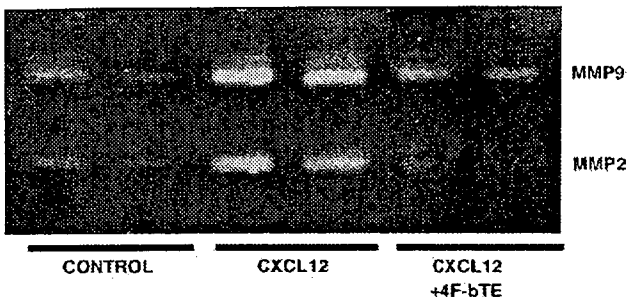


Fig. 5. Effect of CXCR4 antagonist on CXCL12-induced secretion of matrix metalloproteinases. TCCSUP bladder cancer cells were untreated or stimulated with 100 ng/ml CXCL12 for 24 hours. Analysis carried out by gelatin zymography revealed that unstimulated bladder cancer cells expressed low levels of MMP-2 and MMP-9. Notably, a significant increase in the production of MMP-2 and MMP-9 in bladder tumor cells occurred upon stimulation with CXCL12. 24 h co-treatment of TCCSUP cells with 100 nM 4F-bTE resulted in an inhibition of the CXCL12-induced MMP-2 and MMP-9 secretion.

3.5. Effect of 4F-bTE on CXCL12-induced expression of MMP-2 and MMP-9

MMP's are involved in extracellular matrix degradation, which facilitates tumor cell invasion [16]. Unstimulated TCCSUP bladder cancer cells expressed little MMP-2 and MMP-9. In contrast, stimulation with CXCL12 (24 h, 37 °C) significantly increased expression of both MMPs, which were released into the cell medium. Addition of 4F-bTE (24 h, 37 °C) at 100 nM blocked the CXCL12-induced release of MMP-2 and MMP-9 to background levels (Fig. 5).

4. Discussion

Tumor metastasis is not uniquely pathological, as it incorporates mechanisms of normal cell migration such as leukocyte trafficking and homing using the chemokine system [17]. Earlier, we identified CXCR4 as the only chemokine receptor whose expression level in bladder cancer correlates with tumor progression [6]. The chemokine ligand CXCL12, which has specificity for the CXCR4 receptor, exhibits peak expression in lymph nodes, lung, liver and bone marrow, the most common sites of bladder cancer metastasis [2]. In this study, we investigated the ability of 4F-bTE to modulate CXCL12-mediated migration, invasion, cytoskeletal changes and MMP expression in bladder cancer cells.

Results showed that 4F-bTE impaired CXCL12-induced chemotaxis and extracellular matrix invasion of bladder cancer cells in a dose-dependent manner. Similar blocking effects have been reported for pancreas and breast cancer cell lines using T140 analogs TN14003 and TC14012 [19,20]. Our studies additionally showed that treatment of bladder cancer cells with 100 nM 4F-bTE blocked CXCL12-induced actin polymerization and lamellipodia formation. In agreement with our findings, Mori et al., showed inhibition of CXCL12-induced actin polymerization by the CXCR4 antagonist TN14003 in pancreatic cancer [19]. Our studies also showed that 4F-bTE was able to inhibit CXCL12-induced secretion of MMP-2 and MMP-9, findings that are in agreement with the role of the CXCR4/CXCL12 axis in breast cancer and rhabdomyosarcoma cells [16,18].

An important problem in previous studies was the high cytotoxicity of the CXCR4 antagonist T140, which correlated with the number of positive charges on the peptide. The new generation of T140 analogs were modified by substituting the positively charged amino acids, arginine and lysine with the nonbasic polar amino acid L-citrulline, resulting in a reduction of total positive charges [10]. To investigate the possible toxicity of 4F-bTE in our experimental system, we performed an annexin V and propidium iodide assay. This revealed no apoptotic or necrotic effects of the drug.

The *in vitro* effects of 4F-bTE in blocking CXCL12-induced motility, actin polymerization and MMP activity in bladder cancer cells are promising, but *in vivo* studies are needed to support the potential utility of this type of drug in cancer patients. It is interesting in this regard that *in vivo* experiments using breast cancer cells in SCID mice has already shown that 4F-benzoyl-TN14003 could significantly reduce pulmonary tumor metastasis [20].

5. Conclusions

The CXCR4/CXCL12 axis appears to be a crucial feature in the metastasis of bladder cancer. In vitro, the CXCR4 antagonist 4F-bTE blocks CXCL12-induced migration, invasion, actin polymerization and MMP activity of bladder cancer cells. While further Phase I clinical trials are required to confirm the efficacy of 4F-bTE, our results suggest that 4F-bTE may play an important role as an anti-metastatic drug in bladder cancer and represent a novel therapeutic strategy.

Acknowledgments

This work was supported by the Reinhard-Nagel Foundation, Germany.

References

[1] Stein JP, Lieskovsky G, Cote R, Groshen S, Feng AC, Boyd S, et al. Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1, 054 patients. J Clin Oncol 2001;19:666.

[2] Muller A, Homey B, Soto H, Ge N, Catron D, Buchanan ME, et al. Involvement of chemokine receptors in breast cancer metastasis. Nature 2001;410:50.

[3] Nagasawa T, Hirota S, Tachibana K, Takakura N, Nishikawa S, Kitamura, et al. Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1. Nature 1996;382:635.

[4] Bernardini G, Ribatti D, Spinetti G, Morbidelli L, Ziche M, Santoni A, et al. Analysis of the role of chemokines in angiogenesis. J Immunol Methods 2003;273:83.

[5] Zlotnik A, Yoshie O. Chemokines: a new classification system and their role in immunity. Immunity 2000;12:121.

[6] Retz M, Sidhu SS, Blaveri E, Kerr SC, Dolganov GM, Lehmann J, et al. CXCR4 expression reflects tumor progression and regulates motility of bladder cancer cells. Int J Cancer 2005;114:182.

[7] Taichman RS, Cooper C, Keller ET, Pienta KJ, Taichman NS, McCauley LK. Use of the stromal cell-derived factor-1/CXCR4 pathway in prostate cancer metastasis to bone. Cancer Res 2002;62:1832.

[8] Burger M, Glodek A, Hartmann T, Schmitt-Graff A, Silberstein LE, Fujii N, et al. Functional expression of CXCR4 (CD184) on small-cell lung cancer cells mediates migration, integrin activation, and adhesion to stromal cells. Oncogene 2003;22:8093.

[9] Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science 1996;272:872.

[10] Tamamura H, Omagari A, Hiramatsu K, Gotoh K, Kanamoto T, Xu Y, et al. Development of specific CXCR4 inhibitors possessing high selectivity indexes as well as complete stability in serum based on an anti-HIV peptide T140. Bioorg Med Chem Lett 2001;11:1897.

[11] Kolmogoroff A. Confidence limits for an unknown distribution function. Ann Math Stat 1941;12:461.

[12] Smirnoff N. Tables for estimating the goodness of fit of empirical distributions. Ann Math Stat 1948;19:279.

[13] Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Statist 1947;18:50.

[14] Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. J Amer Stat Ass 1953;47:583.

[15] Friedl P, Wolf K. Tumour-cell invasion and migration: diversity and escape mechanisms. Nat Rev Cancer 2003;3:362.

[16] Fernandis AZ, Prasad A, Band H, Klosel R, Ganju RK. Regulation of CXCR4-mediated chemotaxis and chemoinvasion of breast cancer cells. Oncogene 2004;23:157.

[17] Rossi D, Zlotnik A. The biology of chemokines and their receptors. Annu Rev Immunol 2000;18:217.

[18] Libura J, Drukala J, Majka M, Tomescu O, Navenot JM, Kucia M, et al. CXCR4-SDF-1 signaling is active in rhabdomyosarcoma cells and regulates locomotion, chemotaxis, and adhesion. Blood 2002;100:2597.

[19] Mori T, Doi R, Koizumi M, Toyoda E, Ito D, Kami K, et al. CXCR4 antagonist inhibits stromal cell-derived factor 1-induced migration and invasion of human pancreatic cancer. Mol Cancer Ther 2004; 3:29.

[20] Tamamura H, Hori A, Kanzaki N, Hiramatsu K, Mizumoto M, Nakashima H, et al. T140 analogs as CXCR4 antagonists identified as anti-metastatic agents in the treatment of breast cancer. FEBS Lett 2003;550:79.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Antiviral Research 66 (2005) 119–128
Antiviral Research
www.elsevier.com/locate/antiviral

Highly potent anti-HIV-1 activity isolated from fermented *Polygonum tinctorium* Aiton

Yu Zhong [a], Yoshiyuki Yoshinaka [b], Tadahiro Takeda [c], Noriko Shimizu [c], Sayaka Yoshizaki [d], Yoshio Inagaki [a], Shinobu Matsuda [e], Gisho Honda [f], Nobutaka Fujii [f], Naoki Yamamoto [a,d,*]

[a] *Department of Molecular Virology, Bio-Response, Graduate School, Tokyo Medical and Dental University, 1-5-45, Yushima, Bunkyo-ku, Tokyo 113-8519, Japan*
[b] *Human Gene Sciences Center, Tokyo Medical and Dental University, Bunkyo-ku, Tokyo 113-8519, Japan*
[c] *Kyoritsu University of Pharmacy, Minato-ku, 105-8512, Tokyo, Japan*
[d] *National Institute of Infectious Diseases, Shinjuku-ku, Tokyo 162-8640, Japan*
[e] *Institute of Hemorheological Function of Food Co. Ltd, Hyogo, Japan*
[f] *Graduate School of Pharmaceutical Sciences, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan*

Received 13 September 2004; accepted 9 February 2005

Abstract

A water-soluble extract of fermented *Polygonum tinctorium* Aiton (Polygonaceae) called *Sukumo*, exhibited a potent inhibitory activity against HIV type 1 in vitro. The extract potently suppressed acute HIV-1 (III_B) infection in MT-4 cells with EC_{50} values of 0.5 μg/ml but exhibited low cytotoxicity to MT-4 cells even at a high concentration (CC_{50} > 1000 μg/ml). It also inhibited giant cell formation in co-cultures of HIV-infected cells and uninfected Molt-4 cells. *Sukumo* extract was found to interact with both the viral envelope glycoprotein and cellular receptors, thus blocking virus-cell binding and virus-induced syncytium formation. There was a good correlation between the extract's anti-HIV-1 activity and its inhibitory effects on HIV-1 binding. It also suppressed replication of herpes simplex virus type 1 in Vero cells with an EC_{50} of 11.56 μg/ml. On the other hand, there was no appreciable activity against influenza A virus, poliovirus or SARS corona virus when tested at concentrations ranging from 3.2–400 μg/ml as shown by microscopic image analysis for cytopathic effect (CPE). Physico-chemical studies revealed that the anti-HIV activity in the extract was essentially maintained after boiling at 100 °C in 1N HCl or 1N NaOH, and after treatment with 100 mM $NaIO_4$. The inhibitory activity of the extract was also not reduced after pronase digestion. The active factor in the extract is likely to be a novel compound(s) having a polyanionic substructure and a molecular weight of 10,000–50,000.

Keywords: *Polygonum tinctorium*; *Sukumo* extract; HIV-1; HSV-1; Viral entry

1. Introduction

One of the logical targets of the viral life cycle at which to inhibit HIV-1 replication is the step in the process where the infectious virion enters its host cell (Moore and Stevenson, 2000; Lin et al., 2002). Therefore, the identification of HIV entry inhibitors, which can serve as novel anti-HIV drugs, is urgently needed.

Retroviral infection is initiated by the attachment of the virion to the cell surface, which even occurs before glycoproteins on the viral envelope interact with specific receptors on the host cell to trigger fusion. A great variety of polyanionic compounds have been described which act as virus adsorption inhibitors. This class of compounds also comprises the cosalane analogues, containing the polycarboxylate pharmacophore, as well as the sulfated polysaccharides extracted from sea algae (Nakashima et al., 1987a, 1992; Santhosh et al., 2001; Witvrouw and De Clercq, 1997). All of these compounds are assumed to exert their anti-HIV activity by shielding the positively charged sites in the V3 loop region of

* Corresponding author. Tel.: +81 3 5803 5178; fax: +81 3 5803 0124.
E-mail address: yamamoto.mmb@tmd.ac.jp (N. Yamamoto).

0166-3542/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2005.02.003